



HAL
open science

Phenotype plasticity and populations' dynamics : social interactions among cancer cells

Marie André-Ratsimbazafy

► **To cite this version:**

Marie André-Ratsimbazafy. Phenotype plasticity and populations' dynamics : social interactions among cancer cells. Cancer. Université Sorbonne Paris Cité, 2016. English. NNT : 2016USPCB032 . tel-01589235

HAL Id: tel-01589235

<https://theses.hal.science/tel-01589235>

Submitted on 18 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITÉ PARIS DESCARTES
INSTITUT CURIE

Doctoral School **Frontiers in Life Sciences**

University Department **UMR168**

Thesis defended by **Marie** ANDRÉ-RATSIMBAZAFY

Defended on **20th June, 2016**

In order to become Doctor from Université Paris Descartes and from Institut Curie

Academic Field **Interdisciplinary Biology**

Thesis Title

**Phenotype plasticity and
populations' dynamics
Social interactions among cancer cells**

Thesis supervised by François AMBLARD

Committee members

<i>Referees</i>	Emmanuelle CHARAFFE-JAUFFRET	Professor at Centre de Recherche en Cancérologie de Marseille
	Joelle SOBZAK-THEPOT	Professor at Université Pierre et Marie Curie
<i>Examiners</i>	Michel MORANGE	Professor at Centre Cavallès
	Antoine TESNIÈRE	Professor at AP-HP
	Pierre SAVAGNER	Senior Researcher at Centre de Recherche en Cancérologie de Montpellier
	Fabrice LAVIAL	Junior Researcher at Centre de Recherche en Cancérologie de Lyon
<i>Guest</i>	Marc-Henry STERN	Senior Researcher at Institut Curie
<i>Supervisor</i>	François AMBLARD	Senior Researcher at Institut Curie



UNIVERSITÉ PARIS DESCARTES
INSTITUT CURIE

Doctoral School **Frontiers in Life Sciences**

University Department **UMR168**

Thesis defended by **Marie** ANDRÉ-RATSIMBAZAFY

Defended on **20th June, 2016**

In order to become Doctor from Université Paris Descartes and from Institut Curie

Academic Field **Interdisciplinary Biology**

Thesis Title

**Phenotype plasticity and
populations' dynamics
Social interactions among cancer cells**

Thesis supervised by François AMBLARD

Committee members

<i>Referees</i>	Emmanuelle CHARAFFE-JAUFFRET	Professor at Centre de Recherche en Cancérologie de Marseille
	Joelle SOBZAK-THEPOT	Professor at Université Pierre et Marie Curie
<i>Examiners</i>	Michel MORANGE	Professor at Centre Cavallès
	Antoine TESNIÈRE	Professor at AP-HP
	Pierre SAVAGNER	Senior Researcher at Centre de Recherche en Cancérologie de Montpellier
	Fabrice LAVIAL	Junior Researcher at Centre de Recherche en Cancérologie de Lyon
<i>Guest</i>	Marc-Henry STERN	Senior Researcher at Institut Curie
<i>Supervisor</i>	François AMBLARD	Senior Researcher at Institut Curie



UNIVERSITÉ PARIS DESCARTES
INSTITUT CURIE

École doctorale **Frontiers in Life Sciences**

Unité de recherche **UMR168**

Thèse présentée par **Marie** ANDRÉ-RATSIMBAZAFY

Soutenue le **20 juin, 2016**

En vue de l'obtention du grade de docteur de l'Université Paris Descartes et de l'Institut Curie

Discipline **Biology Interdisciplinaire**

Titre de la thèse

**Plasticité des phénotypes et
dynamique des populations
Interactions sociales entre cellules cancéreuses**

Thèse dirigée par François AMBLARD

Composition du jury

<i>Rapporteurs</i>	Emmanuelle CHARAFFE-JAUFFRET	professeur au Centre de Recherche en Cancérologie de Marseille
	Joelle SOBZAK-THEPOT	professeur à l'Université Pierre et Marie Curie
<i>Examineurs</i>	Michel MORANGE	professeur au Centre Cavailles
	Antoine TESNIÈRE	professeur à l'AP-HP
	Pierre SAVAGNER	directeur de recherche au Centre de Recherche en Cancérologie de Montpellier
	Fabrice LAVIAL	chargé de recherche au Centre de Recherche en Cancérologie de Lyon
<i>Invité</i>	Marc-Henry STERN	directeur de recherche à l'Institut Curie
<i>Directeur de thèse</i>	François AMBLARD	directeur de recherche à l'Institut Curie

*Je dédie cette thèse à mon Mari Mamy,
à Maman et Papa, Neneny et Dada, Elise, Hery et Volana,
à ma Famille, Blanche
à Anne-Marie, Pépé et Grand-Mère.*

Acknowledgments

I would like to thank people who contributed to my PhD and encouraged me during my PhD educational time at the Curie Institute.

Firstly, I am grateful towards Dr François Amblard, my PhD tutor, who welcomed me in his team and who allowed me to work on this original project.

I would like to thank Dr Marina Glukhova, Dr Marc-Henry Stern, Dr Gordon Tucker et al. (Dr Olivier Nosjean, Pr. Antoine Tesnière, Dr Isabelle Bonnet, Dr Sylvie Coscoy, Dr Stéphanie Mangenot), who make up my Thesis Committee supervisors, for the time they dedicated to my work, giving me the opportunity many times to discuss about my ideas and my work.

I am very grateful towards Pr. Emmanuelle Charaffe-Jauffret and Pr. Joelle Sobzak-Thepot, my referees, who took time to read and give me feedbacks on my thesis. And I thank Pr. Michel Morange, Pr. Antoine Tesnière, Dr Pierre Savagner, Dr Fabrice Laval and Dr Marc-Henry Stern as being examiners of my work.

I thank Fanny Tabarin-Cayrac, François Xavier Gobert, Sylvie Coscoy, Isabelle Bonnet, Vassily Gurchenkov, François Waharte, Nicolas Carpi, the cytometry platform (Zosia, Annick and Sophie), the lentiviral platform (a very special thanks to François-Xavier and members of the UMR932 team for the derogative accesses on weekends and very late evenings, music session on P3 and joyful and encouraging discussion and laughs!), the Nikon Imaging Center, the molecular/antibody platform (Ahmed and Aurore for your precious advices, enzymes and grateful help to get commercial free samples), the cell culture room people, for your patience, your practical and theoretical advice and for the fantastic time we spend together. I will not forget these long evenings spent searching for antibodies with Mario and Stephanie. A great thank you to Mario, Adeline, Laurence and Andres who loaned to me for long hours bench and cooled centrifuges; thanks to you I gained so much time in pipetting procedures and spent very good moments of laughs and private jokes.

I would also like to present my joyful thanks to the “third floor” people and my dear Master and (actual and past) PhD students and postdoctoral research fellow (Svitlana, Hannah, Mathieu, Jean, Marie, Guillaume, Simon and Randa, Majdoline, Sandrine, Rodrigo, Adrien, Marine, Bastien, Fabrice and Chiara, Renaud, Fahima, John, Anne-Christine...). By your presence you have helped me to go through those strenuous years. I appreciated working in this pleasant and joyful environment. I thank you for accepting me, with my questions, with my nature and with my odd jokes. I wish you all the best and don't forget to follow your dreams and ideas!

I will wish to take time to say how much pleasant it was to shortly but intensely collaborate through various projects with people I am going to name below. With Fanny we spent long hours

in the dark to take pictures of my small cells under the watchful eye of the confocal microscope. André Nicolas and Dr Didier Meseure were in charge of a daunting task consisting in producing in spades immuno-staining of dozen Tissue MicroArray (TMA) and then we spent couple late mornings debating about percentage and how strong was the staining. With Zofia, Annick and Sophie, “patience” was the keyword of our relationship and it was necessary to be composed in order to sort all these cells and even more to get a decent number of CD24-/CD44- cells. Nathalie Brandon and Elodie Manie gave me valuable time and feedback about my “precious” cell lines, indeed thanks to both of them I am plenty sure that the cells which have been used to perform the experiments are my precious SUM149-PT and SUM159-PT cells. With the help of Isabelle, Timo, Marie, Mathieu, Delphine and Mamy we developed few computing programs to facilitate my data analysis. Thanks to you I examined MATLAB™, Python and other programming languages. I am not an ace regarding these stuffs, nonetheless with your support I got a tiny bit improvement.

I will not forget to thank Dr Svitlana Havrylenko, Mamy Ratsimbazafy and Dr Sylvie Coscoy who took time (many times!) to read this report and gave me precious feedbacks. Thank to “Croco” I had a very close look to my data in order to be exhaustive and consistent regarding the experimental data and the conclusions. In parallel, Svitlana and Mamy were tracking against any English language mistakes that I could have written, thank to them this thesis report is much easier to decode. Without you Mamy, this thesis written through L^AT_EX would have failed to take life; thank to your resourcefulness you solved countless numbers of red problems that were blocking the thesis compilation, and programmed small macros that increased my productivity. Thank you for all those hours spent alongside me in the cafés of the Boulevard Saint-Germain in encouraging my thesis delivery.

I thank the entire team of my PhD School Frontiers in Life Science and its staff, Dr François Taddei, Dr Samuel Bottani and Dr Antoine Tesnières for believing in me and my career plan, for your understanding of my dreams and my mind, your precious help in all the administrative procedures, and all your support along these past 4 years.

I can’t forget the important and large backing from the Servier foundation with its PhD fellowship. I would like to thank Dr Olivier Nosjean, Dr Emmanuel Canet et al. for their support throughout my thesis and their challenging discussions on December around the “annual roundtable”.

Finally, I have to mention the important role of the “Ecole de l’INSERM-Liliane Bettencourt” in my MD-PhD academic education and career. I thank Pr Jean-Claude Chottard et al. for their implication, their dynamism. Thank you for having relying upon me since the beginning. I don’t forget the serious help of Anne-Marie Laffaye, Nathalie François and Sow Barrow. I also thank the Liliane Bettencourt-Schueler foundation for its past and future fellowships.

I would like to thank my friends and my large family more precisely “Papa and Maman” (for their assistance in any circumstances), “Dada and Néneny” and Elise my sister, Volana and Hery my (new!) sister and brother, without forgetting my Husband who always believed in me and supported me in my choices. I am totally aware that backing me is not a piece of cake!

**– I could not have carried out this Thesis without all of you,
and I just want to say you “Thank you”! –**

Contents

	vii
Acknowledgments	ix
Contents	xi
Abstract	xix
List of Tables	xxi
List of Figures	xxiii
Acronyms	xxix
Acronyms	xxix
Preface	1
I State of the art	3
1 Between individuality and collectivity, notion of phenotypes.	5
1.1 Organism to cells, a question of scale and interactions	6
1.2 Cell phenotypes, the race for the characterization	7
1.2.1 Various methods to characterize cell phenotypic specificities	7
1.2.2 Cell phenotypes are part of a process	10
2 Key concepts that define cancer and cancer cells	13
2.1 Comparisons between healthy and cancerous cells, similarities and differences .	13
2.1.1 Self-sufficiency to growth signals	14
2.1.2 Evading apoptosis	14
2.1.3 Evading growth suppressors	15
2.1.4 Replicative immortality	15
2.1.5 Genome instability and cumulative mutations	16
2.1.6 Deregulated cellular energetics	16
2.1.7 Inducing neo-angiogenesis	16
2.1.8 Avoiding immune destruction	16
2.1.9 Tumor promoting inflammation	17
2.1.10 Interaction with tumor stroma	17
2.1.11 Tissue invasion and metastasis	17

2.1.12	Stem cells and cell-to-cell communication	17
2.2	The origin of tumor and cancer, a key question.	18
2.2.1	The Intrinsically Disordered Proteins Theory (IDPT)	18
2.2.2	The Somatic Mutation Theory (SMT) and Cancer Stem Cells	20
2.2.3	The Tissue Organization Field Theory (TOFT)	22
2.2.4	Breast tumors, diversity and clusterization	23
2.3	Why are we interested in previous questions?	25
2.3.1	Diagnostic and prognostic to improve patient follow-up	25
2.3.2	From conventional to targeted therapy strategies	26
3	The issue of social relationships among cells	31
3.1	From stimuli to answer, adaptation to the changing environment	31
3.1.1	Cells are sensitive to diverse source of stimuli	32
3.1.2	How is communication established between two cells?	35
3.2	Communication is unflinching of the spatial (self)-organization	37
4	Thesis work schedule and limits	41
4.1	As a rephrasing of the general question	41
4.1.1	Diversity of cells phenotypes within tumors and diversification	41
4.1.2	Keywords of this thesis	42
4.2	How is organized the thesis statement	42
4.3	Little tricks for your convenience	43
4.3.1	Supplementary data, the Ali Baba cavern	43
4.3.2	Apply a color touch that makes sense	43
4.3.3	Hyperlink as ways to the good direction	44
II	Natural history of cells diversity	45
		47
5	Cell line heterogeneity.	49
5.1	SUM149-PT and SUM159-PT characterization for CD24 and CD44 markers	49
5.1.1	Working assumption relative to cells heterogeneity	49
5.1.2	CD24-/CD44-, an undescribed viable cells phenotype	51
5.2	Methodology of flow cytometry data recording and statistical analysis	53
5.2.1	Data acquisition and selection of cells of interest	53
5.2.2	Flow cytometry qualitative representation	53
5.2.3	Calculation of measurement parameters to characterize the fluorescence of the cells of interest	53
5.2.4	Mean fluorescence normalization and data plotting	54
5.3	Follow-up of CD24 and CD44 expression fine-variations within SUM149-PT and SUM159-PT cell lines	55
5.3.1	Experimental workflow to track the overtime expression of CD24 and CD44 markers	57
5.3.2	SUM149-PT cell line and its fluctuations in CD24 and CD44 expression	57
5.3.3	SUM159-PT cell line and its fluctuations in CD24 and CD44 expression	63
5.4	Cell line populations are continuously fluctuating	68
5.4.1	A few remarks with respect to measured fluctuations	68

5.4.2	Contributions of this study	68
5.4.3	On the way to the next chapter	69
6	Cells are plastic. However do they all share the same capabilities?	71
6.1	Plastic abilities of subpopulations	71
6.1.1	Properties of the experiment's implementation	72
6.1.2	In less than three days, subphenotypes have already exchanged to other phenotypic states	75
6.2	Repopulation processes of CD24-/CD44+ and CD24+/CD44+ subpopulations	76
6.2.1	Repopulation dynamics of CD24-/CD44+ and CD24+/CD44+ issued from SUM149-PT	76
6.2.2	Repopulation dynamics of CD24-/CD44+ and CD24+/CD44+ isolated from SUM59-PT cell line	84
6.2.3	As a conclusion of CD24-/CD44+ and CD24+/CD44 steady state	92
6.3	On the way of CD24-/CD44- re-equilibrium	92
6.3.1	Looking for steady state from CD24-/CD44- subpopulation derived from SUM149-PT cell line	92
6.3.2	Repopulation of CD24-/CD44- cells isolated from SUM159-PT cell line	95
6.4	The equilibrium status of the re-populating subpopulations	99
7	Looking for population steady state, influence of cells phenotypic genealogy	103
7.1	Memory of the origin over cells generation, over phenotypes, over time on cell fate decisions	103
7.1.1	Experiment workflow to study cells memory	103
7.1.2	Influence of P1 on P2's future - Case of SUM149-PT	106
7.1.3	Influence of P1 on P2 future - Study on SUM159-PT cell line	110
7.2	Influence of the origin through cells generation, over phenotypes, over time on cell fate decisions	115
III	Clonogenicity, Diversity and Environment influences	117
		119
8	Characterization of clonogenic capabilities of biological models	121
8.1	Description of the experimental methodology followed for <i>in vitro</i> tumorspheres-forming assays	121
8.1.1	Design of the tumorspheres-forming assay for T1 and T2	122
8.1.2	Assessing tumorsphere-forming assay efficiency, characterizing tumorsphere size and phenotypic heterogeneity	122
8.2	Influence of quantity and SUM149-PT and SUM159-PT cell line tumorigenicity	124
8.2.1	Clonogenicity of SUM149-PT and SUM159-PT cell line	124
8.2.2	Influence of the quantity of plated cells on cell proliferation	126
8.3	The CD24-/CD44- subphenotype survives but does not form tumorsphere in 3D	128
8.3.1	The CD24-/CD44+ and CD24+/CD44+ subphenotypes	128
8.3.2	The subphenotype CD24-/CD44-	130

9 Influence of 3D spatial relations on heterogeneity and plastic cell capabilities	131
9.1 Differential expression of phenotypes in 2D and 3D	131
9.2 Influence of the quantity of plated cells on phenotypes	135
9.3 Phenotypic plasticity in 3D culture conditions	137
9.3.1 The methodology that has been followed	137
9.3.2 CD24-/CD44+ and CD24+/CD44+ and re-diversification in 3D . . .	137
9.3.3 CD24-/CD44+ et CD24+/CD44+ subpopulations have different proliferative capacity	139
IV Strategies proposal to search social rules that govern cell phenotypic changes	141
	143
10 Long term follow-up at the single cells level of cell state transitions	145
10.1 Some criteria to model cell state transitions	145
10.1.1 On the way of self-renewing phenotypes	145
10.1.2 Rhythm of cells division and phenotypic transitions	146
10.1.3 Influence of memory on the probabilities to transit from one phenotype to another	148
10.1.4 Expression of phenotypes and spatial organization	148
10.2 The characteristics of the molecular tool	148
10.2.1 Phenotypes fluorescent reporter, a combination of two expressing endogenous promoters	150
10.2.2 A reporter of molecular platform delivery	151
11 Validation of individual functional units forming the PhenotypeTracker molecular platform	153
11.1 Δ LNGFR as reporter of the lentiviral transduction of target cells	153
11.1.1 Δ LNGFR gene is poorly endogenously expressed within SUM149-PT and SUM159-PT cell lines	154
11.1.2 Sur-expression of Δ LNGFR transgene within HEK293T cells	155
11.1.3 Promoters, dependencies of over-expression of transgenes in SUM149-PT and SUM159-PT cell lines	156
11.2 Cell localization of transgenes encoded through the PhenotypeTracker molecular platform	158
11.3 Transgene expression bias within SUM149-PT and SUM159-PT subphenotypes	161
11.4 Validation of some human driving genes promoters	165
11.4.1 Regulatory sequences of human CD44 gene	165
11.4.2 Regulatory sequences of human CD24 gene	166
12 Order or disorder? Resulting of a spatial dependence of cells	169
12.1 Statistical methodology to determine cells autonomy or cells dependence	170
12.1.1 If cells are autonomous	173
12.1.2 If cells are dependent of the others	173
12.2 A first approach thanks to homogeneous systems	173
12.2.1 Method that assesses a local spatial organization within cells phenotypes	173

12.2.2	Preliminary results regarding the expression of the phenotypes depending on spatial localization	177
12.3	Conceptualizing and modeling of interactions among cells	181
12.3.1	The basic features of cellular automaton	181
12.3.2	Features and simulation of cell states transitions	183
12.3.3	Looking for spatial correlation, influence of cells coupling and referential	187
General conclusion		193
Main contributions of this PhD thesis		193
Characterization of the heterogeneity of SUM149-PT and SUM159-PT cell lines regarding CD24 and CD44 markers		193
Phenotypes plasticity and recovery of heterogeneity from homogeneous populations		194
Singular tumorigenic abilities of cell lines subphenotypes		195
Monitoring cell-state transition thanks to the lentiviral “PhenotypeTracker” platform		196
Methodology to measure the level of reliance regarding the neighborhood		196
Research topics that we are currently exploring as a result of this thesis work		198
Whole transcriptomic analysis of sorted CD24/CD44 subpopulations and over-time phenotypic re-equilibrium of cell populations		198
Real time follow-up of cells phenotype transitions		200
Modeling phenotypes plasticity and dynamic steady-state		201
Future prospects for this research project		202
Do cancer cells spatially organize themselves?		203
On the way of other cell lines which are phenotypically heterogeneous		203
V Appendix		205
A Supplementary data		207
A.1 Additional data for part II to part V		207
A.1.1 Supplementary data associated to part “natural history of cells diversity”		207
A.1.2 Supplementary data associated to part “strategies proposal to search social rules that govern cell phenotypic changes”		225
A.2 Looking for heterogeneity and screening of breast cancer cell lines.		227
A.2.1 Breast cancer cell line screening for KRT5/KRT6 markers		229
A.2.2 Breast cancer cell line screening for KRT8/KRT18 markers		229
A.2.3 Breast cancer cell line screening for KRT14 marker		230
A.2.4 Breast cancer cell line screening for CD44 marker		231
B Methodologies developed to segment cells edges from confocal images		233
B.1 Automatic or manual cells segmentation, principle of image segmentation		234
B.2 Segmentation of cells nuclei		235
B.3 Segmentation of cells edges		236
B.3.1 Thresholding methodology		236
B.3.2 Filtering methodology		236
B.3.3 Spreading methodology		237
B.3.4 Deviation around the mean methodology		237
B.3.5 Correction of background lack of homogeneity methodology		238
B.4 Future perspective of cells segmentation		238

B.4.1	Image homogenization	239
B.4.2	Background subtraction	239
B.4.3	Edges detection and thresholding	239
C	Validation of SUM149-PT and SUM159-PT cell lines, an epic adventure	241
C.1	Screening for SUM149-PT and SUM159-PT safety, batch number 1	242
C.2	Authentication of SUM149-PT and SUM159-PT, batch number 2	243
D	Protocols developed and used throughout the experimental work	245
D.1	Protocol – Cell culture – HEK293T	245
D.1.1	Thawing of frozen cells	246
D.1.2	Culture and subculture of the cells	247
D.1.3	Freezing and cryopreservation of cells, DMSO 10%	248
D.2	Protocol – Cell culture – SUM149 PT and SUM159 PT 5%	249
D.2.1	Thawing of frozen cells	250
D.2.2	Culture and subculture of the cells	250
D.2.3	Freezing and cryopreservation of cells, DMSO 10%	251
D.3	Protocol – Cloning – Bacteria Transformation DG1, Stellar, or TOP10	253
D.3.1	Transformation procedure	253
D.3.2	Rapid transformation procedure	254
D.4	Protocol Cloning Blunt Quick Blunting	255
D.4.1	Quick Blunting reaction	255
D.5	Protocol Cloning Dephosphorylation Antarctic Phosphatase	256
D.5.1	Dephosphorylating DNA with Antarctic Phosphatase	256
D.6	Protocol Cloning Dephosphorylation Alkaline Phosphatase Calf Intestinal	257
D.7	Protocol Cloning Gel & PCR Purification	258
D.7.1	DNA purification from gel slice	258
D.8	Protocol Cloning Ligation In-Fusion	259
D.9	Protocol Cloning Ligation LR Clonase II	260
D.9.1	LR cloning reaction	260
D.10	Protocol Cloning Ligation Quick Ligase	261
D.10.1	Quick ligation protocol	261
D.11	Protocol Cloning Oligos Annealing	262
D.11.1	Annealing reaction	262
D.12	Protocol Cloning Oligos Resuspension	263
D.12.1	Resuspension procedure	263
D.13	Protocol Cloning Plasmid DNA Purification Maxiprep	264
D.13.1	Low-copy plasmid purification	264
D.13.2	High-copy plasmid purification	265
D.14	Protocol Cloning Plasmid DNA Purification Midiprep	267
D.14.1	Low-copy plasmid purification	267
D.14.2	High-copy plasmid purification	268
D.15	Protocol Cloning Plasmid DNA Purification Miniprep	270
D.15.1	Low-copy plasmid purification	270
D.15.2	High-copy plasmid purification	271
D.16	Protocol Cloning Sequencing BigDye® 1.1 and 3.1	272
D.16.1	Preparation of the PCR mix	272
D.16.2	PCR protocol	272
D.16.3	Administrative procedure	272

D.16.4 Analysis procedure	273
D.17 Protocol FACS Cells Flow Cytometry – Hoechst 33258 or DAPI Staining . . .	274
D.17.1 Preparation of cells for flow cytometry	275
D.17.2 Direct immunofluorescence staining of cell surface antigens for flow cytometry	275
D.17.3 Indirect immunofluorescence staining of cell surface antigens for flow cytometry	276
D.17.4 Compensation particles settings for flow cytometry	276
D.18 Protocol FACS Cells Sorting Antibodies Staining	277
D.18.1 Preparation of cells for flow cytometry	277
D.18.2 Direct immunofluorescence staining of cells for flow cytometry	278
D.18.3 Indirect immunofluorescence staining of cells for flow cytometry	278
D.18.4 Compensation particles settings for flow cytometry	279
D.19 Protocol Immunofluorescence of Fixed Cells using PFA	280
D.20 Protocol of tumorspheres assays	283
E Design strategies developed to construct the molecular platforms	285
E.1 Design and construction of #B3013#	285
E.2 Design and construction of #B3014#	288
E.2.1 Construction #Bb0015#	288
E.2.2 Construction #B3014#	290
E.3 Design and construction of #B3039#	291
E.3.1 Construction #B3039#	292
E.4 Design and construction of #B3043#	295
E.4.1 Construction #I0041#	296
E.4.2 Construction #B3043#	296
E.5 Design and construction of #B3047#	299
E.6 Design and construction of #B3051#	304
E.6.1 Construction #B3051#	305
E.7 Design and construction of #B3056#	306
E.7.1 Blunting procedure of #Bb0039-I1#	306
F Various tables	309
F.1 Listing of molecular constructs	309
F.2 Listing of oligos	312
F.3 Listing of antibodies	315
G Programs used to analyze experimental data	317
G.1 Code to read.lsm images - Matlab	317
G.2 Code to compute fluorescence intensity of single cells - Matlab	320
G.3 Code to perform dataset from flow cytometry analysis data - Matlab	323
G.4 Code to calculate MFI, CV and SD from flow cytometry analysis based on the CD44 marker threshold - Matlab	325
G.5 Code to consolidate daily data into a single Excel file, formatted for GraphPad's graphs - Python 3	327
Bibliography	331

PHENOTYPE PLASTICITY AND POPULATIONS' DYNAMICS

Social interactions among cancer cells**Abstract**

It is commonly accepted that tumors arise from cells that escape the homeostatic controls which underlie the healthy histological structure and that cell phenotype is not the result of deterministic biochemical and genetic processes, but rather the stochastic and dynamic outcome of multiple intra- and intercellular regulation networks.

This PhD aims to quantitatively study the phenotypic homeostasis of the cell populations and to present an approach to the fundamental question, never heretofore studied, regarding the autonomy versus collective control of cell fate.

We studied in the long run, using flow cytometry and in 2D and 3D conditions, the level of expression of CD24 and CD44 of two breast cancer cell lines (SUM149-PT and SUM159-PT). Three phenotypes were isolated (CD24-/CD44+, CD24+/CD44+, CD24-/CD44-), the latter had not previously been documented in the literature. The phenotypic behavior of CD44-low and CD44-high subpopulations has been characterized by assessing their proportion and analyzing the fluorescence map. Thereby, we observed both a periodic behavior of appearance and disappearance of pool of cells characteristics of each cell lines and a phenotypic re-diversification for each subpopulation. Only the resulting population derived from CD24-/CD44- provided the same balance as the original unsorted population. 3D re-diversification process was observed in tumorspheres from CD24-/CD44+ and CD24+/CD44+. The cells CD24-/CD44- did not have that potential but nonetheless outlived anoikis. These behaviors suggest that there is an inter-cell coordination regulating the balance of phenotypic proportions. To discover the social rules regulating inter-phenotypic spatial organization, we have set up a reporter of the endogenous variations of CD24 and CD44 and developed a theoretical model of cell interactions.

This work has confirmed our hypothesis that inter-cellular social rules are determining the phenotypic expression at both the uni- and multicellular scales.

Keywords: tumor and cell line heterogeneity, cells' phenotype characterization, cell plasticity, stem cells, cell fate decision, cell autonomy *versus* collective decision

PLASTICITÉ DES PHÉNOTYPES ET DYNAMIQUE DES POPULATIONS

Interactions sociales entre cellules cancéreuses**Résumé**

On admet communément que les tumeurs proviennent de cellules échappant aux contrôles homéostatiques qui sous-tendent les structures histologiques saines et que le phénotype d'une cellule n'est pas le résultat de processus génétiques et biochimiques déterministes mais la conséquence stochastique de réseaux de régulation intra- et intercellulaires.

Ce doctorat vise à étudier quantitativement l'homéostasie phénotypique de populations cellulaires et à présenter une approche à la question fondamentale, mais jusqu'alors jamais étudiée, concernant l'autonomie versus le contrôle collectif du devenir des cellules.

Nous avons étudié sur le long terme, par cytométrie de flux et dans des conditions 2D puis 3D, le niveau d'expression de CD24 et CD44 de deux lignées cellulaires de cancer du sein (SUM149-PT et SUM159-PT). Trois phénotypes ont été isolés (CD24-/CD44+, CD24+/CD44+, CD24-/CD44-), ce dernier n'avait pour le moment pas été documenté dans la littérature. Le comportement phénotypique des sous-populations CD44-low et CD44-high a été caractérisé en évaluant leur proportion et en analysant leur spectre de fluorescence. Ainsi nous avons observé des comportements périodiques d'apparition et de disparition de pool de cellules caractéristiques des lignées et une re-diversification des phénotypes pour chacune des sous-population. Seule la population issue de CD24-/CD44- re-diversifiée présente le même équilibre que la population initiale non triée. En 3D, le processus de re-diversification a été observé dans les tumorsphères issues de CD24-/CD44+ et CD24+/CD44+. Les cellules CD24-/CD44- n'ont pas ce potentiel mais survivent néanmoins à l'anoïkis. Ces comportements laissent penser qu'il existe une coordination inter-cellulaire régulant l'équilibre des proportions phénotypiques. Pour découvrir les règles sociales régissant l'organisation spatiale inter-phénotypique, nous avons mis en place un rapporteur des variations du niveau d'expression endogène des marqueurs d'intérêt et élaboré un modèle théorique d'interactions cellulaires.

Ce travail a conforté notre hypothèse selon laquelle il s'établit des règles sociales inter-cellulaires déterminant l'expression phénotypique à l'échelle uni- et pluricellulaire.

Mots clés : hétérogénéité cellulaire et tumorale, caractérisation phénotypique cellulaire, plasticité cellulaire, cellules souches, prise de décision cellulaire, autonomie cellulaire *versus* prise de décision collective

List of Tables

A.1	Data list of cell lines and repopulating cells samples whose transcriptomic analysis was performed. - Part I	223
A.2	Data list of cell lines and repopulating cells samples whose transcriptomic analysis was performed. - Part II	224
A.3	Data list of breast cancer cell lines screened through immunostaining procedures to semi-quantitatively study the level of expression of 6 markers of interest . . .	228
F.1	Data list of molecular constructs. - Part I	310
F.2	Data list of molecular constructs. - Part II	311
F.3	Data list of oligos. - Part I	313
F.4	Data list of oligos. - Part II	314
F.5	Data list of antibodies	316

List of Figures

5.1	Working assumptions with respect to SUM149-PT and SUM159-PT tumor cell lines heterogeneity.	50
5.2	Experiment workflow to characterize CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ subphenotypes	52
5.3	Experiment workflow to monitor CD24 and CD44 expression of SUM149-PT and SUM159-PT cell lines.	56
5.4	Long-term monitoring of changes in SUM149-PT cell line heterogeneity. . . .	59
5.5	Long-term changes monitoring of the proportion of CD44 negative cells among viable SUM149-PT cells.	60
5.6	Long-term monitoring of CD24/CD44 expression profile of SUM149-PT cell line plated into long term 2D culture conditions.	61
5.7	Correlation of CD24 and CD44 expression within SUM149-PT cell line	62
5.8	Long-term monitoring of changes in SUM159-PT cell line heterogeneity. . . .	64
5.9	Long-term changes monitoring of the proportion of CD44 negative cells among viable SUM159-PT cells.	65
5.10	Long-term monitoring of CD24/CD44 expression profile of SUM159-PT cell line plated into long term 2D culture conditions.	66
5.11	Correlation of CD24 and CD44 expression within SUM159-PT cell line	67
6.1	Working questioning with respect to cell lines subphenotypes plasticity	72
6.2	Experiment workflow setup to characterize overtime SUM149-PT and SUM159-PT cell lines subphenotypes plasticity abilities.	74
6.3	Long-term monitoring of the repopulation of CD24-/CD44+ SUM149-PT cells subphenotypes.	77
6.4	Long-term monitoring of CD44 negative expressing cells among viable repopulating SUM149-PT cells population.	78
6.5	Long-term monitoring of CD24 and CD44 expression among viable repopulating CD24-/CD44+ SUM149-PT cells and comparison with unstained and double-stained non-sorted SUM149-PT cells.	79
6.6	Long-term monitoring of CD24+/CD44+ SUM149-PT cells subphenotypes repopulation	81
6.7	Long-term monitoring of CD44 negative cells among viable repopulating CD24+/CD44+ SUM149-PT cells population.	82
6.8	Long-term monitoring of CD24 and CD44 expression among viable repopulating CD24+/CD44+ SUM149-PT cells and comparison with unstained and double-stained non-sorted SUM149-PT cells.	83
6.9	Long-term monitoring of CD24-/CD44+ SUM159-PT cells subphenotypes repopulation.	85

6.10	Long-term monitoring of CD44 negative expressing cells re-expression among viable repopulating sorted CD24-CD44+ SUM159-PT cells population.	86
6.11	Long-term monitoring of CD24 and CD44 expression among viable repopulating sorted CD24-/CD44+ SUM159-PT cells population and comparison with unstained and double-stained non-sorted SUM159-PT cells.	87
6.12	Long-term monitoring of CD24+/CD44+ SUM159-PT cells subphenotypes repopulation.	89
6.13	Long-term monitoring of CD44 negative expressing cells re-expression among viable repopulating sorted CD24+/CD44+ SUM159-PT cells population.	90
6.14	Long-term monitoring of CD24 and CD44 expression among viable repopulating sorted CD24+/CD44+ SUM159-PT cells population and comparison with unstained and double-stained non-sorted SUM-159PT cells.	91
6.15	Long-term monitoring of the repopulation of SUM149-PT cells CD24-/CD44-subphenotype.	93
6.16	Long-term monitoring of CD44 negative cells among viable repopulating initially sorted from CD24-/CD44- SUM149-PT cells.	93
6.17	Long-term monitoring of CD24 and CD44 expression among viable repopulating sorted SUM149-PT cells population and comparison with unstained and double-stained non-sorted SUM149-PT cells.	94
6.18	Long-term monitoring of the repopulation from SUM159-PT cells CD24-/CD44- subphenotypes.	96
6.19	Long-term monitoring of the re-expression of CD44 positive expressing cells among viable re-populating CD24-/CD44- SUM159-PT cells population.	97
6.20	Long-term monitoring of CD24 and CD44 expression among viable repopulating CD24-/CD44- subpopulation and comparison with unstained and double-stained non-sorted SUM159-PT cells.	98
6.21	Long-term tracking of CD24/CD44 expression profile by sorted subphenotypes and unsorted SUM149-PT cells.	100
6.22	Long-term tracking of CD24/CD44 expression profile by sorted subphenotypes and unsorted SUM159-PT cells.	101
7.1	Experimental workflow to monitor SUM149-PT and SUM159-PT cell lines subphenotypes secondary repopulating capacities.	105
7.2	Long-term monitoring of CD44 positive expressing cells re-expression among viable repopulating secondary sorted SUM149-PT subphenotypes cells.	106
7.3	Long-term monitoring of the repopulation of secondary sorted SUM149-PT subphenotypes cells repopulation.	107
7.4	Long-term monitoring of CD24 and CD44 expression among viable secondary repopulating sorted SUM149-PT subphenotypes cells populations and comparison with unstained and double-stained non-sorted SUM149-PT cells.	108
7.5	Long-term monitoring of CD44 positive expressing cells re-expression among viable repopulating secondary sorted SUM159-PT subphenotypes cells.	111
7.6	Long-term monitoring of secondary sorted SUM159-PT subphenotypes cells repopulation.	112
7.7	Long-term monitoring of CD24 and CD44 expression among viable secondary repopulating SUM159-PT sorted subphenotypes cells populations and comparison with unstained and double-stained non-sorted SUM159-PT cells.	113
8.1	Tumorspheres assay workflow.	123

8.2	Measurements to characterize tumorigenic abilities of SUM149-PT and SUM159-PT cell lines and sorted subphenotypes following tumorsphere-forming assays.	124
8.3	Clonogenic abilities of SUM149-PT and SUM159-PT non sorted cells.	125
8.4	Comparison of clonogenic abilities between first and secondary re-plated SUM149-PT and SUM159-PT cell lines after respectively 7 and 14 days in 3D culture conditions	127
8.5	Clonogenic abilities of first and secondary re-plating of sorted SUM149-PT and SUM159-PT cells subphenotypes after respectively 7 and 14 days in 3D culture conditions	129
9.1	Workflow of the markers characterization of tumorspheres assay.	132
9.2	Flow cytometry data analysis gating strategy workflow to characterize tumorspheres expression for CD24 and CD44 markers.	133
9.3	Influence of culture conditions (2D and 3D culture conditions) on CD24 and CD44 expression profile of SUM149-PT and SUM159-PT cells	134
9.4	Comparison of the CD24/CD44 expression profile of cells issued from either first or secondary re-plating of SUM149-PT and SUM159-PT cells after respectively 7 and 14 days in 3D culture conditions	136
9.5	Comparison of the CD24 and CD44 expression profile of cells issued from either first or secondary re-plated sorted SUM149-PT and SUM159-PT cells subphenotypes after respectively 7 and 14 days in 3D culture conditions	138
10.1	CD24 and CD44 phenotypes space within which cells evolve and transit from one phenotype to another one.	146
10.2	Overview of working assumptions with respect to the search for correlations regarding the expression of cell phenotypes over time.	147
10.3	Schematic representation of a real-time phenotype change tracking at the single cell level	149
10.4	Schematic illustration of the PhenotypeTracker lentiviral platform.	150
11.1	Validation of the non endogenous expression of LNGFR protein within SUM149-PT and SUM159-PT cell lines.	154
11.2	Choice of promoter driving Delta Truncated form of Low affinity of Nervous Growth Factor Receptor (Δ LNGFR) reporter whose expression is stable regardless of phenotypes' expression over time.	157
11.3	Confocal image analysis strategy developed by Isabelle Bonnet (UPMC, Institut Curie) to measure transduced single cell LNGFR expression level.	158
11.4	Validation of NLS-Fluorescent protein expression on HEK293T cell line.	160
11.5	Monitoring cell sorting procedures of Δ LNGFR expression within CD24 and CD44 phenotypes of SUM149-PT cell line transduced with CD44 expression PhenotypeTracker platform.	163
11.6	Monitoring the cell sorting procedures of Δ LNGFR expression pattern within SUM149-PT cell line transduced with CD44 expression PhenotypeTracker platform.	164
11.7	Fluorescence microscopy of SUM159-PT cell line after 72h of CD44 expression PhenotypeTracker platform	166
12.1	Questioning about spatial correlations between cells fate decisions and cells phenotypes patterning.	171

12.2	Variance of the distribution of the “average phenotype” calculated per ROI and in function of cells degree of dependence	172
12.3	Analytical method implemented to demonstrate a bias with respect to CD24 and CD44 expression cells distribution between the whole cell population and geographically defined areas	173
12.4	Experiment workflow of tumor cells spatial organization.	175
12.5	Foreign population by nationality group by district of Paris	176
12.6	Spatial organization assessment of repopulating sorted SUM159-PT subphenotypes after 5 days of culture.	179
12.7	Confocal image analysis workflow developed by Isabelle Bonnet (UPMC, Institut Curie) and Romain Cendre (CIC-IT Nancy) to detect single cells edges.	180
12.8	Interconnections between theoretical and experimental approaches to examine the spatial interactions among various cell phenotypes.	181
12.9	Automaton modeling the behavior of 3 cells phenotypes and cells decision-taking as an interacting system.	183
12.10	Study of the change of variance regime depending on the scale of analysis.	188
12.11	Simulations monitoring the expression pattern of two phenotypes (0 and 1) which phenotypic expression may or may not be coupled to previously expressed phenotypes.	189
12.12	Simulations monitoring the expression pattern of two phenotypes (0 and 1) which phenotypic expression may or may not be coupled to previously expressed phenotypes.	190
A.1	FSC and SSC characterization of SUM149-PT and SUM159-PT sorted subphenotypes.	208
A.2	Density plating of cells sorted subphenotypes per assay	209
A.3	Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM149-PT CD24-/CD44- subphenotype cells.	210
A.4	Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM149-PT CD24-/CD44- subphenotype cells.	211
A.5	Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM149-PT CD24-/CD44+ subphenotype cells.	212
A.6	Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM149-PT CD24-/CD44+ subphenotype cells.	213
A.7	Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM149-PT CD24+/CD44+ subphenotype cells.	214
A.8	Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM149-PT CD24+/CD44+ subphenotype cells.	215
A.9	Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM159-PT CD24-/CD44- subphenotype cells.	216
A.10	Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM159-PT CD24-/CD44- subphenotype cells.	217
A.11	Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM159-PT CD24-/CD44+ subphenotype cells	218
A.12	Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM159-PT CD24-/CD44+ subphenotype cells	219
A.13	Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM159-PT CD24+/CD44+ subphenotype cells	220

A.14	Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM159-PT CD24+/CD44+ subphenotype cells	221
A.15	Experiment workflow followed to perform transcriptomic analysis on SUM149-PT and SUM159-PT cell lines, cells sorted subphenotypes and their derivative repopulating cells	222
A.16	Schematic representations of PhenotypeTracker lentiviral platforms.	226
A.17	Large scale screening of breast cancer cell lines for their expression of KRT5 and KRT6 biomarkers	229
A.18	Large scale screening of breast cancer cell lines for their expression of KRT8 and KRT18 biomarkers	230
A.19	Large scale screening of breast cancer cell lines for their expression of KRT14 biomarkers	231
A.20	Large scale screening of breast cancer cell lines for their expression of CD44 biomarkers	232
B.1	Principle of single cell segmentation.	234
B.2	Principle of nuclei segmentation from DAPI pictures	235
B.3	Segmentation of cells edges through threshold methods	236
B.4	Segmentation of cells edges through filters methods	236
B.5	Segmentation of cells edges through spreading methods.	237
B.6	Segmentation of cells edges through deviation around the mean method.	237
B.7	Segmentation of cells edges through correction of picture background non-homogeneity.	238
C.1	Sequencing of HRAS and BRCA1 specific mutations on SUM149-PT (A) and SUM159-PT (B) cell lines, batch 1.	243
C.2	Sequencing of BRCA1 specific mutation on SUM149-PT and SUM159-PT cell lines, batch 2.	244
C.3	STR screening of SUM149-PT and SUM159-PT cell lines, batch 2.	244

Acronyms

A | B | C | D | E | F | H | I | K | L | M | N | O | P | R | S | T | V | W

A

APC Allophycocyanin. 154, 159

AU Arbitrary Unit. 63, 84

B

BCL2 B-Cell CLL/Lymphoma 2. 15

BRCA1 Breast Cancer 1, early onset. 47, 242

C

CAF Cancer-Associated Fibroblast. 17

CAG Chicken β -Actin promoter coupled to the CMV early enhancer. 155

CMV CytoMegaloVirus immediate-early promoter. 155

C-MYC Cellular form of Myc, similar of the Myelocytomatosis Viral oncogene (v-Myc). 47

CNS Central Nervous System. 33

CRISPR Clustered Regularly Interspaced Short Palindromic Repeats. 167

CSC-LIKE Cancer Stem Cell-like. 20, 21, 27

D

DAPI 4',6-Diamidino-2-Phenylindole. 53, 282

DNA DeoxyriboNucleic Acid. 199

E

ECM Extra-Cellular Matrix. 17, 33

EF1 α human Elongation Factor 1 α . 155, 158, 161, 165

EGF Epidermal Growth Factor. 14, 165

eGFP enhanced Green Fluorescent Protein. 165

EGFR Epidermal Growth Factor Receptor. 14, 24, 47

EPO Erythropoietin. 34

ER Estrogen-Receptor. 24, 47

ERBB Erythroblastic Leukemia Viral Oncogene with ERBB-1 named EGFR, ERBB-2 named
HER-2, ERBB-3 named HER-3 and ERBB-4 named HER-4. 47

ESA Epithelial Surface Antigen. 41

F

FGF Fibroblast Growth Factor. 14, 16

FSC Forward Scatter. 50, 53

H

HGF Hepatocyte Growth Factor. 14

HIF Hypoxia-Inducible Factors. 166

HLA Human Leukocyte Antigen. 241

HRAS Harvey Rat Sarcoma Viral Oncogene Homolog. 242

I

IDPT Intrinsically Disordered Proteins Theory. 18, 20, 22, 23

IL Interleukin. 14

iPS Induced Pluripotent Stem. 116

K

KRT Cytokeratin. 203

L

LIF Leukemia Inhibitory Factor. 2

Δ LNGFR Delta Truncated form of Low affinity of Nervous Growth Factor Receptor. xxv, 153–158, 161, 162, 165

M

MFI Mean Fluorescence Intensity. 155, 156, 158

MOI Multiplicity of Infection. 156

N

NF κ -B Nuclear Factor Kappa-light-chain-enhancer of activated B cells. 165

NGF Neuronal Growth Factor. 15

NLS Nuclear Localization Sequence. 158

NSC-LC Non-Small Cell Lung Cancer. 14

O

ORF Open Reading Frame. 167

P

P1 Sub-phenotypic population issued from primary sorting procedure. 103

P1' Repopulating population derived from P1. 104

P2 Sub-phenotypic population issued from secondary sorting procedure. 103, 104

pA Action potential. 34

PDGF Platelet-derived Growth Factor. 14, 17

PGK PhosphoGlycerate Kinase 1 derived from mouse strain. 155, 158

PR Progesterone-Receptor. 24, 47

pRB Protein of Retinoblastoma. 15

R

RNA RiboNucleic Acid. 199
ROI Region Of Interest. 170
RSV Rous Sarcoma Virus. 169

S

S₁ Primary Sorting. 103
SD Standard Deviation. 58, 63
SEM Standard Error of the Mean. 101, 127, 129
SMT Somatic Mutation Theory. 20–23
SNP Single Nucleotide Polymorphism. 166
SP-1 Specificity Protein 1. 166
SSC Side Scatter. Also called 90-degree scatter or right angle scatter. 50, 53
STR Short Tandem Repeats. 241, 242
SV₄₀ Simian Virus 40 early-promoter. 150, 155

T

T₁ Tumorspheres derived from primary plating of cells. 122, 128
T₂ Tumorspheres derived from secondary plating of cells derived from T₁ cells. 122, 128
TGF β Tumor Growth Factor β . 15, 47
TMA Tissue MicroArray. x, 203
TOFT Tissue Organization Field Theory. 22, 23
TSP₁ Thrombospondin. 16

V

VEGF Vascular Endothelial Growth Factor. 16, 34

W

WT Wild Type. 161

Preface

«L'esprit d'observation procure un bien plus grand avantage que celui de satisfaire une curiosité raisonnable : en nous apprenant combien en général les hommes sont injustes, il nous prépare à l'injustice. Ainsi, lorsqu'elle nous atteint, nous en sommes moins blessés, et la considérant comme une infirmité de l'espèce plutôt que comme le tort d'un individu, nous sommes moins exposés au malheur de haïr.»
- Pierre-Marc-Gaston de Lévis - *Maximes et préceptes* (1808)

Dear Reader,

These lines surround the beginning of the writing time that I take before my thesis submission and, more specifically, these first words sign the beginning of a period of reflection, distance with my work to draw a conclusion on what have been three years of my life ; *i.e.* some 1000 days of personal investment, searches, experiments, readings... To sum up, a Doctoral work on a research topic that matters to me.

Through this first somewhat informal introduction page I would write down what prompted me to get interested in cell plasticity and especially in highlighting the factuality of a social relationship network and in quantifying spatial organization rules holding sway/governing cells' life. I hope this will help you, dear reader, to get quicker into my work spirit, the essence of my thought.

In this way I will take you with me and date back four years earlier. We are in November 2011 and I am conducting a research topic on the reprogramming of human differentiated cells. If you are not an expert on that revolution which has shaken the biological community all around the world in the beginning of the year 2006, I would recommend you to type on Pubmed Central[®] 4 letters, just 4, iPSC which correspond to "induced Pluripotent Stem Cells". This unpretentious investigation within the world biggest library dedicated to research studies would help you to embrace how high ranking the discovery of reprogramming was.

At this stage of my writing I do not wish, in any case, to dwell on the various methods developed to improve the efficiency of the reprogramming process using several animal species differentiated cells, nor even on the extensive therapeutic hopes arising with this new technology, or on the multiple studies seeking to depict the reprogramming process. Quite the reverse, I just try to make you understand that on the context of my previous work, this technological revolution was the starting point of an almost philosophical reflection about cells sociology. Since then I have never stopped making the parallel between observations about cells behavior (origins, development, organization, influences ...) I could make in petri dishes and human behavior in its great whole I could notice each day long.

Therefore, the next lines linger on the observations I made during cellular reprogramming experiences. And I will tell you how, thanks to them, I realized that reprogramming events were

following unknown underlying determinism. I still remain convinced that undefined social rules were the underlying causes and this is a belief that I will not cease to proclaim throughout this writing.

Particularly laborious and costly in time and patience, cells reprogramming process remained nearly effective. And throughout my experimentation, I noticed that there was a likelihood gradient of the reprogramming effectiveness depending on the cell's location within the dish. It seemed obvious that reprogrammed cell colonies were arising through spatial rules governing cell behavior. In my mind, this phenomenon could only come from the presence of well predefined cell social rules which could be modeled.

Intrigued about these latter observations, I wondered if I could develop a strategy to put in evidence this spatial correlation. Even more, I was attached to highlight the causal source for that spatial organization. It remained to determine some signal transmissions officers that should bring up a de-differentiated cells. And I was particularly thinking about chemical or nuclear acid substances directly related to reprogramming (Leukemia Inhibitory Factor (LIF), micro-RNA perhaps?).

At the time the great project of this work was to optimize the cell reprogramming process to make it compatible with a clinical procedure. Thus, the vision that I had can be formulated as :

- If one highlights social relationships between cells that benefit to the reprogramming process.
- And if one establishes one or some causal elements in this chain.

Then, we will be merely able to play on the establishment of some clue social relationships in order to optimize the efficiency of reprogramming process.

A few months later, I joined the Curie Institute and the team of François Amblard who proposed me to study on the question of autonomy versus social rules underlying phenotype dynamics multiplicity. I do not have ceased to be impregnated by my previous thoughts :

- Many cells or even each cell features a plastic capacity allowing it to naturally evolve and change.
- There are strong cell interactions with their neighborhood and with their nearby cells.
- Cells social relationships must strongly influence these latter decisions making.
- Cells should spatially be organized with respect to each other's in order to facilitate and to improve their capacities of evolution.

And my premium reflection found echoes within an article published few months earlier [113] and related to cancer cells plasticity of cancer cells. We were perfectly interested in the biological model which was outlining very significant phenotypic transitions. Nevertheless, we remained critical concerning the theoretical model in which this article was evolving. So, François and myself wondered about the drivers of cells plasticity capacities with the supreme aim of quantifying social interactions of cells. There stand out the beginning of this exciting research project and the adventure of my PhD project.

Therefore, let me give you an appointment with cells plasticity and heterogeneity within the world of cancer. I wish you a great reading!

Part I

State of the art

Chapter 1

Between individuality and collectivity, notion of phenotypes.

«I am myself; you are yourself; we are two distinct persons, equal persons. What you are, I am. You are a man, and so am I. God created both, and made us separate beings. I am not by nature bound to you, or you to me. Nature does not make your existence depend upon me, or mine to depend upon yours. I cannot walk upon your legs, or you upon mine. I cannot breathe for you, or you for me; I must breathe for myself, and you for yourself. We are distinct persons, and are each equally provided with faculties necessary to our individual existence. In leaving you, I took nothing but what belonged to me, and in no way lessened your means for obtaining an honest living. Your faculties remained yours, and mine became useful to their rightful owner»

- Frederic Douglas, Letter to Thomas Auld (1848) - Frederick Douglass : Selected Speeches and Writings (1999)

The first section of this PhD thesis manuscript aims at presenting the state-of-the-art regarding :

- The concepts of individuality and collectivity, and the notion of phenotypes
- Key concepts that define cancer and cancer cells
- The issue of social relationships among cells

Therefore, this first chapter will briefly introduce the spirit of this PhD work. Therein, we have consistently examined the link, sometimes difficult to draw, between individuality and collectivity. In other words, we have tried to understand both the relationship and the organization between individuals in a group :

- Is it based on individualism and autonomy ?
- Or, is it based on collectivism and interdependence ?

Does the individual life or choice or behavior belong to him - or does it belong to the group, the community, the society, or even the state ? At the human scale, this issue challenges considerations of sociological, epistemological, political and even ethical natures [27]. Individualism plans that each individual is acting on its own, making its own choices, and to the extent they interact with the rest of the group as individuals. Whereas collectivism recognizes the group as the primary entity, with the individuals lost along the way. At the cellular level, is it relevant to consider both

the place of the individual within a society and the extent of its freedom to make decisions on its future ?

We will endeavor to discuss the different levels of organization within a complex corporation. And we will pay a particular attention to the smallest units of this society : the cells. The diversity of cells within an organism is a reflection of the diversity of functions which are necessary for the survival of the whole part. Thus, we will question the link between cellular phenotype and function in a cell group. Finally, this discussion will lead us to consider cell characteristics along a temporal and spatial scale.

1.1 Organism to cells, a question of scale and interactions

Multicellular organisms display five levels of organization which interact all together. Each level is attained by merging formerly separate individuals from a lower level [91]. The surroundings level of organization from the simplest to the broader scale are :

- **Cells** correspond to the basic biological units of both structure and function in living things.
- **Tissues** are made up of cells which share distinguishing features of structure and function and which work together to perform a specific activity. Four different tissues are composing the human body such as epithelial, connective, muscular and nervous tissues.
- **Organs** made up of tissues that work together to perform a specific activity.
- **Organ systems** grouping two or more tissues that work together to perform a specific function for the organism.
- **Organism** which defines the entire living things. It can carry out all basic life processes such as releasing wastes and energy from food, growing, responding to the environment, and reproducing.

Multi-cellular organism comprises self-organizing cell beings which live as a cooperative community. And they together ensure that the body can operate globally. Cooperation is certainly fundamental to the emergence of new levels of fitness in the biological hierarchy[300]. Since cooperation increases the fitness of the group, and new units of selection start out as groups of previously existing units [212]. Indeed, multi-cellular organism could be stated as “the whole is being greater than the sum of its part”. The common understanding is that the “whole” of something, like a car, automatically encloses the idea that the “whole” (the car) is composed of “parts” (engine, doors, chassis...). In other words, the sum of “parts” working together makes up the “whole”.

The smallest “part” of the “whole” organism is the cell (from Latin *cella*, meaning “small room”). This later was described and named by Robert Hooke in 1665 in reference to cells of a honeycomb [137]. Two centuries later, Matthias Jakob Schleiden (1804–1881), Theodor Schwann (1810–1882) and Rudolf Virchow (1821–1902) stated the “cell theory” which main message consist on :

- All organisms are composed of one or more cells.
- Cells are the fundamental unit of structure and function in all living organisms.
- All cells come from preexisting cells.
- All cells contain the hereditary information necessary for regulating cell functions and for transmitting information to the next generation of cells.

In complex multicellular organisms, cells specialize into different cell types (keratinocytes, myocytes, neurons, red blood cells, fibroblasts, stem cells...) that are adapted to particular functions. Yet all cells contain nearly identical copies of the genome, and even so exhibit drastically different phenotypes. Phenotype diversity is partly due to the differential expression of the genes that cells contain.

1.2 Cell phenotypes, the race for the characterization

Cell phenotype can be defined as the conglomerate of multiple traits, characteristics and processes that can be made visible (and then quantified and compared) by some technical procedures. Most molecules, metabolites, structures, behavior involving genes and proteins expression are not always visible at first glance. Nonetheless, they are observable, and thus, are part of the phenotype characterization [192]. This latest results in the elaboration of a cell particular function and role. So, there is a huge diversity of phenotypes in an organism (such as the human body), which leads to suggest that almost each cell has its own phenotype. Even a single neuron, for example, has a set of phenotype characteristics that distinguish it from other neurons as well as from other cell types, such as the nearby astrocytes [76].

1.2.1 Various methods to characterize cell phenotypic specificities

The concepts of “genotype” and “phenotype” have been introduced by Wilhelm Johannsen in his textbook on heredity research (1909), titled “Elemente der exakten Ererblichkeitslehre” (The Elements of an Exact Theory of Heredity). Then, it had been more deeply developed in a 1911 paper titled “The Genotype Conception of Heredity” [149]. Through these writing, he made clear the difference between an organism genotype and what heredity produces. Moreover, he clearly defended that the genotype of an organism gives rise to the organism’s phenotype through the process of development, under the influence of the environment. This formulation means that each characteristics of individual organisms are not directly transmitted to their offspring. Thus, the relationship can be formulated as follows :

Genotype (G) + environment (E) + genotype & environment interactions (GE) → phenotype (P)

Thus, cell biology studies attempt to :

- Understand the cellular and molecular mechanisms which regulate the broad phenotype diversity within both tissues and organs.
- Appreciate the similarities and differences between cell types and cell clusters which share the same characteristics to both extrapolate and generalize the principles learned from one cell type to other cell types.

Morphological aspect

The first knowledge we get on a cell comes from careful observations and descriptions of their morphology. These findings are closely related to the invention of the microscope (and initially the optical microscope). It is hard to connect the invention of “the first tool for observing invisible elements to the naked eye” to a specific inventor, however, it seems that in 1624 Galileo Galilei (1564–1642) had already presented a microscope to European Royalties. The microscope name (from the Ancient Greek : *mikrós*, “small” and *skopéîn*, “to look” or “to see” is attributed to Giovanni Faber (1574–1629).

This invention enabled studies, descriptions and identifications of shape, structure, form and size of cells. The first advocates of this discipline were :

- Marcello Malpighi (1628–1694) baptized as the “Father of microscopic anatomy, histology, physiology and embryology” for his studies of skin, kidneys, liver, lung and blood cells.
- Robert Hooke (1635–1703) who published in 1667 “Micrographia” known for its rich descriptions and illustrations of cells morphology and tissues organization [137].

The early 20th century allowed the development of alternative microscopes, using electrons rather than light to generate the image, and later on able to image surfaces at the atomic level (like the scanning-tunneling or the atomic-force microscopes).

During the last decades of the 20th century, particularly riding the wave of the genomic era, many techniques for fluorescent labeling of cellular structures were developed to analyze cell structure at a molecular level. Thus, recent developments in light microscope largely focused their effort on fluorescent microscopy (as for example the development of scanning confocal microscope) which are approaching the resolution of electron microscopes (as the development of the stimulated emission depletion (STED) microscopes)

Thus, improved microscopy techniques coupled with growing interest for the molecular level (DNAs, RNAs, proteins) or even the atomic scale, drove the loss of interest in morphological and structural studies of the cell. Indeed, the bond between cell morphology and function has been moved to a direct relationship between the molecule (mutation, 3D structure, maturation default, quantity, temporality of expression) and cellular function. Yet this viewpoint separated even more the study of cells (dis)-function from the (dis)-organization of a tissue.

Molecular, transcriptomic and protein characteristics

The central dogma of molecular biology, stated by Francis Crick in 1956, explained the flow of genetic information within a biological system and highlighted the link between genes, cells function and morphology [61]. Thus, following a linear pathway, the transfer of the genetic information is carried out as follows :

- DNA codes for RNAs, thanks to the transcription process.
- mRNA codes for proteins, thanks to the translation process.
- DNA is the molecule of heredity that passes from parents to offspring.
- Proteins make up the structure of the body and carry out most of its functions.

Classically, the genetic research has been restricted to both the sequencing of protein-coding genes and correlating genetic mutations with cellular phenotype (healthy or even pathological one) in a binary digital fashion. To make this notion clear, this signifies that a direct link between marker expression and cells ability to function, like proliferation, migration, metabolism or catabolism, molecule synthesis, signal transmission, interaction, contraction, stimulus response, carcinogenesis, stress response and drug resistance..., had been established.

Slowly, this genocentric view has been modulated with the addition of the epigenetic modulation [263, 303]. Epigenetics tries to understand how variations of physiological phenotype are caused by external or environmental factors. These later both modulate genes expression and affect how cells read genes without changes in the DNA sequence.

Thus, non-genetic inter-cellular heterogeneity is caused by intrinsic noise and differences in gene expression and extrinsic micro-environmental factors, which induce variation and noise in gene transcription machinery [305, 338]. This results in differential gene expression levels that can lead to phenotype alterations in a cell.

Most experimental methodologies for the characterization of cellular phenotype provided ensemble-averaged data which are based on the analysis of bulk samples comprising 10^5 to 10^7 cells. However, the most direct way to address inter-cellular heterogeneity is to perform studies at the single-cell level. Indeed, information with single-cell resolution is certainly crucial to understand the behavior of complex multi-cellular organisms. Studies with individual cells challenge existing technologies (small volumes of around 500 fl and low number of biomolecules per cell). In this way they require methodologies with high-detection sensitivity and specificity. Nevertheless, these expression studies are carried out of usually without consideration for the cellular environment, *i.e.* without observing neither the cell-to-cells interactions nor the spatial organization (within a group of cells forming a cell tissue) of inter-cellular heterogeneity. How does the phenotype of one individual cell differ from another? What implication may these differences have in the context of cell population?

Network of interactions : the Interactome

The term of “interactome” often refers to both physical interactions among molecules (*i.e.* protein-protein interactions), and to indirect interactions among genes (*i.e.* genetic interactions) [269]. Nonetheless, we should be naive to use DNA sequence as a barcode to identify or define cells phenotype and even less the phenotype of a whole organism. Even if genes play a fundamental and even unique role in living things, we clearly need to pay attention to biochemical, chemical and physical molecules when they interact with DNA molecules. Thus, we would like to discuss other types of interactions that may occur within biological networks. Thus, we will be able to understand why and how a cell exhibits a specific phenotype which can differ from the surrounding cells.

A cell phenotype has meaning only because it is expressed within or according to a particular environment (regardless of whether organism is uni- or multi-cellular). Thus, observed in isolation, the actual phenotype of a cell could be under-determined. Search in the coming years will have to focus deeper on cell-cell interactions to formulate answers to these following questions concerning the geometry, their strength and consequences [247, 275] :

- Which cell is next to which one and which cell is communicating with which one?
- Which phenotype do influence the behavior of other cells? And following which social rules?
- What are the (temporal and spatial) characteristics of this relationship? Cells are they under mutual and symmetrical influences?
- What is the consequence on cell phenotype expression, regarding some precise features?

Therefore, the motto is to no longer to focus our attention on the genetic pathways but rather to have a look at the consequences. Indeed, the main goal is to understand both why a cell phenotype is changing to another one (not at the molecular level, but) at the population level and the social relationships which govern cells decisions as an individual and as a population.

Who is interested in the interactions must consider the notion of time scale ; *i.e.* it is necessary to monitor changes in phenotypes during interactions, over time. The follow-up over time of extreme phenotype (meaning cells for which phenotypic characteristics are entirely understood) at the single cell level is not yet within reach.

1.2.2 Cell phenotypes are part of a process

Cell phenotypes are not static features and evolve through the incidence of various stimuli. Indeed, Conrad Hal Waddington (1905–1975) proposed through his classical epigenetic landscape model that genetically predetermined cells can follow any specific permitted trajectories (and which eventually lead to different cellular phenotypes). Thus, the genome serves as a repository of dynamic control information, whose state can be reprogrammed to match the stable phenotypic states [307]. In 1967 and 1969, this approach has been completed by the attractor theory of Stuart Alan Kauffman which proposes that [158] :

- Cell phenotypic states represent dynamical attractors in the complex adaptive landscape represented by gene expression state space.
- These attractors represent the activation and coordinated regulation of pathways.
- Cell differentiation steps are transitions between attractors.

In healthy conditions, diversification and plasticity of phenotypes seems as being under the control of rules that pay attention to homeostasis within tissues, organs and organism.

Time takes advantage of cell state phenotypes

Time is a valuable parameter during which all molecular, cellular, tissue and individual processes express themselves, interact and regulate one to the others through positive and negative feedback [5]. These specific biological mechanisms coordinate internal and exterior events.

Each level of interaction previously defined present its proper temporality of actions. Thus, at the scale of cells processes, several biological clocks range from a few picoseconds (for the mechanisms of light detection by retinal cells for example) to several hours (as it is for the circadian clock) through some intermediate length of mechanisms (such as the movement of molecules, the passage of ions into a channel, the steps of the DNA polymerase during DNA replication) [5]. Some processes across the individual have a constant oscillation rate, as for example the daily rhythm of melatonin secretion by the pineal brain [334]. Moreover, oscillatory temporal phenomenon, such as cyclic gene expression, sometimes come out as a spatially periodic event, as it will be discussed on section 3.2 on page 37.

Spatial organization to enhance specialization of cells function

Depending on the tissues, cells are more or less specialized to realize a specific task. These issues play a key importance on all properties that cells express (*i.e.* the morphology, the expressed molecules, the behavior and the interactions). One of the most notable illustration concerns the specialization of cells of the intestinal crypt. Within this environment, cells function is highly governed by the position of the cells along the crypt [62]. The rapid turnover of the mammalian intestinal epithelium is supported by 5 or 6 stem cells located around the basis of the crypt. During each mitosis, one of the two daughter cells remains in the crypt as a stem cell, while the other differentiates and migrates up the side of the crypt [261].

We will dive deeper in this discussion in chapter “The issue of social relationships among cells” through which we will present the influence of communication in the spatial organization on the manifestation of cell phenotype.

To summarize our thoughts, a cell phenotype corresponds to a set of criteria characterizing a cell at a time “ t ”. This state is inherently unstable because of [293] :

- Stochastic expression of a network of molecules within each cell.
- Multiple interactions that each cell maintains with both the neighborhood (consisting of cells and molecules) and the different scale of organization within an organism.

Furthermore, the expression of a cell phenotype depends on both vertical (interactions though various scale) and horizontal (cells interactions over the same scale) processes among which do prevail multiple interactions. We are interested in both the different modes of interaction available to the cells (chemical and physical) and the consequences concerning the spatial organization of the cells within chapter 3 on page 31.

The pathological framework for this PhD thesis “cancer” did not escape the frantic search to characterize the several phenotype cells components. Indeed, cancer involves abnormal cell displaying potentialities to invade or spread to other parts of the body. Thus, for a long while researchers focused on the characterization of the phenotypes of cancerous cells, *i.e.* through molecular screening to compare to healthy cells. This discussion will be held in next chapter entitled “Key concepts that define cancer and cancer cells”.

Key concepts that define cancer and cancer cells

*«J'avais l'impression d'avoir un cancer de l'amour dont les métastases se propageaient dans mon coeur.»
- Nicolas Carteron - Elle était si jolie (2012)*

Egyptian papyrus (the Edwin Smith Papyrus) dating from 3000 to 2000 BC already described breast and bones tumors [228]. In the 5th century BC, Hippocrates designated under the term cancer (through analogy to their crab shape), malignant tumors observed for various organs (stomach, rectum, uterus, breast) [143, 168, 183, 184]. Then, Aulus Cornelius Celsus (25 BC - 50 AD) and Galen (130 to 220 AD) described the different degrees of evolution of cancerous diseases and promoted excision as treatment of first grade tumors.

We have to wait until the middle of the 19th century for the discoveries regarding cancer cells, carcinogens and proposed treatments, to accelerate.

2.1 Comparisons between healthy and cancerous cells, similarities and differences

In the past, research on tumor cells, meaning their characterization, sorting and manipulations, has enhanced passions. In 1838, Johannes Muller established that malign tumors were composed of cells. Following him, his student, Rudolph Virchow argued after extensive histologic analysis that all cells, including cancer cells are derived from normal cells [7]. These two key affirmations have driven the thinking of what would be the key actors of cancer disease progression, meaning the tumor cells, and the definition of what is cancer, meaning a disease involving abnormal cell growth. And this is how year 1952 has marked a turning point with Margaret Gey, who established the first malign tumor cell line from uterine cancer. This biological cancer tool has played a key role in the characterization of cancer cells allowing comparisons with healthy cells [19, 324].

In years 2000 and 2011, Douglas Hanahan and Robert Weinberg enumerated the different characteristic features of tumor (both nonmalignant and malignant) cells which, up to date, continue to guide our thinking regarding carcinogenesis [119, 121].

2.1.1 Self-sufficiency to growth signals

Normal tissues and healthy cell growth is under the dependence of various exogenous mitogenic signals synthesized by the neighboring cells (paracrine regulation) or the body (endocrine regulation) [173]. These growth signals contain both growth factors (as Hepatocyte Growth Factor (HGF), Epidermal Growth Factor (EGF), Fibroblast Growth Factor (FGF), Wnt), cytokines (as CXCL12, Interleukin (IL)6) and hormones (as insulin, estrogens...) [81]. Tumor tissues gain independence to these external factors through diverse ways.

By modifying their surface receptors

- Either by increasing the number of receivers (like chemokines receptors). This phenomenon has been well observed in carcinomas of the brain, the stomach and mammary glands [231].
- Either by enhancing receptor sensitivity to growth factors through multiple point mutations. Also, RAS protein is regularly found in a mutated form enabling it to be constantly self-activated and thus independent of upstream regulatory programs (as for example in colon carcinoma) [209, 239]. Other proteins also play key role in the activation pathways of cell growth can be mutated as B-RAF [65, 336] in melanoma or Epidermal Growth Factor Receptor (EGFR) in Non-Small Cell Lung Cancer (NSC-LC) [100].
- Or in continuously activating signal transmitters of growth factors [227].

By increasing growth factor production

- By self-producing growth factors to sub-serve autocrine activation, as for example the Platelet-derived Growth Factor (PDGF) [35].
- By managing neighboring healthy and tumor cells (that form the tumoral micro-environment) to increase growth factor production [25, 290].

In such cases, the abnormal production of growth factors leads to continual activation of cell division, and the tumor cells proliferate even in the absence of growth factors from physiological sources.

2.1.2 Evading apoptosis

This form of programmed cell death (firstly described by Kerr et al. in 1972 [160]) follows a particular time schedule in all body cells (beginning with membrane disruption, deterioration of membranes and nuclear cytoskeleton, cytoplasm exclusion leading to nucleus fragmentation) [342]. To escape signals of apoptosis, cancer cells play on three different fields.

Sensors

Which determine life or dead conditions of cells depending on several signals :

- From the environment, meaning the extrinsic apoptotic program (cell-to-matrix or cell-to-cell adherence based survival signals involving Fas-ligand for example) [12, 131] or anoikis defining the abnormal cells survival to the deprivation or deficient cell-to-cell or cell-to-matrix contacts [111].
- Internal to the cells, also named the intrinsic apoptotic program that is sensitive to DNA-damaged signals (as TP53 regulatory system for example), unbalanced growth and hypoxia [299].

Apoptotic triggers

Which conveys signals between sensors and effectors. They lean upon the balanced expression of pro- (like B-Cell CLL/Lymphoma 2 (**BCL2**) proteins family) and anti-apoptotic (as for example BAX, BAK proteins families) signaling [210].

Effectors

Which concentrate all apoptotic signals that converge on mitochondria. These latter release a catalyst of apoptosis named the cytochrome C which then activates various proteases from the Caspases family (Caspase 8 and 9), thereby allowing inexorably the activation of the chain of cell death process [12, 197].

Mutations and inhibition of proteins expression (C-MYC, BCL2, TP53 for examples) at these several steps of signal processing, highly increase cell survival abilities [1, 134], as described for example in familial breast cancer [98, 258] or B-cell non-Hodgkin lymphoma [262, 270].

And dependence receptors

Physiologically and pathologically, dependence receptors mediate programmed cell death by monitoring the absence of certain trophic factors as it is regarding the androgen receptor (withdrawal of testosterone leads to the death of prostate epithelial cells), insulin receptor (lack of insulin leads to the apoptosis of pancreatic cells), or p75 neurotrophin receptor (leading to neurons suicide when it lacks Neuronal Growth Factor (NGF) binding) [34, 251].

2.1.3 Evading growth suppressors

In normal tissues, anti-proliferative factors maintain tissue homeostasis and stop proliferative signal through two different ways [171] :

- In forcing healthy cells to go on quiescence, meaning the G0 cell cycle stage.
- Or in forcing cells to differentiate through entrance in post-mitotic stages.

Tumor cells inhibit anti-proliferative signals to keep their multiplicative abilities and are thus actively cell cycling. Anti-proliferative proteins such as Rb, P107 and P130 are the cornerstone that cancer cells must bypass. For example, Protein of Retinoblastoma (**PRB**) circuit, which is regulated by both the Tumor Growth Factor β (**TGF β**) and other extrinsic factors, is dysregulated in a variety of tumors either directly at the TGF- β receptor or at the different down- or upstream factors [72, 144, 165]

2.1.4 Replicative immortality

The set of dysregulations of cell proliferation programs causes the generation of a very large number of cells. Progressively cells divide and microscopic tumors constitute a macroscopic tumor. To escape the intrinsic limitation of “the replication potential” of each healthy cell (meaning around 60 to 70 cells divisions) [124, 125], tumor cells have developed a set of resources to attain the replicative immortality. The key issue consists of maintaining at any cost the telomeres length [278]. Two main strategies have also been formulated either by up-regulation of telomerase expression [280] or by inter-chromosomal exchange of telomeres [37].

2.1.5 Genome instability and cumulative mutations

Acquisition of the multiple hallmarks depends on the succession of alterations in the genome of neoplastic cells which enable the outgrowth of selectively advantageous cells sub-clones. During cell division, the cells accumulate point mutations, deletions and duplication within their genome [40, 188]. A system of genome maintenance and repair ensures that the mutation rate is rare. Thus, the number of mutations found in non-tumor cells and accumulated throughout cell life, remains far lower than the multiple mutations found in tumor. Mutability is achieved through increase of sensitivity to mutagenic agents and deregulation of the machinery of genomic maintenance [198]. So, there are caretakers of the genome at different stages in order to detect, repair and inactivate DNA damages [350]. Genome alteration depends on tumors type, genome maintenance and repair defects. Thus, genome instability is inherent to human tumor cells, and the acquisition of hallmark capabilities [221, 350].

2.1.6 Deregulated cellular energetics

In order to fuel cell growth and division, proliferating tumor cells need to adjust their energetic metabolism [97]. Under aerobic conditions, normal cells process glucose to pyruvate via glycolysis in the cytosol, and then pursue oxidative breakdown of pyruvate to carbon dioxide in the mitochondria through the citric acid cycle. However, tumor cells exhibit a higher use of glycolysis and lactate fermentation (under anaerobic conditions), and which process is named the Warburg effect [332]. To compensate for the loss of energy relatively to mitochondrial oxidative phosphorylation, tumor cells considerably increase the uptake and use of glucose through the up-regulation of the synthesis of glucose transporters [167]. This anaerobic synthesis capability enables rapid adaptation of cells to hypoxic conditions prevailing in a majority of tumors [64, 94, 132].

2.1.7 Inducing neo-angiogenesis

Cell functions (growth, speciation...) and cell survival require a constant supply of oxygen and nutrients, as well as the discharge of metabolic wastes and carbon dioxide. Thus, virtually every cell is located less than 100 microns from a nourishing blood capillary. Early during the multi-stage development of invasive cancer, neoplastic tumors develop neo-angiogenesis in up-regulating key inductors of angiogenesis (*i.e.* Vascular Endothelial Growth Factor (VEGF)-A and its receptors, Thrombospondin (TSP1), FGF-2 ...) through extra-cellular matrix remodeling, hypoxia and inflammation activation [84, 120]

2.1.8 Avoiding immune destruction

The immune system performs a constant monitoring of the organism in order to recognize and eliminate the invaders from the body. Thus, the vast majority of cancer cells and tumor formations are particularly well recognized by either CD8+ cytotoxic T lymphocytes, or CD4+ Th1 helper T cells, or natural killers [141, 197]. However, few cells manage to pass through the cracks and to escape immune surveillance. The mechanisms that tumor cells have implemented operate through either the paralyzing of infiltrating natural killers and cytotoxic lymphocytes through the secretion of TGF- β and other suppressor factors [318], or the recruitment of inflammatory cells as regulatory T cells and myeloid derived suppressor cells [253].

2.1.9 Tumor promoting inflammation

Tumors are densely infiltrated by innate and adaptive cells of the immune system and even more under inflammatory conditions which trigger tumor growth with opportunism [96, 357]. Indeed, inflammation supplies activators of tumor growth to the tumor environment. As for example growth factors sustain proliferating signaling, or survival factors that limit cell death. Moreover, pro-angiogenic factors and extra-cellular matrix modifying enzymes support angiogenic invasion and metastasis or other inductive programs [120].

In parallel, inflammatory cells release chemicals (reactive oxygen species or other actively mutagenic actors) which help in activating common hallmarks abilities. Tumor inflammation is the core issue for the acquisition of almost all tumoral hallmarks [58].

2.1.10 Interaction with tumor stroma

Tumor development and opportunistic growth is triggered by both the presence of tumor cells and also the environmental context [245, 335]. As the cancer progresses, the surrounding microenvironment (stromal components) co-evolves into an activated state through continuous paracrine communication, thus creating a dynamic signaling circuitry that promotes disease evolution and ultimately leads to a fatal disease. The tumoral stroma is surely recruited from adjacent tissues and hosts a variety of both cellular and molecular components. Tumor associated stromal cells, derived from bone marrow [170], are composed for example of Cancer-Associated Fibroblast (CAF), mesenchymal cells secreting signal (which synthesize potent activators of stromal and tumor cells such as the PDGF), pericytes, pro-tumorigenic immune cells (derived from myeloid progenitors), inflammatory Extra-Cellular Matrix (ECM). Thus, oncogenic stromal compartment provides stimulatory signals and protection to tumor cells, and plays key role in supporting tumor growth and transformation (as invasion and metastasis processes) [24, 25, 290].

2.1.11 Tissue invasion and metastasis

Few cells are able to escape from primary tumors and colonize other neighboring tissues, and even less to migrate to distant sites to undergo new tumor colonies. These secondary located tumor, named metastasis tumors, are responsible of up to 90% of cancer linked death [210]. Tissue invasion and metastasis rely on highly complex processes which are absolutely subjected to a strong environmental pressure. Stages of the invasion-metastasis domino-effect may be summarized as : local invasion, followed with intra-vascularisation into nearby blood or lymphatic vessels, cancer cells transit, ex-vascularisation and formation of micro-metastasis which can grow up to a macroscopic tumour. Indeed, this process results on the intrinsic coupling of both effects as the modification of cells micro-environment (as for example molecule of cell-to-cells adhesion like E-cadherin [45], β -catenin, immunoglobulin [53], integrin...) and the activation of extra-cellular proteases [283].

2.1.12 Stem cells and cell-to-cell communication

The diverse studies on embryonic and normal adult stem cells, have pushed the stem cell theory of cancer to gain momentum [310, 319]. This model is established on the statement that malignant tumors should arise from a peculiar population of cells. These later should exhibit abilities to indefinite self-renew and to give rise to all differentiated cell types of the tissue of origin (non-stem cells) [56]. These cells have been potentially enriched, firstly, in leukemia [31] and then in various solid tumor tissues as colon, breast, prostate, neuro-ectoderm [75, 101, 117, 235, 279]. However specific and universal markers are still under research and cell identification remains on functional tests (as further developed in section 2.2.2 on page 20).

Hallmarks, features that distinguish something from others

The hallmarks, as previously listed, are a set of criteria whose purpose is to define the characteristics that distinguish cancer cells from other cells. However, as noted by Yuri Lazebnik, the above-mentioned features characterize every tumor cells both cancerous and non-cancerous [174]. Cancer cells hallmarks should by definition portray the core mechanisms that underlie the proficiency of cancer cells to invade surrounding tissues and metastasize. Currently, characterization of cancerous cells remains complex. Indeed, the main limitation is the lack of proper controls, normal cell lines are used to compare their characteristics to cancer cells while we should be using tumoral non-cancerous cells. We largely lack knowledge with respect to strict criteria that lead to malignancy. We must, henceforth, compare tumor cells which are similar at $t = 0$ but differ in their invasiveness at $t + 1$ to highlight differences in morphology, gene expression, representativeness of phenotypes within cancerous and non-cancerous cells and other stromal cells.

Over 100 years of arduous cancer research have passed, and we would expect a thorough knowledge of the disease and particularly of its genesis. As underlined by John Cairns, the study of molecular biology in cancer has expanded our knowledge with respect to the healthy pathways and processes. Nonetheless, these knowledges have not yet provided an answer about the origin of cancer, its causes and treatments [40]. How does it develop in a healthy body? Where does it come from, that is, from which cells does it grow? How does a healthy tissue accept major divisions of certain cells and the growth of a tumor mass then its transformation into potentially invasive cancerous tumor? are all questions whose answers remain controversial. The following section intends to identify some theories that have gradually emerged over the years and discuss the origin of tumors and cancer.

2.2 The origin of tumor and cancer, a key question.

Since the earliest times, the physicians have puzzled over the causes and origins of cancer. So since Hippocrates, and for nearly 1300 years thereafter, physicians believed that the body had four humors (the blood, the phlegm, the yellow bile and the black bile). When humors were balanced people were healthy, however disequilibrium of one or many humors caused cancer. Only in the 18th century did this theory evolve and give way to various other theories such as the infectious disease theory, the lymph theory, then the chronic mutation and the trauma theory. The so-called modern biology has brought many arguments and with them a diversity of causes that are still at the heart of debates.

How does appear the first tumor cell more or less able to progress up to cancer stage? Is this cell originally present and potentially pro-invasive? How does a cell having a limited number of divisions turn into a cell capable of dividing potentially indefinitely? Several models have been proposed to understand how at “ $t = 0$ ” a tissue showing no tumor cell, displays some at “ $t + 1$ ”.

2.2.1 The Intrinsically Disordered Proteins Theory (IDPT)

It is widely held that cancer is a genetic disease and that each tumor development arises from changes that have occurred in cellular genome. Indeed, such modifications are believed to mark the genome, such that the history of cancer life is encrypted in the somatic mutations present in the tumor cells. The Intrinsically Disordered Proteins Theory (IDPT) focuses on instabilities affecting the normal network of protein interactions, either spontaneously or externally triggered [15]. Following the initial trigger, this theory postulates that a cascade of genetic mutations and molecular pathways dis-regulations leads to tumor formation and malignancies.

External causes of tumor formation

Although mutations can commonly accumulate through bad luck over cell division, we wanted to discuss the importance of so-called external causes which have strong oncolytic potential.

In 1915, Katsusaburo Yamagima and Ichikawa demonstrate for the first time that chemicals of environmental origin were able to induce tumor development, for example on rabbit by applying coal on their skin [95].

Other environmental factors can induce tumor formation. In 1911, Peyton Rous described sarcoma on chicken caused by infectious agent known as the Rous sarcoma virus [264, 265]. Several other viruses are now linked to tumor development like the hepatitis B and C virus (cancer of liver) [337], the herpes virus EBV (Kaposi sarcoma and non-Hodgkin lymphoma) [157, 242], HPV (cervix, vulva, vagina, and penis cancer) [68]. Today, the World Health Organization's International Agency for Research on cancer has identified and listed more than 100 chemical, physical and biological carcinogens which include a wide range of exposure :

- Lifestyle factors taking into account nutrition, tobacco and physical activities
- Natural occurring exposure as UV light, radon gas and infectious agent
- Medical treatment as radiations and others (chemotherapy, hormone, immune suppressor chemicals)
- Workplace exposure and household
- Pollution entities

The key point of these exposures is that they trigger evident DNA damages. And through the next section we will develop the key genetic basis of cancer development and progression.

DNA damages and accumulation of mutations through cells generation

Cancer is currently defined as an evolutionary process that results from the accumulation of somatic mutations, genetic alterations at different scales, in the progeny of normal cells and leading to dysfunctions within proteins interactions :

- Either aneuploidy, in which entire chromosomes are gained or lost [72].
- Or intra-chromosomal instability, characterized by insertions, deletions, translocations, amplifications, and other forms all sharing the feature of utilizing DNA breakage as an early step.
- Or usual replicative errors as any point or oligo-base mutations affecting pro or anti-oncogene, which are rare except for DNA replication error inherited syndromes (as for example hereditary nonpolyposis colorectal cancer) and a small fraction of sporadic cancers [350].

The genome of tumor cells is unstable, and this instability results in a cascade of mutations. Some of which enable tumor cells to bypass the host regulatory processes and randomly lead to cell phenotype carrying a selective growth advantage. These were arbitrarily classified as either putatively causal (driver mutations) or irrelevant (passengers' mutations), although this classification remains unverifiable because of the use of transformed cells [316].

The purpose of any mutation at the genetic scale is to enable the establishment of a phenotype exhibiting at minimum the expression of the cancer hallmarks presented in section 2.1 on page 13.

Thus, genetically unstable cells take selective advantage over the others while accumulating mutations over divisions [188]. Although a single mutation appears to be sufficient for a cell to become tumoral, some studies have identified number of genome defections per tumor and cancer cells. Indeed, they can amount to several hundred ; as for example, the mean number of genomic events per colorectal carcinoma cell totals approximately 11 000 [302].

These identifications have led to a better understanding of molecular and biochemical pathways involved in cell proliferation abilities, in tumorigenesis and malignant transformations. The changes seem to be the cause of increased proliferation of certain cells. However, it is still necessary to determine how many cells there are ? Which cell phenotype is the most sensitive or susceptible to proliferate without any spatial environmental constraints ? Finally, it should be noted that some somatic mutations regarding the same highlighted genes (nonetheless silent) have been identified in the cells of normal tissues [198].

2.2.2 The Somatic Mutation Theory (SMT) and Cancer Stem Cells

Normal tissues (either brain, or breast, blood, or intestine) are maintained by multipotent or unipotent stem cells and progenies (named as committed progenitors) which amplify and can transit to low or non-dividing terminally differentiated cells which accomplish tissue functions [323]. The inherent property of cancer cells results in the creation of multiple daughter cells or subclones with different phenotypic properties. In the same manner, increasing evidences suggest that cells phenotypes of both non-solid and solid tumors are hierarchically organized, enhancing heterogeneity among tumor cells [29, 88, 233, 234]. The Somatic Mutation Theory (SMT) combined the IDPT and its consequences that initially occurs in a single somatic cell.

Influence of genetic and epigenetic status on specific tumor formation

Genetic lineage tracing experiments in mammal models have challenged this issue and try to point out cell hierarchy and the influence of over-expression of oncogenes or deletion of suppressor-gene on the behavior of cell hierarchy. Therefore, for example, these same studies at the skin level enabled confirming that cancer development is influenced by both mutations and the specificity of targeted cell phenotypes [29]. Furthermore, oncogene expression in a given tissue led to the transformation of some cells and not others suggesting that the epigenetic transcriptional state of oncogene-targeted cells plays an important role in tumor initiation. Thus, tumor-cell heterogeneity is the result of both genetic and epigenetic influence [199].

Initiation and maintenance of tumors, a distinct subpopulation if Cancer Stem Cell-like (CSC-LIKE)

Therefore, CSC-LIKE theory proposes two complementary concepts :

- Tumor malignancies should arise and be maintained from cancer cells which present stem cells features and quite similar of the one present in normal tissues.
- Cancer tumors are phenotypically heterogeneous and composed of the same type of cells as found in normal tissues (stem cells, transit amplifying cells, terminally differentiated cells [31]). Close to adult stem cell characteristics, their presence in tumors are emerging from :
 - Either stem cells themselves which undergo mutations and become cancer stem cells [166].
 - Or from the de-differentiation of already differentiated cells [164, 260].

- Or transfer of pro-oncogenic molecules between stem cells and other more differentiated cells (like mutated cells) [2, 348].

Nevertheless, these cells were previously observed in the late 90's in hematopoietic cancer (leukemia) and they have been subsequently discovered in various solid tumors (brain, breast, cancer, prostate). The latter exhibit the following characteristics [126] :

- Self-renewal capacities
- Chance to differentiate
- Great expression of telomerase enzyme
- Evasion from apoptosis
- Increase in transportation of molecules through the membrane
- Ability to migrate and generate metastasis
- Sensitive to micro-environment
- Unlimited division abilities

Markers and special features of CSC-LIKE are still under high-stake research in order to monitor this particular phenotype among all the cells forming the tumor masses. However, as mentioned above, no universal marker could be determined. Thus, stemness features are predominantly identified on functional assays through tumorigenic assay both *in vitro* and *in vivo* with xenotransplantation assays of human tumor biopsy-derived cancer cell subpopulation.

Nonetheless, the cancer stem cell theory stays controversial because evidence for the existence of cancer stem cells relies solely on functional assays that involve breaking down a tumor, selection of particular cells and transplantation. This process does not exactly mirror natural cancer growth. To prove that the transplanted cells are really cancer stem cells, scientists will need to look at individual cells and find direct evidence that the cells do indeed feed tumor growth within the natural environment of the body.

Cancer expansion through stochastic mutations and clonal selection processes

Tumor can be viewed as an aberrant organ initiated by a single or more tumorigenic cancer cells that have acquired mutations, gain growth advantage and therefore present indefinite proliferative abilities. The SMT theory also implies that the individual clones of cancer cells evolve independently from each other competing for space and resources such as oxygen and nutrients, accumulating over divisions genetic and epigenetic mutations, and thus acquiring the necessary hallmarks for tumorigenesis [139].

Because some tumors have a clonal origin, dividing tumorigenic cells must give rise to phenotypically diverse cell progenies with possibly indefinite or limited proliferative potential. If phenotype diversification in normal tissues and tumor bear some resemblance or share common mechanisms, tumors should reflect some aspects of the differentiation patterns that normally occurs in normal tissues. In parallel, variable expression of markers and an heterogeneous combination of diverse cell phenotypes (as expressed in a healthy tissue) are found in tumor. This strongly suggest that the observed morphological and phenotypic heterogeneity results of the differentiation of a subset of rare tumor cells : the cancer stem-like cells. These later share common features, behavior and regulation pathways with normal adult stem cells but which are expressed in a distorted way (as for example in leukemia) [31].

2.2.3 The Tissue Organization Field Theory (TOFT)

The tumors carry the identities of the bodies they originate from, and exhibit behaviors similar to those of normal tissue. Their characteristics are :

- Altered tissue organization
- Excessive local accumulation of cells (hyper- or neoplasia)

Giovanni Battista Morgagni wrote in the 17th century the foundations of modern anatomopathology and observation of tissues, on which still rests the diagnosis of several non-malignant and malignant tumors. However, the advent of molecular biology, which sees in a tumoral pathologies a molecular or genetic disease, swears to diagnose cancer through genomic screens.

Reductionism versus organicism

The advances in molecular biology over the last three decades have paved the way for a genetically anchored theory of cancer. An alternative view comes from the tradition of developmental biology, placing the organism at the focal point of research.

This theory probably, qualified as ‘mainstream theory’, assumes that cancer has its causes rooted in genetic malfunctioning at the cellular level. This reductionist approach was made explicit in Weinberg’s book, *One Renegade cell* (1998) and emphasizes the key role of a single mutated cell in carcinogenesis. In disruption with previous view, proponents of the Tissue Organization Field Theory (TOFT) support that the causes of cancer have to be searched not at the genetic level but at the tissue level. Indeed, cancer cell should originate from the disruption of tissue organization [193].

From Johann Mendel (1822–1884) to Friedrich Miescher, passing through the discovery of the DNA molecule by Rosalind Franklin, Maurice Wilkins, Jim Watson and Francis Crick, genetic determinism has imposed itself in biological thinking. This scientific thought conceptualizes both embryonic and tumor development as the deployment of a program written in our genes that takes place within each cell of our body in a defined temporality. However, this view ignores the differences that exist between gene expression (taking into account both of the gene sequence and the regulatory elements carried by the DNA molecules subject to epigenetic mechanisms) in a cell, organ or individual and phenotypic complexity found at all scales.

Furthermore, cell theory states that the cell is the unit of life, but that unit cannot be fully autonomous in a multi-cellular context. The organismic vision registers the various scales in open systems where the causalities are multiple and different levels of biological organization (cell, tissue, organ, individual) are intertwined.

Tissue alteration at the basis of tumor development

Former tumoral theories, *i.e.* IDPT and SMT, centered on both the molecular and the cellular level of tumoral tissue organization, have mainly dictated cancer research and therapies proposals. Nevertheless, another point of view regarding tumor initiation have been defined in 1999 by Carlos Sonnenschein and Ana Soto [193, 296]. The tissue organization field theory, is contesting these traditional reductionism approaches and claims that cancer disease is much more complex than just genetic and molecular components to trigger the development of tumors. According to TOFT theory, the breakdown of tissue organization involving many cells (leading to damage of their structure and architecture), is the only cause of cancer (and not mutated genes or cells proliferation). This means that deregulated genes are the consequence and not the cause of tumor initiation and growth.

Conversely, the TOFT posits that tumorigenesis, as the histo- or organo-genesis, are phenomena that are not taking place at the cellular level but at the tissue level. Thus, cancer can be seen as development gone wrong. TOFT explicits that carcinogens generate a disturbance in the interaction between cells that maintain cellular organization, tissue repair and local homeostasis : the social tissue. This altered stromal architecture would enable cells to express their default state, that is to say the proliferation and motility, leaving the emergence of tumors. And tumor progression, according to the Darwinian model, would be under the influence of tumor phenotypes cooperation (either commensalism, in which one individual of a pair benefits but not the other ; or mutualism, in which both benefit, resulting in synergy) more or less beneficial to tumor cells [14]. Thus, new properties may emerge in cooperating groups that single individuals do not exhibit.

The TOFT has long been touted by its main defenders as the main alternative to both SMT and IDPT theories. However, the latter clearly supplements, but does not supplant, the traditional views of tumorigenesis in which one clonal population of cells develops all the necessary molecular and genetic traits to form a tumor. Cooperation among tumor cells raises new questions about tumorigenesis and has subsequent implications on the development of new therapeutic approaches. Indeed, under cooperation conditions, some tumor cells may provide resources that other cells require, leading to the survival of more cells and the subsequent accumulation of mutations faster than traditional theories allow.

2.2.4 Breast tumors, diversity and clusterization

This thesis work fits into the framework of a study on the diversity of phenotypes found in two cell lines of human breast cancer (namely SUM149-PT and SUM159-PT). Thus, through this section, we will focus on the diversity found at different scale levels : pathological, anatomo-pathological and cellular.

Breast cancer is a heterogeneous disease, according to its histo-pathological and biological features, and to the different clinical outcomes. Based on such heterogeneity, breast cancer cannot be viewed as a single clinico-pathological entity. Solid tumors with mammary origin have been subject to the phenotypic characterization of different cell actors, both from tumor tissue excision and from cell lines *in vitro* models of tumor pathologies. Various methodologies (based on (pan-)genomic, transcriptomic, proteomic, or histopathological datasets, the course and the grade of the tumors) have enabled to initially classify tumor disorders (malignant and non-malignant). This enabled to set a hierarchy of cell phenotypic diversity within tumors. The main purpose of these classifications is to select the best treatment for the patients, and update of knowledge related to cell biology, or to response to treatment lead to changes in classifications [295].

The histo-pathological classification is based upon characteristics of biopsy specimens. The World Health Organization (WHO) propose the classification of breast of both benign and malignant tumors [340] mainly on the basis of their histo-pathological origin : epithelial, myoepithelial, mesenchymal, fibroepithelial and the nipple. Three most common histo-pathological types, which collectively represent around three-quarters of diagnosed breast cancers, are :

- Invasive ductal carcinoma which represent around 55% of breast cancers in America.
- Ductal carcinoma in situ diagnosed in 13% of breast tumors.
- Invasive lobular carcinoma which represents 5% of breast tumors.

Grading of a breast cancer depends on the microscopic similarity of breast tumor cells to normal breast tissue, it means the closer the appearance of cancer cells is to normal cells, better is the prognosis. Classification of cancer distinguish well differentiated (low-grade) tumors between

moderately differentiated (intermediate-grade) tumors, and poorly differentiated (high-grade) tumors. Also, this grading reflects progressively less normal appearing cells that have a worsening prognosis. If cells are not well differentiated, they will appear immature, will divide more rapidly, and will tend to spread. The grading criteria are as following :

- The tubule formation assesses what percent of the tumor forms normal duct structures, this is a measure of the order of the tissues structures.
- The nuclear pleomorphism assesses whether the cell nuclei are uniform (in shape and in size) like those in normal breast cells.
- The mitotic count assesses number of cells which are dividing, more cells divide, the cancer is more aggressive.

Staging is the process of determining how many tumors are developed within the body and their location. This is quite a measure of the extent or the severity of an individual's cancerogenesis [6], when considered along with other classification aspects such as Estrogen-Receptor (ER) and Progesterone-Receptor (PR) levels in the tumoral tissue, the human EGFR-2 (HER2/neu) status, the menopausal status, and the person's general health.

The receptor status of breast cancers categorize the cells based on the presence of estrogen receptors (ER), progesterone receptors (PR), HER2 and androgen receptor. Receptor status is a critical assessment for all breast cancers as it determines the suitability of using targeted treatments. To illustrate, estrogen receptor positive (ER+) tumor cells (which usually have a better prognosis) depend on estrogen for their growth, and treatment should be based on reducing either the effect of estrogen (*i.e.* tamoxifen) or the actual level of estrogen (*i.e.* aromatase inhibitors). Conversely, triple negative cancer (*i.e.* which do not express positive ER, PR and HER2), are still lacking targeted treatments and has poorer prognosis. Finally, androgen receptor is expressed in 80–90% of ER+ breast cancers and 40% of “triple negative” breast cancers. Activation of androgen receptors appears to suppress breast cancer growth in ER+ tumor cells, while in ER- tumor cells it appears to act as growth promoter, and thus could be utilize as a prognostic marker. The identification through the receptor status has been coupled with transcriptomic analysis, as presented thereafter.

Six to seven main groups of breast disease, mainly determined by a **hierarchical clustering of transcriptomic data**, are accepted and presented as the molecular portrait of human breast tumors [109, 244, 297] :

- Basal-like tumors (ER-, PR- and HER2-), are characterized by the over-expression of basal compared to luminal gene, cytokeratin 5 and 17, laminin and fatty acid binding protein.
- ERBB2+ tumors (HER2+), characterized by the over-expression of basal compared to luminal genes and genes carried by the ERBB2 amplicon at position 17q22.24 (with for example the ERBB2 and GRB7 gene).
- Normal-like tumors, characterized by the over-expression of basal compared to luminal genes and genes usually expressed by adipose tissue and non-epithelial cell types.
- Luminal sub-type A tumors (ER+ and low grade), showed the highest expression of both luminal and the ER α genes, the GATA binding protein 3, the X-box binding protein 1, the trefoil factor 3, the hepatocyte nuclear factor 3, and the estrogen-regulated LIV-1.
- Luminal sub-type B tumors (ER+ but often high grade), which showed low to moderate expression of the luminal-specific genes including the ER cluster.

- Luminal sub-type C tumors (ER-/AR+), distinguished from luminal subtypes A and B by the high expression of a novel set of genes whose coordinated function is unknown (feature they share with the basal-like and ERBB2+ subtypes).
- Claudin-low tumors, present low expression of cell-cell junction proteins (including E-cadherin) and is frequently infiltrated with lymphocytes.

2.3 Why are we interested in previous questions ?

This question amounts to asking how the collaborative work between basic research and clinical medical research is a key element to improve the management and maximize the quality of patient care. Thus the interest in the differential characterization between normal and tumor cells, to understand how tumor cells arise from a normal tissue, what are the mechanisms that tumor cells are activating to grow... enables us to :

- Understand the underlying mechanisms of cancer biology.
- Integrate the basic knowledge to clinical applications.
- Develop new diagnostic strategies and therapeutic approaches.
- Test, refine and validate innovative therapies to patients.

“Clinic to bench” and “bench to bedside” approaches are the secret of therapeutic success.

2.3.1 Diagnostic and prognostic to improve patient follow-up

Improved methods of detection, diagnostic and stratification of cancers to determine the prognosis (probability that cancer will be cured or reappear) are at the heart of the stakes of management of any patient. Therapeutics have diversified and tools of diagnosis and prognosis arouse more and more interest. Indeed, it is necessary to better understand the profile of the tumor to provide the most appropriate treatment and correlate the phenotype of the individual to tumor characteristics (both genetic, molecular, phenotypic diversity and spatial organization...).

Thus, when a person experiences disturbing symptoms or gets worrying results after a physical examination or a screening test, it is important to be able to diagnose or rule out cancer as soon as possible. If a cancer is diagnosed, it is necessary to establish precisely the stage, *i.e.* if it has spread to other parts of the body locally or remotely, to decide what better treatment and ask a prognosis. Finally, following the treatments carried out, it is necessary to know whether tumors responded correctly to treatment, *i.e.* if they have disappeared and monitor relapse potential.

A set of biological tools or indicators allow the diagnosis, stratification and tracking of tumors :

- **Dosage of tumoral biomarkers** naturally present in the body which can, if they are present in larger quantities than normal, indicate the presence of cancer. Some tumor markers are specific to one type of cancer, while others are associated with several types of cancers as for example [138] :
 - α -foeto-protein (AFP), tracker of hepatocellular carcinoma, testis and ovarian tumors [26]
 - Tumor antigen 125 (CA 125), follow-up of ovarian cancer treatment and relapse [320]
 - Tumor antigen 15-3 (CA 15-3), follow-up of breast cancer treatment and relapse [154]

- Carbohydrate antigen 19-9 (CA 19-9), follow-up of pancreatic cancer treatment and relapse [314]
 - Carcinoembryonic antigen (ACE), tracker of colorectal, breast, lung, pancreas, stomach and ovarian tumors [55]
 - Human chorionic-gonadotrophin (HCG), tracker of testis, ovarian, breast, stomach... tumors [238]
 - Human prostate-specific antigen (PSA), tracker of testis tumors [16]
- **Follow-up of genes expression and tracking of mutations** to validate both the mutations and the level of expression of DNA, RNA or proteins through, for example, high-throughput and single cells sequencing technologies.
 - **Techniques of medical imagery** helps in finding tumor origin to orientate surgery procedures and other treatment strategies such as positron emission tomography (PET), single photon emission computed tomography (SPECT), the computed tomography (CT) in 2D and 3D to observe changes in molecular synthesis or chemical structures, in organs or tissues function and structure, and in blood flux.

2.3.2 From conventional to targeted therapy strategies

Significant progress has been made in the diagnosis and treatments in cancer over these three latest decades. Treatment options depends on the type and stage of cancer (*cf.* section 2.2.4 on page 23), possible side effects, and the patient's preferences and overall health. In cancer care, a multi-disciplinary team of medical professionals often work together to create a patient's overall treatment plan that combines different types of treatments :

- **Traditional medical therapy** which will be developed in section 2.3.2.
- **Complementary therapy** are used alongside conventional medical treatments. It can help people with cancer to feel better, may improve patient's quality of life and help to cope better with symptoms caused by the cancer or side effects caused by cancer treatment (as for example aromatherapy, acupuncture, herbal medicine, massage therapy, visualization, yoga).
- **Alternative therapy** is used instead of conventional medical treatment (as for example laetrile administration, fasting, Gerson therapy, shark cartilage instruction...).

Conventional therapies to help in reducing tumor height and growth

The treatments offered to patients have continued to evolve along with scientific progress (for example regarding genes functionalities, pathways, markers and cell types), and theories of cancer evolution. And even if therapeutic support seemed slow and very painful for a long time before modern treatments we are familiar with.

In the 18th century, Scott John Hunter, one of the founder of modern surgery declared that if tumor had not invaded nearby tissues, it had to be removed. The discovery of general anesthesia and later antisepsis and asepsis have improved operations procedures and patient recovery and, in the same way boost surgical innovation.

Together with other complementary therapies, surgery is part of our therapeutic panel for the support of solid cancers. One of the first related therapy comes from Emil Grubbe and Wilhelm Roentgen in 1896 who showed that X-rays by creating DNA damage are able to kill cells less than one year after their discovery. Although X-ray do not allow any cell selectivity (both cancerous and

non-cancerous), cancer cells remain easier to kill because they are dividing much faster. However, a major limitation of this technology is that it is very inefficient once cancer had spread among nearby original tissues.

It was not until the 40's that the monitoring in the First World War of soldiers gassed with mustard gas shows that they suffered of a peculiar loss of lymphocyte cells. This targeted action on a particular cell type has inspired the use of nitrogen mustard to treat lymphoma. This treatment became the first licensed chemotherapy agent. Other drug has rapidly broadened cancer chemotherapy techniques. Especially since the discovery of oncogenes and tumor suppressors have shown how genes networks could be manipulated and that changes in their expression level caused major consequences in cell behavior and tumor growth. Discoveries lead way to molecules capable of inhibiting tumor growth (such as the Glivec [315]...).

Effectiveness of cancer therapy is frequently impaired by intrinsic or acquired tumor resistance to conventional therapeutics as the cytotoxic agents or ionizing radiation. Cellular mechanisms of therapeutic resistance are due to [191, 353] :

- Increase DNA damage recognition and repair
- Alteration of cell cycle checkpoints
- Impairment of tumor apoptotic pathways
- Reduced accumulation of cytotoxic chemotherapeutic agent through drug efflux mechanism (like ABC transporters [107])

To cope with solitary chemotherapy resistance, a number of therapeutic protocols have become accustomed to combine or mixed at several specific interval of times the provided drugs. The basis of chemotherapy is in principle to seek out cancer cells even if they are spread throughout the body, however, the high peripheral toxicity to cancer cells led to improve the specificity of action.

Targeted therapy to specifically block defective processes in cancer cells

CSC-LIKE subpopulation have been suspected to be at the heart of resistance to conventional cancer therapies [271]. To reduce the risk of relapse and dissemination, cancer stem cells-directed targeting strategies are currently under development and clinical studies as for example targeted immuno-therapy (monoclonal antibodies), small molecules, engineered oncolytic viruses, activated immune cells, blockade of CSC function, CSC-directed differentiation... Nonetheless, several factors should be carefully taken into account to design CSC-directed targeting strategies :

- Limit adverse affects on self-renewal and maintenance of physiological tissue due to similarities between CSC-LIKE and adult stem cells.
- Heterogeneity of CSC-like pool and potential difference of sensitivity among CSC pool.
- Development of resistance to targeted therapies.
- Efficiency will depend on prior reduction of tumor bulk population.

Most targeted therapies help treat cancer by interfering with specific proteins that help tumors grow and spread throughout the body. They treat cancer in many ways in :

- Helping the immune system destroy cancer cells.
- Stopping tumor cells from dividing and thus, slowing cancer uncontrolled growth.

- Stopping signals that help form blood vessels.
- Causing tumor and cancer cell death.
- Starving cancer of the hormones it needs to grow in preventing either the body from making specific hormones or the hormones from acting on cells.

To achieve this goal, some targeted therapies are already available on the market, and many others are currently under development. They include [8, 237] :

- Hormone therapies, as for example, regarding breast cancer, the aromatase inhibitors (anastrozole, ...), some selective estrogen receptor modulators (tamoxifen, ...), or estrogen receptor downregulators (fulvestrant).
- Signal transduction inhibitors, as for example the EGFR inhibitors (as for example cetuximab used to treat some lung, colorectal cancer), HER2 inhibitors (as trastuzumab to treat breast or stomach cancer), BRAF inhibitors (as vemurafenib to treat melanoma), ALK inhibitors (as crizotinib to treat some lung cancers) [178].
- Gene expression modulator.
- Apoptosis inducer such as bortezomib or carfilzomib which inhibit the proteasome.
- Angiogenesis inhibitor as for example bevacizumab or ramucirumab.
- Immunotherapies which include drugs that target the PD-1 and PDL-1 proteins, such as pembrolizumab, and drugs that target CTLA-4 protein, such as ipilimumab.
- Toxin delivery molecules as for example brentuximab vedotin or ado-trastuzumab emtansine.
- Cancer vaccines.
- Gene therapy [63].

To limit the apparition of a resistance to treatments (by neo-mutation of the therapy target or use of other pathways to achieve tumor growth), targeted therapy is usually prescribed in combination with one or more traditional chemotherapy drugs.

Thus, immunotherapy has been initiated from the collaboration of two researchers from Cambridge University, César Milstein and Georges Köhler (1975), who perfected the process of synthesis and purification of antibodies. Specifically, able to target cells in the body, man-made antibodies could be used to target tumors in carrying radio-active particles or chemotherapy drugs, or block growth of blood vessels that tumors need to grow. Thus, the panel in curative vaccine material consists in either vaccine based on whole tumor cells, or antigenic vaccines, or anti-idiotypic vaccines, or vaccines based on dendritic cells, or DNA vaccines.

To illustrate, Sipuleucel-T (Provenge[®]) was the first therapeutic vaccine for the treatment of prostate cancer in the metastatic stage and resistant to hormone therapy. Approved in April 2010 by the FDA and then by the EMA, it relies on the use of dendritic cells presenting antigens able to detect prostate tumor cells expressing the antigen Prostatic Acid Phosphatase (PAP). Another therapeutic vaccine developed by Olivier Adotevi and Christophe Borg, is in clinical trial phase. Named UCPVax (Universal Cancer Peptide) because of targeting the enzyme telomerase which is universally expressed by each cells, it will be offered initially to fight against recidivist metastatic lung cancer [105]. This method is now one of the most promising approaches in cancer.

Other therapeutic strategies derived from immunotherapy have gradually emerged to prevent the appearance of tumor. The concept is therefore based on stimulation and preventive education of the patient's immune system so that it can gradually fight ahead of any pathology. Thus, the administration of vaccines helps fight against an oncolytic virus, such as :

- Vaccination against hepatitis B continues to largely prevent liver cancer in areas of the world where the rate of infection by the hepatitis B virus is important [205].
- The vaccine against the most virulent strains of papillomavirus (HPV) 6, 11, 16 and 18 (Cervarix and Gardasil) for the fight against cancer of the womb [313].

Finally, when the cancer is located, some unconventional therapies are sometimes set up upon patient's request. For example, some have been designed at first in ex-USSR and are based on general inflammation of the body and detoxification mechanisms naturally activated by a body subject to drastic fasting. Short-term starvation associated to chemotherapeutic agents retards tumor progression and reduces side effects in cancer patients, and thus, extends the survival of advanced-stage cancer patients [177].

Progress continues with limiting side-effects of oncology therapies on adjacent tissues (and thus favoring personalized medicine and spatially targeted therapies) as one of the main objectives. However, we must highlight that all devised strategies today aim either to indiscriminately kill all tumor cells, to eliminate a particular phenotype, or to push towards phenotypes transformation. Up to now, there is still no plans to modify tumor and non-tumor cell interactions. In order that healthy rules of interaction and communication lead to the inhibition and the reversion of tumor growth towards a healthy tissue environment.

Theories presented above do not agree about the root causes of the emergence of potentially proliferative and migratory tumor cells. However, they share in common the idea that tumor cells are phenotypically diverse (*i.e.* for their expressed markers, their temporal expression, their morphology, their function ...) and they will need to interact and communicate (and for some with a healthy environment). This is the topic that we will develop in the next chapter.

Chapter 3

The issue of social relationships among cells

« Tout groupe humain prend sa richesse dans la communication, l'entraide et la solidarité visant à un but commun : l'épanouissement de chacun dans le respect des différences. »

- Françoise Dolto -

Communication is a process used to convey a message, *i.e.* uni/bidirectional transfer of information, between two entities in response to an internal or external stimuli with two main steps :

- A transmitter which transmits information to a receiver via a broadcast channel
- A recipient which indicates receipt or not of the message and the way it has been perceived

This oversimplified scheme applies to all communication systems for which the actor can be either animated or unanimated. If we focus on the principles of communication at the cellular level, the latter mechanisms enable one cell to influence the behavior of another. Indeed, this is the basis for functional coordination between unicellular organism as well as between cells in multicellular organisms. We position this chapter with respect to the study of the media of communications used among cells communities. Then, we will discuss the role of communications and present several examples for which communication among cells is essential to trigger tissue function and spatial organization.

3.1 From stimuli to answer, adaptation to the changing environment

Modifications of the internal or external environment trigger stimuli which activate receptors, cells, change in the activity of a cell, or an organism. Survival is based on this process which allows millions of cells to communicate and coordinate their activities to perform and regulate key body processes. And through this section we will detail the different types of communication processes between living cells.

Processes from the smallest to the highest scale (molecule to organism) are closely regulated so that internal conditions, *i.e.* at each scale of the organization, remain relatively constant. This concept of “internal medium”, namely homeostasis, have been proposed by Claude Bernard (1813–1878) and popularized by Walter Bradford Cannon (1871–1945) through its book entitled “The Wisdom of the Body” (1932) [43]. The general features of homeostasis regroup four main purposes :

- Steady-state in a system is not spontaneous, it requires mechanisms that act to maintain this constancy.
- Healthy steady-state conditions require that any tendency toward change automatically meets with signals that counteract changes.
- The regulating system consists in a number of cooperating mechanisms acting simultaneously or successively.
- Homeostasis does not occur by chance, but is the result of self-organization.

Homeostasis is often ill-defined as a state of internal constancy. In reality homeostasis is not a static state ; it is a dynamic ever-changing state which maintains core processes within a range [103]. This mechanism of actions is based on a set of steps that will be detailed thereafter.

3.1.1 Cells are sensitive to diverse source of stimuli

The cells of our organisms are sensitive to a variety of stimuli resulting from both the external and the internal environments. Some cells have been specialized in detecting, processing and transmitting signals. They are therefore equipped with special sensors of several kinds :

Sensitiveness to physical stimuli

Some receptors are sensitive to **mechanical stimuli** [147]. And such forming the group of mechano-receptors they can be found within the skin structures, the vestibular system of the inner ear, alveolar walls in juxtaposition to the pulmonary capillaries or the ligamentous system. These latest receptors are providing information to the nervous system about touch, pressure, vibration, and cutaneous tension.

The skin and the hair follicles harbor a variety of morphologically distinct mechano-receptors and classified following their threshold of activation (time-dependent opening of ion channels in response to a stimulus, typically membrane depolarization) and the rate of adaptation (phenomenon of sensory receptor adjustment to different levels of stimulation ; critical for allowing sensory systems to operate over a wide dynamic range). They include the Pacinian and the Meissner corpuscles, the Merkel discs, the Ruffini endings and the root hair plexus [30].

Proprioception, described as the “position-movement sensation” or “the sense of locomotion and muscles” by Julius Caesar Scaliger (1484–1558), means the ability to sense both the relative position of the different parts of the body and the strength of effort needed to perform a movement. Diverse proprio-receptors are found within the skeletal striated muscles, the tendons and the joints. Muscles splindles, located within the belly of striated muscles, detect any change of the length and help both the brain to determine the position of the body parts and the motor-neurons to regulate the contraction of the muscles through activation of the stretch reflex. Golgi tendon organs, also named neuro-tendinous spindle, are proprioceptive sensory receptor organs which take place at the connection between muscle fibers and tendon. During muscle contraction, this proprioception receptor senses any exceptional large increases of tension of the collagen fibrils within the tendon organ and sets up a negative feedback loop to protect muscles’ integrity.

Some cells are sensitive to **electromagnetic stimulation**, like light (visible electromagnetic radiation) [273]. These photo-receptor cells are specialized neurons present within the retina. They are capable of photo-transduction, meaning that they are able to convey light (through the absorption of photons) into signals (leading to membrane hyper-polarization of rod cells membrane and the activation of bipolar cells and then ganglion cells). Three types of photo-sensitive cells have been described in humans, namely the rods, the cones and the intrinsically photosensitive retinal ganglion cells. These photo-receptor cells contain specialized photoreceptive pigments (such as rhodopsin) which change their conformation through signals and therefore transform the physical energy from light into electrical signals thanks to the activation of the transducing pathway.

Changes of innocuous environmental temperature is monitored by both cold and warm-sensitive thermo-receptors that are principally expressed. These nociceptors particularly innervate (facial) skin, cornea and urinary bladder and manage temperature sensitivity thanks to feedback response of the hypothalamus. The pathways and mechanisms that regulate body temperature are not well understood. However, studies on transient receptor potential channels (TRP channels), located mostly on the cells plasma membrane are thought to behave like microscopic thermometers to feel hot and cold temperatures [215, 327].

All cells living in contact with others cells tissue are sensitive to **tissue tension, pressure and geometrical coherence of the cells** which induce the correct formation of cell polarity and tissue structure [306]. Few cells apart from the hematopoietic and immune system are free to move, indeed most of the cells are linked through cell-to-cell or cell-to-matrix junctions which form a network of strong anchoring adhesions (as for example within epithelial cells, muscular fiber and cardiac cells). These dynamic protein complexes are responsible to the organization of cells cytoskeleton and its connection with either the ECM or other surrounding cells adhesive structures. They are also involved as regulator centers to monitor the availability of signaling molecules and thereby, participate to fundamental processes such as cell proliferation, differentiation, tissue homeostasis, cell migration during wound healing phenomena, and morpho-genesis [32].

Sensitiveness to chemical stimuli

Chemical signals are used by many types of plants and animals and are released into the air, or in some cases into water, in very small quantities. These chemicals are extremely specific stimuli. Cells respond to these stimuli however little is known about the core characteristics of chemicals and how their stimuli trigger electrical signals to be processed by the Central Nervous System (CNS).

One famous chemical stimuli is the **olfactive molecules** which activates the olfactory system (same results are obtained through activation of the gustative system). This latest can be divided into two components :

- The peripheral system sensing an external stimulus and encoding it as an electric signal in neurons
- The central system integrating and processing signals in the CNS.

Thanks to the peripheral system, odorants bind to the olfactory receptors (of the olfactive epithelium) which do not present so much specificity, *i.e.* and thus display affinity to a wide range of odorant molecules. Once activated, olfactory receptors trigger activation of neurons through electrical signals up to the central system (for further processing).

Some cells, and more precisely neuronal cells, are activated through **endogenous neurotransmitters** released within the chemical synapse. Synapses are neuron-to-targeted cell

junctions (which are either neurons, muscle cells, or gland cells). Also, pre-synaptic neuron release chemical neurotransmitters within the synaptic cleft and which further bind to the receptors of post-synaptic target cell. This phenomenon was discovered by Ramón y Cajal (1852–1934) and enriched by Otto Loewi (1873–1961). A neurotransmitter influences trans-membrane ion flow (cell depolarization) either to increase (excitatory synapse of type I) or to decrease (inhibitory synapse of type II) the probability that the targeted cell will produce an Action potential (pA). Type I and II neuron-to-target cell junction differ on their appearance, their location and the shape of their released synaptic vesicles. Different types of neurotransmitters are synthesized and released by pre-synaptic neurons and can be classified thanks to their chemistry characteristics (amino acids, neuro-peptides, mono-amines, gaso-transmitters, purine-based molecules). Effects of a neurotransmitter system depends on the connections of the neurons (excitatory or inhibitory) and the properties of the binding receptors, either excitatory or inhibitory.

The release of hormones can be triggered by **humoral and hormonal stimuli** which are regulated by feedback mechanisms to maintain homeostasis, meaning to regulate body internal environment and to maintain the stable conditions necessary for the body to operate. These stimuli are either changes in the blood of certain minerals or nutrients, actions of other hormones leading to rhythmic releases of hormones, or neurological stimuli leading to hormone release in short bursts or spurts as required. The correct range of hormones concentration must be maintained through feedback mechanisms, because they can trigger powerful effects, even at low concentration, on the body.

Some soluble molecules as carbon dioxide (and therefore by the same occasion change of pH level) or dioxygen stimulate chemo- and baro-receptors to regulate breathing for example [93]. Indeed, we find many nerve cells which collect information on chemical substances circulating within the blood. Regarding the breathing phenomenon, it exists two types of chemo-receptors, the central one located in the brain bulb and the peripheral one located in the carotid and aortic corpuscles [115, 326, 333]. Oxygen deficit is not simply a patho-physiological incident, indeed in the same time this is a physiological event. Adaptations to high altitude or physical endurance exercise are examples of “physiological oxygen deficit” which lead to increased expression of hematopoietic hormones (such as Erythropoietin (EPO)) or angiogenesis factors (such as VEGF) [333]. EPO production is controlled at the transcriptional level and maintains red blood cell mass by promoting the survival, the proliferation and the differentiation of erythroid progenitors. The relationship between the O₂ content of the blood and erythropoiesis was first described by Francois-Gilbert Viault in 1890 (1849–1918), who observed an increase in red blood cells number after a trek in the highlands of Peru at 4500 of altitude [185].

Specialized cells sense changes in the **extracellular concentration of specific ions** (such as Ca²⁺, Na⁺, Cl⁻... to maintain homeostasis of charged ions because they are key actors of most signal transducing pathways [47].

Extracellular calcium-sensing receptors (CaSR) are members of the family of G-protein coupled receptor which are expressed at the surface of parathyroid and kidney cells. These receptors play a major role in the maintenance of the concentration of physiological calcium (Ca²⁺) by regulating the circulating levels of parathyroid hormone ((PTH) for both synthesis and secretion steps and parathyroid cells proliferation) [259]. This phenomenon has been identified in 1993 by Edward M. Brown [36]. Ca²⁺ ion plays several roles in signal transduction, muscular fiber contraction and neuronal transmission. Sodium is an essential nutrient that regulates blood volume, blood pressure, osmotic equilibrium and pH. Moreover, sodium ions are key actors in neuron function and osmo-regulation between cells and the extracellular fluid. The distribution of sodium ions is mediated by the active transporter pumping ions Na⁺/K⁺-ATPase.

Through cell-to-cell transfer of nucleic molecules or organites

Finally, inter-cellular exchanges of genetic information are newly described mechanisms that cells are using to communicate.

Horizontal gene transfer has previously been described in bacteria. Indeed, naturally competent bacteria are able to take up exogenous DNA and undergo genetic transformation thanks to three different means [51, 287] :

- The conjugation mechanism, discovered in 1946 by Joshua Lederberg and Edward Tatum [176]. It consists on transferring genetic material between bacterial cells either by direct cell-to-cell contact or by a bridge-like connection between two cells.
- The transduction process consists on transferring DNA from a bacterial cell to a second one through a bacteriophage virus vector. This mechanism has been discovered by Norton Zinder and Joshua Lederberg in 1951 in the bacterium *Salmonella Typhimurium*.
- The transformation was first demonstrated in 1928 by Frederick Griffith on *Streptococcus Pneumoniae* bacteria that could become highly virulent when putting in contact with heat-killed virulent strain of bacteria. It consists on modification of the genetic inheritance of one bacterium thanks to its direct uptake of exogenous genetic material from surroundings bacteria through cell membranes transfer.

Regarding human cells, similar genetic transfer has been newly characterized, mainly small RNAs such as microRNAs, mitochondrial DNA... [214]. These exchanges of short or longer distance by entering the bloodstream (small RNAs can be detected in blood and other body fluids such as urine, saliva and milk) occur through either lipid or lipo-protein complexes, extra-cellular vesicles such as apoptotic bodies or exosomes, intimate membrane contacts, or tubules transfer such as gap junctions [78, 322].

This mechanism of cell-to-cell behavioral regulation is particularly important in cancer. Indeed, tumor cells produce abundant extra-cellular vesicles, with the potential to influence the behavior of surrounding healthy cells in order to facilitate tumor growth, metastasis or immune evasion.

We have just pointed out a set of molecules that ensure communication between cells in order to allow the cells to adapt to changes in their environment, the tissues to regulate their organization and homeostasis and whole organisms to globally function and to survive. We will now discuss the physical methods underlying inter-cellular communication.

3.1.2 How is communication established between two cells ?

The transmission of a message (from either the outside environment or the internal environment of the body) can be achieved only in two ways : through direct contact or at distance. For these purposes, a set of structures exist to allow all physical, chemical and genetic stimuli to influence inter-cellular communications.

Local communication between plasma membranes

Cell membrane separates the interior of cells from the outside. They are formed of phospho-lipid bilayer with embedded proteins and are involved in diverse cellular processes such as cell adhesion, ion conductivity and cell signaling, and serve as the attachment surface for several extra-cellular structures and intra-cellular cyto-skeleton.

Specialized cell junctions occur at points of cell-cell and cell-matrix contact in all tissues, and they are plentiful in epithelia. Cell junctions can be classified into three functional groups :

- The tight or occluding junctions, also named *zonula occludens*, that form a seal between the membrane of two neighboring cells. This seal consists in trans-membrane proteins such as claudin and occludin family proteins which establish para-cellular barriers to control transcellular and paracellular transport of molecules, but prevent passive flow between cells.
- The adherent or anchoring junctions maintain cell-to-cell pressure and tissue tension :
 - Some, made of cadherin (a trans-membrane glyco-protein), are named *zonula adherens* and their function is to strengthen and stabilize the occluding bands and to link cyto-skeletons of adjacent cells.
 - Others consist of desmosomes or *macula adherens* and permit strong adhesion among the cells.
 - Finally others correspond to hemi-desmosomes which connect cells to the underlying basal *lamina*.
- And the communicating or gap junctions composed of connexin proteins which per group of 8 units form a complex called connexon. Their main functions consist in both the regulation of direct transfer and the exchange of nutrients and signal molecules between neighboring cells.

Cell projections represent another mode of communication that is used by various cell types.

Cytonemes and tunneling nano-tubules projections consist on thin elongated intercellular membrane bridges that function in cell–cell signaling and inter-cellular transport [281].

Cytonemes or filopodial bridges are thin filopodia-like structures extensions of cell membrane and composed of tightly bundled parallel actin filaments. They connect neighboring cells via mechanisms of adhesion, which enable ligand-receptor-mediated transfer of surface-associated cargoes from cell to cell. They have been proposed to operate as long-range structures to enable a morphogen-patterning gradient for signaling membrane-associated molecules.

By contrast, tunneling nanotubes represent a subset of F-actin-based transient tubular connections that allow direct communication between distant cells, meaning that they mediate continuity between the cytoplasm of remote cells. Also, tubular conduits are established between cells to provide exchange of both cell-surface molecules and cytoplasmic content such as organites (as for example mitochondria [298]). They have been observed in various cell types including neurons, myeloid, lymphocyte T and endothelial progenitor cells...

Local to long-range communication through receptors on the cell

Cell signaling has been most extensively studied in the context of cells communication among single organism. However, it also occurs between the cells of two different organisms in a “symbiotic” context, as for example :

- During pregnancy between embryonic and uterus cells.
- In the gastrointestinal tract between bacteria and epithelial/immune system cells.

Many cell-to-cell signals are water-soluble and cannot penetrate membranes. Instead, these signals are received by trans-membrane proteins protruding out from the cell surface that change shape when bound by specific molecules. Cell surface receptors are responsible for the specific binding of an extra-cellular signaling molecule. These later trigger a chain of events within the cell and activate one or more intra-cellular signaling molecules, these events inducing changes in cells behavior. Cell surface receptors fall into three main classes : namely the ligand-gated ion

channel receptors, the enzyme-coupled receptors and the G-protein-coupled receptors. Receptor proteins embedded in the plasma membrane possibly undergo conformational changes when they bind specific signal molecules, triggering a chain of events within the cell. Signaling between cells is subdivided as follows :

- Intracrine signals produced by and staying within the target cell.
- Autocrine signals produced by the target cell. They are then secreted and affect the target cell itself or similar cells via receptors, as for example immune cells.
- Juxtacrine signals are transmitted along cell membranes via protein or lipid components and are capable of affecting either the emitting or neighboring cells.
- Paracrine signals target cells in the vicinity of the emitting cell such as neurotransmitters or histamine for example.
- Endocrine signals are triggered by hormones which travel through the blood to reach distant cells among the body.

Passive *versus* active message transfer

The cell membrane is selectively permeable to ions and organic molecules and controls the entrances and exits of molecules needed for survival. The cells employ a number of transport mechanisms across the membranes, and thus the movement of substances can be :

- Either “passive”, meaning that it occurs without the input of cellular energy. Some substances (such as small molecules like carbon dioxide or oxygen, ions) can move across the plasma membrane by diffusion. Because membranes act as a barrier for certain molecules and ions, this might trigger differential concentrations on the two sides of the membrane. The resulting gradient drives passive transport and its rate is limited by the permeability to the message.
- Or “active” which require the cell to expend energy to pump across the membranes (such as nutrients like sugars or amino acids and products of metabolism) inside and outside through trans-membrane protein channels and transporters. Proteins channel, also called permeases, are usually quite specific, recognizing and transporting only a limited group of chemical substances, often even only a single substance.

3.2 Communication is unflinching of the spatial (self)-organization

Spatio-temporal properties of communication pathways are essential.

One example takes place through organogenesis, and particularly through somitogenesis process (during which a pair of somites arise every 30 minutes in zebrafish, 90 minutes in the chick embryo and 100 minutes in snakes). Indeed, somitogenesis is regulated by the oscillatory expression of a network of both genes and gene products. These consistently timed-expressions cause cells to oscillate between a permissive and a non-permissive state. Indeed, the molecular clock is provided by oscillatory expression of several homeotic genes members of the FGF family, Wnt and Notch pathway, as well as targets of these pathways. Timing output informs segments of proper length and number, which characteristics are specific to species (such for example 30 in

zebrafish, 295 in corn snake) [106]. In conclusion, temporal rhythms and the level of expressed homeotic regulators are used to spatially organize organisms through periodic patterning.

Within both vertebrates and the major phyla of invertebrates, basic regulators and triggers of organism development are roughly the same. These homologous molecules (among species) both define specialized cell types, body region and prime body patterning. Selective genes expression controls the four essential processes by which an organism is constructed :

- Cell proliferation to produce many cells from one.
- Cell specialization to create cells with various characteristics at several positions.
- Cell interactions to coordinate the behavior of one cell with that of its neighborhood.
- Cell movement to rearrange the cells and form structured tissues and organs.

Thus, cell-to-cell interactions and connections play meaningful role in :

- The spatial organization of complex and precise patterns during the development of multi-cellular organisms.
- The expression of a genotypic and/or phenotypic diversity within the tissues [70].

One outstanding example concerns the stem cell niche and its architecture. Close interactions between stem and progenitors' cells regulate stem cell rate of division [28]. Understanding how the role of each cells is defined requires to identify which cell phenotypes are regulated by cell-to-cell interactions and phenotypic regulator. This step will further necessitate to investigate the spatial organizations and the social rules which govern phenotypic states [118]. However, conventional methods for assaying cell–cell interactions are mainly applicable at a cell population level. Nonetheless, these measures are limited in elucidating how cell phenotypes organize themselves. This lack results in an incomplete view of the geometry of cell–to-cell interactions [28]. Thus, assessing how cells interact at the single cell level is a notable objective to provide new insights into developmental biology that may be missed by traditional population-averaged studies.

Finally, spatial organization fit also into sub-cellular level. Indeed, correct co-locations of enzymes and reactive molecules, products, and transporters are necessary for proper initiation and triggering of the genetic regulatory network. It results in sub-cellular compartmentalization (*i.e.* within organelles and micro-compartments formed via the interactions of enzyme groups), spatially restricted micro-domains enriched in molecules or active proteins to create inhomogeneous signaling over (either animal or plant) cell's cytoplasm [161, 226, 308]. This may allow the cell to spatially segregate inputs and outputs processing, as part of its signal-handling repertoire. On example corresponds to regulation of NF κ B expression which is achieved through sophisticated spatio-temporal coordination of ubiquitylation events. This orchestration is essential to the development of specific tissues and cell types, especially among nervous cells and axonal differentiation [10, 108]

Thus, the spatial organization of multi-cellular organism relies on a differentiated spatial organization of intracellular components.

Self-organization process was formulated in 1947 by the cybernetician William Ross Ashby (1903–1972) and was at the basis of the “epigenetic landscape model” mentioned in section 1.2.2 on page 10. He stated that “any deterministic dynamic system will automatically evolve towards a state of equilibrium that can be described in terms of an attractor in a basin of surrounding states. Once there, the further evolution of the system is constrained to remain in the attractor. This constraint on the system as a whole implies a form of mutual dependency or coordination between

its constituent components or ‘subsystems’ [11]. In other words, lower levels of organization depend on the control from upper organizations.

Recent years have been focused on processes at smaller scales (*e.g.* sub-cellular, molecular, atomic level). However, changes are on the road and new integrative studies are arguing another objective. These later consist on making use of knowledge from molecular biology and combine them with the characteristics of whole organisms, and thus to measure the outcomes on developmental and behavioral processes. More and more, the decision-making and the potential adaptability of the organisms is recognized as an important driver of evolution. Thus, this renewed focus on the whole organism (as an alternative of the gene-focused older view) is changing the face of evolutionary biology.

Moreover, homeostasis is not a static process which would hinder any evolution, any changes. Rather, it allows the regulated system to evolve in order to deal novelties and manage potential disruption [339]. However huge disruptions in homeostasis can affect biological systems at all scales (either molecules, pathways, cells, and even ecosystem of cells *i.e.* organisms)...

Defects within the communication framework trigger tumors development and cancer dissemination

Cells are at the center of communication processes, constantly sending out and receiving signals. However, defects on cell-to-cell communications can arise at all stages :

- Transmitter cells can fail to send out signals at the proper time.
- Signal can fail to reach the target cell.
- Targeted cells can be unable to properly respond to a signal.
- Some deficient cells can respond even if they do not have received any signal.

In fact, most diseases (as for example, in type I diabetes for which pancreatic cells that produce insulin are lost leading to the accumulation to toxic levels of sugar in the blood, or in multiple sclerosis for which the protective myelin sheath around nerve cells in the brain and spinal cord are destroyed. The affected nerve cells can no longer transmit signals from one area of the brain to another) involve at least one breakdown in cell communication.

Cancer development is no exception to this rule [266]. As depicted in chapter 2 on page 13, cancer cells might arise from flawed intra-cellular communication pathways, defective and differential inter-cellular exchanges of small molecules like micro-RNA through gap junctions for example [87]

Once tumor cells are well implanted, they possess or opportunistically use powerful mechanisms to “anticipate”, and even “learn” from variations and challenges of the environment. For example, cancer cells develop drug and immune resistance to exhibit traits of a persistent phenotype. Moreover, they display many other collective behavior capabilities and cooperative strategies necessary for survival under extreme stress (as for example from the immunity, several drugs...). These characteristics present cancer cells as smart communicating cells and tend to portray tumors as societies of cells capable of making decisions [22, 42]. But how are organized tumor cell phenotypes? What determines cells dynamics (*e.g.* cell movements, phenotypes changes, cell division...)? And how does this affect cells adaptation ?

Thesis work schedule and limits

*«L'homme raisonnable s'adapte au monde ; l'homme déraisonnable s'obstine à essayer d'adapter le monde à lui-même. Tout progrès dépend donc de l'homme déraisonnable»
- George Bernard Shaw - Maximes pour révolutionnaires*

This chapter entitled “Thesis work schedule and limits” will be particularly enriched with questions that will be tackled through the next few chapters. Then I will introduce the parts and chapters layout which correspond to the main body of the research work. Thus, the rationale which has guided all of this past three years of work will be the conducting wire of the first section. And the final section will list some small tricks to make it easier to understand and read this manuscript. I will take this opportunity to clarify the color codes which guide the presentation of the results and small adds-on that will enhance your reading through the electronic version.

4.1 As a rephrasing of the general question

4.1.1 Diversity of cells phenotypes within tumors and diversification

Regarding cell phenotypic diversity described in mammary tumors, a panel of surface markers was widely used to clustered different functional cell subpopulations : the CD24, CD29, CD44, CD49f, CD133 and Epithelial Surface Antigen (ESA) (also named CD326 or EpCAM) markers [90, 117, 246]. The expression of these markers, especially their level of expression, permitted to describe some cell groups, to study the expression of certain genes and functionality (*cf.* part II on page 47).

Cell subpopulations of interest, in this work, have been identified and isolated. Thus, we reduced the overextended phenotypic characterization complexity by only monitoring two markers : CD24 and CD44.

One of the primary objectives of this thesis was to quantify, two cell populations clustered according to the level of expression of the CD44 marker (CD44-low and CD44-high). Indeed, in these early studies, the view we adopted was to work at the scale of the cell population (and not on the scale of the cell taken individually). Thus, the measurement parameters that we quantified and monitoring the short length are the proportions occupied by the above-mentioned clusters and the level of expression of both CD24 and CD44 markers in the cell clusters. We took the

opportunity to vary the initial conditions of the studies for :

- The degree of heterogeneity (*i.e.* diversity in the representation of cellular phenotypic subgroups) initial (at $t = 0$ of the study) of culture.
- The conditions of cells culture in 2D or 3D.

However, these studies across populations have shown that the composition of cell populations varies with the phenotypic expression of initial cultured cells. These observations have gradually led us to wonder about the role and influence of an individual within a population.

The second part of our study can be formulated as follows : how is the phenotype of individual cells influenced by the way the phenotype of the group? And conversely, how the phenotype of an individual influence the expression of the group? We abstracted ourselves from the regular reasoning which fails to take into account the relationships between cells and their potential impact in the decision of individual and collective decisions. To caricature, the orthodox point of the view is to study the phenotypes of cells, either collectively or individually, and regardless of the organization of cells relative to each other. This means : omitting the possibility that cell mutually determine each other and only undergo self-determinism over time (and in a defined temporality) a number of characteristics (such as markers, one or several functions) based on information initiated or taken up by neighboring cells. So throughout this second part, we wanted to reinstate the pressure related to the community to understand the plastic behavior of both cell populations and individuals.

4.1.2 Keywords of this thesis

This work is part of a context of characterization of the degree of **cell phenotypic heterogeneity**, and the **plasticity** of cells to take decisions and **transit from one phenotypic state to another**. A particular phenotype, called **stem cells**, presents a special interest, by definition, because of its potential ability to reform a heterogeneous population and to self-renew. Cell fate question to what extent cells, taken as individual, are **autonomous or mutually interdependent** to take decisions regarding their future.

4.2 How is organized the thesis statement

This thesis reports three different part of work and which are entitled as following :

- Natural history of cells diversity
- Clonogenicity, Diversity and Environment influences
- Strategies proposal to search social rules that govern cell phenotypic changes

First part “Natural history of cells diversity” This first part consists in three chapters in which we study the equilibrium of cell populations in the expression of markers CD24 and CD44, all in a context of 2D cell culture. We will vary the culture settings in order to understand at first the degree of influence held by the phenotypic nature of the plated cells at the initial state on the expression of CD24 and CD44 marker of cell populations at equilibrium. Thus our studies will take place starting from an initially heterogeneous cell population in chapter 5 on page 49 and initially homogeneous in chapter 6 on page 71. Finally, the last chapter will measure the degree of influence exerted by the phenotype of the parental population in the behavior of a daughter population on the expression of CD24 and CD44 markers of the latter.

Second part “Clonogenicity, Diversity and Environment influences” This second part consists in two chapters in which we study the behavior of SUM149-PT and SUM159-PT cells grown in 3D conditions. The first chapter will review the clonogenic capabilities of CD24-/CD44, CD24-/CD44+ and CD24+/CD44+ cell lines and subpopulations described in the previous section. The second chapter will characterize the phenotypic expression (through the characterization of the markers CD24 and CD44) of cell populations from 3D cultures.

Third part “Strategies proposal to search social rules that govern cell phenotypic changes” Finally, the third part consists in three chapters. It will question the notion of independence or dependence in decision-making. The first chapter will explain the issues and working hypotheses raised following the previous results, with the intention to monitor in real-time the expression of phenotypes across each cell. We will use this chapter to present the characteristics of the tools we hope to set up and which will be validated in chapter 2. Then the final chapter will present our methodology to quantify the degree of influence of the social rules on the choice of phenotypic expression of cells.

4.3 Little tricks for your convenience

This report is intended to be as easy as possible to read. And to achieve this purpose few color codes, and some interlinks and hyperlinks for the electronic version have been added. Therefore, I itemized them so that you could take into account these extras as early as possible during your reading.

4.3.1 Supplementary data, the Ali Baba cavern

Appendices contain a number of experimental data analysis and additional informations :

- The validation of two cell lines we studied (namely SUM149-PT and SUM159-PT).
- The library of protocols because I considered that separating the technicalities from the overall methodology would make the speech more understandable and digest.
- Methods put in place to build the molecular tools presented in Strategies proposal to search social rules that govern cell phenotypic changes.
- Various tables bringing together the references of the oligos, antibodies and molecular constructs, and finally.
- The programming codes written by very kind friends to facilitate the analysis of collected data.

4.3.2 Apply a color touch that makes sense

While it is true that colors brighten the environment, touches of color that stakes out this writing are also intended to facilitate the reading and the understanding of the graphics. The note below falls particularly within the framework of ??.

A color per phenotype population

In our experiments we focused on the expression behavior of three cell subpopulations derived from both SUM149-PT and SUM159-PT cell lines. Thus we linked to each phenotype (and this independent of the cell line) a specific color, as follows :

- CD24-/CD44- phenotype will be in **orange**.
- CD24-/CD44+ phenotype will be in **blue**.
- CD24+/CD44+ phenotype will be in **deep green**.

A color code for flow cytometry quantification

A plurality of quantitative measurements are associated with the data from flow cytometry. To ease the reading of the data, we have developed a color code and graphics system associated with the data of interest. Thus, during each acquisition, the cells were clustered in two groups according to their level of expression of the CD44 marker, as follows :

- CD44-low expressing cells will be in **orange**.
- CD44-high expressing cells will be in **light green**.

Then, each cluster (CD44-low and CD44-high) was tested against both CD24 and CD44 markers, and data for these analyzes were differentiated in such a way that :

- Monitoring the CD24 marker as an open circle, and the curves, as dashed line.
- Monitoring the CD44 marker as a full circle, and curves have full line.

4.3.3 Hyperlink as ways to the good direction

Likewise, we tried to limit repetitions to make the smoothest possible playback. That is why we have set up a hypertext links framework to facilitate the comings and returns between various sources of information.

Figures to text and *vice versa*

We added references to figures in order to illustrate the text. So by clicking the figure ID inserted as reference, you will automatically be redirected to related figure (and this even if it is part of the additional data).

Link to previous discussions

Furthermore, some sections (such as protocols, experimental methodologies or analysis, and the main conclusions) were exploited repeatedly in this manuscript. To avoid repetition that might be boring, we preferred to make referrals. Thus, each section or part, to which we referred, is associated with a hyperlink through which you can easily navigate to find the information of interest.

Bibliography is available in one click

Finally, we made extensive use of the literature, each referenced after supplementary data. References ID corresponds to the numbers in brackets (by alphabetical order of bibliographical references). A click on the corresponding numbers will instantly direct you to the sources used including the internet link.

Part II

Natural history of cells diversity

*«A mesure que l'homme progresse, il a de plus en plus conscience qu'il est maître de lui-même»
- Amanda W.P. Guruge -*

We will investigate the intricacies of cancer cells, of phenotypes heterogeneity, and heated debates on stem cells, phenotype transitions and population dynamics. And this is directly through the introduction to this section entitled “Natural History of Cells diversity” we will get acquainted with two cellular models that we will follow throughout these pages.

The study which is hereby presented involves two breast cancer cell lines named SUM149-PT and SUM159-PT. These latter cell lines have been derived by Steve Ethier in 1993 and are currently available from the Asterand Company [20, 81, 82, 224].

The SUM149-PT cell line was developed from invasive ductal carcinoma from a patient with ER negative and PR negative inflammatory breast cancer. SUM149-PT cells are very sensitive to the $TGF\beta$ -response and present an Erythroblastic Leukemia Viral Oncogene with ERBB-1 named EGFR, ERBB-2 named HER-2, ERBB-3 named HER-3 and ERBB-4 named HER-4 (ERBB) receptor status as EGFR++ ERBB-2,3 +ve, ERBB-4 -ve. However, no oncogene amplification could be found despite the fact that a deletion of a thymine residue at position 2288 of Breast Cancer 1, early onset (BRCA1) gene has been identified [20, 80]. The cell line is immortal and expresses luminal cytokeratins 8, 18, and 19 consistent with its origin from luminal breast epithelial cells [225]. SUM149-PT has been shown to form tumors in nude mice [89, 109, 113, 200].

The SUM159-PT cell line was developed from a primary tumor from a patient with ER negative and PR negative anaplastic carcinoma of the breast. SUM159-PT cell line is very sensitive to the $TGF\beta$ -response too and presents an ERBB receptor status as EGFR +ve ERBB-2 +ve. These cells present large cytogenetic recombinations, deletions and translocations resulting in major gene amplifications as the oncogene Cellular form of Myc, similar of the Myelocytomatosis Viral oncogene (v-Myc) (C-MYC) [20, 92]. This cell line has been immortalized and expresses luminal cytokeratins 8, 18, and 19 consistent with its origin from luminal breast epithelial cells. SUM159-PT is as well known to form tumors in nude mice [89, 109, 113].

Recent genomic studies have characterized the cell heterogeneity of human breast tissues based on cell surface expression of CD24 and CD44 markers, using epithelial cells derived from the pleural effusions of primary tumors (reduction of mammary tissues) [90, 117, 285]. Also, two subgroups of cells phenotypes have depicted differences on gene expressions :

- CD24-/CD44+ cells exhibit features of basal cells and express genes that are involved in motility.

- Whereas, CD24+/CD44+ cells display characteristics of more differentiated luminal epithelial cells and express genes involved in hormone response.

Using flow cytometry, breast cell lines can be grouped into three classes based on CD44 and CD24 expression profiles [89] :

- Class 1 cell line referred to luminal mainly expresses CD24 markers and exhibits a more differentiated morphology and luminal cytokeratin expression, which is consistent with their luminal-type classification.
- Class 2 cell line referred to basal/mesenchymal consisting of SUM159-PT cell line for example presents more than 90% CD24-/CD44+ cells and exhibits a spindle-like appearance, consistent with their basal-type classification.
- Finally, the class 3 cell line comprising the SUM149-PT cell line exhibits two distinct CD24/CD44 populations (*i.e.* CD24-/CD44+ and CD24+/CD44+) and contain distinct populations of cells with either basal-like or luminal-like features.

SUM149-PT and SUM159-PT breast cancer cell lines present a well-established phenotypic heterogeneity. Indeed, in its 2008 paper, Fillmore et al. could separate three subphenotypes from these two cell lines by combining the cell staining of three different surface markers, respectively CD24, CD44 and CD326. Thus, CD24+/CD44+/CD326- and CD44+/CD24+/CD326+ populations are depicted as luminal cells, CD24-/CD44+/CD326- populations are mentioned as basal cells and the CD24-/CD44+/CD326+ may be enriched in stem-like cells [89, 113, 117].

A hierarchy is intended to govern the rules of cells functionality thanks to criteria of differentiation or stemness both in healthy and tumor tissues [234, 311]. And thus according to the stem cells model, also named “the initiating-cells model” (*cf.* section 2.2.2 on page 20), a stem cell is at the basis of the whole cell’s organization to compose a tissue. These stem cells are able of infinite self-renewal and show forceful phenotype plasticity characteristics [89, 171]. Depending on the rules of symmetric and asymmetric cell divisions, from the so-called «cell-of-origin», and then the committed progenitors they are derived, up to the last cells issued from the speciation tree, cells are differentiating and lose their abilities to indefinitely divide, to self-renew, and to transit from one phenotype to another one [208].

However, new data challenges this irreversible and linear theory of “cell differentiation tree” with the observation of transitions from differentiated cells to stem-like cells (in two breast cancer cells) [48, 113]. Both of these publications, somewhat defiant respect to cell differentiation knowledge and innovative therapeutic strategies, have largely sparked my interest to start.

And in order to understand more deeply the dynamics of cells populations’ plasticity, I will firstly present a hitherto non-described cell phenotype in a cancerous environment. We will then look at the dynamics of cell plasticity of both homogeneous and heterogeneous populations. Moreover, we will focus on the dynamics that characterize how cell populations once again become heterogeneous. Finally, we will question how a balance is set up within a dynamic system based on previous theoretical models.

Cell line heterogeneity. Definition of a separable phenotype : CD24-/CD44-

« Tout le monde savait que c'était impossible à faire. Un jour est arrivé quelqu'un qui ne le savait pas, et qui l'a fait. »

- Churchill -

We will firstly characterize the behavior of SUM149-PT and SUM159-PT cells within the CD24/CD44 space of expression. We will start by qualitatively describing how cells populations move within this surrounding expression space.

This discussion will lead us to talk about a new cell phenotype. This cell subpopulation could be isolated from both SUM149-PT and the SUM159-PT cell lines. Then we will define a quantitative methodology to characterize the behavior of these two cell populations and follow their evolution over time. We will eventually discuss the degree of cell (in)dependence regarding probabilities of the triangle of phenotype transition decisions.

5.1 SUM149-PT and SUM159-PT characterization for CD24 and CD44 markers

SUM149-PT and SUM159-PT are two cell lines derived from breast cancer tumors, and numerous teams have investigated their characteristics using two markers CD24 and CD44 [20, 48, 82, 89, 113, 356]. A characterization using flow cytometry tools enables us to qualitatively assess the heterogeneity of expression of SUM149-PT and SUM159-PT cell lines for these two markers as described in the literature (*cf.* fig. 5.1 on the following page). Thus, the two cell lines significantly differ in their expression behavior within the space of CD24 and CD44 markers.

5.1.1 Working assumption relative to cells heterogeneity

In SUM149-PT cell line, more than 85% of the cells strongly express both CD24 and CD44 markers. The expression intensity for the CD44 marker is in average 300 times greater compared to unstained cells, while CD24 marker is expressed with a mean intensity 24 times higher compared

to an unstained population. The expression level of CD44 exhibits a very weak relative dispersion, while the expression of CD24 marker is much more spread out. Besides the dominant population, we observed a cell strip of almost 10% of the population decreasing of nearly one log its level of expression for the CD24 marker. In parallel to these cell populations, a cell population containing on average 1% of the population of living cells stands out, *i.e.* CD24-/CD44-. Indeed, the latter, according to my knowledge, has never been described in the literature in breast cancer and does express few or even no CD44 and CD24 markers (*cf.* fig. 5.1 on the left).

The cell line, named SUM159-PT, is characterized by about 85 to 90% of cells for which both the expression of CD44 marker is on average 1400 times higher than unstained cells and the mean intensity of CD24 expression is approximately 18 times more important than in unstained cells. This local maximum is highly condensed for both the CD24 and the CD44 markers. Nonetheless, this cell line presents a high level of heterogeneity for the expression of the markers of interest and thus two spits of cells stands out. The first one stands for around 5 to 10% of cells which expression varies for CD24 marker and expresses this marker 1.5 decade more to the large majority of cells while the high level of CD44 level of expression remains stable. The second one cells overhang consists of less of 5% of cells which level of CD44 expression decrease by half a log whereas CD24 expression stays stable. Finally, as previously described for the SUM149-PT cell line, a small population (around 0.5%) CD24-/CD44- expressing cells could be isolated. This latter clearly corroborates with the CD24 and CD44 expression level of unstained cells (*cf.* fig. 5.1 on the right).

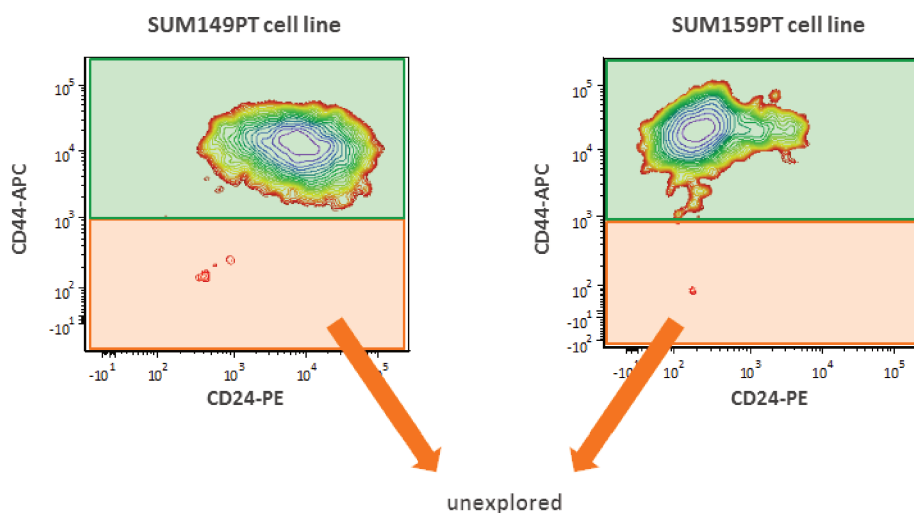


Figure 5.1 : Working assumptions with respect to SUM149-PT and SUM159-PT tumor cell lines heterogeneity.

orange selection : CD44 negative population and green selection : CD44 positive population.

The newly isolated CD24-/CD44- subpopulation was further characterized for its size parameter (Forward Scatter (FSC)) and internal granularity (Side Scatter. Also called 90-degree scatter or right angle scatter (SSC)). According to the parameters measured in flow cytometry, the CD24-/CD44- viable cells subpopulation appears to be somewhat small (low FSC value) and slowly grainy (without objects inside the cell to scatter the laser light) (low SSC value) (*cf.* fig. A.1 on page 208). These latest morphological characteristics are close to those of cellular debris.

We further investigated if the observed small and grainy population was not an artifact of cells selection, and we ensured the viability of these cells. We decided to sort that peculiar cell subpopulation and to look for its ability to recover and multiply within a standard SUM149-PT and SUM159-PT cell culture medium.

5.1.2 CD24-/CD44-, an undescribed viable cells phenotype

CD24-/CD44+ (in blue), CD24+/CD44+ (in green) and CD24-/CD44- (in orange) subpopulations were isolated and sorted from both SUM149-PT and SUM159-PT cell lines (*cf* fig. 5.2 on the following page subfig. A). Thereafter, sorted cells populations were re-plated with SUM149-PT and SUM159-PT cell culture media as described in appendix D.2 on page 249. This sorting experiment has been performed at three independent times, and for each replicative sorting 3 to 5 intra-experimental replicas were studied.

Then pictures of SUM149-PT and SUM159-PT subphenotypes were taken through a transmission microscopy (x10 magnification) after 1 day and 5 days (*cf* fig. 5.2 on the following page subfig. B and C). As a first observation, all cell subpopulations survived within the cell culture medium conditions and could multiply as we clearly observed that the density of cells increased between day 1 and day 5 of culture. This observation had previously been postponed by other colleagues regarding CD24-/CD44+ and CD24+/CD44+ subphenotypes [89, 113]. The culture of CD24-/CD44- sorted phenotype ensured that this latter isolated population contained viable cells able to adhere and divide as observed on fig. 5.2 on the next page subfig. C.

The shape of CD24-/CD44- cells is different from the two other phenotypes (*cf* fig. 5.2 on the following page) :

- The shape of both CD24-/CD44+ and CD24+/CD44+ cells are elongated and spindle-like.
- The morphology of CD24-/CD44- subpopulation is rounded with notable nucleus and a high nucleo-cytoplasmic ratio (data not shown) and a lamellopodia-like structure.

We decided to phenotype track this CD24-/CD44- cell population in long term culture. So we set up a method of quantitative analysis allowing us to easily isolate CD44-high from the CD44-low expressing cells by relying on a threshold of positivity for CD44 marker.

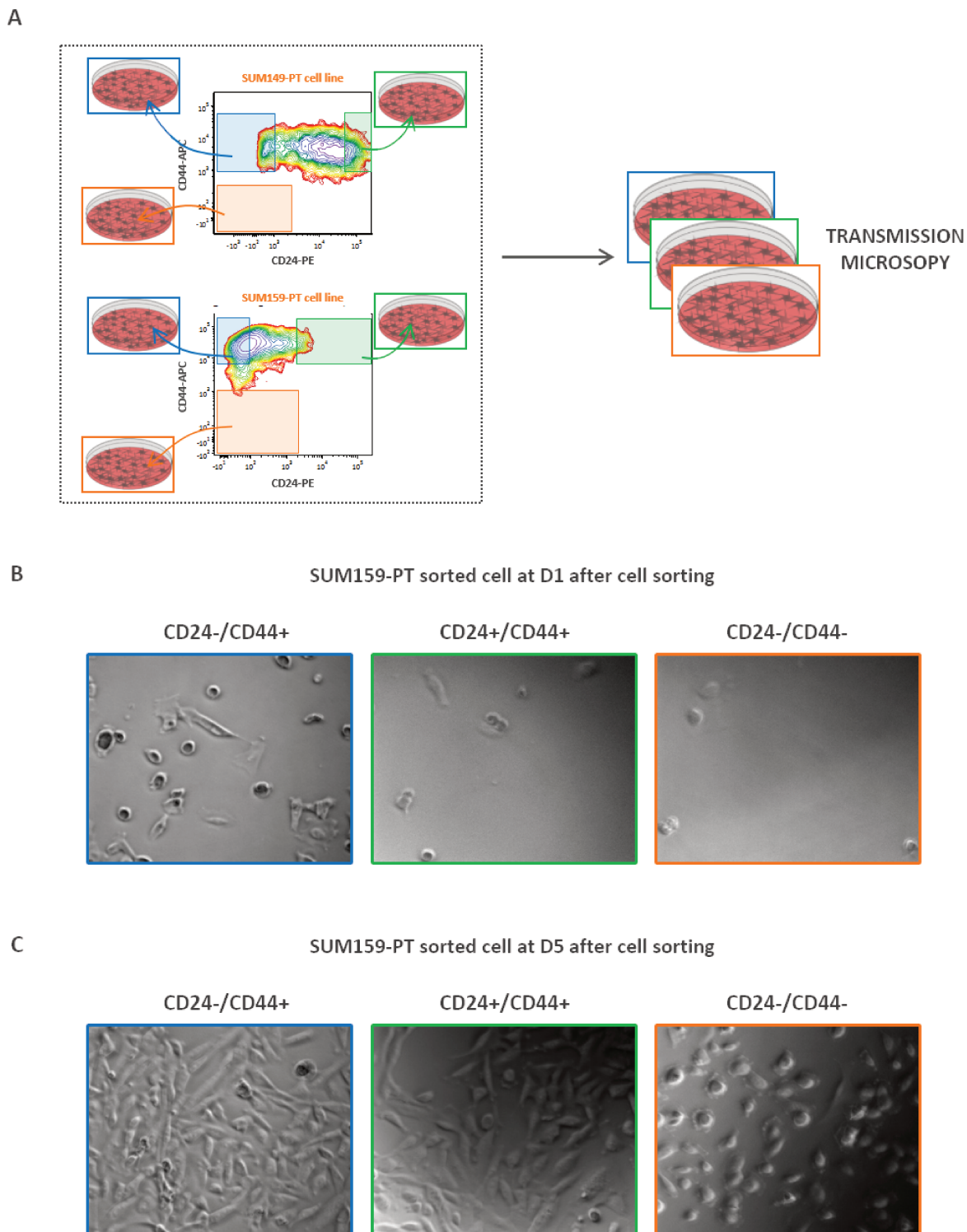


Figure 5.2 : Experiment workflow to characterize CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ subphenotypes

A. Experiment workflow including SUM149-PT and SUM159-PT cells subphenotypes sorting strategy to assess cells line phenotypic heterogeneity. B. Non calibrated pictures of SUM159-PT sorted subphenotypes after 1 days of culture in 2D ($\times 10$ magnification) phase contrast microscope. C. and of SUM159-PT sorted subphenotypes after 5 days of culture in 2D ($\times 10$ magnification).

in blue : CD24-/CD44+ cells sorted phenotype, in green : CD24+/CD44+ cells sorted phenotype, in orange : CD24-/CD44- cells sorted phenotype

5.2 Methodology of flow cytometry data recording and statistical analysis

Let describe the conditions to record flow cytometry data, the methodology to settle on the cells of interest and the calculations made to characterize CD24 and CD44 expression of SUM149-PT and SUM159-PT cells populations.

5.2.1 Data acquisition and selection of cells of interest

The flow cytometry acquisitions of SUM149-PT and SUM159-PT cells were performed on a LSR2 flow cytometer (BD Biosciences©) following standard procedures. Briefly forward scatter (FSC) and side scatter (SSC) parameters were manually adjusted before acquisition so that cells were clearly delineated, and the voltages for fluorescence channels were set on unstained cells. Compensation features between PE and APC fluorochromes were set in live via automatic Beckton Dickinson FACSDiva™ software wizards on single color compensation beads. Both single-color-stained compensation beads and cells were acquired to adjust post-acquisition compensation. Afterwards, cells were acquired through a maximum flow rate of 2000 events per second [127].

After the acquisition, flow cytometry compensated data were processed through FCS DeNovo Software© (FCS Express 4 Flow Research Edition) following the procedure illustrated on fig. 5.1 on page 50. Also briefly, cells of interest were selected on FSC and SSC criteria to exclude cells debris. Then single cells and viable cells were further selected on 4',6-Diamidino-2-Phenylindole (DAPI) negativity staining. Flow cytometry data of living cells were extracted for each acquired parameters and saved for further data analysis. In parallel, a positivity threshold for CD44 expression was determined on the basis of CD24-PE antibody single staining and unstained cells. Thus, the value of the CD44 positivity threshold was registered into a file for further use.

5.2.2 Flow cytometry qualitative representation

Flow cytometry data are represented as bi-parameters cytogram contour plot. In this way, plot shows all the viable cells and the different colors represent the density of the cells (as contour lines) at a given position. The density line mean that each line encloses a decreasing fraction of the population are taken as a geometric series whether the number of cells is less than x percent of the total number of cells, where x successive fraction is the value in the value field. Color scaling was specified as logarithmic and 25 discrete color levels (from red to blue through green colors) were used to generate gradients color.

Cell fluorescence intensity were plotted on 'logicle' displays (also named bi-exponential axis) in order to improve the representation of cells with minimal fluorescence, Thus CD24-/CD44- cells could be visualized with 'logicle' displays when they should have been piled up on the axis with logarithmic displays [127, 232].

5.2.3 Calculation of measurement parameters to characterize the fluorescence of the cells of interest

To define the signature of each cell lines, we described both qualitatively and quantitatively the variations of markers expression. Thus, we had to establish means of measurements and to monitor them over time.

To automatically analyze flow cytometry data, we developed a program to separate cells which strongly express CD44 marker from weakly expressing cell thanks to the CD44 positivity threshold previously defined.

We calculated the proportion of CD44-high (in green on figures) and CD44-low (in orange on

figures) expressing cells within the viable cells. And then we continued by measuring the expression of CD24 and CD44 markers of the determined cell populations.

We firstly calculated the pseudo-geometrical mean of the fluorescence of CD24 and CD44 markers to characterize both CD44-high and CD44-low expressing cells. Due to compensation, some low expressing cells showed negative value for their fluorescence intensity which precluded a simple calculation of the mean of the extracted values. Thus, we firstly approximated that :

To characterize a distribution $\{x_i\}$, one can use the arithmetic mean $\frac{\sum(x_i)}{n}$, or the geometric mean $(\prod x_i)^{\frac{1}{n}}$. In the present situation the former is not appropriate because of the huge dispersion, and the latter is not appropriate either because of possibly negative values. We thus used a novel mapping of the data as :

$$x \mapsto \sinh^{-1}(x_i)$$

and used the arithmetic mean of $\{\sinh^{-1}(x_i)\}$.

Because $\sinh^{-1}(x)$ behaves like x for $x \leq 1$ and $\sinh^{-1}(x)$ behaves like $\frac{1}{2} \log(x)$ for $x \gg 1$, the arithmetic mean of $\{\sinh^{-1}(x_i)\}$ behaves like $\frac{\sum(x_i)}{n}$ for small values of x and like $(\prod x_i)^{\frac{1}{n}}$ for large values.

The pseudo-geometrical mean is obtained as :

$$\sinh\left(\frac{\sum_{i=1}^n \sinh^{-1}(x_i)}{n}\right)$$

Then, we studied how stretched or squeezed was the distribution of the fluorescence intensity measured for CD44-high and CD44-low expressing cells populations for CD24 and CD44 markers. For that purpose, we calculated the pseudo-geometrical standard deviation value of the fluorescence distribution. Based on previous approximation the formula was defined as :

$$\sinh(\sigma(\sinh^{-1}(x_i)))$$

And finally this calculated pseudo-geometrical standard deviation value was graphed for each sample that we named ***“fluorescence intensity standard deviation”***.

The MATLAB program that was developed to handle these calculations is available in appendix G.4 on page 325.

5.2.4 Mean fluorescence normalization and data plotting

Pseudo-geometrical mean were firstly normalized in order to normalize for fluorescence variation from day to day which would be due to :

- Possible changes in quality of staining antibodies.
- Changes of the sensitivity of the flow cytometer equipment.

For each flow cytometry acquisition, we recorded the emitted fluorescence from a mix of compensation beads where 50% of the beads were stained with one specific antibody and 50% others did not carry any antigen and thus could not fix the antibody of interest. Then, for each experiments, we calculated the pseudo-geometrical mean of the positively stained beads thanks to the procedure that we previously described in section 5.2 on page 53. We named this value as the ***“beads measured fluorescence mean” of the positive beads***.

By relying on the set of all the “beads measured fluorescence mean” values issued from each

experiments, we calculated the arithmetic mean derived from the pool of “beads measured fluorescence mean”. We named the latter value as the “*beads typical fluorescence mean*” value of positive compensation beads. Also, this “typical fluorescence mean” used as a reference for the standardized measure of the positive signal of CD24 and CD44 markers.

We finally calculated for each “beads measured fluorescence mean” the “*fluorescence cells sample correction coefficient (i.e. ratio)*” in order to normalize cells fluorescence intensity in function of experimental procedures fluctuations. These values relative to each cell sample were calculated based on the :

$$\text{ratio} = \frac{\text{“beads measured fluorescence mean”}}{\text{“beads typical fluorescence mean”}}.$$

Then the fluorescence of each sample were normalized in assigning the “fluorescence cells sample correction coefficient” to the respective “sample measured fluorescence mean”. We named the corrected fluorescence as “*sample corrected fluorescence*”.

Then we defined a secondary normalization in defining the expression levels of the cells according to the basal fluorescence parameters of the cell : the auto-fluorescence. We also calculated the ratio of fluorescence between the unstained cells and the stained cells. As calculated by this formula :

$$\frac{\text{“sample corrected fluorescence”}}{\text{“unstained sample corrected fluorescence”}} = \text{“normalized mean fluorescence intensity”}.$$

We finally graphed this value to monitor the “*normalized mean fluorescence intensity*”.

5.3 Follow-up of CD24 and CD44 expression fine-variations within SUM149-PT and SUM159-PT cell lines

Cell populations are described as stable on average for the expression of their phenotypes [113], regarding

- The proportions of their several subphenotypes
- The level of expression of characteristic phenotypes markers (CD24 and CD44 in our case)

To rephrase what Gupta et al. proposed as a model, each single cell defined at $t = 0$ by its own phenotype is able at $t + 1$ to transit to another phenotypic state according to a constant probability “p”. Changes in cells phenotype behavior have been modeled according to the Markov chain and with essential following characteristics [113] :

- Cells have no memory of their former phenotypic state ($t - 1$), and therefore their decision-making is independent of their former experiences.
- Neighboring cells do not intervene in the decision-making of the cell. That is to say that, each cell, regardless of the phenotypic states of the neighboring cells at time $t = 0$ and their potential influence, is autonomous regarding its decision-taking.

If the system, defined at the scale of the cell population, meets the transition laws defined above, then the phenotype proportions and the corresponding level of expression of phenotypic markers CD24 and CD44 should rise to a stationary regime. Otherwise, if the average expression of a particular marker fluctuates over time, this implies that other factors intervene in the expression of the markers of interest.

The main issues can be reformulated as followed :

- A cell population which is characterized by its phenotypic heterogeneity, is the proportion of expressed phenotypes stable overtime ?
- Does the level of expression of CD24 and CD44 markers fluctuate over time within cell populations ?

The expression level for CD24 and CD44 markers of newly thawed SUM149-PT and SUM159-PT cells lines was monitored at several time points by flow cytometry techniques for approximately 130 days long (*cf.* fig. 5.3).

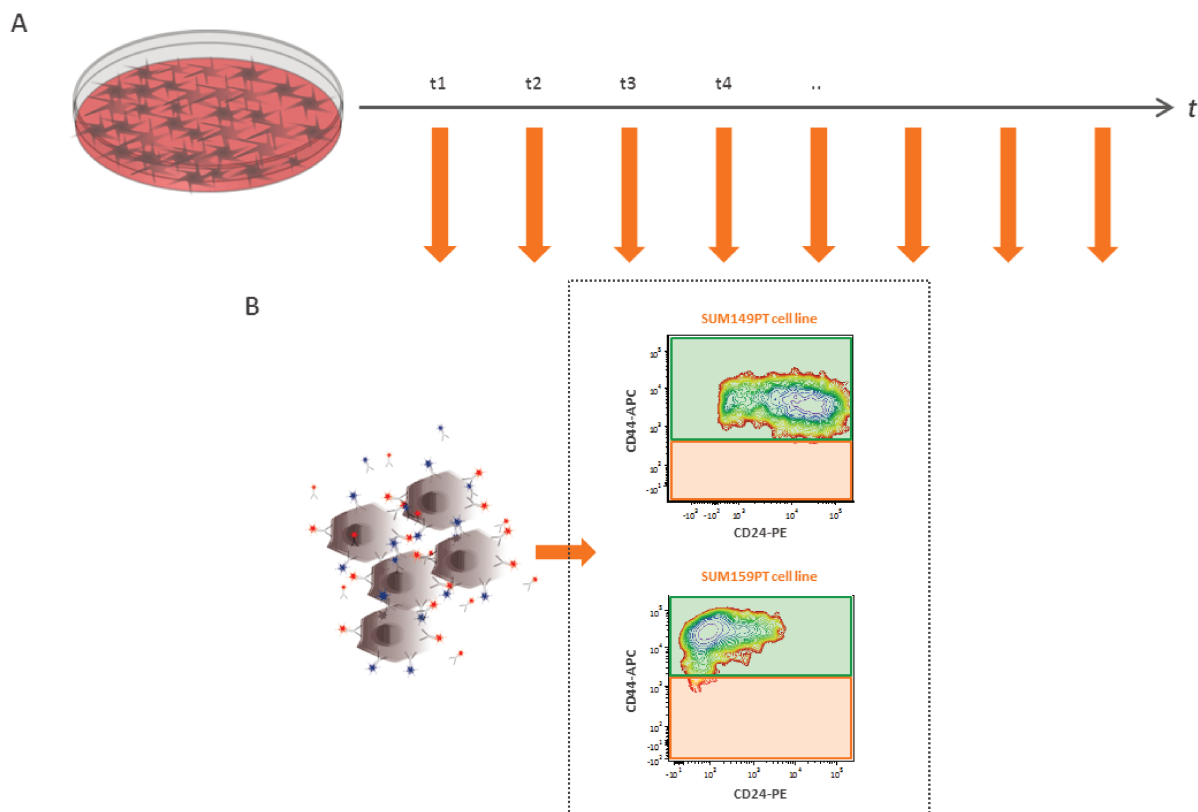


Figure 5.3 : Experiment workflow to monitor CD24 and CD44 expression of SUM149-PT and SUM159-PT cell lines.

A. Workflow to assess CD24/CD44 expression profile of SUM149-PT and SUM159-PT cell lines and its evolution overtime. B. Cytogram contour plot of CD24 and CD44 expression markers to characterize the heterogeneity of SUM149-PT and SUM159-PT cell lines.

Through the next paragraph we will describe the conditions of SUM149-PT and SUM159-PT cell culture and the staining protocol that was followed to track CD24 and CD44 expression within SUM149-PT and SUM159-PT cells populations. Afterwards we will monitor CD24 and CD44 within these cell lines and the resulting variations of CD44-low expressing and CD44-high expressing cells populations. And we will further discuss the significance of the observed fluctuations of :

- The geometrical mean of the fluorescence emitted from CD24-PE and CD44-APC markers within SUM149-PT and SUM159-PT cells populations
- The dispersion of the fluorescence map for both markers.

5.3.1 Experimental workflow to track the overtime expression of CD24 and CD44 markers

Cells were cultured as previously described in appendix D.2 on page 249, passed when cells reached between 70 to 80% of confluency and re-plated at a tenth of the confluence in a new flask. This procedure was followed every 3 to 4 days (*cf.* fig. 5.3 on page 56).

For each date we took advantage of cells sub-culture to monitor the expression of CD24 and CD44 markers. Thus, between 300.000 and 500.000 single dissociated cells were stained for CD24 and CD44 antibodies and acquired with a LSR2 flow cytometer (BD Biosciences©) following the protocol previously described in section 5.2 on page 53.

As a quick note, each staining procedures (of either the compensation beads or the cells) were performed through saturating staining conditions. Such that each surface antigen of the cells or of the beads is occupied by a primary antibody directly coupled with a fluorochrome. Thus, to efficiently anticipate flow cytometry experiments we first optimized staining protocols, concentrations of antibodies and cell staining time (data not shown). Protocols are available at the end of this manuscript (*cf.* appendix D.18 on page 277 and appendix D.17 on page 274).

Four different analyses were performed for each time point on cells sample :

- A qualitative analysis that present the cell fluorescence with a bi-parameter (CD24 *versus* CD44) cytogram contour plot.
- A measurement of CD44-low and CD44-high proportion within viable recorded cells.
- The normalized mean fluorescence intensity as previously described.
- The standard deviation of fluorescence intensity.

5.3.2 SUM149-PT cell line and its fluctuations in CD24 and CD44 expression

As described previously in section 5.1 on page 49, the expression profile of SUM149-PT cell line in the CD24/CD44 space is noteworthy. We monitored its qualitative and quantitative evolution over 128 days through 17 time points (*cf.* fig. 5.4 on page 59).

The contour plots of our flow cytometry analysis highlight that CD24-/CD44- subpopulation has changing proportions within the cell line (*cf.* fig. 5.5 on page 60). Five major increases happened at 5 different time points (T1, T44, T64, T85, T106) at which the proportion of this rare subpopulation goes over 1%. Those increases are followed by decreases (sometimes down to 0.1% for T4, T10, T128). We can observe that those increases and decreases seem to follow a cycle of 21 days (From D1 to D44, about 2x21 days, from D44 to D64 : 20 days, from D64 to D85 : 21 days, from D85 to D106 : 21 days). To sum up, we highlighted the periodicity in the proportion of CD44-low and CD44-high cells.

Then, we investigated if we could quantify changes in the average level of expression of CD24 and CD44 markers and this regardless of fluctuations in cellular compartments proportions of CD44-low and CD44-high (*cf.* fig. 5.6 on page 61). For this, we measured the average of expression of CD24 and CD44 markers of CD44-low and CD44-high cell populations as mentioned in section 5.2.3 on page 53. These are approximately regarding :

- CD44-high cell population : 433.8 ± 430.4 for CD44 and 27.0 ± 22.9 for CD24
- CD44-low cell population : 2.91 ± 2.04 for CD44 and 0.92 ± 0.99 for CD24

Cells expressing CD44-high show frequent fluctuations of the average expression level of both markers. Some of those fluctuations have particularly large amplitude :

- 1 log increase (UA) affecting CD44 marker between T44 and T64
- 1.5 log decrease (UA) between T106 and T128 for CD44 marker
- 1 log decrease (UA) between T15 and T64 for CD24

Nonetheless, we did not find any correlation between the exponential levels of CD24 and CD44 markers within this population (*cf.* fig. 5.7 on page 62).

Regarding CD44-low population (in orange), we also noted several fluctuations that are significantly affecting the expression level of both markers as notably illustrated in fig. 5.6 on page 61 :

- 1 log increase (UA) affecting CD44 marker between T85 and T106
- 1.5 log decrease (UA) between T85 and T106 for CD24 marker

For this population, variations of markers seem to be correlated (*cf.* fig. 5.7 on page 62).

Then, we measured the heterogeneity level of expression of the cell population by measuring the dispersion level of the fluorescence map of CD24 and CD44 markers (*cf.* fig. 5.6 on page 61). CD44-high subpopulation has a Standard Deviation (SD) of 0.72 ± 0.05 for CD44 marker and a SD of 1.43 ± 0.23 for CD24 marker. During this measuring period, these values are almost stable, we do note nonetheless small fluctuations affecting CD24 marker that reflect shape changes previously mentioned in fig. 5.4 on the next page.

Regarding CD44-low subpopulation, fluctuations related to the dispersion of CD24 et CD44 markers are much larger :

- For CD44 : 4.80 ± 6.69 AU
- For CD24 : 4.8 ± 4.3 AU

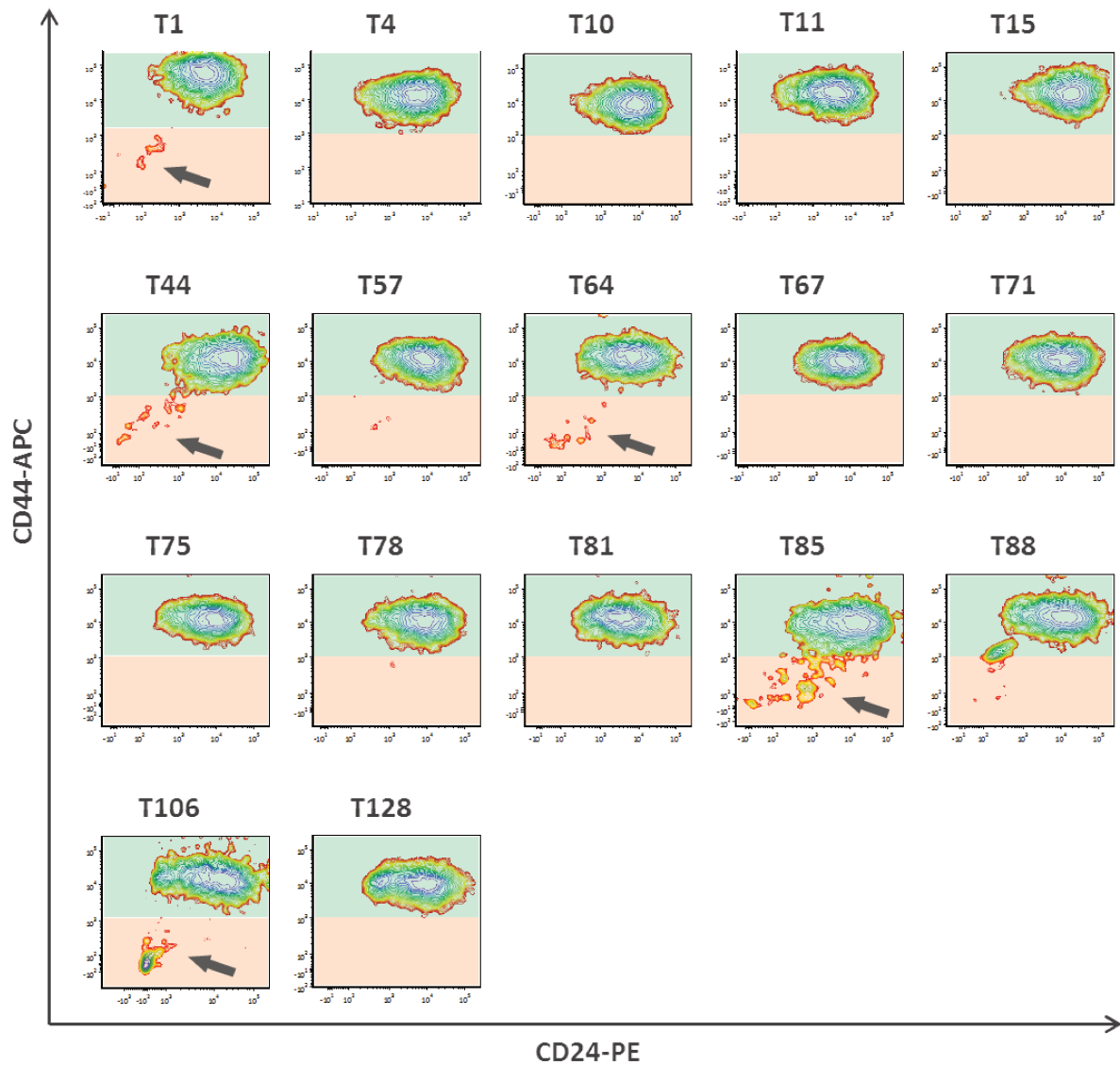


Figure 5.4 : Long-term monitoring of changes in SUM149-PT cell line heterogeneity.
 Cytogram contour plot of CD24 and CD44 expression markers characterizing changes in SUM149-PT cell line heterogeneity.
 Grey arrows indicate that CD24-/CD44- population is more than 1%

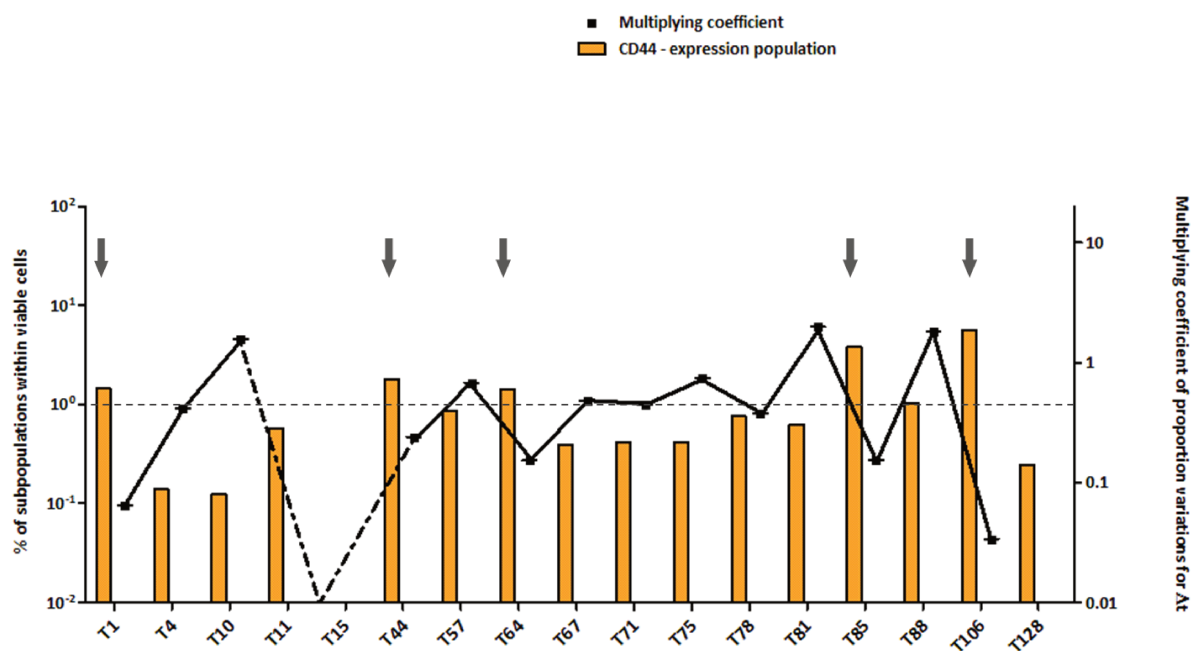


Figure 5.5 : Long-term changes monitoring of the proportion of CD44 negative cells among viable SUM149-PT cells.

Vertical stacked bars graph of the proportion of CD44 negative SUM149-PT cells

Grey arrow indicate that CD24-/CD44- population is more than 1%

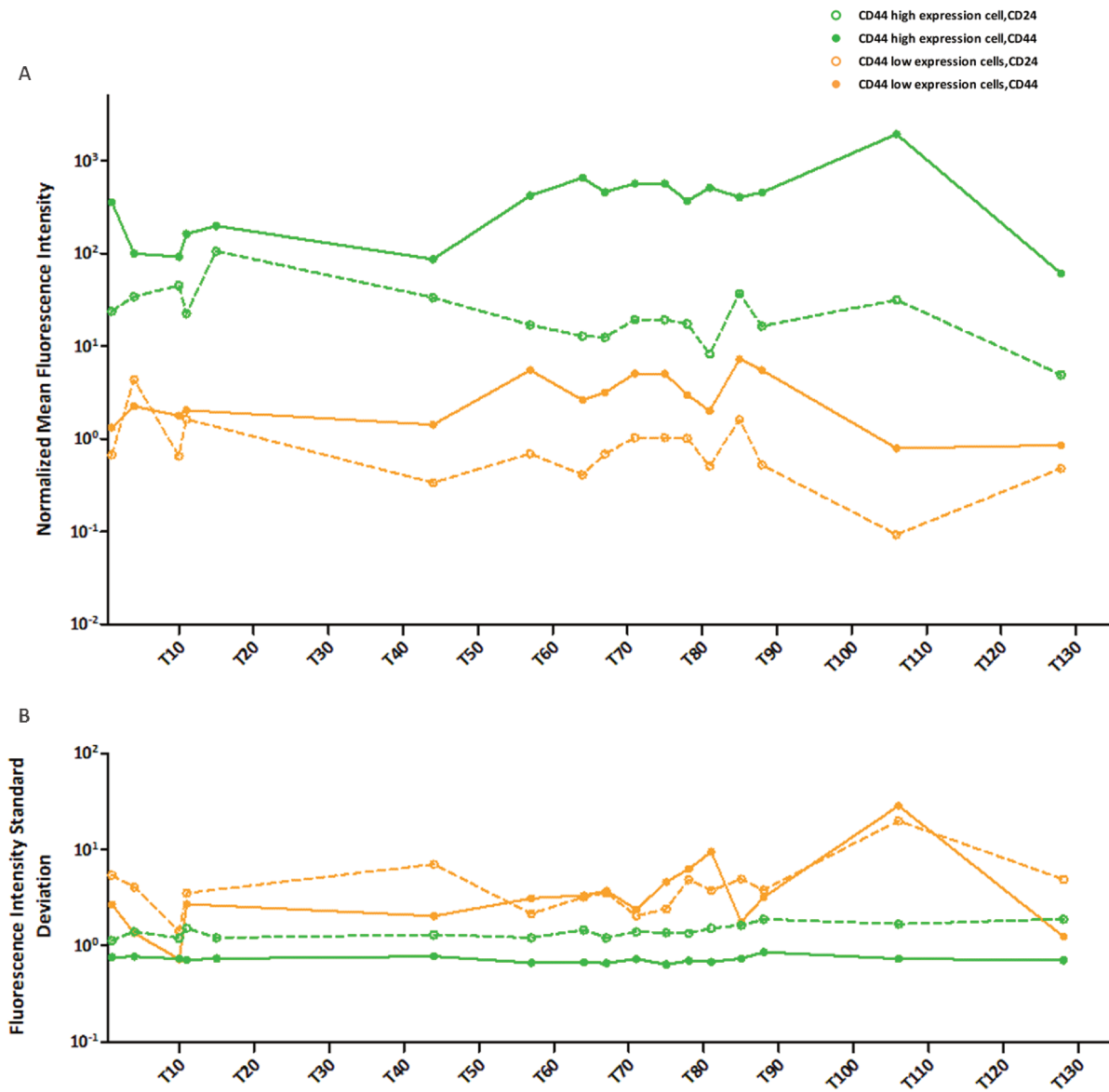


Figure 5.6 : Long-term monitoring of CD24/CD44 expression profile of SUM149-PT cell line plated into long term 2D culture conditions.

A. Grouped column scatter of long term tracking of CD24 and CD44 normalized pseudo-geometrical mean fluorescence of SUM149-PT viable cells. B. and of standard deviation of SUM149-PT viable cells

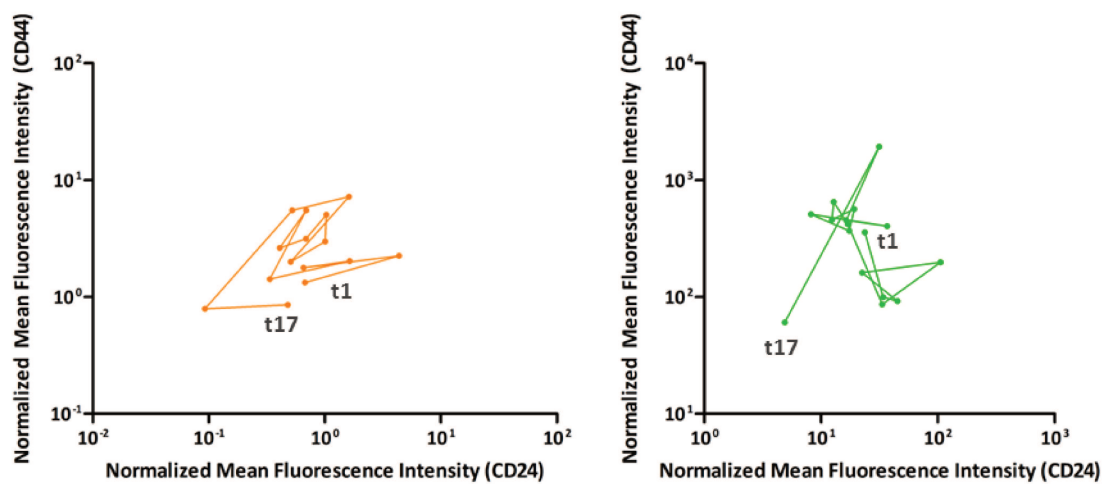


Figure 5.7 : *Correlation of CD24 and CD44 expression within SUM149-PT cell line*
in orange : CD44-low expressing cells and in green : CD44-high expressing cells

5.3.3 SUM159-PT cell line and its fluctuations in CD24 and CD44 expression

We also monitored the expression profile of SUM159-PT cell line in CD24/CD44 expression space. This study was conducted over 134 days with 19 time points (*cf.* fig. 5.8 on the following page).

As previously described in section 5.1.1 on page 49, CD24-/CD44+ cells are very well represented in SUM159-PT cell line. Compared with SUM149-PT cell line, there is no time point where CD24-/CD44- phenotype reach 1% of the population (*cf.* fig. 5.9 on page 65). As the sensitivity threshold is at 0,5% to 1%, this cell line does not appear in bi-parametric representation of flow cytometry data.

The qualitative and quantitative analysis by contour plot of SUM159-PT cell line highlights two cell strips :

- CD24-/CD44med
- CD24+/CD44+

The latter appear and disappear at a pace of their own. The CD24-/CD44med cell strip was observed 5 times (T1, T7, T16, T49, T69) following a cycle of about 8 days. CD24+/CD44+ cell strip was observed 6 times (T10, T20, T62, T72, T81, T93) following a cycle of about 10 days. When the expression level of CD24 marker rises within the population, we do not detect any CD24-/CD44med cell strip. We also note that CD24-/CD44med cell strip disappears simultaneously to CD24-/CD44med strip widening (T1-T3, T7-T10, T17-T20, T69-T71) (*cf.* fig. 5.8 on the next page).

The analysis of average variations for CD44-high (green) subpopulation confirms those qualitative observations in large part. Indeed, between T1 and T4, there is an increase in CD44 marker average (586 to 1419 Arbitrary Unit (AU)), likewise between T16 and T20 (755 and 1625 AU). Regarding CD24 marker, the average rises between T7 and T10 (7.04 to 13.3 AU) and a decrease between T20 and T49 (4.87 to 4.12 AU) (*cf.* fig. 5.10 on page 66).

However, we unexpectedly cannot confirm a link between the average expression level of CD24 and CD44 markers (*cf.* fig. 5.11 on page 67). This might be because CD24-/CD44med and CD24+/CD44+ cell strip fluctuations represent a small percentage of the total cell population and those variations are not significant enough to affect the average.

Similarly, the study of dispersion of the fluorescence map for the CD24 and CD44 markers allowed us to quantitatively validate the changes in length of CD24-/CD44med and CD24+/CD44+ cell strips :

- Between T10 and T16, CD24-/CD44med elongation results in an increase of SD of 1.64x (from 0.5 to 0.82 AU) for CD44 marker.
- Between T20 and T49, the elongation of CD24-/CD44med cell strip causes an increase of SD of 1.5x (from 0.49 to 0.75 for CD44 marker).
- Simultaneously, the shrinking of CD24+/CD44+ cell strip results in a decrease of SD of 4x (from 3.08 to 0.73 AU).

Generally, the elongation of CD24+/CD44+ cell strip is coupled to the shrinking of CD24-/CD44med cell strip. This statement is confirmed by the quantitative study of the dispersion for the fluorescence map of both markers of interest. Thus, the variation of dispersion for CD24 marker are inversely correlated to the variation of dispersion for CD44 marker as shown by fig. 5.10 on page 66.

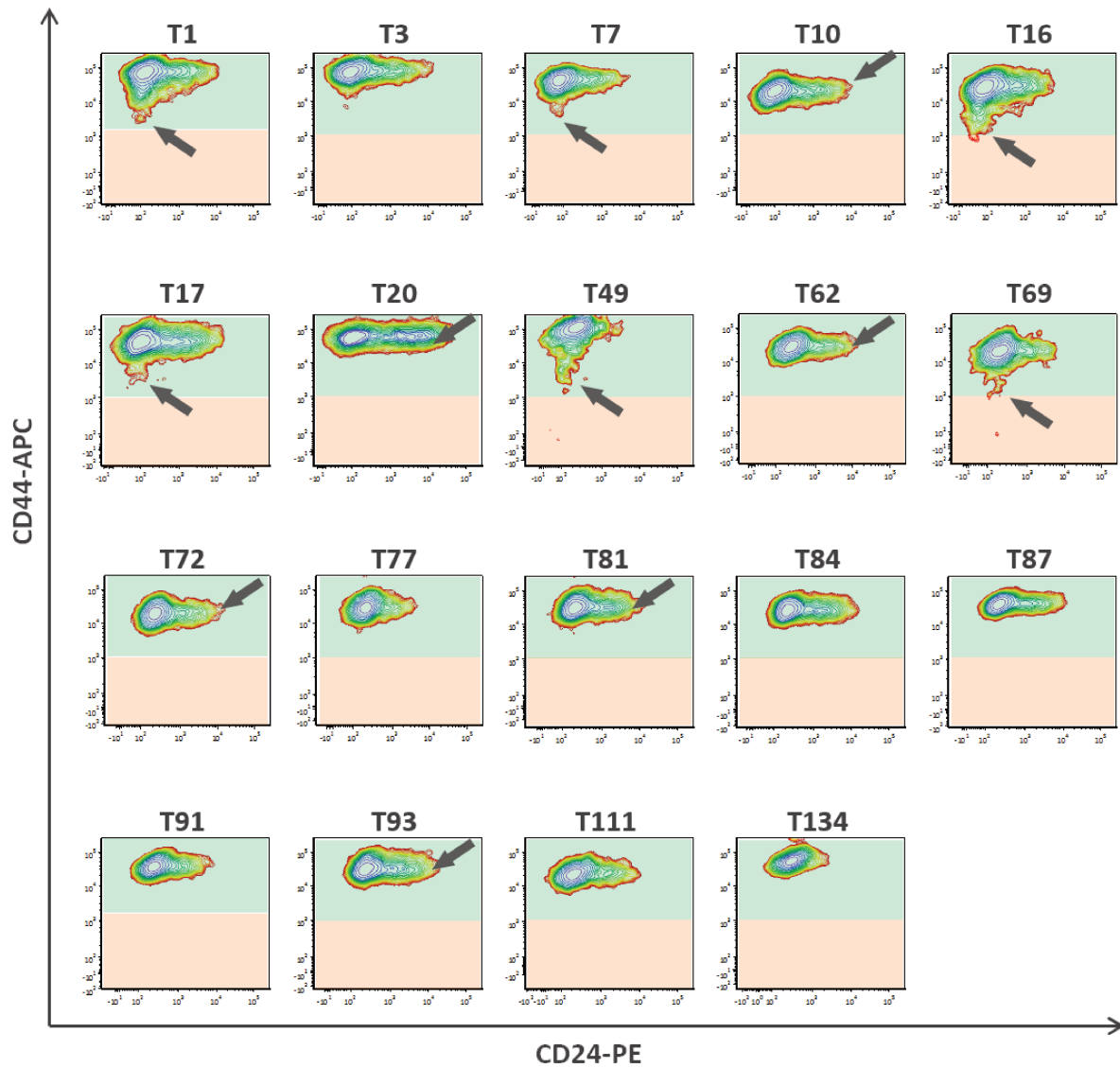


Figure 5.8 : Long-term monitoring of changes in SUM159-PT cell line heterogeneity.

Cytogram contour plot of CD24 and CD44 expression markers characterizing changes in SUM159-PT cell line heterogeneity.

Grey arrows indicate CD24-/CD44med subphenotypes split

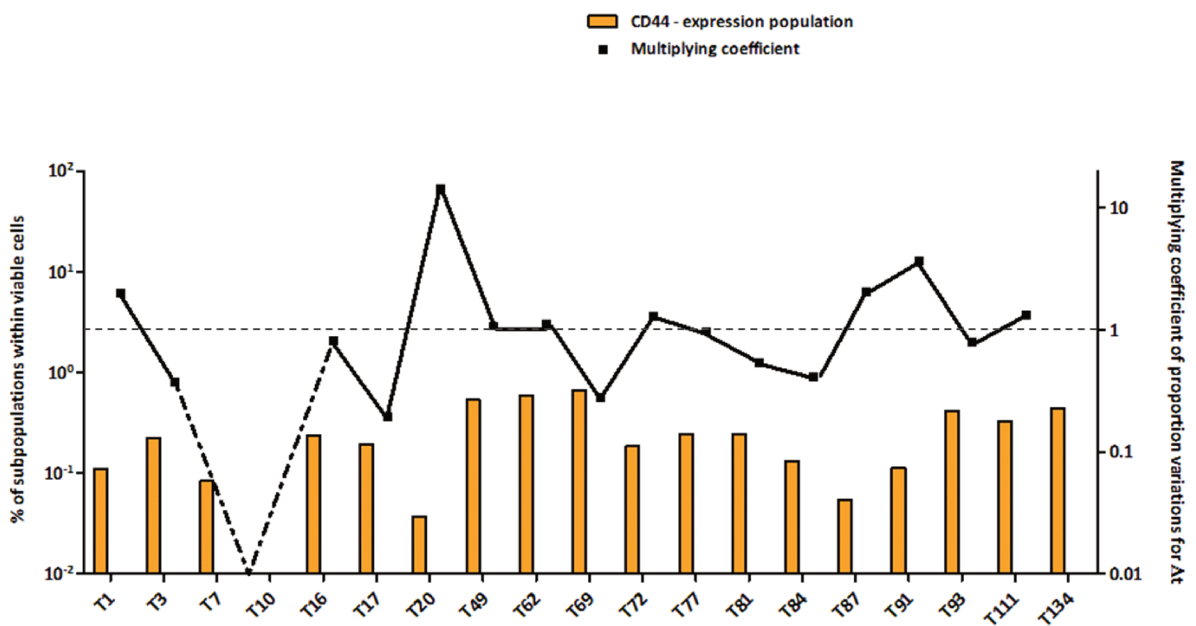


Figure 5.9 : Long-term changes monitoring of the proportion of CD44 negative cells among viable SUM159-PT cells.

Vertical stacked bars graph of the proportion of CD44 negative SUM159-PT cells

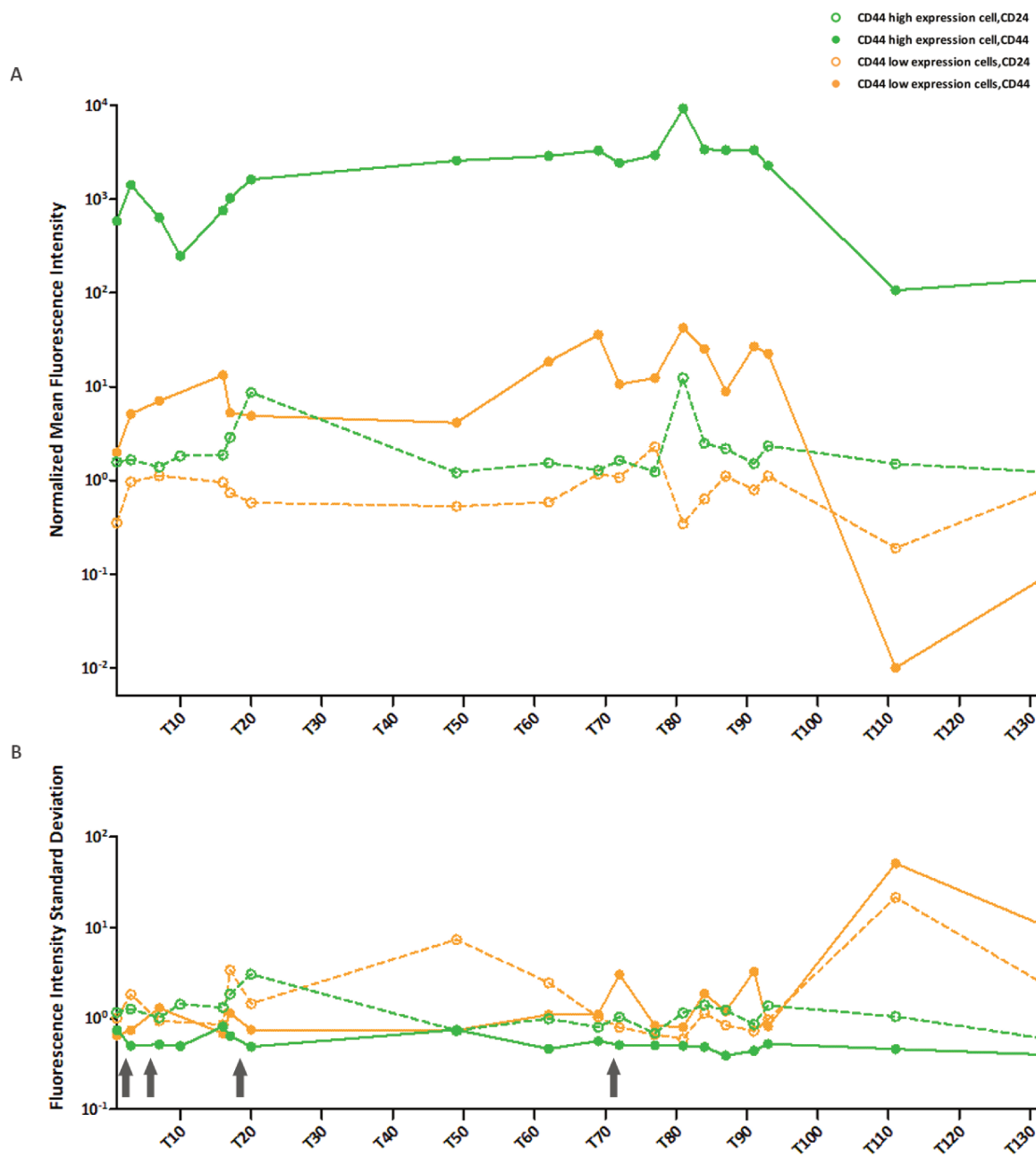


Figure 5.10 : Long-term monitoring of CD24/CD44 expression profile of SUM159-PT cell line plated into long term 2D culture conditions.

A. Grouped column scatter of long term tracking of CD24 and CD44 normalized pseudo-geometrical mean fluorescence of SUM159-PT viable cells. B. and of standard deviation of SUM159-PT viable cells

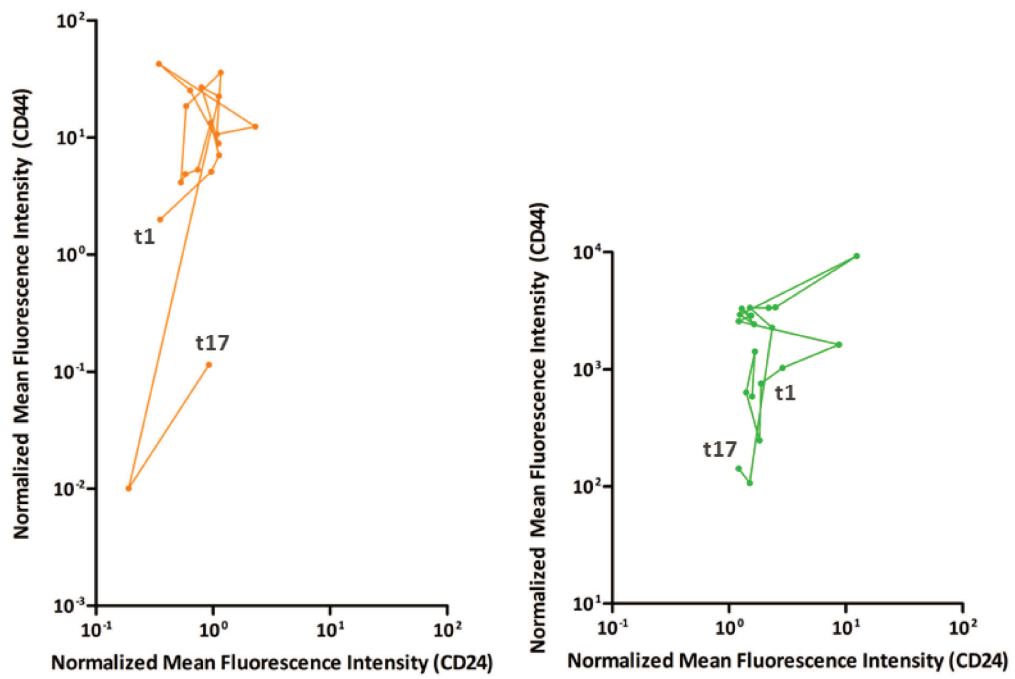


Figure 5.11 : Correlation of CD24 and CD44 expression within SUM159-PT cell line
in orange : CD44-low expressing cells and in green : CD44-high expressing cells

5.4 Cell line populations are continuously fluctuating

5.4.1 A few remarks with respect to measured fluctuations

Before summarizing the key points that we had made through this monitoring of two breast cancer cell lines (SUM149-PT and SUM159-PT), we will focus on the limitations arising from the techniques used to monitor fluctuations (proportions, level and dispersion of expression).

The first limitation pertains to the manual gating. This process necessarily entails bias in the selected cells. To solve this problem, we should perform this procedure using a method of automatic gating and directly fit distributions. This is one of the key challenges in flow cytometry and we will overview some strategies that alleviate this issue in section 12.3.3 on page 193.

The second limitation pertains to CD24-/CD44- cell selection. Those are gated very closely to cellular debris which may cause unwanted noise and proportion bias. To combat this issue, we may ideally visualize selected cells. ImageStream (or Imaging flow cytometer) seems to be the proper technology to this end as a combination of a flow cytometer and quantitative image analyzer. It has the major advantage of allowing visual distinction between normal cells and cell fragments or apoptotic bodies.

Finally, our last concern pertains to the quantification of fine fluctuations observed in contour plots (*cf.* fig. 5.4 on page 59 and fig. 5.8 on page 64) that are not reflected in the mean. Indeed, when the latter varies it means that :

- Either numerous cells undergo minor changes.
- Or fewer cells are significantly affected.

To highlight quantitatively both small and large fluctuations of expression of markers, it would be best to work across the fluorescence map, keeping the two-dimensional parametric measurements and proportions of cells. This methodology was developed by Pyne et al. with whom we are intending to work for subsequent analysis [130, 248].

5.4.2 Contributions of this study

To conclude this chapter, several key points may be underlined.

First of all, we validated the heterogeneity of expression of SUM149-PT and SUM159-PT cells for CD24 and CD44 markers, with an expression profile similar to previously reported profiles for both markers. We note three phenotypes of interest. Two of them have been described previously in the literature : CD24-/CD44+ corresponding to a pool of both basal-like and stem-like cells and CD24+/CD44+ composed of more differentiated cells [48, 89, 113].

Our work allowed us to identify a new cell phenotype (CD24-/CD44-) in breast cancer cell lines (*cf.* fig. 5.4 on page 59 and fig. 5.8 on page 64). This phenotype is characterized by small cells presenting a lamellipodia-like structures and, whose internal components present very low granularity (*cf.* fig. A.1 on page 208) and a high nucleo-cytoplasmic ratio (*cf.* fig. 5.2 on page 52). Viable and representing less than 1% of the global cellular population, this phenotype has been followed over 130 days of analysis (*cf.* fig. 5.2 on page 52).

This cellular phenotype, never observed before in a breast cancer cell population, was nevertheless described in the past in a cellular model derived from normal human breast tissue [71] and was also reported in an ovarian cancer cell line (A2780) [112].

Finally, the two biological systems are in a dynamic stable state characterized by the following parameters :

- The proportions of CD44-low and CD44-high phenotypes.

- The average expression level of CD24 and CD44 markers.
- The dispersion of the expression scatter plot which reflects the degree of heterogeneity of markers studied.

Our observations highlighted a behavioral signature for both cell lines. Thus :

- For SUM149-PT cell line, the CD24-/CD44- cell pool exceed 1% of the total cell population according to a 21-day cycle (*cf.* fig. 5.4 on page 59).
- SUM159-PT cell line undergoes cycles during which the dispersion of CD24 expression marker shows large increase and decrease as a time period of almost 10 days. Moreover, this cells population withstands both the dynamic appearance and disappearance of the CD24-/CD44med overhang according to an 8-day time cycle (*cf.* fig. 5.8 on page 64).

These latest observations illustrate how the expression of markers, as a function of time, and more particularly the periodicity of the dynamic expression, sign and characterize the behavior of the population. Thus, time, formalized as the unit which describes the time interval between two similar states, clearly needs to be distinctly appreciated in the sequence of changes of the state-wide of the cells. Indeed, “time shapes cells” and the equilibrium of cell populations [5].

5.4.3 On the way to the next chapter

Analysis of breast cancer cell lines enabled the identification of three distinct cellular phenotypes : CD24-/CD44+, CD24+/CD44+ et CD24-/CD44-. Significant variations of the average expression level of the population lead us to believe that, at the scale of individuals, some cells are subject to changes of the expression level of the markers. Thus, these cells, characterized by their phenotype at $t = 0$, must have the capability to transit to another phenotype at $t + 1$. These cells are plastic but are all cells, regardless of their phenotype, capable of plasticity?

Cells are plastic. However do they all share the same capabilities ?

«À chaque changement de point de vue correspond un changement de lois. Il existe donc une infinité de lois.»
- Frank Herbert - *Les yeux d'Heisenberg* (1966)

Previous chapter demonstrated that the cell populations as a whole present a dynamic proportion variables in the expression space of CD24 and CD44 markers. This feature implies that cells are individually able to change within this space of markers. Therefore, some cells are able to pass from a phenotypic state to another.

At any time points, some cells share the same level of expression for CD24 and CD44 markers and can be grouped into subsets based on their level of expression. Three different cell subpopulations were isolated by selecting the extreme levels of expression for these markers : CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- (*cf.* fig. 5.2 on page 52 subfig. A). In this chapter, we sought to study the dynamics of repopulation of each subpopulation.

Throughout this work, we asked ourselves whether the new generations re-reached the same balance as defined within the non-sorted population. To this end, we isolated the various subphenotypes based on the expressed markers and followed their repopulation using flow cytometry. Even though the plastic ability of CD24-/CD44+ and CD24+/CD44+ subpopulations has previously been demonstrated by Gupta et al. and Fillmore et al. [89, 113], it has not been assessed at long time and it is not known for the newly defined phenotype. Indeed, studies of the phenotypes CD24-/CD44+ and CD24+/CD44+ had been carried out after 6 days without sorting and monitoring over time of the phenomenon of re-expression/de-expression of markers. Our aim was therefore to study the repopulation dynamics by tracking the three phenotypes over time (*cf.* fig. 6.2 on page 74).

6.1 Plastic abilities of subpopulations

Cell plasticity is the ability of a cell to acquire a new identity and to adopt an alternative fate. Many examples of cellular plasticity have been described in physiological, experimental or pathological contexts [57, 113, 151, 243, 260, 345]. We will strive to determine whether the three

cell populations, we previously defined, are able to transition between one phenotypic state to another. For example, can CD24-/CD44- cells population produce CD24-/CD44+ and/or CD24+/CD44+ cells again while self-renewing its pool (*cf.* fig. 6.1)?

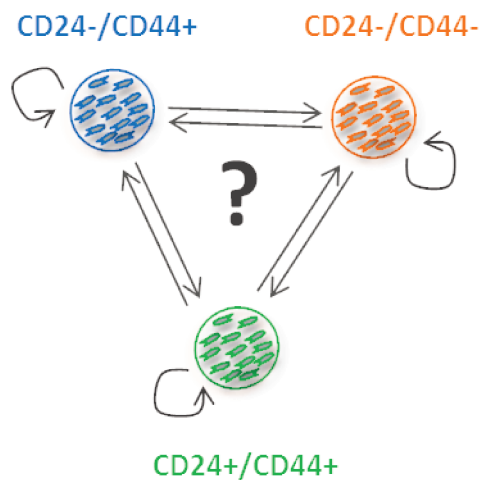


Figure 6.1 : Working questioning with respect to cell lines subphenotypes plasticity

6.1.1 Properties of the experiment's implementation

To answer this question, the three phenotypic subpopulations of both SUM149-PT and SUM159-PT cell lines (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-) were removed by following the procedures described in section 5.1.2 on page 51. Then each of these subpopulations was put back into culture. And the level of expression of the CD24 and CD44 markers was monitored over a long period (between 70 and 130 days of culture) by flow cytometry as previously described in section 5.2 on page 53.

This sorting experiment has been performed at three independent times, and for each replicative sorting 3 to 5 intra-experimental replicas were studied (*cf.* fig. 6.2 on page 74). These three independent experiments have been pooled and results will be presented together.

Sorted cell populations have been plated at a density limited by the efficiency of the cell sorting procedure. However, we took care that inter-experiment density of cells plating remains similar as depicted by fig. A.2 on page 209 in order to allow us to compare experiments among several sorting procedures. Thus regarding SUM149-PT cell line, cells were plated at a density of 20 ± 5 cells/mm², CD24-/CD44+ cells were plated at 158 ± 28 cells/mm² and CD24+/CD44+ cells were plated at 188 ± 23 cells/mm². In parallel, SUM159-PT subpopulations, cells were plated at a density of 16 ± 3 cells/mm², CD24-/CD44+ cells were plated at 506 ± 126 cells/mm² and CD24+/CD44+ cells were plated at 47 ± 2 cells/mm².

Sorted cells were cultured as described in appendix D.2 on page 249, passed when cells reached between 70 to 80% of confluency and re-plated at a third of the confluence in a new petri-dish. This procedure was followed every 3 days (*cf.* fig. 6.2 on page 74).

For each date we took advantage of cells sub-passage to monitor the expression of CD24 and CD44 markers. Thus, between 100 000 and 150 000 of single dissociated cells were stained for CD24 and CD44 antibodies and acquired with a LSR2 flow cytometer (BD Biosciences©) following the recommendations that I previously described in section 5.2 on page 53.

Four different analyses were performed for each time point on cells sample (*cf.* section 5.2.3 on page 53) :

-
- A qualitative analysis with a bi-parameter (CD24 *versus* CD44) cytogram contour plot
 - A measure of CD44 low and CD44 high proportion within viable recorded cells
 - The normalized mean fluorescence intensity as previously described
 - The fluorescence intensity standard deviation

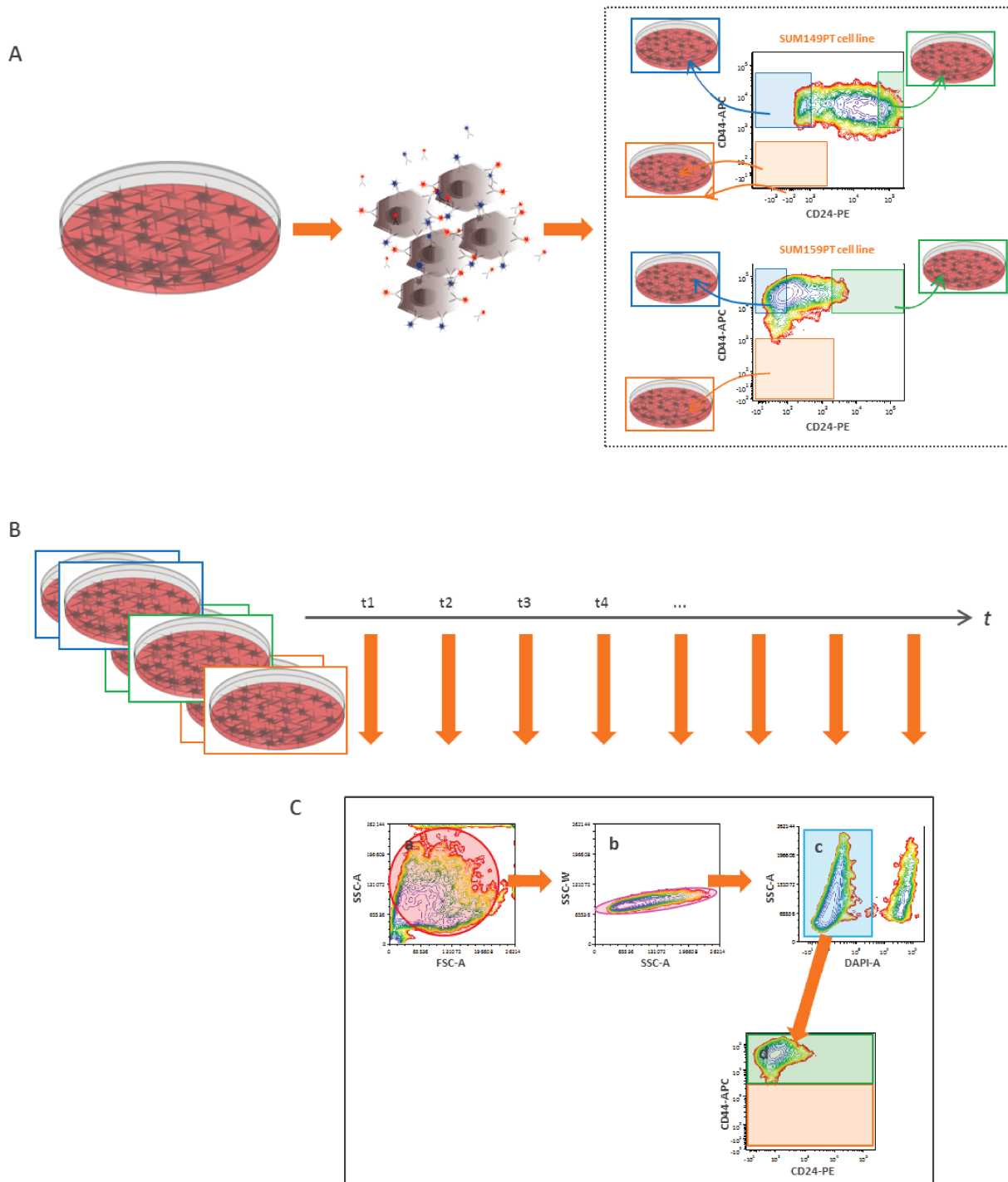


Figure 6.2 : Experiment workflow setup to characterize overtime SUM149-PT and SUM159-PT cell lines subphenotypes plasticity abilities.

A. Sorting strategy workflow of SUM149-PT and SUM159-PT cells subphenotypes. In blue : CD24-/CD44+ cells sorted phenotype, in orange : cells sorted phenotype, in green : CD24+/CD44+ cells sorted phenotype. B. Experimental strategy to monitor sorted cells subphenotypes over time. C. Flow cytometry data analysis gating strategy to characterize repopulating cells subphenotypes (a) Cytochrome contour plot of Light scattered cells characteristics. (b) Cytochrome contour plot of Side Scatter characteristics. Pink selection stands for single cells population. (c) Cytochrome contour plot of cells viability parameter. Light blue selection stands for viable cells population. (d) Cytochrome contour plot for CD24 and CD44 expression markers.

6.1.2 In less than three days, subphenotypes have already exchanged to other phenotypic states

Cell subpopulations express one time again CD24 and CD44 markers

It turns out that the subpopulations (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-) derived from SUM149-PT cell line, and homogeneous at D0, rebuilt heterogeneous populations. These later consist in about 83% CD44-high (CD24-/CD44+ and CD24+/CD44+ cells) (*cf.* fig. 6.15 on page 93 and fig. 6.16 on page 93). In the markers space, the expression profile of CD44-high cell population is very similar to the one of the non-sorted population, as illustrated by the fluorescence map of CD24 and CD44 markers of fig. 6.17 on page 94.

The same analysis was performed for the subpopulations derived from SUM159-PT cell line (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-). These latest also demonstrated very high plastic capabilities. Indeed, at D6 SUM159-PT derived subpopulations already rebuilt 97% of cells positively expressing CD44 marker (*cf.* fig. 6.18 on page 96 and fig. 6.19 on page 97). This CD44-high population presents an expression profile very similar to the non-sorted population it derived from. Indeed, the majority of cells are expressing both CD24-/CD44+ markers and two cell strips (CD24-/CD44med and CD24+/CD44+), which are some characteristics of SUM159-PT cell line (*cf.* fig. 6.20 on page 98).

CD24-/CD44+ and CD24+/CD44+ subpopulations transit to other phenotypic states

As already pointed out in previous publications [89, 113] both CD24-/CD44+ and CD24+/CD44+ subpopulations display plastic capabilities. Nevertheless, it was interesting to know whether the dynamics of plasticity triggered before day 6 (date of the analysis performed within previous publication of Gupta et al.). Thus, we realized our measurements from D3 and followed up over 130 days (*cf.* fig. 6.3 on page 77 and fig. 6.4 on page 78).

At D3, the CD24-/CD44+ subpopulation derived from SUM149-PT already displays a cell strip stretching to a positive expression of the CD24 marker. Besides, we also note the presence of 0.34% of CD24-/CD44- cells.

The same subphenotype derived from SUM159-PT cell line displays similar characteristics : the presence of a cell strip, also stretching to a positive expression of CD24 marker which signals the re-expression of CD24+/CD44+ phenotype. We also note the presence of 0.1% of CD24-/CD44 double negative cells (*cf.* fig. 6.9 on page 85 and fig. 6.10 on page 86).

The CD24+/CD44+ subphenotype derived from SUM149-PT is also affected and displays after 3 days a decline of CD24 marker and the presence of 0.5% of double negative cells for the monitored markers (*cf.* fig. 6.6 on page 81 and fig. 6.7 on page 82).

This phenotype state transition is particularly notable for the CD24+/CD44+ phenotype derived from the SUM159-PT cell line due to the presence of two cell strips. Both show a decline of CD24 and CD44 markers associated with the presence of 0.2% double negative cells (*cf.* fig. 6.12 on page 89 and fig. 6.13 on page 90).

In conclusion, the three subphenotypes of interest are not only able to pass from a phenotype to another, they also re-express very quickly all the diversity of phenotypes in a pattern of their own :

- Except regarding the proportion, the phenotype mimics the fluorescence map of the non-sorted population from D6 onward.
- Initially, CD24-/CD44+ subpopulation expresses its heterogeneity by stretching along the CD24 marker axis.

- CD24+/CD44+ subpopulation display two cells strips associated with a decline of CD24 and CD44 marker expression.

6.2 Repopulation processes of CD24-/CD44+ and CD24+/CD44+ subpopulations

The previous study presented by Gupta et al. [113] strove to demonstrate that after 6 days of culture, the repopulation phenomenon had re-reached balance of characteristic proportions of the non-sorted cell population. By experimenting over a much longer period (130 days with 30–35 measuring points), we found out that CD24-/CD44+ and CD24+/CD44+ population exhibited variations of both the phenotypic proportions and the average level of expression of the markers (especially for the CD24 marker). Furthermore, the re-equilibrium of both sorted cells populations provides CD24 and CD44 spectra which clearly differ from the ones characterizing the non-sorted population, and this regardless of the cell line we study.

6.2.1 Repopulation dynamics of CD24-/CD44+ and CD24+/CD44+ issued from SUM149-PT

Regarding the **CD24-/CD44+ subpopulation** derived from SUM149-PT cell line (*cf.* fig. 6.3 on the facing page), the relative proportions of CD44-low (orange) and CD44-high (green) subpopulations are varying with a tendency to slightly decrease over time by almost 1 log compared to the SUM149-PT non-sorted cell line (*cf.* fig. 6.4 on page 78).

The expression map along the CD44 axis remains stable and complies with the fluorescence map of the non-sorted population. The observed variations in the average level of expression relate to the CD24 marker. Its fluorescence map becomes bimodal from D6 with a ratio CD24-low/CD24-high = 20/80. Then we detect large variations in the relative importance of each mode. At D16, the CD24-low/CD24-high modes ratio is almost 25/75. At D59 however the ratio is about 46/54. This ratio continues to vary over time and CD24-low cell pool is still very large (between 30 and 70% of the viable population). The fluorescence map of the CD24 marker never goes back to that of the non-sorted population (*cf.* fig. 6.5 on page 79).

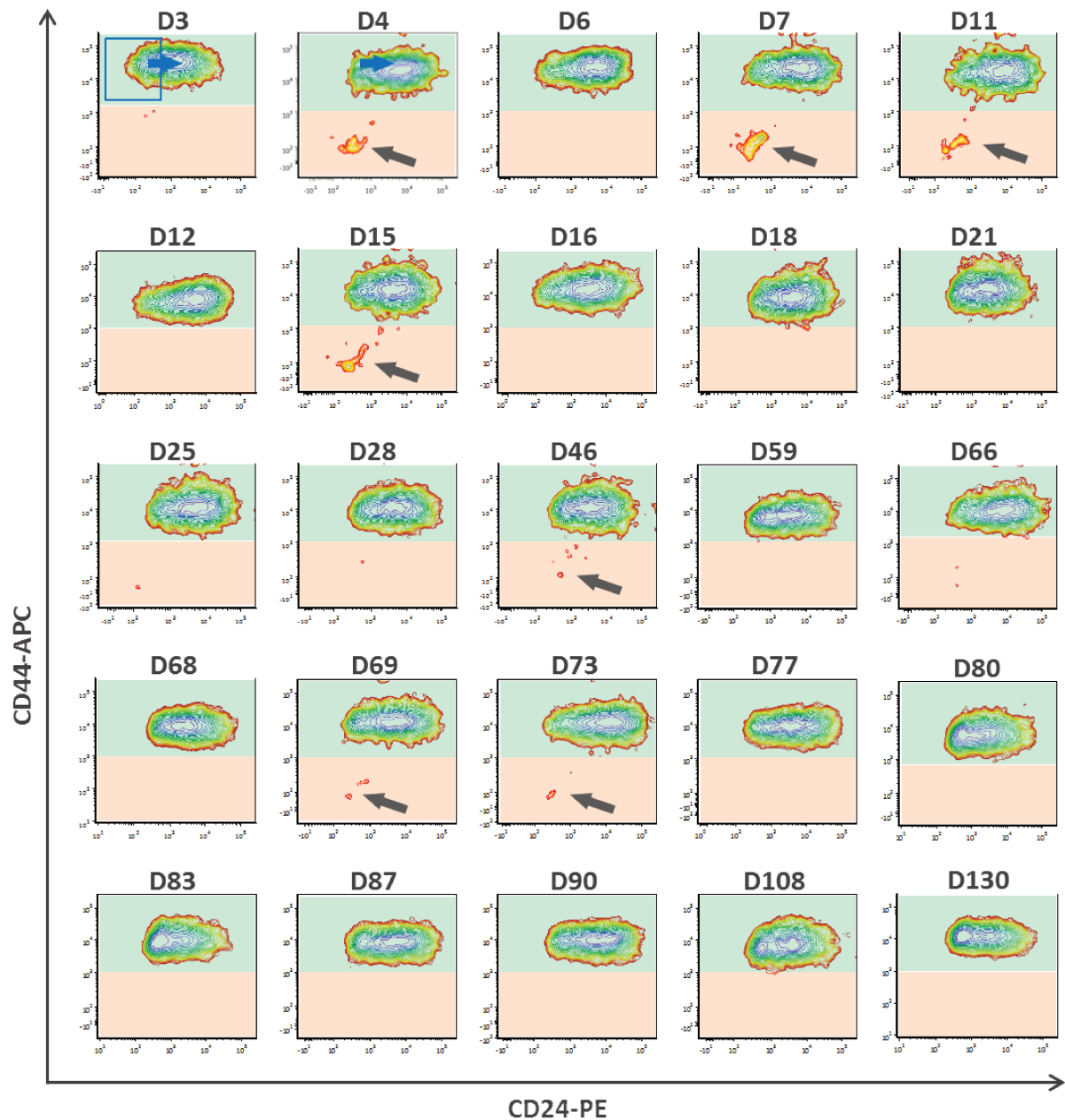


Figure 6.3 : Long-term monitoring of the repopulation of CD24-/CD44+ SUM149-PT cells subphenotypes.

Cytogram contour plot of CD24 and CD44 expression markers characterizing overtime repopulation of sorted CD24-/CD44+ subphenotype cells

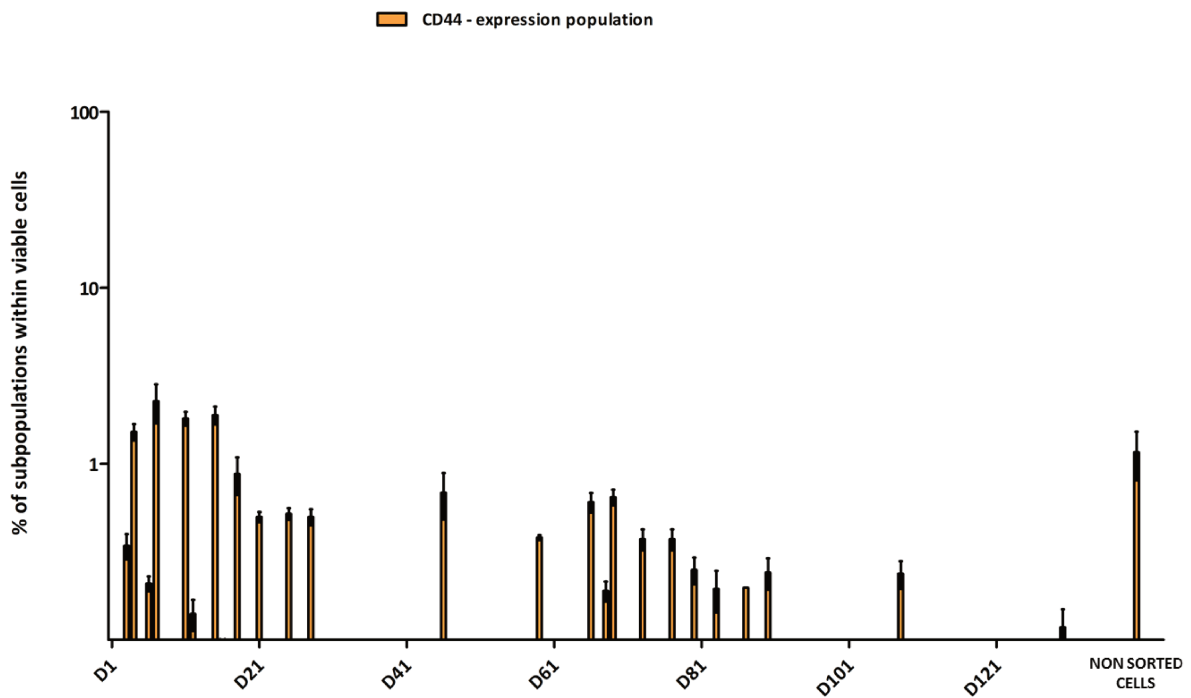


Figure 6.4 : Long-term monitoring of CD44 negative expressing cells among viable repopulating SUM149-PT cells population.

Vertical stacked bars graph presenting the percentage of CD44 negative cells within repopulating SUM149-PT CD24-/CD44+ subphenotype.

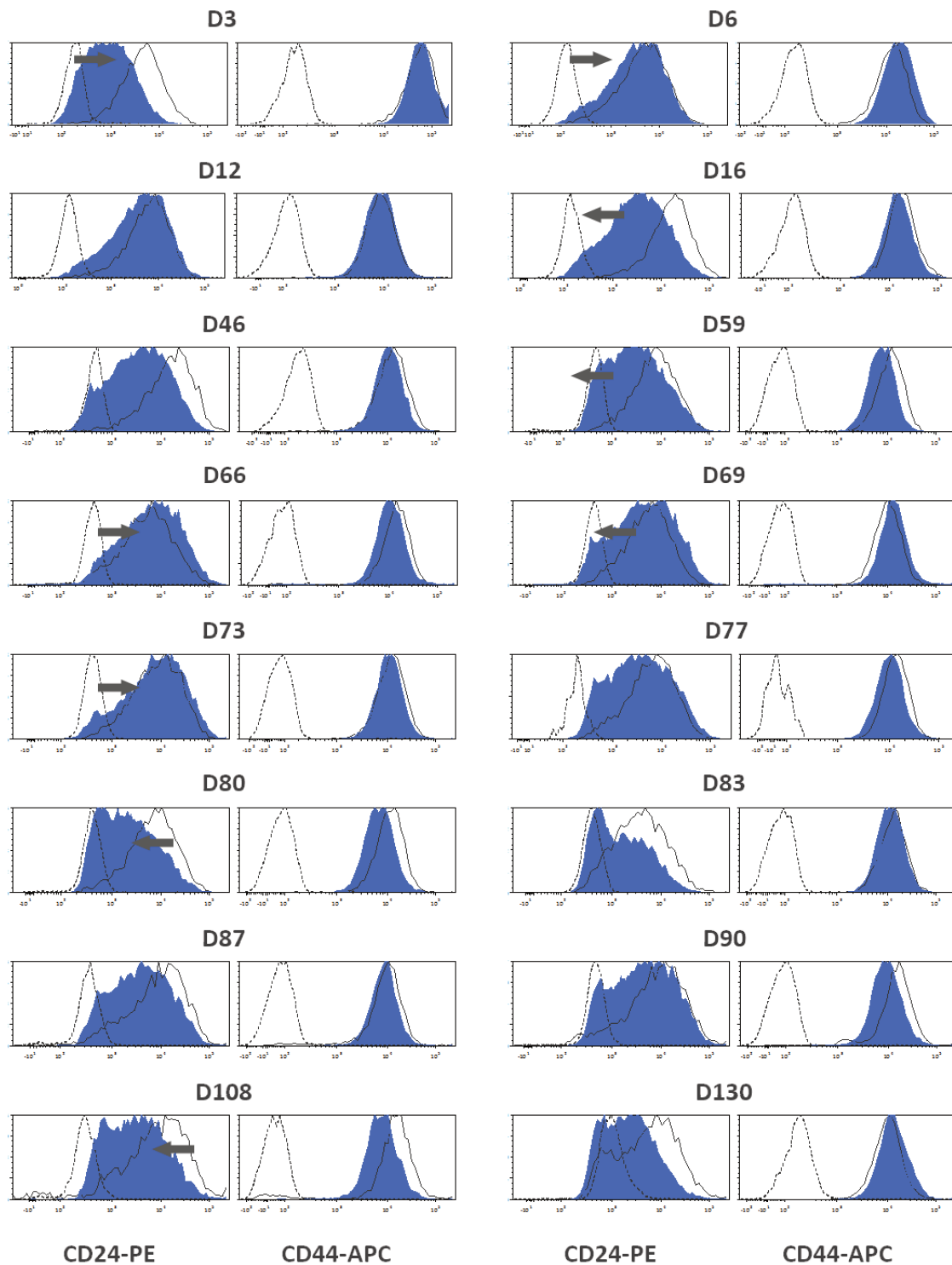


Figure 6.5 : Long-term monitoring of CD24 and CD44 expression among viable repopulating CD24-/CD44+ SUM149-PT cells and comparison with unstained and double-stained non-sorted SUM149-PT cells.

Flow cytometry histograms presenting distribution curves of CD24-PE and CD44-APC fluorescence. blue area selection : fluorescence distribution curve of CD24-/CD44+ SUM149-PT subphenotype repopulation, dashed gray curve : fluorescence distribution curve of unstained non-sorted cell line, full gray curve : fluorescence distribution curve of double-stained non-sorted cell line

CD24+/CD44+ subpopulation derived from the same cell line re-generate the CD24-low cell pool in proportions similar to the non-sorted population (*cf.* fig. 6.6 on the next page). Those proportions vary slightly as previously described in chapter section 5.3.2 on page 57. Like CD24-/CD44+ cell phenotype, CD24+/CD44+ subpopulation exhibits a fluorescence distribution on CD44 axis similar to the fluorescence map of the non-sorted cell population (*cf.* fig. 6.7 on page 82). The fluorescence distribution on CD24 axis retains the same shape as that of the non-sorted population but with a shift of the mode (+0.5log at D66, D73, D77 and D83) (*cf.* fig. 6.8 on page 83).

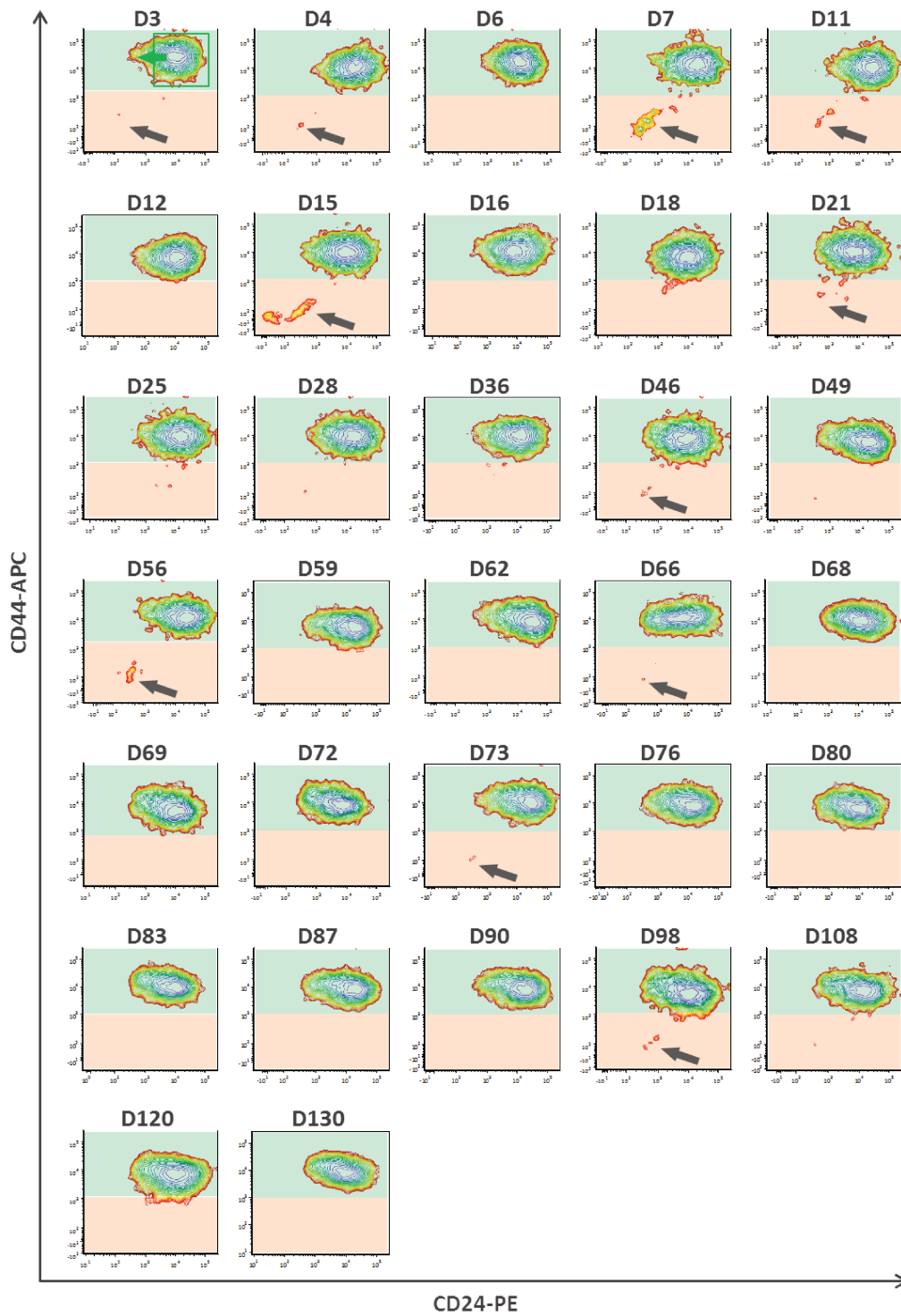


Figure 6.6 : Long-term monitoring of CD24+/CD44+ SUM149-PT cells subphenotypes repopulation

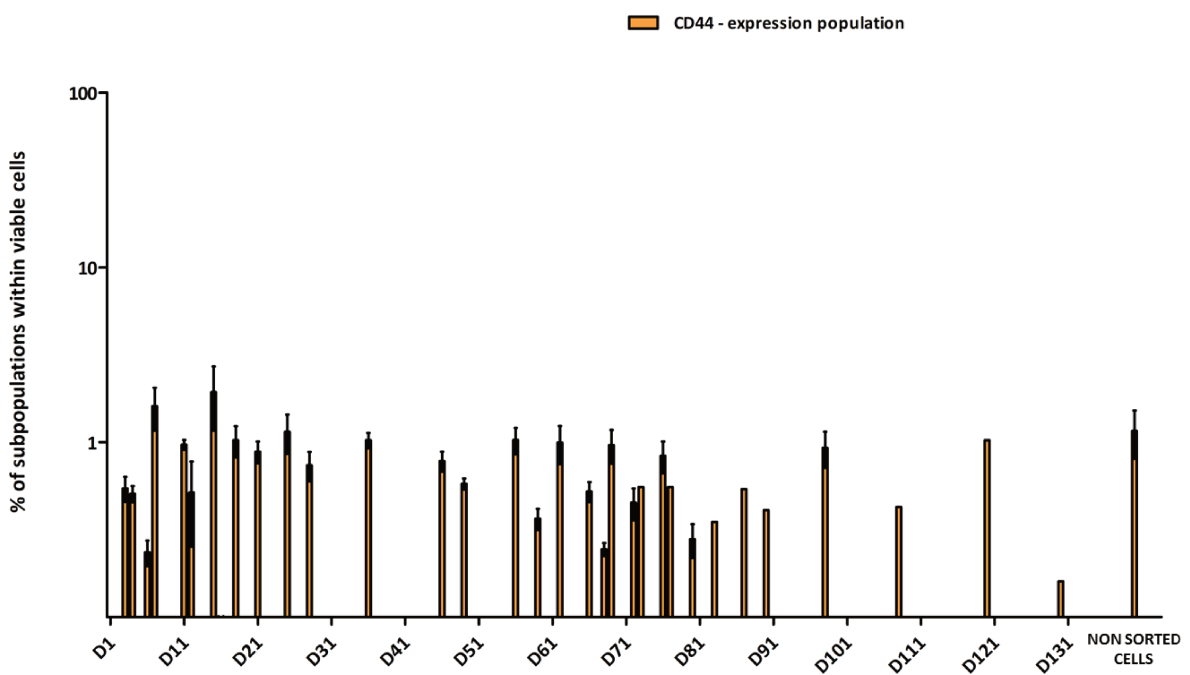


Figure 6.7 : Long-term monitoring of CD44 negative cells among viable repopulating CD24+/CD44+ SUM149-PT cells population.

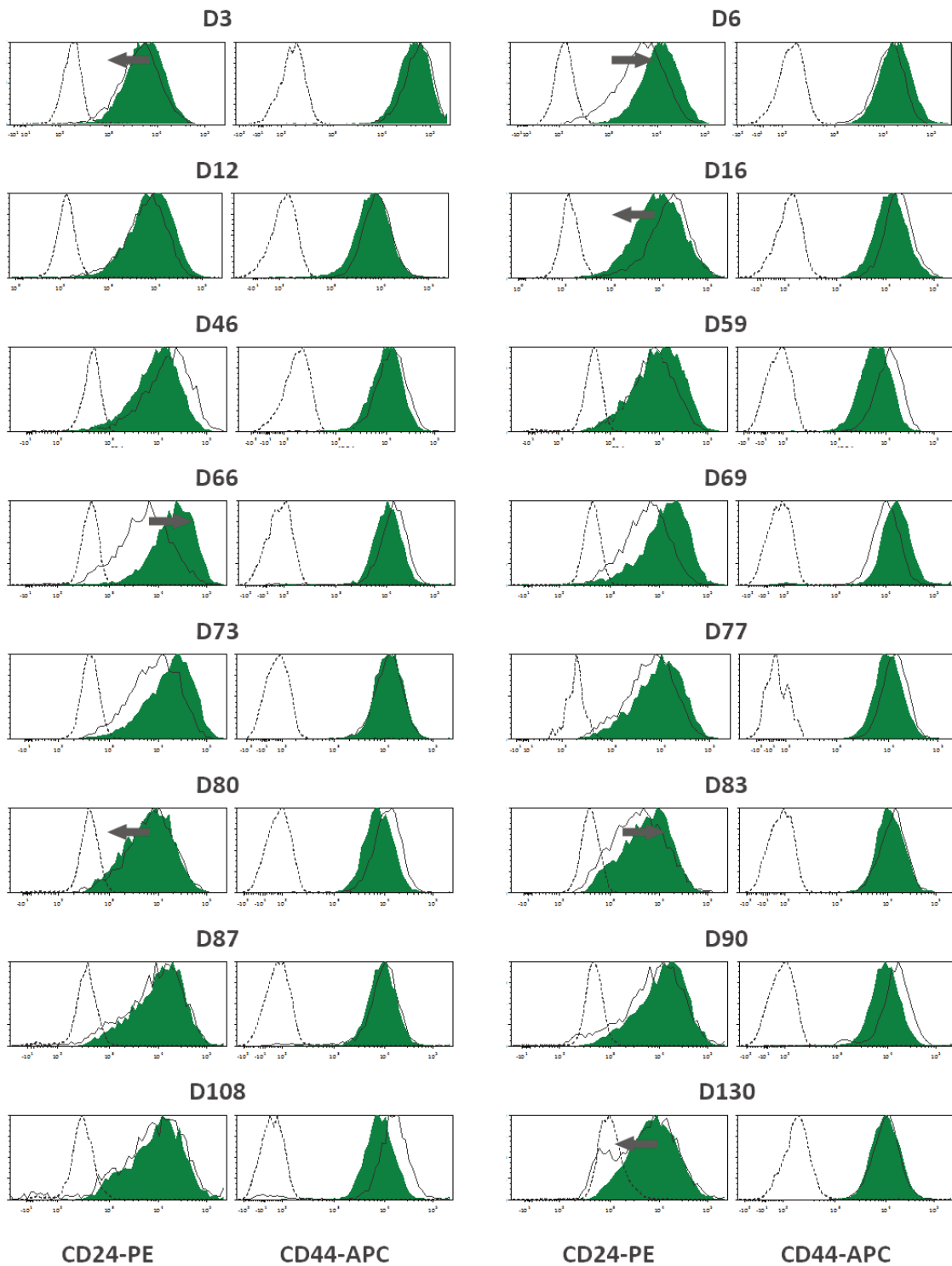


Figure 6.8 : Long-term monitoring of CD24 and CD44 expression among viable repopulating CD24+/CD44+ SUM149-PT cells and comparison with unstained and double-stained non-sorted SUM149-PT cells.

Green area selection : fluorescence distribution curve of CD24+/CD44+ SUM149-PT subphenotype repopulation, dashed gray curve : fluorescence distribution curve of unstained non-sorted cell line, full gray curve : fluorescence distribution curve of double-stained non-sorted cell line

6.2.2 Repopulation dynamics of CD24-/CD44+ and CD24+/CD44+ isolated from SUM59-PT cell line

Variations, such as we described for SUM149-PT, are echoed in the SUM159-PT cell line and are even more pronounced.

The **CD24-/CD44+ subpopulation** re-expresses CD44-low phenotype in proportions similar to that of SUM159-PT cell line (*cf.* section 5.3.3 on page 63 and fig. 6.10 on page 86).

The fluorescence map on CD44 axis follow a mono-modal distribution similar to the non-sorted population (*cf.* fig. 5.10 on page 66). It exhibits variations of the dispersion in the same manner, becoming bimodal at certain time points, D21 and D54 for example (as previously described in section 5.3.3 on page 63). Nevertheless, there is a lag of this fluorescence map of -0.2 to 0.5 log AU compared to the non-sorted population at several time points : D8, D12, D54, D92, D96. The distribution of the population for CD24 marker is unimodal at D6, and then exhibits a right-tailed from D8 to become bimodal at D21. From this date, the fluorescence distribution for CD24 experiences major variations in the relative proportion of two modes. Thus, at D25 CD24-low/CD24-high ratio is 85/15 while at D67 this later is 48/52 fig. 6.11 on page 87.

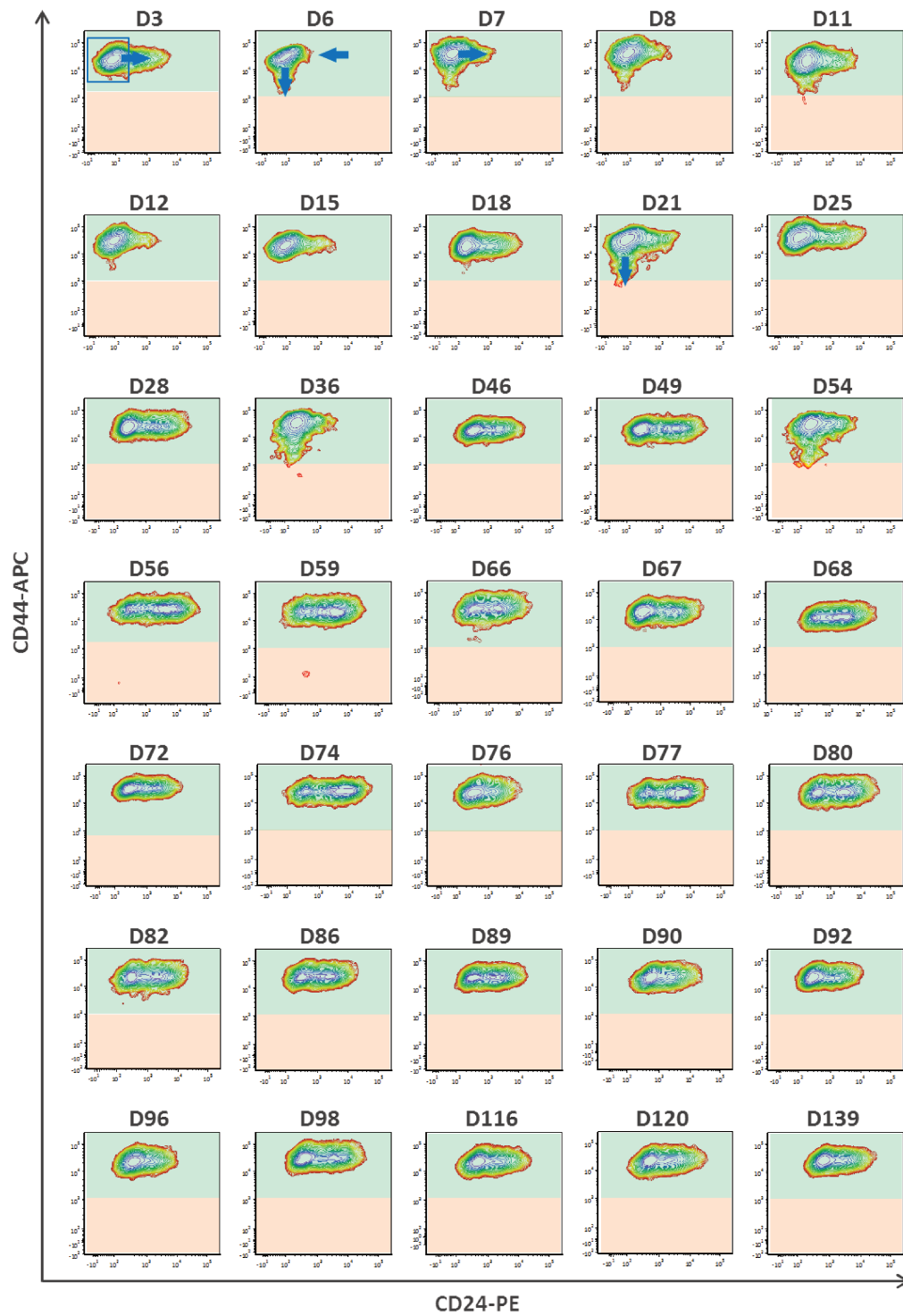


Figure 6.9 : Long-term monitoring of CD24-/CD44+ SUM159-PT cells subphenotypes repopulation.

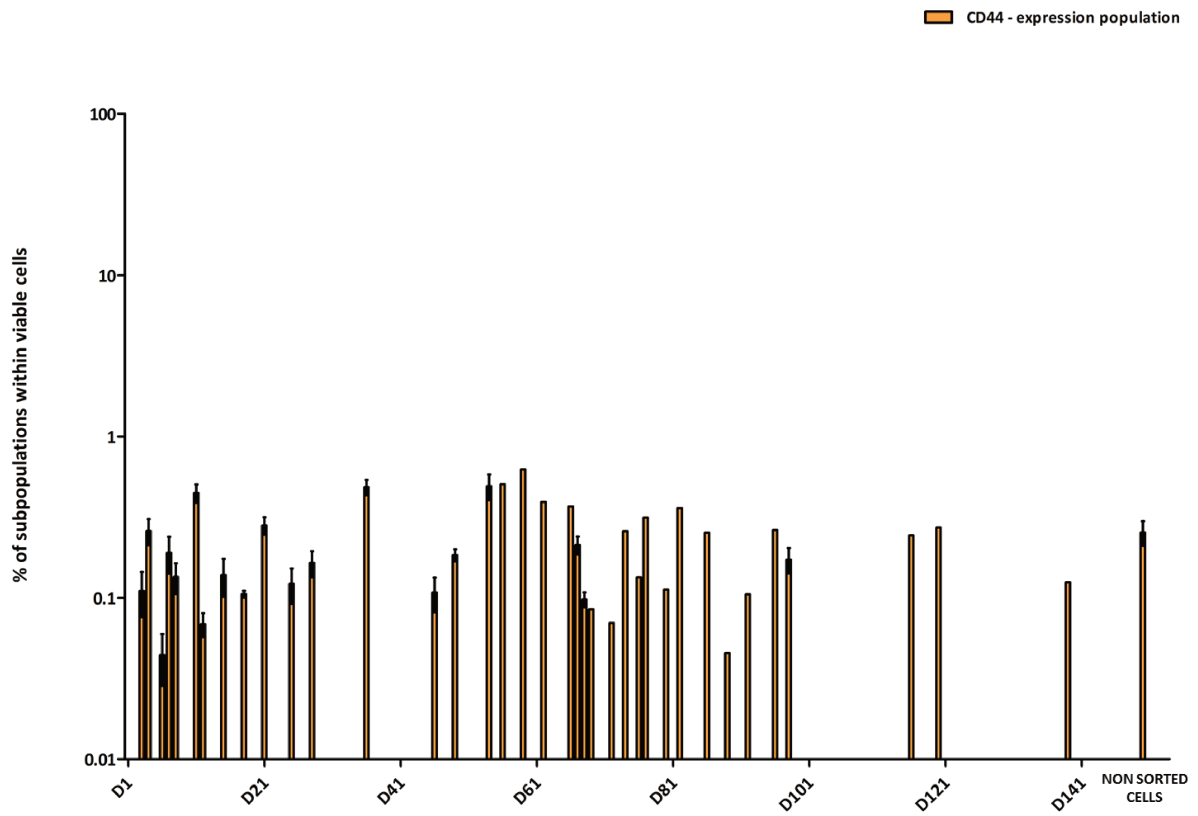


Figure 6.10 : Long-term monitoring of CD44 negative expressing cells re-expression among viable repopulating sorted CD24-CD44+ SUM159-PT cells population.

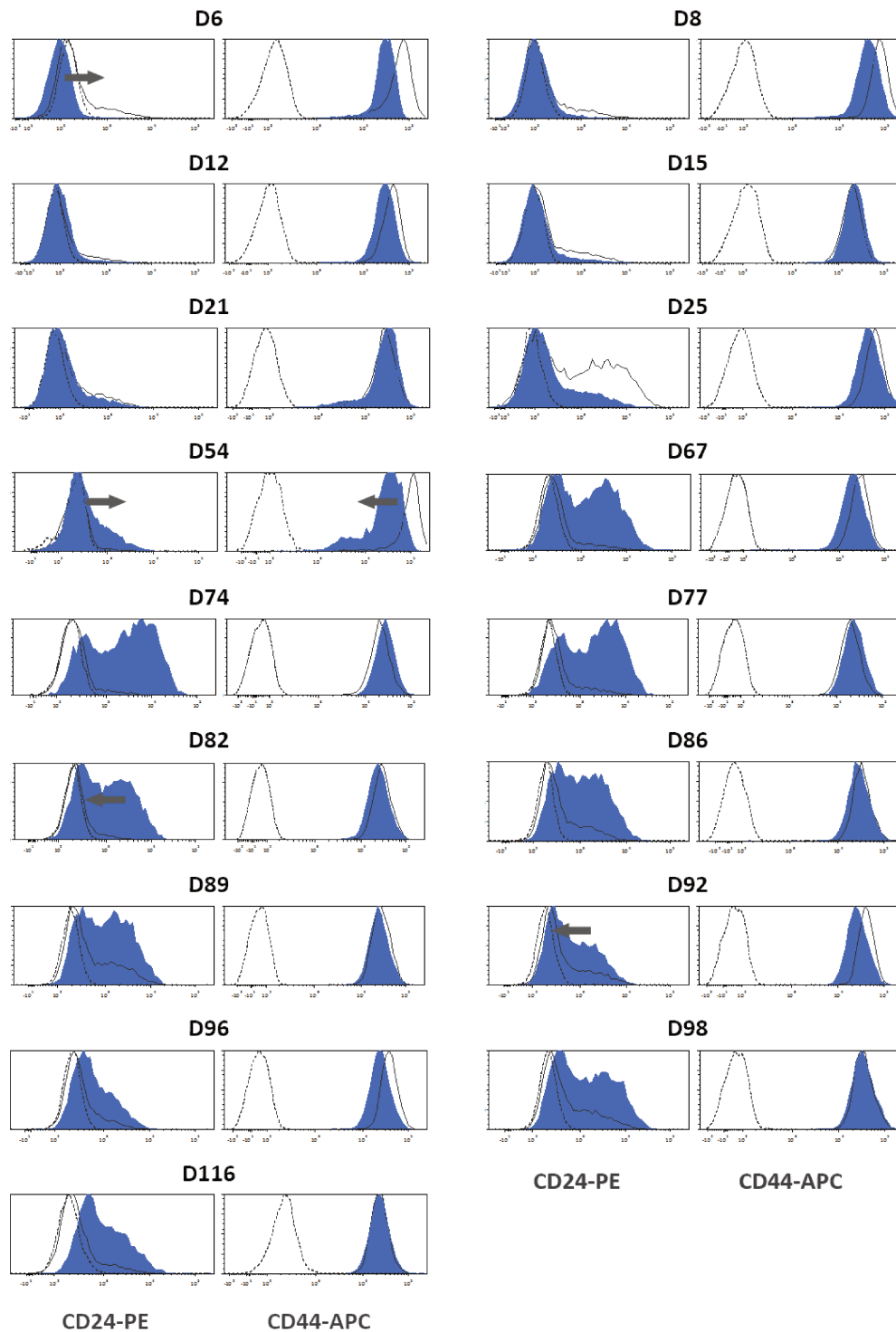


Figure 6.11 : Long-term monitoring of CD24 and CD44 expression among viable repopulating sorted CD24-/CD44+ SUM159-PT cells population and comparison with unstained and double-stained non-sorted SUM159-PT cells.

Blue area selection : fluorescence distribution curve of CD24-/CD44+ SUM159-PT subphenotype repopulation, dashed gray curve : fluorescence distribution curve of unstained non-sorted cell line, full gray curve : fluorescence distribution curve of double-stained non-sorted cell line

Finally, the CD24⁺/CD44⁺ subpopulation, derived from SUM159-PT cell line, re-expressed in the same way the CD44-low phenotype in similar proportions to that of the non-sorted line (*cf.* fig. 6.12 on the facing page and fig. 6.13 on page 90).

The fluorescence map along the CD44 axis does not exhibit any major variations compared to the non-sorted cell population. Therefore, all observed fluctuations occur on the axis of CD24 marker. The distribution of the population follows a bimodal law whose modes ratio established at D6 varies at each following measurements (*cf.* fig. 6.14 on page 91) :

- At D6 CD24-low/CD24-high ratio is 62/38
- At D25 the ratio is 29/71

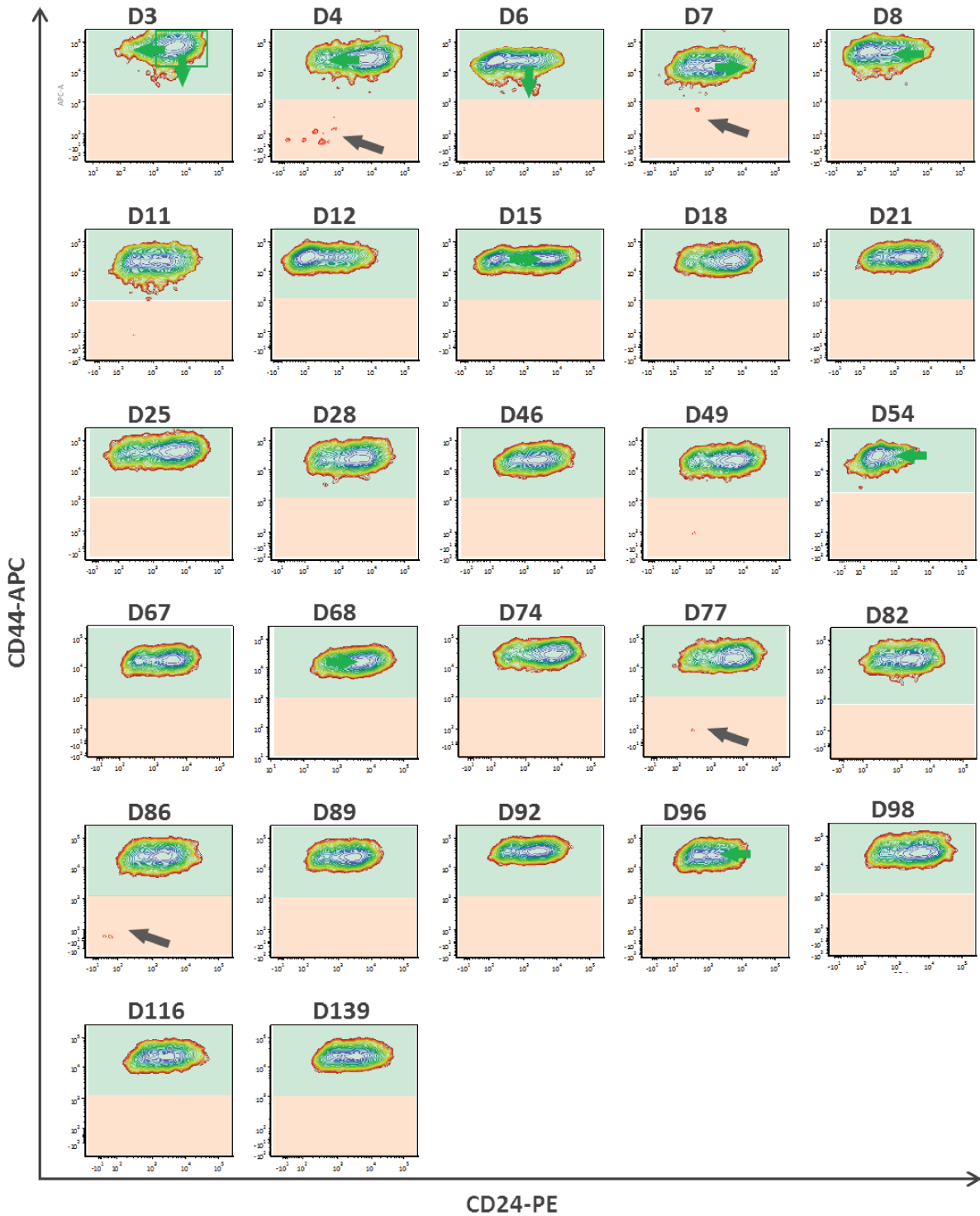


Figure 6.12 : Long-term monitoring of CD24⁺/CD44⁺ SUM159-PT cells subphenotypes repopulation.

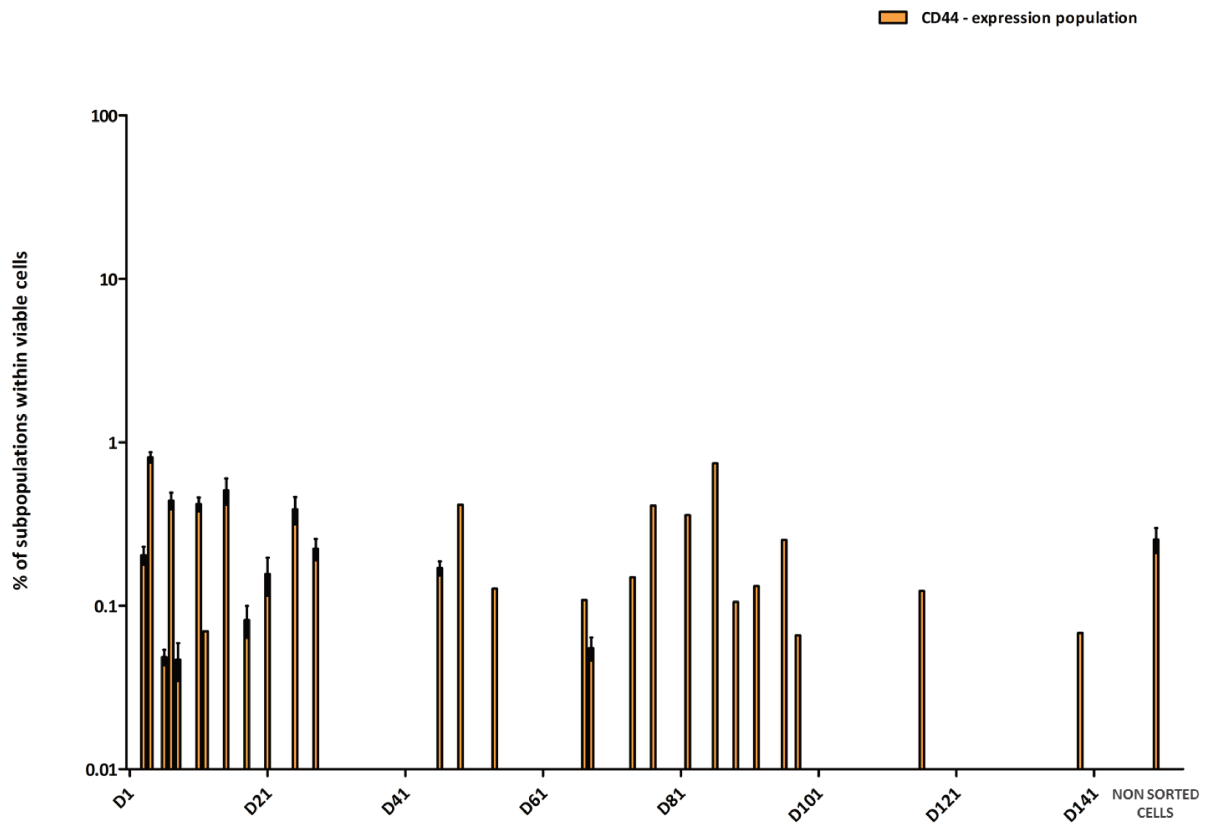


Figure 6.13 : Long-term monitoring of CD44 negative expressing cells re-expression among viable repopulating sorted CD24⁺/CD44⁺ SUM159-PT cells population.

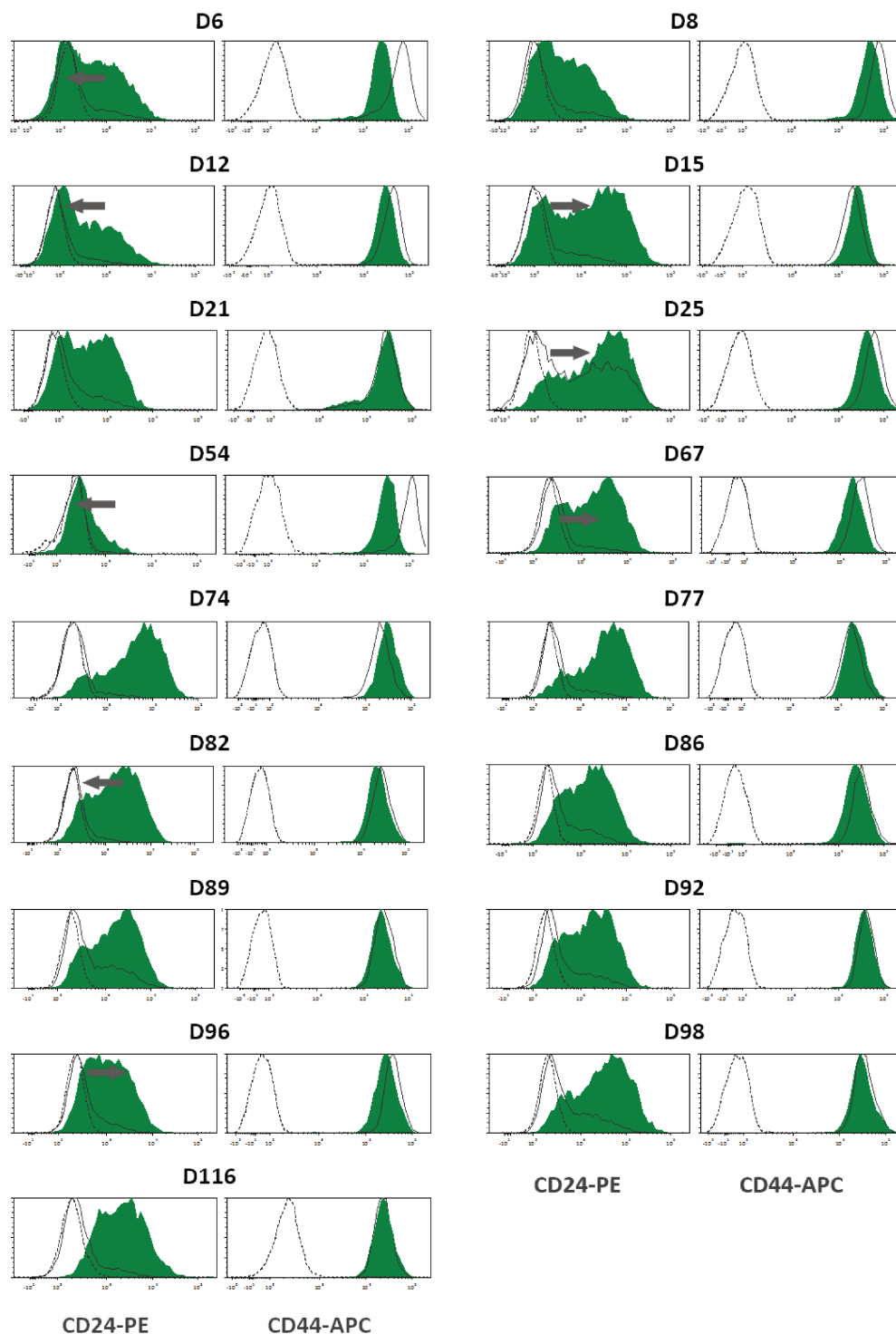


Figure 6.14 : Long-term monitoring of CD24 and CD44 expression among viable repopulating sorted CD24+/CD44+ SUM159-PT cells population and comparison with unstained and double-stained non-sorted SUM-159PT cells.

Green area selection : fluorescence distribution curve of CD24+/CD44+ SUM-159PT subphenotype repopulation, dashed gray curve : fluorescence distribution curve of unstained non-sorted cell line, full gray curve : fluorescence distribution curve of double-stained non-sorted cell line

6.2.3 As a conclusion of CD24-/CD44+ and CD24+/CD44 steady state

To conclude this study on the subpopulation CD24-/CD44+ and CD24+/CD44+, we can highlight the following :

- All sorted phenotypes reproduce CD44-low and CD44-high subpopulations which are quite similar in proportions to the non-sorted cell line.
- The different subpopulations need more than 6 days to return to an equilibrium that stays nevertheless dynamic.
- The dynamic equilibrium depends on the original phenotype and differs from the expression profile of non-sorted cell lines :
 - For phenotype CD24-/CD44+, CD24-low/CD24-high ratio skewed towards CD24-low.
 - For phenotype CD24+/CD44+, CD24-low/CD24-high ratio skewed towards CD24-high.

6.3 On the way of CD24-/CD44- re-equilibrium

The population was highlighted in the SUM149-PT and SUM159-PT lines in different proportions :

- An average of 1% for SUM149-PT (*cf.* fig. 5.5 on page 60)
- An average of 0.5% for SUM159-PT (*cf.* fig. 5.9 on page 65)

The first experiences to isolate this rare phenotype validated its plasticity. Indeed, from D6 onward, this subpopulation strongly re-expressed the CD24 and CD44 markers (*cf.* fig. 6.15 on the next page).

Both the diversification of the subpopulation and its steady state present a major interest.

6.3.1 Looking for steady state from CD24-/CD44- subpopulation derived from SUM149-PT cell line

The phenomenon of diversification of CD24-/CD44- subpopulation was monitored over 68 days (*cf.* fig. 6.15 on the facing page).

Falling from 100% at D0 to 16.5% at D6, the proportion of this pool increases again until it reached 78% around D12–D15. Then, it gradually decreases until the last day of analysis (D68) to around 2% (slightly higher than the non-sorted population rate) (*cf.* fig. 6.16 on the next page). Right from D6, CD44-high cells (newly reappeared) exhibit a fluorescence distribution of CD24 and CD44 markers with the same shape as that of the non-sorted CD44-high population. Nevertheless, we note, until D12, an overshoot of 0.5 to 0.7 log of the mode of the distribution on the axis of CD44. After 21 days of culture, the new population reached its steady state. The fluorescence distribution of CD24 and CD44 markers, for the CD44-high population, is the same as the one characterizing the original population.

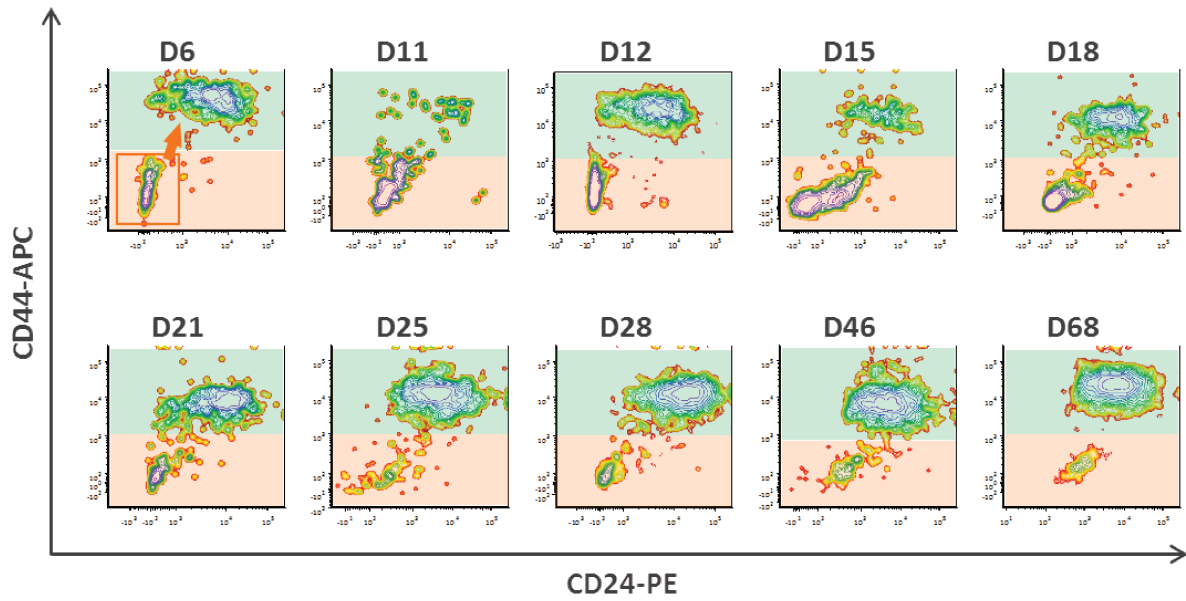


Figure 6.15 : Long-term monitoring of the repopulation of SUM149-PT cells CD24-/CD44-subphenotype.

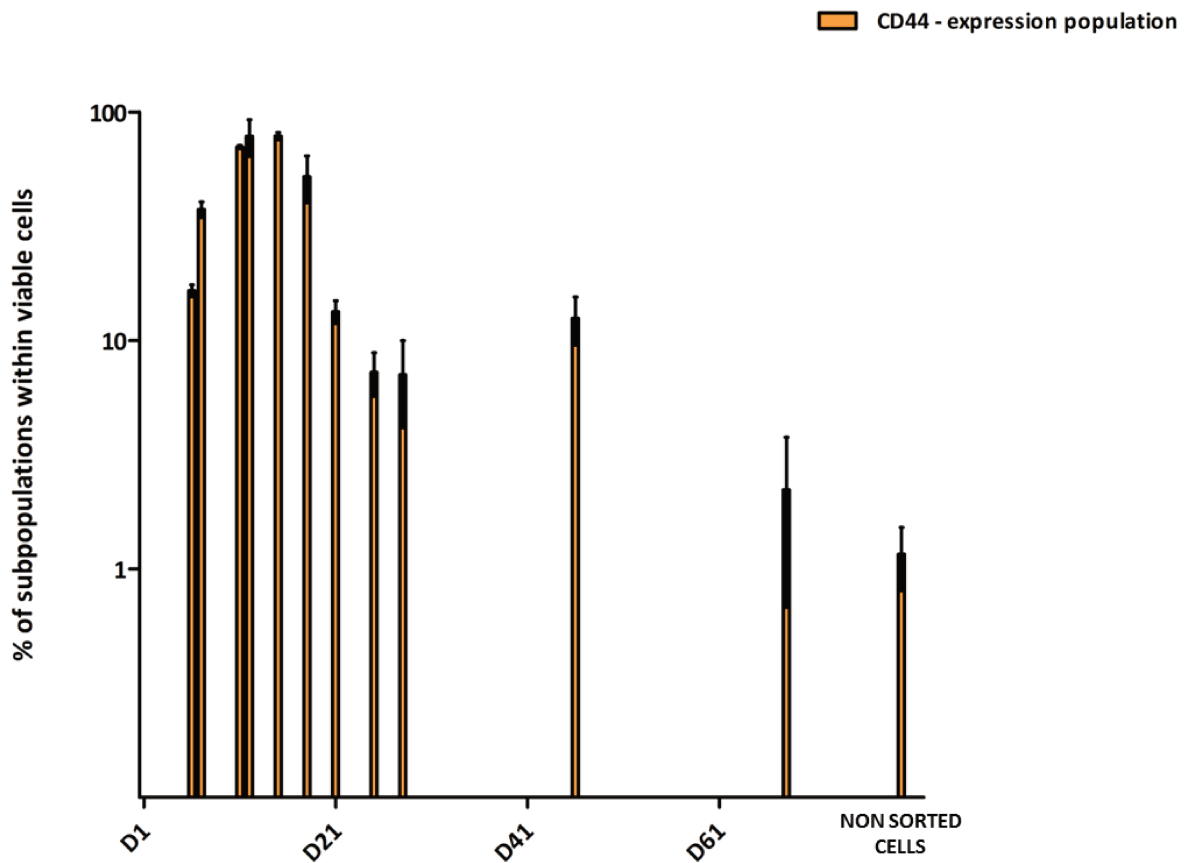


Figure 6.16 : Long-term monitoring of CD44 negative cells among viable repopulating initially sorted from CD24-/CD44- SUM149-PT cells.

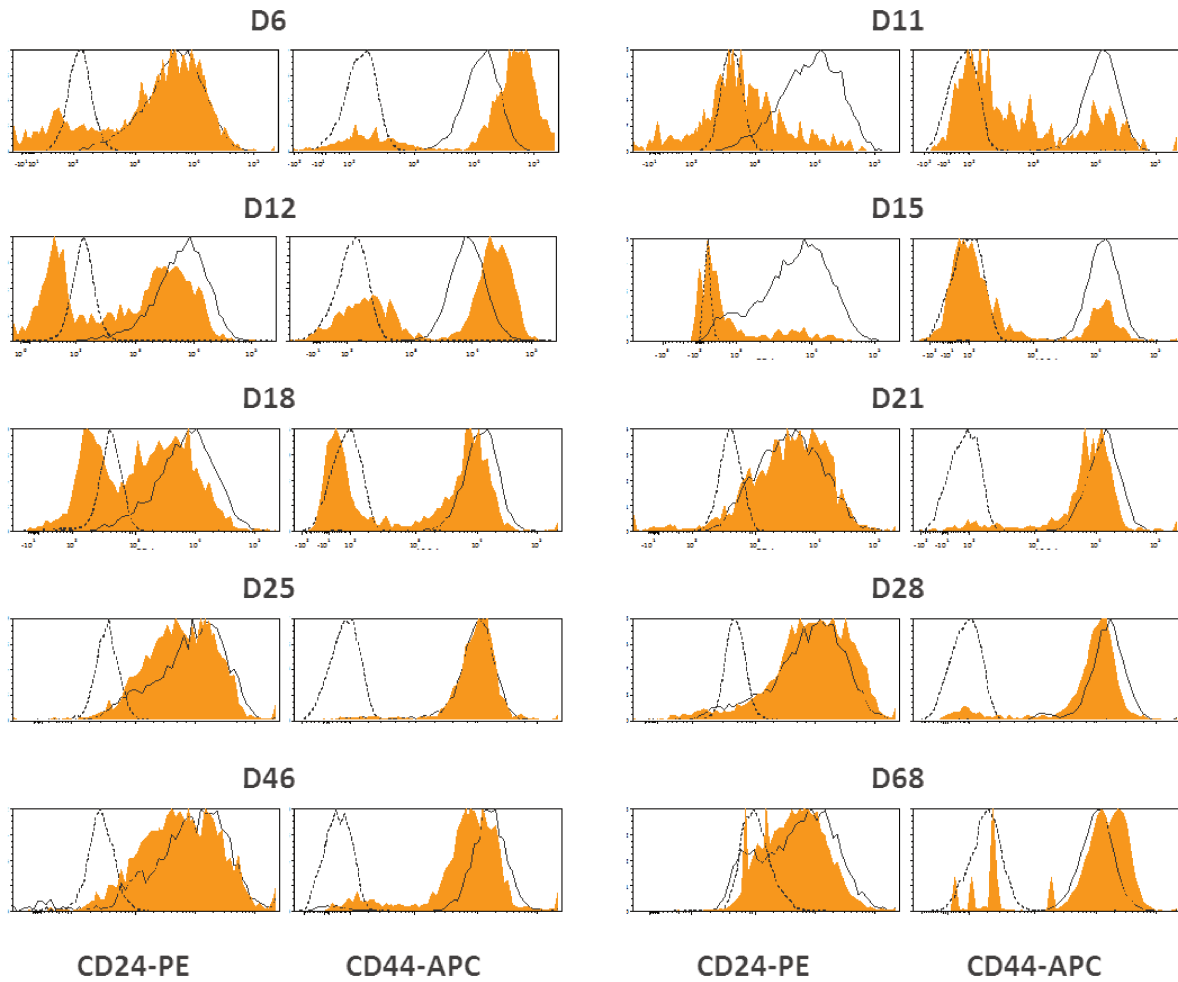


Figure 6.17 : Long-term monitoring of CD24 and CD44 expression among viable repopulating sorted SUM149-PT cells population and comparison with unstained and double-stained non-sorted SUM149-PT cells.

Orange area selection : fluorescence distribution curve of SUM149-PT subphenotype repopulation, dashed gray curve : fluorescence distribution curve of unstained non-sorted cell line, full gray curve : fluorescence distribution curve of double-stained non-sorted cell line

6.3.2 Repopulation of CD24-/CD44- cells isolated from SUM159-PT cell line

Repopulation of the subpopulation derived from SUM159-PT line was studied for 68 days through 16 measurement points (*cf.* fig. 6.18 on the following page).

SUM159-PT derived subpopulation follows the same dynamics of SUM149-PT. Indeed, falling from 100% at day 0 to 3.2% at D6, then the pool of CD24-/CD44- cells grows to 6.5 at D7. Finally, the proportion of the subpopulation decreases and stabilizes from D12 onwards at 0.5% (identical proportion to that of the non-sorted SUM159-PT cell line) as shown on fig. 6.19 on page 97.

At D6, the re-diversified population already exhibits fluorescence maps for the markers CD24 and CD44 that are identical to those of the original cell line :

- The fluorescence map of the CD24 marker has a bimodal distribution with a ratio of CD24-low/CD24-high of 86/14.
- The fluorescence map for CD44 exhibits a monomodal distribution.

However, we can observe at D7 and D15 an overshoot of +0.5 log on the axis of the CD44 marker compared to the non-sorted cell-line (*cf.* fig. 6.20 on page 98). ²

The study related to the dynamics of repopulation of the CD24-/CD44- phenotype brings light to the following points :

- The subpopulation derived from SUM149-PT takes longer than its counterpart SUM159-PT to stabilize (68 days for SUM149-PT against 12 days for SUM159-PT).
- Re-stabilization of proportions follows the following process : decrease, then increase then a last decrease until proportions are similar to the non-sorted population. Nevertheless, the duration of these steps may vary depending on the original cell line (21 days for SUM149-PT against 11 days for SUM159-PT).
- During the first 12 days of repopulation, the mode of CD44-high subpopulation distribution present an overshoot on the axis CD44 compared to that of non-sorted cell line.
- The newly recovered dynamic equilibrium is identical regarding both proportions and characteristics (mode, average, dispersion) to the fluorescence distribution of non-sorted cell line for CD24 and CD44 markers.

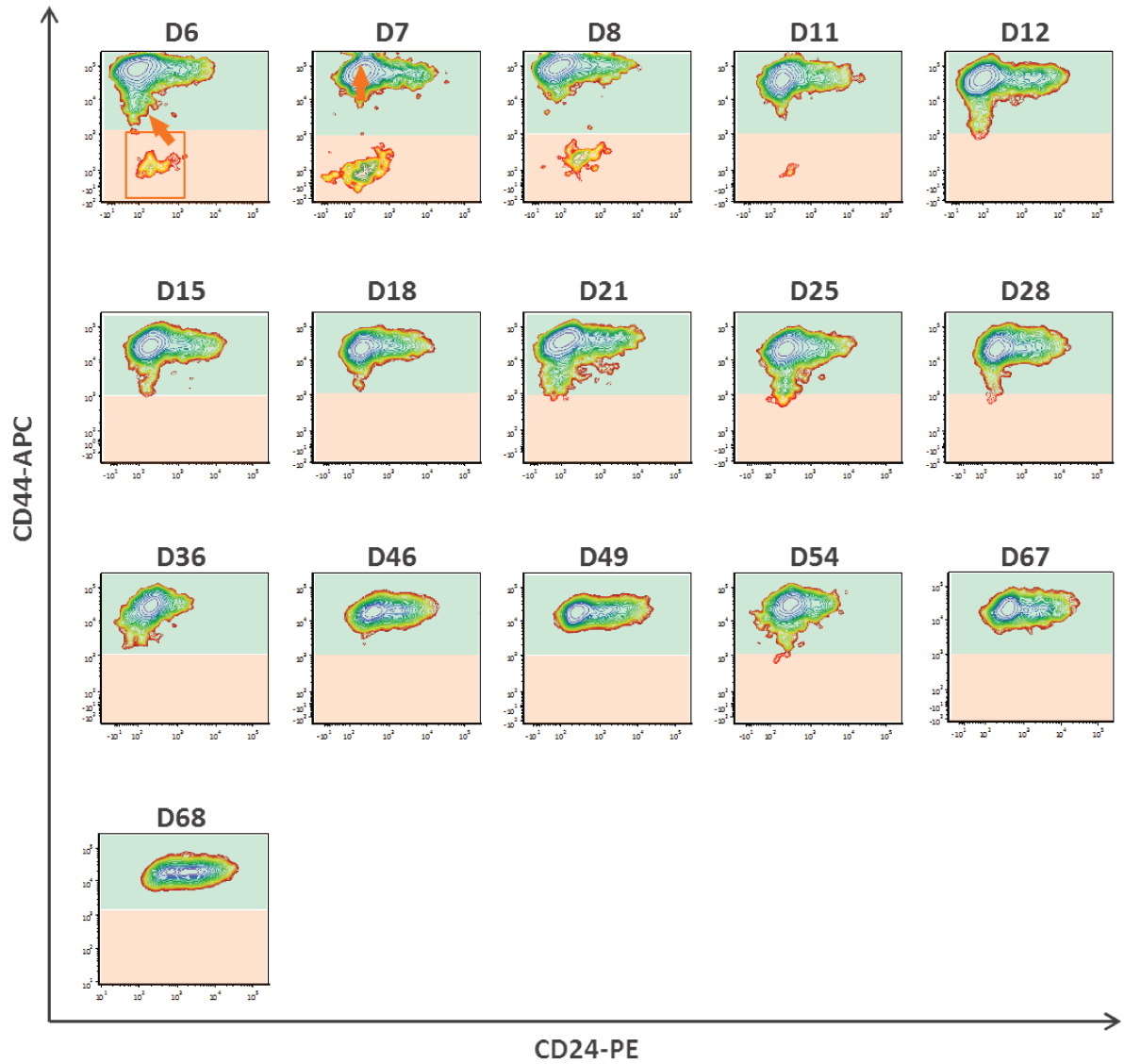


Figure 6.18 : Long-term monitoring of the repopulation from SUM159-PT cells CD24-/CD44-subphenotypes.

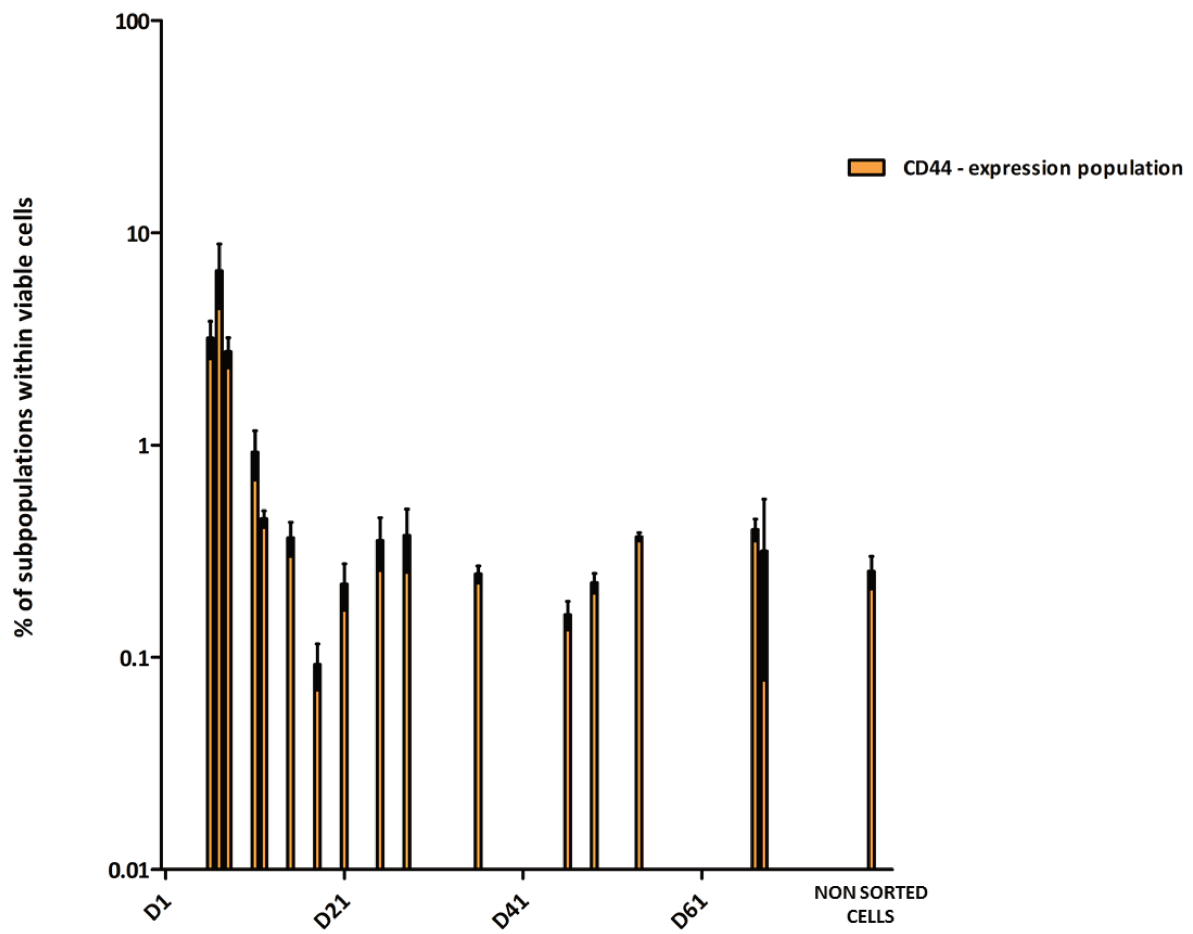


Figure 6.19 : Long-term monitoring of the re-expression of CD44 positive expressing cells among viable re-populating CD24-/CD44- SUM159-PT cells population.

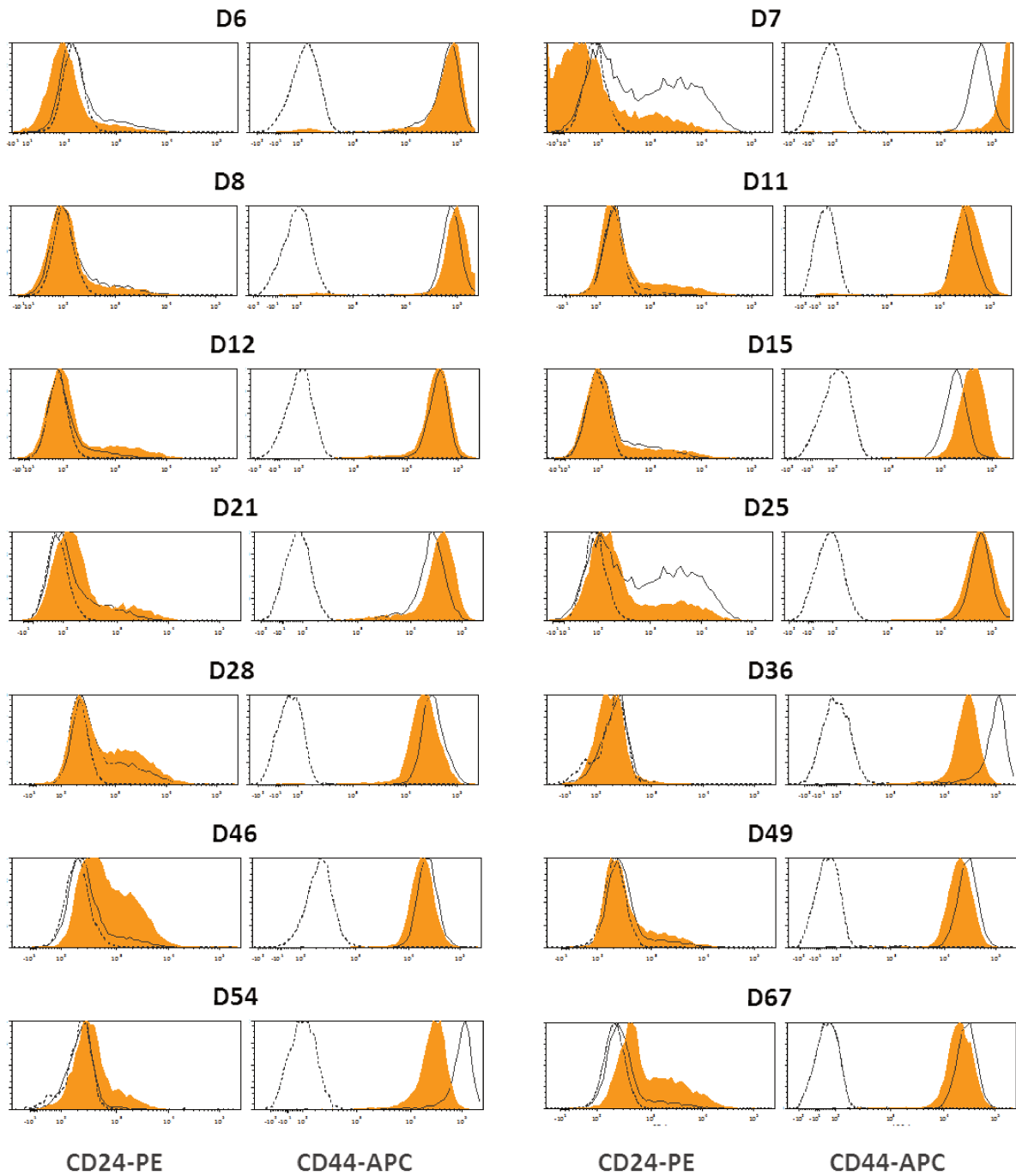


Figure 6.20 : Long-term monitoring of CD24 and CD44 expression among viable repopulating CD24-/CD44- subpopulation and comparison with unstained and double-stained non-sorted SUM159-PT cells.

Orange area selection : fluorescence distribution curve of the repopulation of SUM159-PT CD24-/CD44- subphenotype, dashed gray curve : fluorescence distribution curve of unstained non-sorted cell line, full gray curve : fluorescence distribution curve of double-stained non-sorted cell line

6.4 The equilibrium status of the re-populating subpopulations

A fundamental aspect has to be highlighted following these experiments of phenotype repopulations. Only the CD24-/CD44- phenotype is able to replenish a population whose characteristics (proportion and distribution of markers) are identical to that of the non-sorted population it originates from.

Besides, in average, the monitoring of the several subpopulations during their re-diversification shows that the heterogeneous repopulations retain traces of its origin (*cf.* fig. 6.21 on the next page and fig. 6.22 on page 101). Also,

- The average expression level of marker CD24 in the population issued from the diversification of CD24-/CD44+ subphenotype is generally lower than that of the CD24+/CD44+ subphenotype.
- The average expression level of marker CD24 in the population issued from the diversification of CD24-/CD44- subphenotype derived from SUM149-PT cell line generally ranges from the mean of the population issued from the diversification of CD24-/CD44+ subphenotype and that of CD24+/CD44+. While this subpopulation derived from SUM159-PT expresses more faintly, on average, the CD24 marker with respect to the other repopulations.

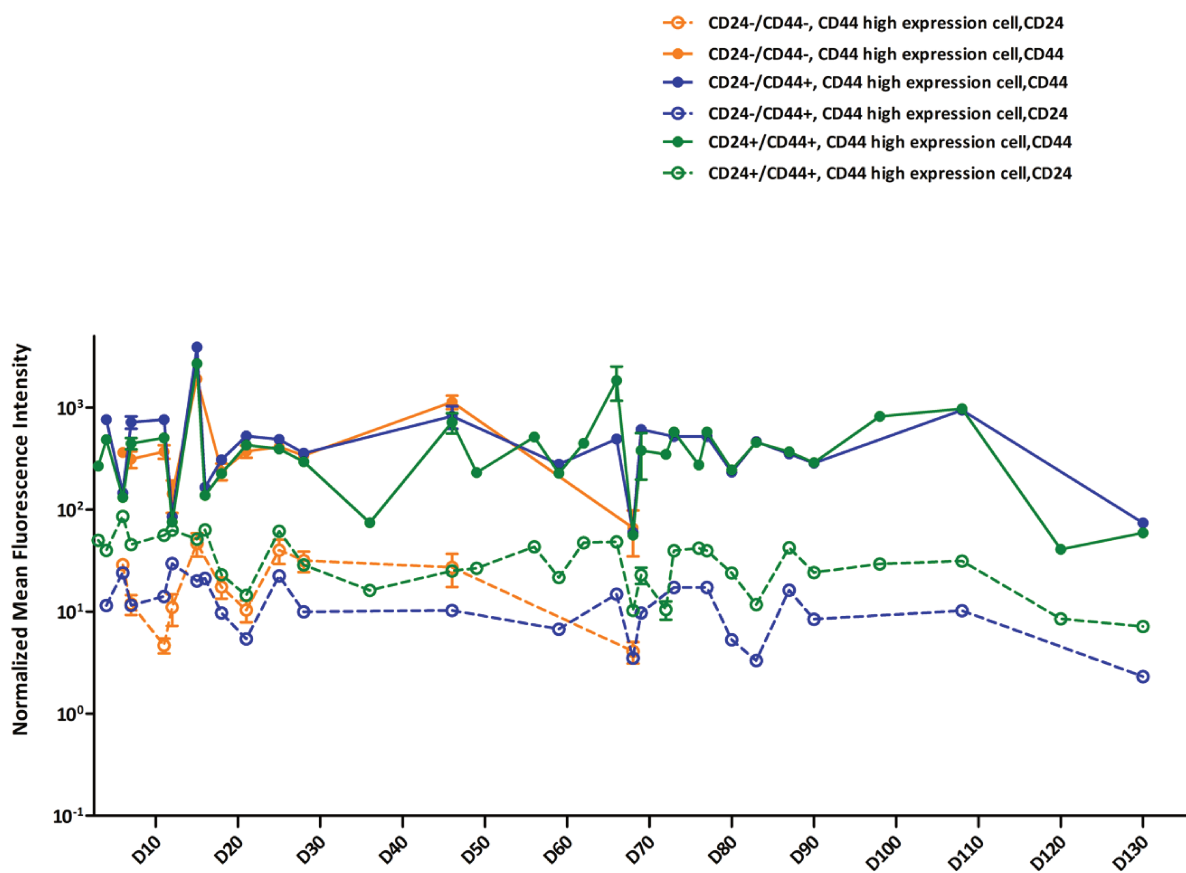


Figure 6.21 : Long-term tracking of CD24/CD44 expression profile by sorted subphenotypes and unsorted SUM149-PT cells.

Grouped column scatter graphing long term following of CD24 and CD44 normalized geometric mean fluorescence intensity of SUM149-PT viable cells.

For each time point, the arithmetic mean of pooled data are plotted on the graph as a bullet point and associated to the calculated SEM (error bar).

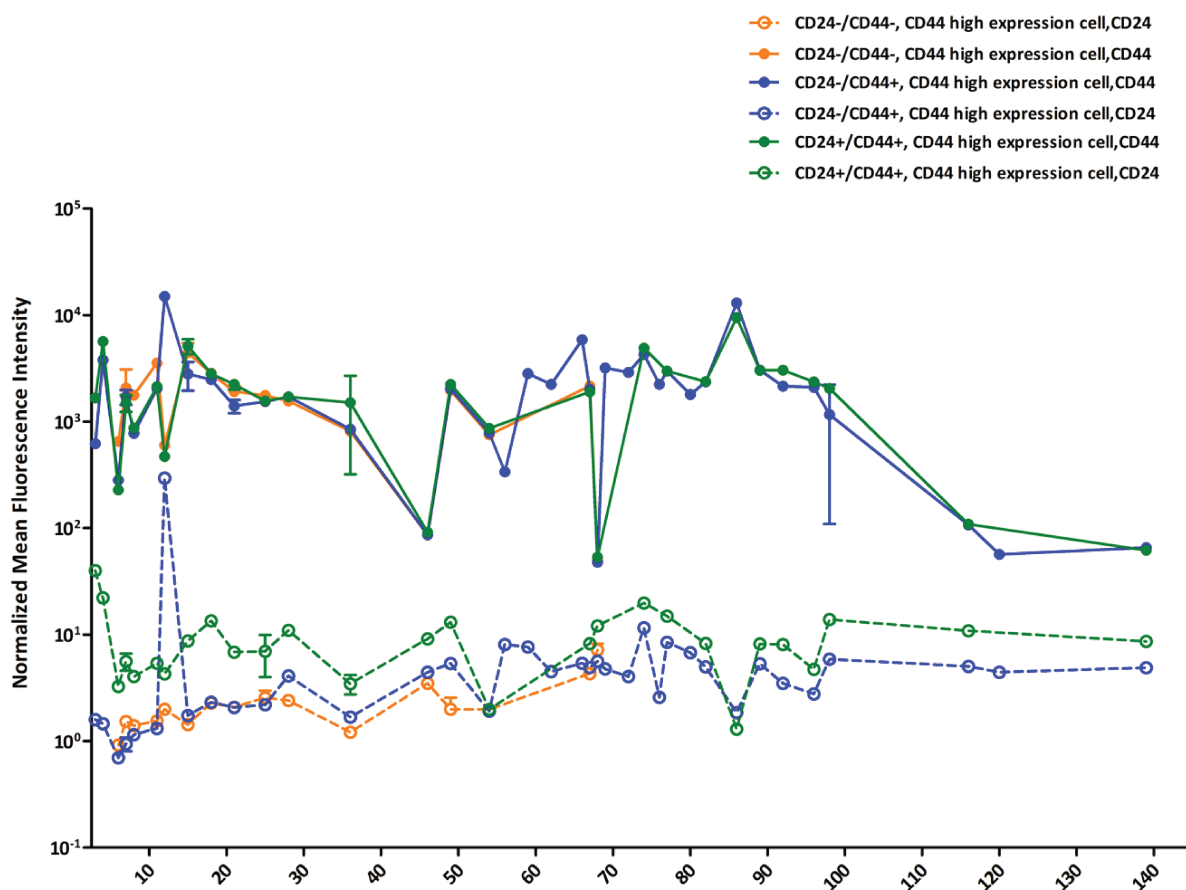


Figure 6.22 : Long-term tracking of CD24/CD44 expression profile by sorted subphenotypes and unsorted SUM159-PT cells.

The arithmetic mean of all replica is plotted on the graph as a bullet point per time point and error bar indicates Standard Error of the Mean (SEM).

Looking for population steady state, influence of cells phenotypic genealogy

*«Qui a jamais demandé à la thèse et à l'antithèse si elles étaient d'accord pour devenir synthèse?»
- Stanislaw Jerzy Lec - Nouvelles pensées échevelées*

On the basis of results discussed in the previous chapter, we had an overlook on dependence or independence from the past regarding the dynamics and outcome of the repopulations. This is what we will discuss in this chapter.

7.1 Memory of the origin over cells generation, over phenotypes, over time on cell fate decisions

Each Sub-phenotypic population issued from primary sorting procedure (P₁) derived from the Primary Sorting (S₁) gave birth to a diverse population consisting of CD24-/CD44+, CD24+/CD44+, CD24-/CD44- (Sub-phenotypic population issued from secondary sorting procedure (P₂)) at D14. Will P₂ cell phenotypes have an identical behavior to that of P₁ ?

- If the cells future phenotype is independent of their past, then whatever the P₁ phenotype which resulted in P₂ (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-), then the repopulation dynamics of P₂ will only be subject to the transition probability laws governing its phenotypic state (and which previously governed the repopulation of P₁).
- If the cells future phenotype is dependent of their past then the repopulation of P₂ will vary depending on the P₁ phenotype from which it derived. Thus, the repopulation of P₂ will reach a different pattern depending on its original condition P₁ (CD24-/CD44+, CD24+/CD44+ or CD24-/CD44-).

7.1.1 Experiment workflow to study cells memory

To address this point, we set up an experimental protocol based on the study of the behavior of second generation cells from the repopulation. The three phenotypic subpopulations of SUM149-PT and SUM159-PT cell lines were isolated via the previously described procedures (P₁). Each

subpopulation was grown for 32 days following the standard culture protocol (P₂). Then the different phenotypes (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-) were isolated from each P₂ and put back in culture. The level of expression of CD24 and CD44 markers was monitored at D14 and D36 by flow cytometry. This experiment was performed once including 2 to 4 replicates per conditions (*cf.* fig. 7.1 on the next page).

In the same manner, the non-sorted cell populations were cultured at a density consistent with the effectiveness of sorting procedure. We have ensured that the plating cell density remains similar for the different samples. However, due to difficulties in sorting, the CD24-/CD44- cells derived from SUM149-PT CD24-/CD44- and CD24-/CD44+ (Repopulating population derived from P₁ (P₁')) population were plated at a density two times lower than that of other CD24-/CD44-cells.

The sorted cells were cultured as described in appendix D.2 on page 249 and passed when they reached 70 % confluence. They were returned to culture at 1/3 of confluence in a new Petri-dish every 3 to 4 days. Three different analyses were performed for each time point on cells sample as previously reported in section 5.2.3 on page 53 :

- A qualitative analysis with a bi-parameter (CD24 *versus* CD44) cytogram contour plot
- A measurement of CD44 low and CD44 high proportion within viable recorded cells
- An analysis of CD24 and CD44 fluorescence map

Acronyms in use are provided here for your convenience :

- P₀= non-sorted cell line
- P₁= phenotypic subpopulation (CD24-/CD44+ or CD24+/CD44+ or CD24-/CD44-) derived from S1 sorting of the original cell line
- P₁'= repopulation from P₁
- P₂= phenotypic subpopulation derived from S2 sorting of P₁'
- P₂'= repopulation from P₂

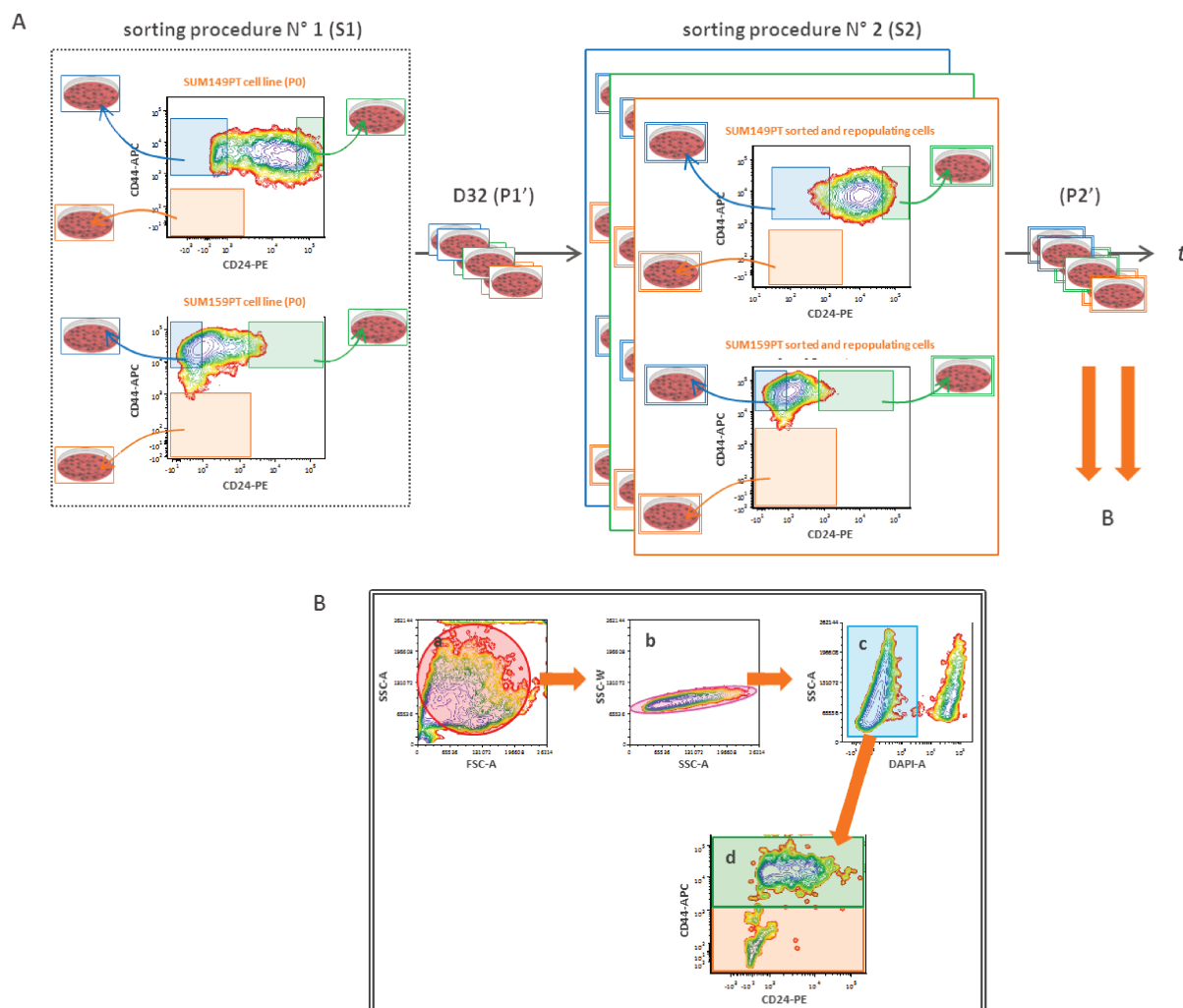


Figure 7.1 : Experimental workflow to monitor SUM149-PT and SUM159-PT cell lines subphenotypes secondary repopulating capacities.

A. Workflow of SUM149-PT and SUM159-PT cells subphenotypes first and secondary sorting strategy. In blue : CD24-/CD44+ cells sorted phenotype, in orange : cells sorted phenotype, in green : CD24+/CD44+ cells sorted phenotype. B.

Flow cytometry gating strategy to characterize secondary repopulating subphenotypes cells culture on CD24 and CD44 markers expression (a) Cytochrome contour plot of light scattered cells characteristics. (b) Cytochrome contour plot of Side Scatter characteristics. Pink selection stands for single cells population. (c) Cytochrome contour plot of cells viability parameter. Light blue selection stands for viable cells population. (d) Cytochrome contour plot for CD24 and CD44 expression markers

7.1.2 Influence of P1 on P2's future - Case of SUM149-PT

We will continue our analysis, taking into account the fluorescence distribution of CD24 and CD44 marker for each P2' populations.

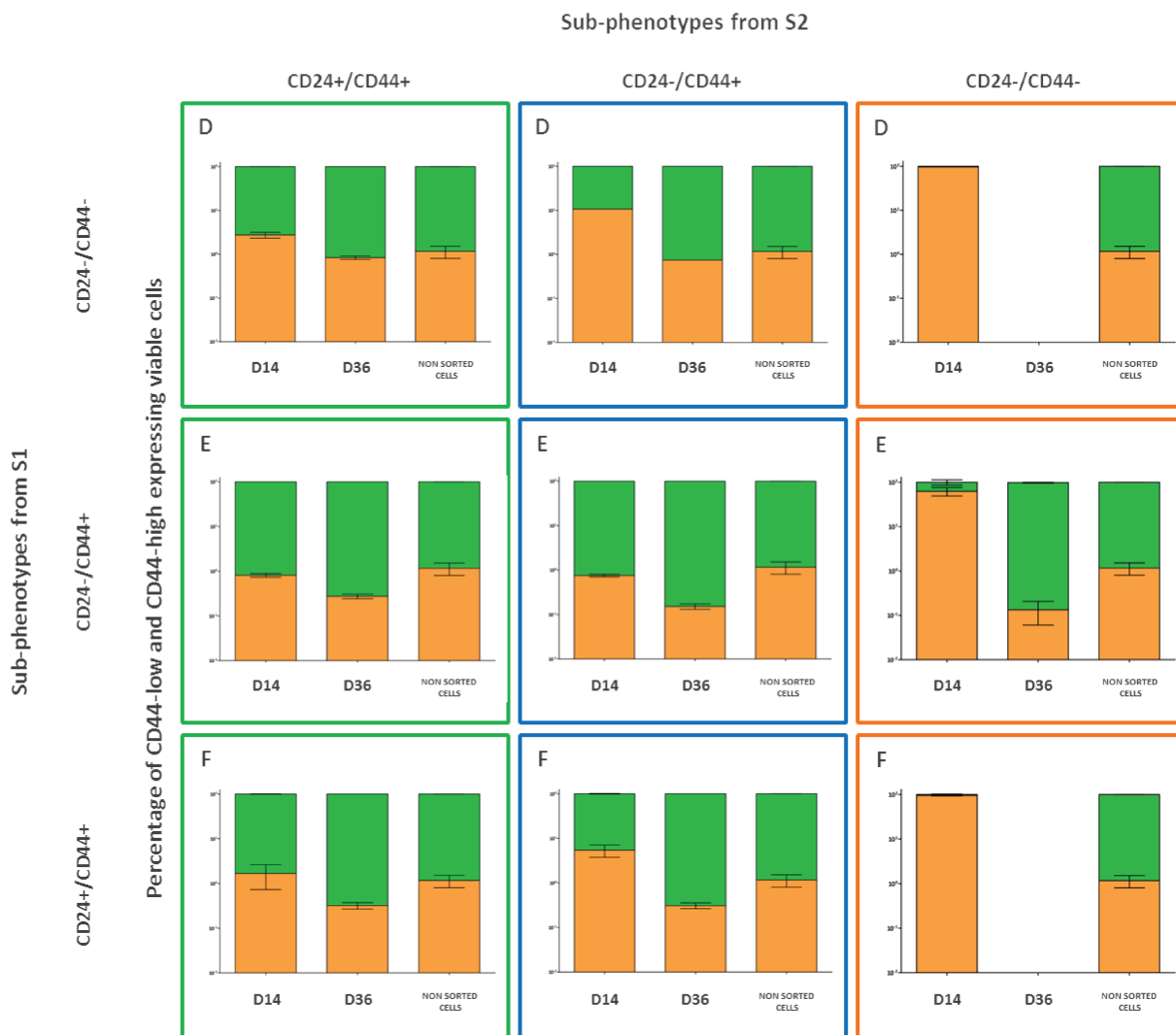


Figure 7.2 : Long-term monitoring of CD44 positive expressing cells re-expression among viable repopulating secondary sorted SUM149-PT subphenotypes cells.

D. Vertical stacked bars graph presenting the percentage of CD44 negative (in orange) and CD44 positive (in green) expressing cells characterizing secondary repopulation of primary sorted CD24-/CD44- subphenotype cells. *E.* of primary sorted CD24-/CD44+ subphenotype cells. *F.* of primary sorted CD24+/CD44+ subphenotype cells
 blue color framing : CD24-/CD44+ cells sorted phenotype, green color framing : CD24+/CD44+ cells sorted phenotype and orange color framing : CD24-/CD44- cells sorted phenotypes

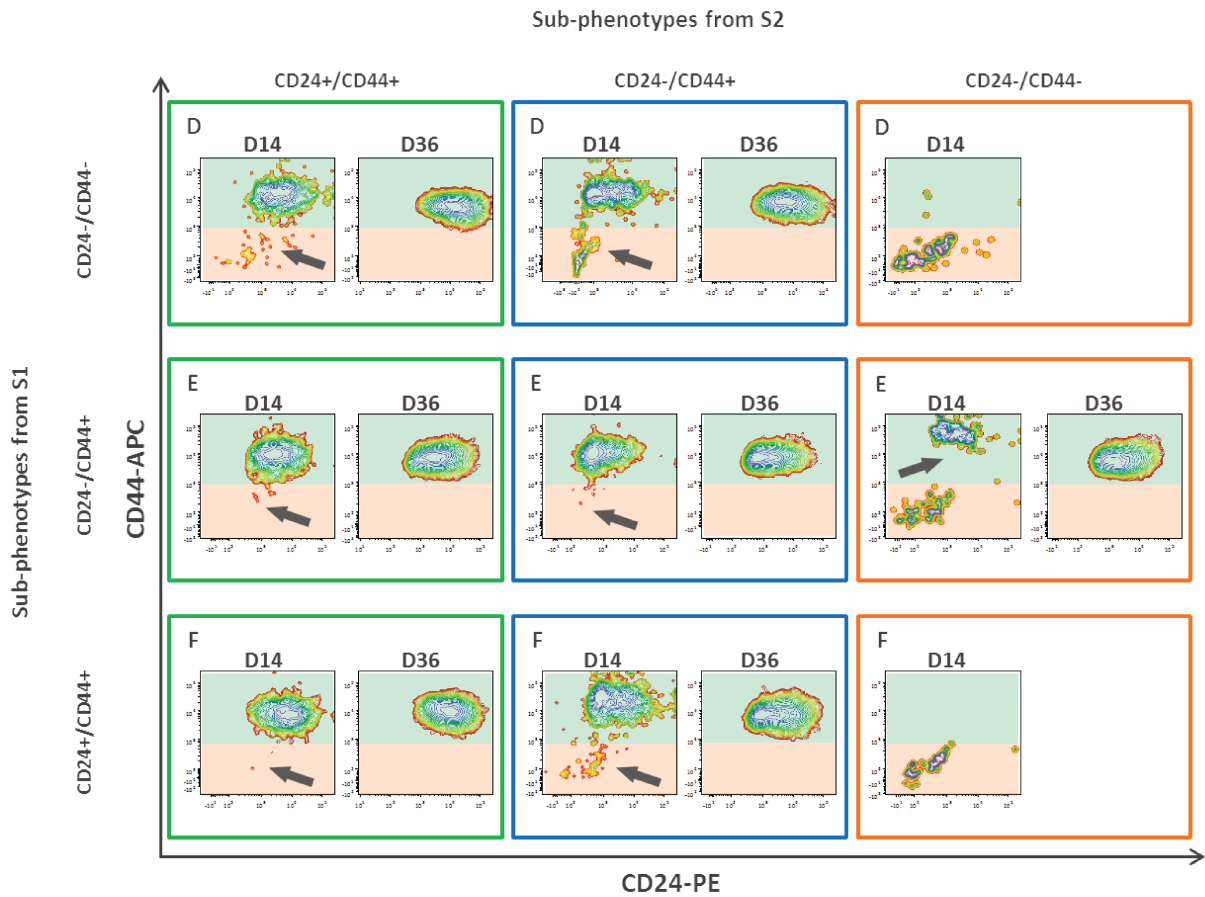


Figure 7.3 : Long-term monitoring of the repopulation of secondary sorted SUM149-PT subphenotypes cells repopulation.

D. Cytogram contour plot of CD24 and CD44 expression markers characterizing secondary repopulation of primary sorted CD24-/CD44- subphenotype cells. *E.* of primary sorted CD24-/CD44+ subphenotype cells. *F.* of primary sorted CD24+/CD44+ subphenotype cells
 blue color framing : CD24-/CD44+ cells sorted phenotype, green color framing : CD24+/CD44+ cells sorted phenotype and orange color framing : CD24-/CD44- cells sorted phenotypes

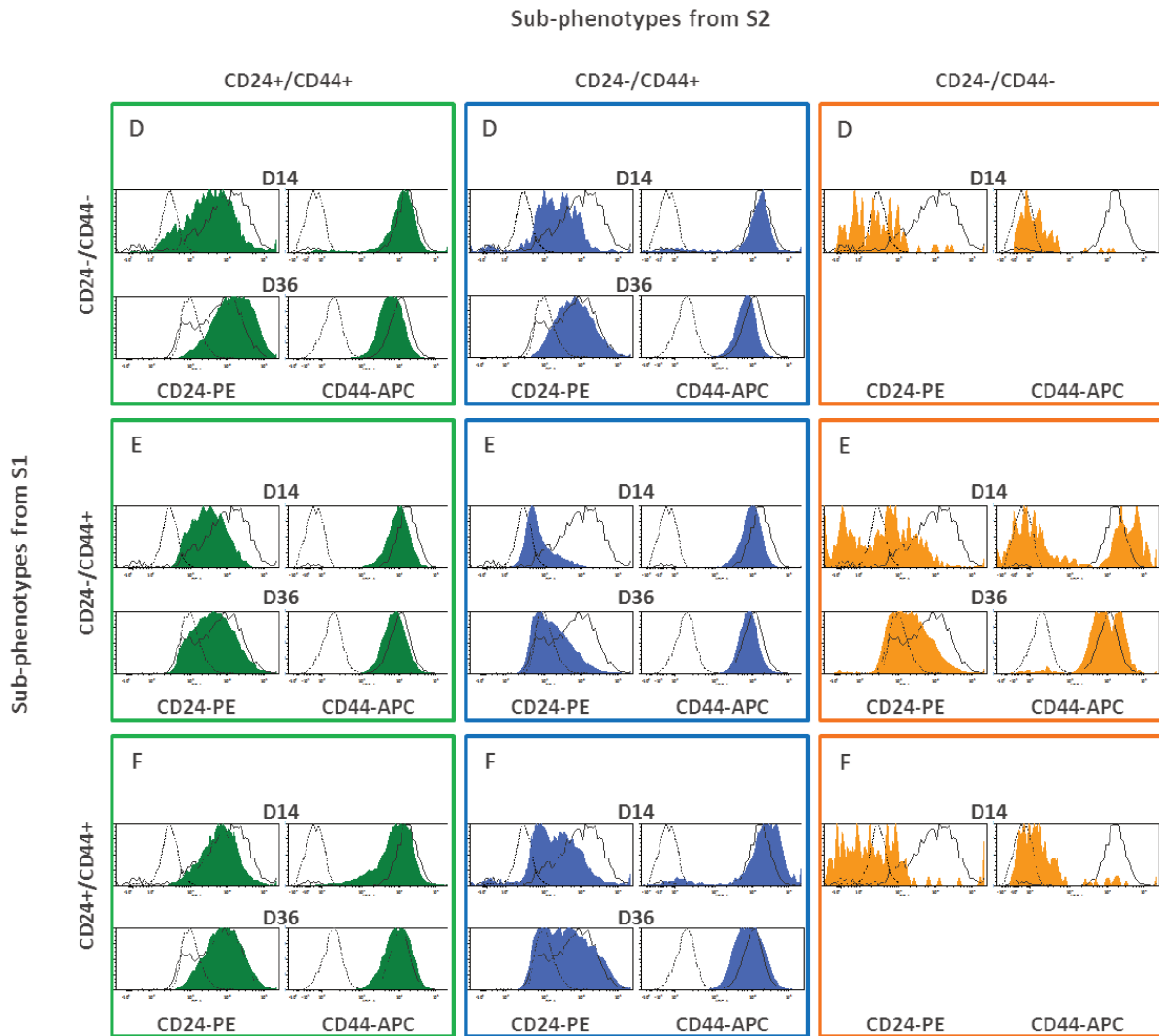


Figure 7.4 : Long-term monitoring of CD24 and CD44 expression among viable secondary repopulating sorted SUM149-PT subphenotypes cells populations and comparison with unstained and double-stained non-sorted SUM149-PT cells.

D. Flow cytometry histograms presenting CD24-PE and CD44-APC fluorescence distribution curves expressed overtime within secondary repopulation of primary sorted CD24-/CD44- subphenotype cells. *E.* of primary sorted CD24-/CD44+ subphenotype cells. *F.* of primary sorted CD24+/CD44+ subphenotype cells

green area selection : fluorescence distribution curve of CD24+/CD44+ SUM149-PT subphenotype secondary repopulation, blue area selection : fluorescence distribution curve of CD24-/CD44+ SUM149-PT subphenotype secondary repopulation, orange area selection : fluorescence distribution curve of SUM149-PT CD24-/CD44- subphenotype secondary repopulation, dashed gray curve : fluorescence distribution curve of unstained non-sorted cell line, full gray curve : fluorescence distribution curve of double-stained non-sorted cell line

Repopulation of P2 CD24+/CD44+

We will begin our study by focusing on phenotypes P2' derived from the CD24+/CD44+ subpopulation and taking into account their origin P1 (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-)

All these populations P2' exhibit a re-diversification of their phenotype at D14. The population CD24-/CD44- appears in proportions higher (3%) than that of non-sorted cell line (P0) whatever its origin P1. At D36 the proportion of CD24-/CD44- decreases and becomes lower than P0 for populations derived from P1 CD24-/CD44+ and CD24+/CD44+ phenotypes. For populations groups derived from P1 CD24-/CD44- phenotype, the proportion of P2' CD24-/CD44- is identical to that of P0 (1%).

At D36, the distribution of the CD44 marker is consistent with that of the original population (P1' CD24+/CD44+) regardless of the original P1 phenotype (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-).

On another hand, the distribution for the CD24 marker exhibits differences :

- The distribution derived from both P1 CD24-/CD44- and P1 CD24+/CD44+ is unimodal and slightly left-tailed. It resembles the distribution population P1' CD24+/CD44+. Nevertheless, we remark that the fluorescence map is more dispersed for the population derived from P1 CD24-/CD44- than the population derived from P1 CD24+/CD44+.
- The distribution derived from P1 CD24-/CD44+ has a bimodal spectrum whose CD24-low/CD24-high mode ratio is equal to 50/50. It strongly resembles the population P1' CD24-/CD44+ at equilibrium.

Repopulation of P2 CD24-/CD44+

The study of P2' phenotypes derived from the subpopulation P2 CD24-/CD44+ similarly shows a phenomenon of rediversification whatever the original P1 phenotype.

The CD24-/CD44- subpopulation is very present for populations derived from P1 CD24-/CD44- and P1 CD24+/CD44+ (between 8 to 10 % of the total population). At D36 the proportion of CD24-/CD44- decreases in the same way as P2 CD24+/CD44+ phenotype as mentioned in section 7.1.2.

The distribution of CD44 at D36 is also consistent with that of the original population (P1' CD24-/CD44+).

As for the distribution for CD24, the three fluorescence maps differ radically depending on the origin P1 :

- The distribution derived from P1 CD24-/CD44- exhibits a unimodal spectrum. It is close to the balanced distribution of population P1' CD24-/CD44- (and of P0 population) whose mode are focused on CD24-high.
- The distribution derived from P1 CD24-/CD44+ is unimodal and skewed (right-tailed), the mode favors CD24-low. Similar to the population P2' which re-expressed CD24-high in small proportion, we concluded that this distribution shows the dual influence of the two sorting S1 and S2 (selecting the phenotype CD24-/CD44+).
- The distribution derived from P1 CD24+/CD44+ is trimodal with the presence of a CD24-low population, a CD24-high population and a CD24-double high population. This profile corresponds to the combination of the fluorescence distribution of CD24+/CD44+ and

CD24-/CD44+ phenotypes as we described them for P1 repopulation in section 6.2.1 on page 76.

So there is an influence from the P1 CD24+/CD44+ population on repopulation of P2 CD24-/CD44+ phenotype P2.

Repopulation of P2 CD24-/CD44-

As for P2 phenotype derived from the subpopulation P2' CD24-/CD44-, we only observed a rediversification of phenotypes for the population derived from P1 CD24-/CD44+.

At D14, the subpopulation CD24-/CD44- represents more than 90% of the population derived from P2 CD24-/CD44+ and 100% of the populations derived from P2 CD24-/CD44- and CD24+/CD44+ phenotype. At D36, only the population derived from P1 CD24-/CD44+ has survived. Also, the proportion of CD44-low phenotype decreased and became lower than that of P0 (0.1% or 1 log difference). **It is interesting to note that only P2 phenotype which, at D14 had already re-diversified its population (with cells positively expressing CD24 or CD44 markers), survived.**

At D36, the distribution derived from P2 CD24-/CD44+ is monomodal for the CD44 marker similarly to population P1 CD24-/CD44-. It is the same for CD24 distribution. The latter is asymmetrical (right-tailed) with a mode centered between CD24 and CD24-low-high. This profile combines the characteristic of P1 CD24-/CD44- (monomodal distribution which is centered on CD24-high) and that of P1 CD24-/CD44+ (whose bimodal distribution has a ratio of CD24-low/CD24-med modes = 50/50).

Thus, the population P1 CD24-/CD44+ influences the repopulation of phenotype P2 CD24-/CD44-.

7.1.3 Influence of P1 on P2 future - Study on SUM159-PT cell line

To analyze the future of the population P2 derived from SUM159-PT cell line we will use the same reasoning as above.

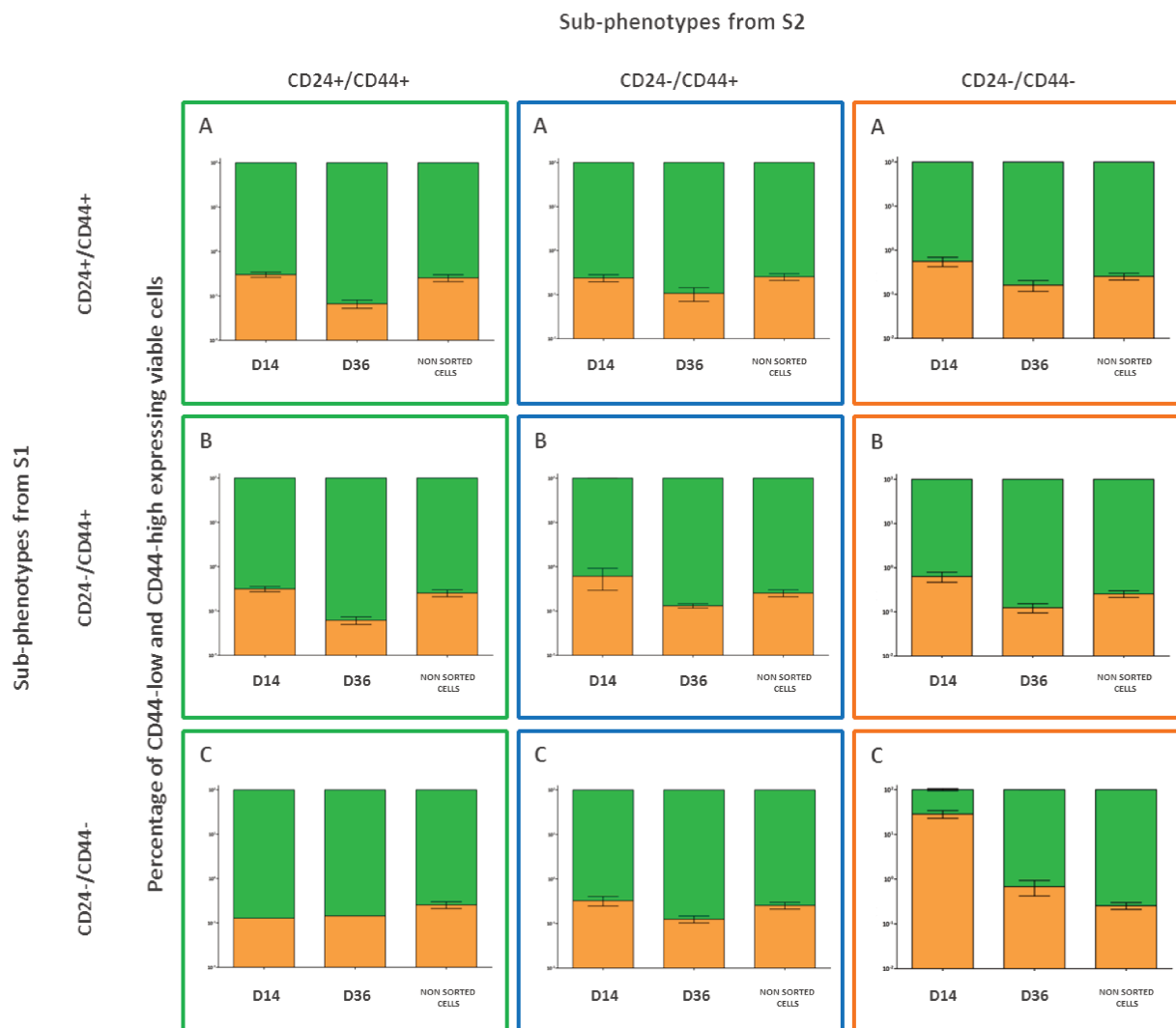


Figure 7.5 : Long-term monitoring of CD44 positive expressing cells re-expression among viable repopulating secondary sorted SUM159-PT subphenotypes cells.

A. Vertical stacked bars graph presenting the percentage of CD44 negative and CD44 positive expressing cells characterizing secondary repopulation of primary sorted CD24+/CD44+ subphenotype cells. B. of primary sorted CD24-/CD44+ subphenotype cells. C. of primary sorted CD24-/CD44- subphenotype cells

blue color framing : CD24-/CD44+ cells sorted phenotype, green color framing : CD24+/CD44+ cells sorted phenotype and orange color framing : CD24-/CD44- cells sorted phenotypes

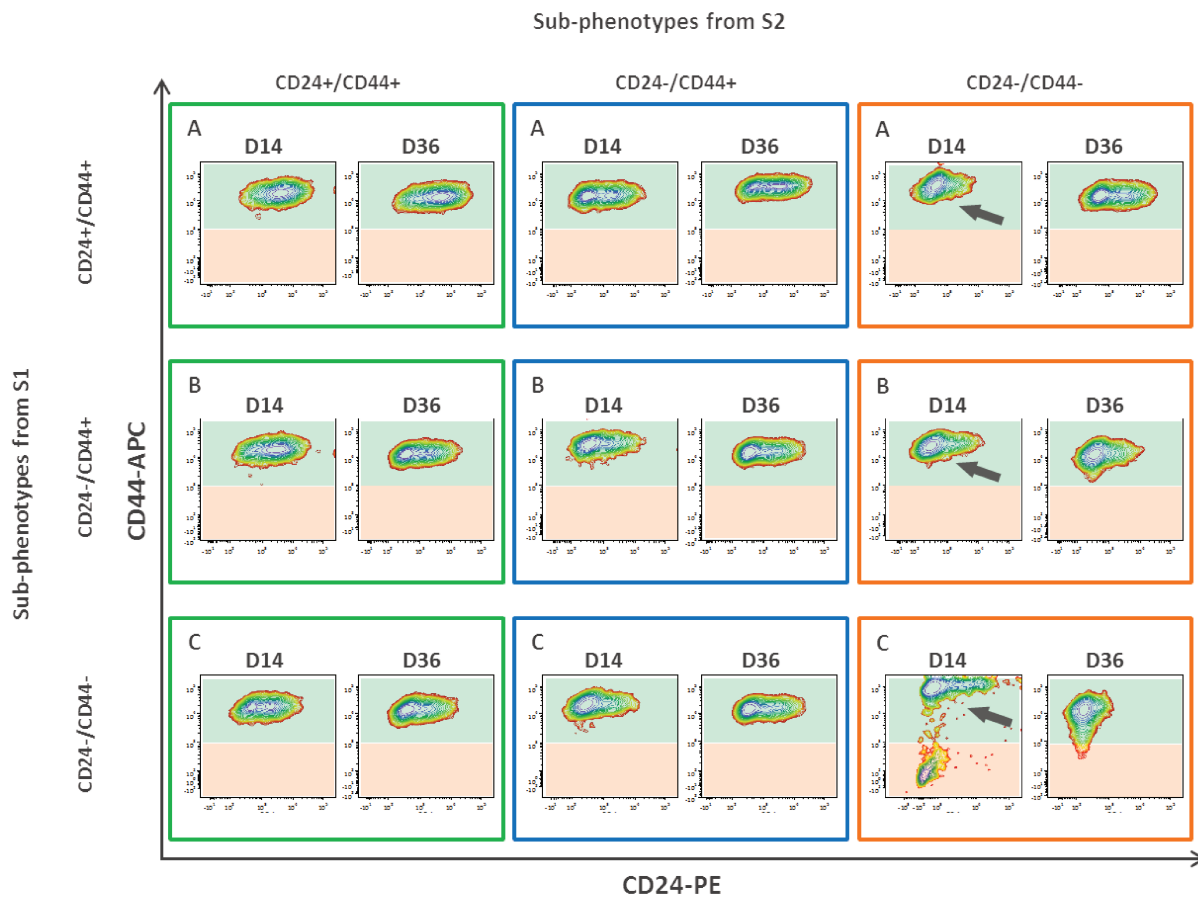


Figure 7.6 : Long-term monitoring of secondary sorted SUM159-PT subphenotype cells repopulation.

A. Cytogram contour plot of CD24 and CD44 expression markers characterizing secondary repopulation of primary sorted CD24+/CD44+ subphenotype cells. B. of primary sorted CD24-/CD44+ subphenotype cells. C. of primary sorted CD24-/CD44- subphenotype cells
blue color framing : CD24-/CD44+ cells sorted phenotype, green color framing : CD24+/CD44+ cells sorted phenotype and orange color framing : CD24-/CD44- cells sorted phenotypes

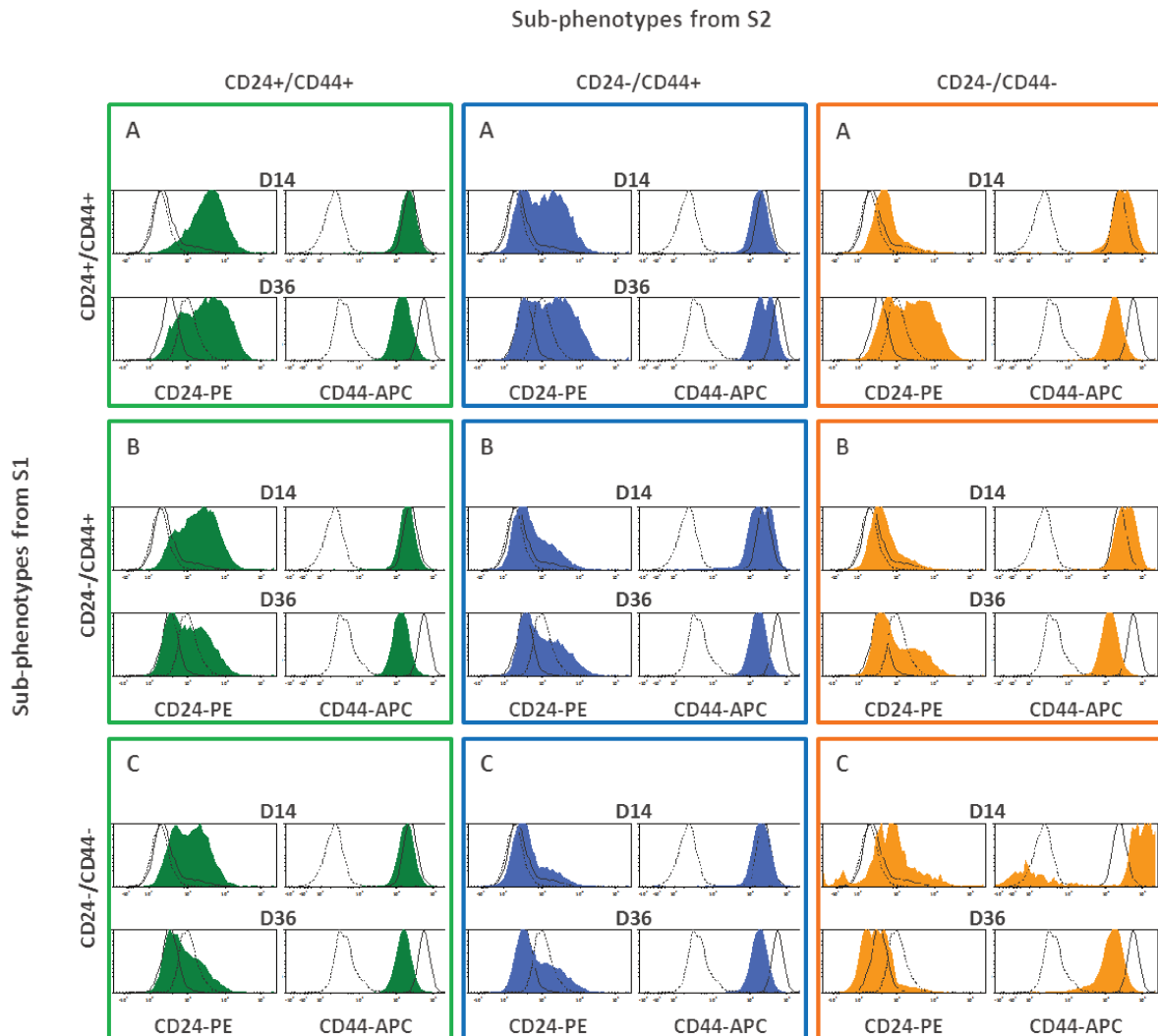


Figure 7.7 : Long-term monitoring of CD24 and CD44 expression among viable secondary repopulating SUM159-PT sorted subphenotypes cells populations and comparison with unstained and double-stained non-sorted SUM159-PT cells.

A. Flow cytometry histograms presenting CD24-PE and CD44-APC fluorescence distribution curves expressed overtime within secondary repopulation of primary sorted CD24+/CD44+ subphenotype cells. B. of primary sorted CD24-/CD44+ subphenotype cells. C. of primary sorted CD24-/CD44- subphenotype cells

green area selection : fluorescence distribution curve of SUM159-PT subphenotype secondary repopulation, bleu area selection : fluorescence distribution curve of CD24-/CD44+ SUM159-PT subphenotype secondary repopulation, orange area selection : fluorescence distribution curve of SUM159-PT CD24-/CD44- subphenotype secondary repopulation, dashed gray curve : fluorescence distribution curve of unstained non-sorted cell line, full gray curve : fluorescence distribution curve of double-stained non-sorted cell line

Repopulation of P2 CD24+/CD44+

The P2' phenotype derived from the P2 CD24+/CD44+ subpopulation rediversifies its population.

At D14, the CD24-/CD44- phenotype represents the same percentage as that observed in the unsorted population. At D36, that proportion decreases for population derived from P1 CD24+/CD44+ and CD24-/CD44+ while it remains stable for the population derived from P1 CD24-/CD44-.

At D36, the distribution for CD44 has the same shape as that of the P1 CD24+/CD44+ population. Nevertheless, we note an overshoot of $-0.5 \log$ (as we have observed in the population P1' CD24+/CD44+ at D6 and D54

The distribution of the CD24 marker has a bimodal shape. The CD24-low/CD24-high ratio differs depending on the origin P1 of the phenotypes :

- It is 35/65 for the distribution derived from P1 CD24+/CD44+. Therefore, the distribution of fluorescence resembles strongly to the latter.
- It is 65/35 for the distribution derived from P1 CD24-/CD44+. This profile combines the fluorescence map of CD24+/CD44+ and CD24-/CD44+ phenotypes (as described in the repopulation of P1). This demonstrates the influence of the P1 CD24-/CD44+ population on the P2 CD24+/CD44+ phenotype.
- It is 60/40 for the distribution derived from P1 CD24-/CD44-. This latter observation shows the particularly strong influence of P1 CD24-/CD44- population on the capabilities of evolution of CD24+/CD44+ population.

Repopulation of P2 CD24-/CD44+

P2 CD24-/CD44- populations derived from P1 subpopulations (CD24-/CD44+, CD24+/CD44+ et CD24-/CD44-) rediversified themselves.

The CD24-/CD44- phenotype exhibits at D14 a proportion equivalent to that of the non-sorted P0 population. At D36, the proportion decreased and is below that of the P0 SUM159-PT population.

CD44 marker has a monomodal distribution equivalent to that of the P1' CD24-/CD44+ population at equilibrium.

As for the bimodal distribution of the CD24 marker, the fluorescence map varies according to the origin P1 phenotype. So :

- For the population derived from P1 CD24+/CD44+, the CD24-low/CD24-high ratio is 50/50. This profile corresponds to the sum of CD24+/CD44+ and CD24-/CD44+ fluorescence maps in equilibrium, as described in the repopulation of P1. It highlights the influences of CD24+/CD44+ population on the dynamics of repopulation P2 CD24-/CD44+.
- CD24-/CD44+ population ratio is 60/40. The profile is thus similar to that of the population P1'.
- For the population derived from P1 CD24-/CD44-, the ratio is 70/30. There is a slight variation compared to P2' CD24-/CD44+ which shows the influence of P1 CD24-/CD44- phenotype on the repopulation of P2 CD24-/CD44+ phenotype.

Repopulation of P2 CD24-/CD44-

Finally, P2 CD24-/CD44- population derived from P1 populations also go through a repopulation process that produces the whole panel of phenotypes (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-).

Indeed, right from D14 there is a strong decrease in the CD24-/CD44- pool of over 99% regarding populations derived from P1 CD24+/CD44+ et CD24-/CD44+ and about 72% for the population derived from P1 CD24-/CD44-. The proportion of this pool of populations derived from P1 CD24-/CD44+ and P1 CD24+/CD44+ decreases and reaches at D36 less than 0.2%, the percentage of the non-sorted SUM159-PT population (P0). As for the population derived from P1 CD24-/CD44-, this percentage reached 0.7% which is slightly lower than that of the P0 population.

The distribution for CD44 has the same profile as that of the population P1' CD24-/CD44-. However, there is an overshoot of -0.5 log on the value of the distribution mode.

The distribution of CD24 is bimodal. The P2' CD24-/CD44- derived from P1 CD24+/CD44+ exhibits a ratio of CD24-low/ CD24-high modes of 50/50. This distribution is identical to that of P1' CD24+/CD44+ population derived from P1 CD24+/CD44+. The distribution of the population P2' CD24-/CD44- derived from P1 CD24-/CD44+ is characterized by a ratio of 70/30, like the fluorescence distribution of P1' CD24-/CD44+ population. And finally, the fluorescence ratio on the CD24 axis of P2' CD24-/CD44- population derived from P1 CD24-/CD44- and evaluated at 90/10 exhibits a fluorescence distribution very similar to that of P1' CD24-/CD44- population. The latter is equivalent to the original P0 SUM159-PT population.

These results highlight that P1 population deeply influences the future of the P2 population (nearly 68 days, about 45 potential cell divisions at the rate of cell division every 36 hours). This suggests that there is a memory of the original phenotype that influences cells decisions across generations. This implies that at the individual level, and depending on its genealogy, a cell in state A does not have the same probabilities of transitioning to a state B or C based on the state history of its ancestor. So if its grandmother was B, then its mother A then cell A' does not have the same probability of becoming B, C or A than if its grandmother and mother were B.

This thought is usefulness only if the CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ sub-phenotypes of P2' are genetically similar to the CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ phenotype of population P1', themselves equivalent to the CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ phenotypes characterized in P0 population. To ensure this essential point we implemented a transcriptomic characterization project of various phenotypic states from different generations of sorting (S1 and S2). We will describe the project in section 12.3.3 on page 193 in more details.

7.2 Influence of the origin through cells generation, over phenotypes, over time on cell fate decisions

This chapter provided a new perspective on the expression dynamics of markers (and phenotypes) that we were able to observe and describe in the chapter chapter 5 on page 49.

Our data conjure up that phenotypically homogeneous populations deriving from heterogeneous populations (depending on the characterization given by the expression of both markers), produced

back very quickly (as soon as three days) heterogeneous populations. This renewed heterogeneity is characterized by the re-emergence of phenotypes that were present in the original heterogeneous population. The characteristics of the renewed diversity depend on the initial phenotype. A particular phenotype CD24-/CD44-, is the only one capable of restoring a diverse population that mimics the original non-sorted population. This diversity is influenced by history of phenotypic states throughout the various previous generations of cells. It is therefore a “dependent memory”.

Associating the importance of the past on the future is a concept widely explored in the field of evolution and development, and this at different levels : society, individual, cell. Regarding the cell, it is part of a differentiation tree and the phenotypic state (which corresponds to the close addition of genetic and environment factors) that characterizes it at time “t” will influence future generations [38, 110, 159]. Thus, a stem cell, following its symmetric and asymmetric cell divisions, gives daughter cells which gradually differentiate themselves and become progenitors more or less committed, more or less differentiated [163, 223, 347]. Through successive cell divisions give rise to differentiated cells that fill a very specific role in the organization and functionalization of tissues [211, 216].

The most obvious example of the influence of memory and of the decisions taken throughout cellular generations, is probably the hematopoietic differentiation tree. Quickly, a hematopoietic stem cell (whose antigenic characteristics are characterized by CD34+, CD59+, CD90/Thy1+, CD38neg/low, c-Kitneg/low and Lin- [155, 207]) with the potential to give the erythroid and myeloid cell lines, gives rise to daughter cells that gradually engage in some branch of differentiation to achieve very precise functions : erythrocytes, platelets, granulocytes... [274]. Finally, the fate of a cell in this tree depends on the degree of differentiation that was accumulated during the generations that preceded it with progressively less and less opportunities to change direction [151]. The other evidence that the past leaves an inerasable mark on the cell and influences its long-term future [67, 217], is demonstrated when technologically trying to rewind the story line from a more or less differentiated cell. This dream is possible through the technology of Induced Pluripotent Stem (iPS) cells that, using a cocktail of transcription factors, is theoretically able to go back in time to a differentiated cell to the stage of pluripotency [142, 236, 346]. Nevertheless, many more evidence are being accumulated in the last five years that show that epigenetic imprint (of the initiating cell) remains permanently recorded in the genome of induced pluripotent cells [250, 304].

The results we have presented in this chapter are part of this approach. And that is why we support the idea that associating the weight of the past in cancer cells is an idea to explore further. Moreover, these outcomes show strong non linearity of the phenotype transition mechanism, and mostly indicate the role of social determinants.

Part III

Clonogenicity, Diversity and Environment influences

«Lâcher prise signifie l'acceptation de la vie en tant que Vie, c'est-à-dire : insaisissable, libre, spontanée et illimitée»
- Pratique Zen -

The first part of this work brought out the large plastic abilities which characterize cell behaviors. Each cell, regardless of its phenotype, is able of flexibility and adapts its own phenotype to let emerge another one. It is clear that cells, each defined by a specific phenotype at a given time, may continuously change and adapt to their environment in order to engender a balance. This balance is dynamic, adaptive to change and information. And more importantly, at each instant this balance is composed of elements with specific characteristics and very diverse identities. Nevertheless, it is impressive to note that these elements, rich with their diversity, agree to find a way of understanding to form a balance, a living society.

Diversity is the master word. This diversity is only possible through a stunning adaptive cell's capacity and rapid changes so that a phenotypic diversity balance could emerge. Nevertheless, each unit of the Society has its own changing capabilities.

Indeed, a selected group of cells features faster flexibility and plasticity. Such as the isolated CD24-/CD44- cells population who gave birth to CD24-positive and CD44-positive cells at an astonishing speed. And more importantly this latter specific cell population is able to reform within 6 days a nearly similar society to the one cellular society which it originates from.

Some cells have different adaptive capacities, no more or less rapid, no more or less effective but whose final outcome (perhaps is it the goal?) is to reform a different societal equilibrium from which they arise. They have the ability to change a pre-existing balance. In human societies they would be characterized and defined as change leaders, able to evolve pre-established models. This is the case for both CD24-/CD44+ and CD24+/CD44+ phenotypic populations whose diversification reform a cell society whose proportions and phenotypic expression characteristics within the phenotypic space differ from those characterizing the original cell Society. To conclude, nature tends to diversity. Nature is rich in diversity, this diversity creates an environment which shapes and makes cells phenotypes what they are, makes cell and human society what it is.

I took advantage of this evidence and go ahead to understand the weight of the environment in shaping the behavior. This reason prompted me to study cells behavior when cells are forced to grow in a three-dimensional configuration and this is the adventure I am about to tell you all along this upcoming part. Nonetheless, you will find out that I did not merely settle for a tri-dimensional growing cells environment. I decided to alter specifications in growing conditions and to set up an environment which would be favorable to the survival of the so-called undifferentiated cells and unfavorable for the survival of the so-described differentiated cells.

So this third part will make explicit the biological, environmental and societal issues in a third dimension. This chapter will describe some interesting differences in the clonogenic abilities of the biological models that I studied. I will argue, in this regard, the impact that some phenotypes

may have regarding the survivability of others. Then my work will bring me to demonstrate that the environment plays a significant role in the phenotypic diversity and the plastic capacities of cells. We will conclude this section by inquiring into the influence of social relationships regarding the cellular behavior.

Characterization of clonogenic capabilities of biological models

« Ce qui relève de la biologie : découvrir, si possible, des règles universelles d'évolution des espèces, mais aussi des règles universelles sur lesquelles sont fondées ces possibilités d'individuation, poussés au plus haut point chez l'Homme. »

- Alain Prochiantz -

Previous studies highlighted that three subphenotypes (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-), each displaying plasticity capabilities, are generating a heterogeneous CD24-/CD44+, CD24+/CD44+ et CD24-/CD44- population in quite similar ratio (CD44-low and CD44-high populations) to the original cell line P0 (*cf.* chapter 6 on page 71). Subphenotype CD24-/CD44- is peculiar as it is able to re-diversify its phenotypes according to an expression distribution for CD24 and CD44 markers similar to P0 cell line as presented in section 6.3 on page 92.

Each of our experiences were conducted in culture conditions of 2D cell adhesion. Our work continued in characterizing the plasticity of each of these phenotypes in an environment requiring them to organize themselves in 3D. We based our study on the fact that cancer cells have large clonogenic capability enabling the production of a tumorsphere from a single cell.

8.1 Description of the experimental methodology followed for *in vitro* tumorspheres-forming assays

As an introduction to the experimental methodology section I would like to warmly thank Dr Marina Glukhova and Dr Aurélie Chiche from the Curie Institute (UMR144) who helped me to design and conduct the following tumorspheres-forming assays. Indeed, I relied on their broad expertise concerning the art of tumorspheres and mammospheres characterization and thanks to their recommendations, donating equipment and time I am able to present this study.

8.1.1 Design of the tumorspheres-forming assay for T1 and T2

Prior to tumorspheres culture, SUM149-PT and SUM159-PT breast cancer cell lines were grown as monolayers according to the protocol in appendix D.20 on page 283 (*cf.* fig. 8.1 on the next page subfig. A). Following the tumorsphere-forming assay protocol adapted for breast cancer cell lines studies by Smart et al. [286, 291], cells were trypsinized from adherent starter cultures, and single cells suspension was washed in large volumes of calcium-magnesium-free PBS in order to remove as much serum as possible.

With regards on tumorsphere-forming assays conducted from specific sorted subpopulations, at that time single cells suspensions were stained for CD24-PE and CD44-APC antibodies as described in protocol appendix D.18 on page 277. Following sorting procedure I have already described in chapter 6 on page 71 (*cf.* fig. 6.2 on page 74) CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ subpopulations were sorted using either an Aria III or an Astrios flow cytometry sorter thanks to the flow cytometry platform of the Curie Institute.

Then cells were filtered through a 40 μm cell strainer and cell concentrations were determined using a Scepter™2.0 Cell Counter (Millipore). Afterward cells were seeded in sphere promoting culture medium at densities of 1×10^4 to 5×10^4 cells/ml in ultralow-adherent 12-well plates (*cf.* fig. 8.1 on the next page subfig. A). After 7 days of culture, the presence of spheres (Tumorspheres derived from primary plating of cells (T1)) was assessed by light microscopy and further analyses were performed (*cf.* fig. 8.1 on the facing page subfig. B).

Afterwards a secondary passage was assessed and tumorspheres or cells aggregates were chemically dissociated with pre-warmed TrypLe and consequent single cell suspension and seeded in sphere promoting culture medium according to the previous protocol at densities of 1×10^4 to 5×10^4 cells/ml. After 7 or 14 days of culture the presence of Tumorspheres derived from secondary plating of cells derived from T1 cells (T2) spheres was assessed (*cf.* fig. 8.1 on the next page subfig. B).

8.1.2 Assessing tumorsphere-forming assay efficiency, characterizing tumorsphere size and phenotypic heterogeneity

Each analysis of T1 and T2 tumorspheres involved firstly the quantification of the tumor-forming assay efficiency, the measurement of the tumorspheres area and the tracking of CD24 and CD44 tumorspheres cells expression. Tumorsphere-forming assays were assessed at least twice for each of the condition I tested.

Sphere-forming efficiency assays T1 and T2 wells were imaged using a digital camera coupled to a dissecting microscope (Leica) with a x6.3 magnification (*cf.* fig. 8.2 on page 124 subfig. Aa). Each digital picture size was scaled up to 1pixel per 6.23 μm and the ROI was manually determined around every tumorspheres using ImageJ software environment (*cf.* fig. 8.2 on page 124 subfig. Ac). The number of spheres up to 40 μm diameter (to discriminate from cell clumps) was counted and used to assess the proportion of spheres formed relative to the number of single cells initially seeded.

Sphere size determination For sphere sizing (*cf.* fig. 8.2 on page 124 subfig. Ba) three random fields per well were digitally imaged using a digital camera coupled to a light microscope (Leica) with a x10 magnification objective. Each picture size was scaled up to 1 pixel per 0.35 μm and sphere area was calculated on scaled image within the manually drawn ROI using ImageJ software (*cf.* fig. 8.2 on page 124 subfig. Bb). I defined the presence of spheres as clearly three-dimensional spherical-like groups of cells growing as dense, floating and compact clusters and clearly distinguishable from loose aggregates of cells.

CD24 and CD44 tumorsphere-forming cells expression T1 and T2 tumorsphere cultures were harvested and dissociated into single cell suspension, then concurrently stained with CD24 and CD44 fluorescent antibody following protocol (*cf.* appendix D.17 on page 274). Raw fluorescent data was collected on LSRII flow cytometer (Becton Dickinson) using FACSDiva acquisition software (Becton Dickinson). Particles and dead cells were excluded based on low light scatter and DAPI negativity. Automatic fluorescence compensation was performed on each occasion and further analysis using FCS Express software (version 4, DeNovoSoftware) (*cf.* fig. 9.2 on page 133). For each experiment, gate based on CD44 expression level was placed based on unstained adherent samples to account for variation in the auto-fluorescence of cells grown in different format. Single color stained adherent samples was checked both to ensure best gate placement and accurately compensation of the fluorescent panels. Further statistical analysis was performed as I previously described in section 5.2 on page 53 using the MATLAB™ environment previously described (*cf.* appendix G.4 on page 325) and data were plotted with Prism 5 for Windows (version 5.01).

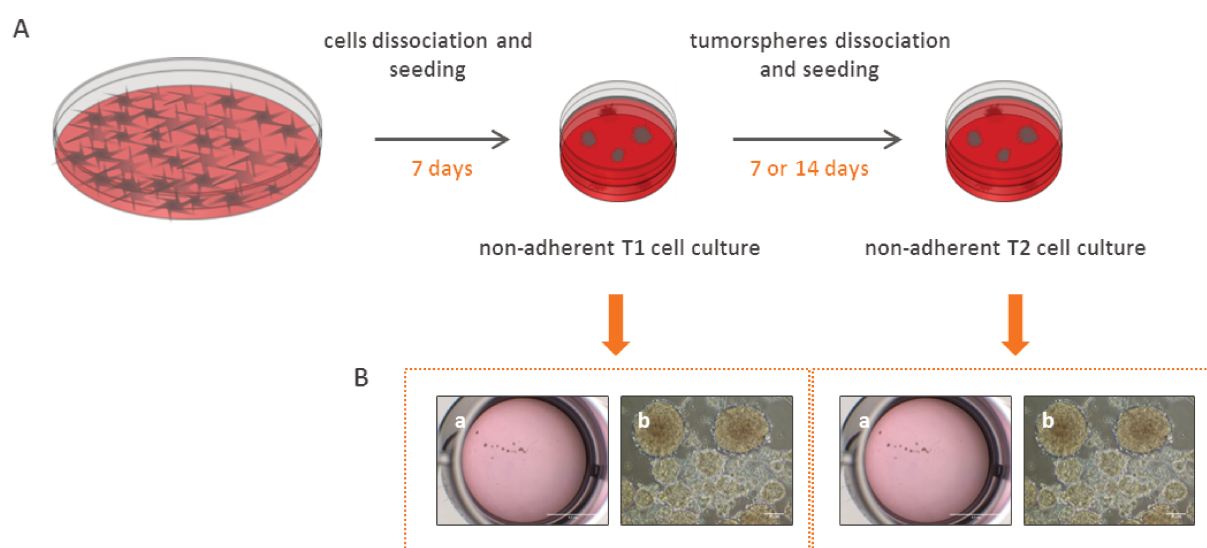


Figure 8.1 : Tumorspheres assay workflow.

A. Experimental workflow to assess tumorigenic potentialities of SUM149-PT and SUM159-PT cell lines and their respective subphenotypes. B. Measurements to characterize tumorigenic abilities of seeded cells (a) count of tumorspheres with a diameter up to 40 μm , (b) measurement of tumorspheres area. (T1 : first plating and T2 : secondary plating)

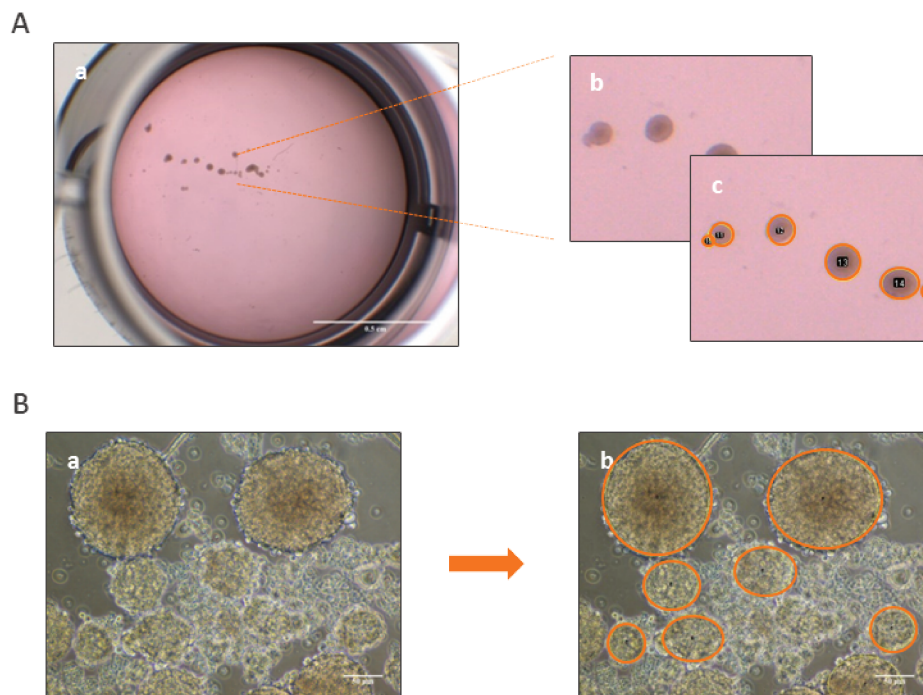


Figure 8.2 : Measurements to characterize tumorigenic abilities of SUM149-PT and SUM159-PT cell lines and sorted subphenotypes following tumorsphere-forming assays.

A. Digitally imaged binocular observations to count tumorspheres with a diameter up to 40 μm (white scale equal 0.5 cm). (a) Calibrated picture taken through a binocular loop x6.3 magnification (picture pixel size= 1pixel/6.23 μm). (b) Zoom of picture (a). (c) Illustration of manually drawn ROIs area measurement. B. Digitally imaged microscopic observations to measure tumorspheres diameter (white scale equal 50 μm). (a) Calibrated picture taken through a transmission microscope x10 magnification (pixel size=1pixel/0.35 μm). (b) Illustration of manually drawn ROIs area measurement

8.2 Influence of quantity and SUM149-PT and SUM159-PT cell line tumorigenicity

8.2.1 Clonogenicity of SUM149-PT and SUM159-PT cell line

Our first work was to validate the experimental protocol for each of the three-dimensional culture of SUM149-PT and SUM159-PT lines. 10 000 cells/ ml in single cell were grown in the NSA medium for each of the cell lines. Then, we looked at their ability to develop clones after seven days in these non-adherent culture conditions.

These experiments were performed independently three times including for each 3 replicates.

Following 7 days of culture we found out that SUM149-PT cell line never gave tumorspheres (*cf.* fig. 8.3 on the facing page subfig. A). However, the cells had multiplied and had grouped together without aggregating and formed shapeless mass. These cells were small, very round (diameter varying between 4 μm and 12 μm) and were refractive. These results contradict some previously published papers that indicate that this cell line would present 3.5 tumorspheres at day 6 for 1000 plated cells [89]. However, they match with pictures published by Debeb et al. [69]. In SUM159-PT cell line, we observed spherical clusters comprised of highly aggregated cells (tumorspheres) [150]. These were of different sizes (*cf.* fig. 8.3 on the next page subfig. B). The smallest, about 40 μm , consisted of particularly refractive cells (in their entirety). The largest however (about 100–150 μm) had an increasing density gradient from the periphery to the center with highly refractive cells and low density outside, low refractive (brownish) and compact inside. These tumorspheres derived from the SUM159-PT lines have been described many times by our

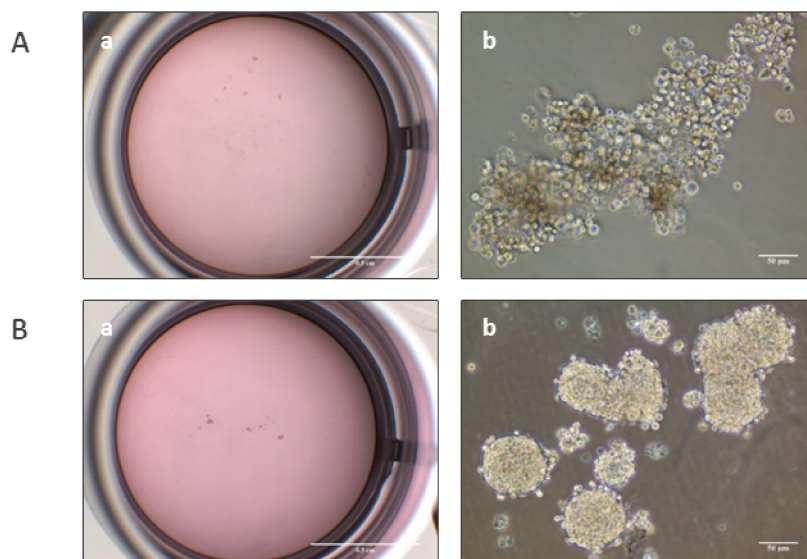


Figure 8.3 : Clonogenic abilities of SUM149-PT and SUM159-PT non sorted cells.

A. Representative calibrated pictures obtained through tumorigenic assay performed on SUM149-PT cell line. B. and from SUM159-PT cell line. (a) Representative calibrated pictures of growing tumorspheres photographed through a binocular loop x6.3 magnification after 7 days in 3D culture conditions (picture pixel size= 1pixel/6.23 μm and white scale equal 0.5 cm) (b) Calibrated picture of tumorspheres issued from SUM149-PT or SUM159-PT cells and taken through a transmission microscope x10 magnification (picture pixel size=1pixel/0.35 μm and white scale equals 50um)

colleagues [50].

8.2.2 Influence of the quantity of plated cells on cell proliferation

To determine the ideal concentration of cells, we cultured 10 000 cells/ml on one hand, and 50 000 cells/ml on the other hand, for 7 days (generation T1 of cell clusters). This experiment was performed 3 times for 3 replicates intra-experiment for the 10 000 cells test and once for 3 replicates for the 50 000 cells test. Then the clusters of cells were dissociated by trypsinization and single cell suspension were cultured again for 7 days (T2 generation). This secondary re-plating was performed on each of the inter-experiment T1 replicates by pooling tumorspheres of all intra-experiment replica. Then, between 1 and 2 intra-experiment T2 replicates were monitored. In the second replating the number of cells plated per experimental condition was :

- For SUM149-PT, 10 000 cells : replating T2 of 10 000 cells
- For SUM149-PT, 50 000 cells : replating T2 of 20 000 cells
- For SUM159-PT, 10 000 cells : replating T2 of 10 000 cells
- For SUM159-PT, 50 000 cells : replating T2 of 50 000 cells

The T1 and T2 cell clusters were photographed using the methodology previously described in section 8.1 on page 121.

The concentration of the original plated cells has no influence on the clonogenic property of SUM159-PT cells (*cf.* fig. 8.4 on the next page subfig. C) across the generations. These observations are consistent with studies published in 2011 by our colleagues [44, 267]. However, variations in concentration strongly influence the proliferative potential of tumorspheres cultured in 3D conditions in both the T1 and the T2 generations as well. Increasing by 5 the concentration of plated cells causes an average increase of the tumor surface (as defined by the projection of tumorspheres on plane x,y) of almost 0.85 log for the T1 generation.

Indeed, when the T1 concentration is 10 000 cells/ml, then the average number of estimated cells for each tumorspheres is approximatively 765 while it is about 18 000 in the T1 tumorspheres whose initial concentration was 50 000 cells/ml. These estimations of the mean number of cells per tumorsphere are based on the mean size of 3D cultured single cells which should probably change when aggregated ; and unfortunately, current transmission microscopic images did not let us to fully enumerate the number of tumorsphere cells. Nevertheless, these observations, regarding change of tumorspheres size in function of the initial concentration of plated cells, are consistent with those presented in the *in vivo* work of Charafe et al. [50].

Since a tumorsphere comes from dividing a single original cell [203, 286], we estimate (based on the previous estimations of cells number) cell cycle time for the 10 000 cells/ml condition at about 17 hours whereas we estimate it at 12 hours for the 50 000 cells/ml condition. Thus, the time between two cell divisions is anti-correlated with the number of cells plated in the initial state and the potential of interaction between cells (proliferative signals...).

However, in the T2 generation, this correlation reverses following a major decrease of proliferative potential in cells that presented a very high rate of division at the T1 generation. This could be explained by a fall in the renewal of cells with proliferative capacity or the development of resistance to the surrounding cell regarding some proliferation signals [41, 241].

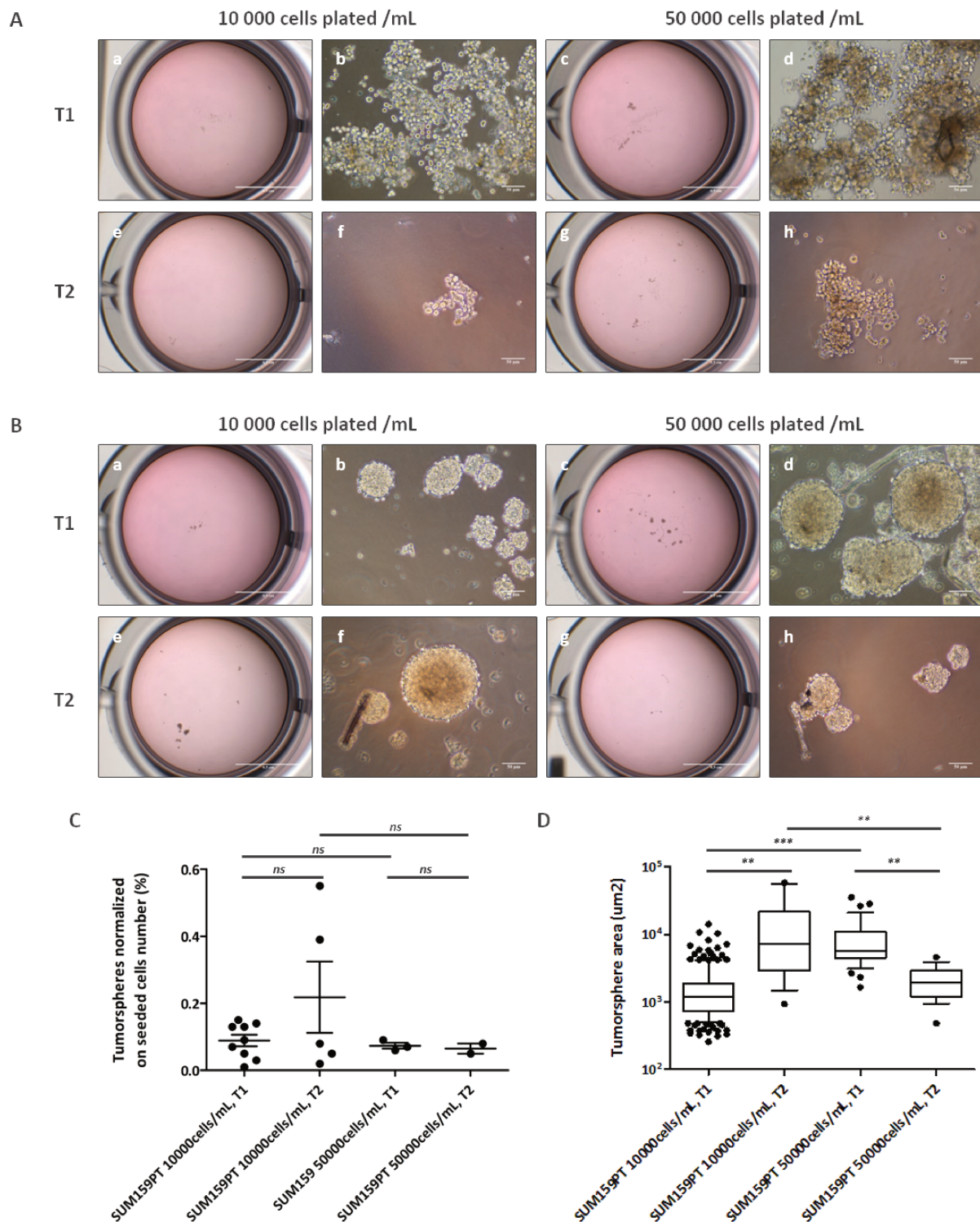


Figure 8.4 : Comparison of clonogenic abilities between first and secondary re-plated SUM149-PT and SUM159-PT cell lines after respectively 7 and 14 days in 3D culture conditions

A. Representative calibrated pictures of tumorigenic assays on SUM149-PT cell line. (a, c, e, g) Representative calibrated pictures of SUM149-PT derived tumorspheres photographed through a binocular loop (x6.3 magnification) in 3D culture conditions (picture pixel size= 1pixel/6.23 μm and white scale equal 0.5 cm) (b, d, f, h) and through a transmission microscope (x10 magnification) (picture pixel size=1pixel/0.35 μm and white scale equal 50 μm) (a, b, e, f) Tumorspheres issued from SUM149-PT cells plated at a density of 10 000 cells/mL (c, d, g, h) and plated at a density of 50 000 cells/mL (a, b, c, d) Tumorspheres developed from SUM149-PT cells first plating in 3D conditions. (e, f, g, h) and secondary plating in 3D conditions. B. Representative calibrated pictures obtained through tumorigenic assay performed on SUM159-PT cell line. c.f. subfig. A. for detailed informations. C.

Grouped data of the normalization of tumorspheres number on seeded SUM159-PT cells number. Calculation is based on tumorspheres up to 40 μm in diameter and each manip replica on the graph is represented thanks to a bullet black point. D. Grouped data of tumorspheres area derived from SUM159-PT 3D cultured cells. Each tumorsphere area is plotted on the graph as a bullet black point and whiskers regroup 10 to 90 percentile.

Statistical analyses were performed using paired 2-tailed Student-t test, 99% of confidence interval and error bars represent SEM.

8.3 The CD24-/CD44- subphenotype survives but does not form tumorsphere in 3D

As in the previous section (*cf.* part II on page 47), we moved from the study of the population to the study of each phenotypic subpopulation (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-) and were careful to ensure their clonogenic capacity. The point that mattered most to us was to point out which subphenotype(s) triggered the formation of cell clusters and thus could be described as “tumor initiating cell” [257, 343].

To do this, we isolated each subpopulation as previously described in fig. 5.2 on page 52 subfig. A. We subsequently cultured 10 000 cells/ml (for each subphenotype) in NSA medium for 7 days (T1). We then quantitatively characterized the emerging cell clusters according to the procedure of analysis already described in section 8.1 on page 121. The cell clusters of each replicate were then pooled and dissociated with trypsin. The single cell suspensions were replated at a concentration of 10 000 cells/ml and cultured for 14 days in the same 3D conditions (T2). This experiment was realized twice (T1 and T2) with 2 to 3 replicates per experiment.

8.3.1 The CD24-/CD44+ and CD24+/CD44+ subphenotypes

Following 7 days (T1) and 14 days (T2) of culture, CD24-/CD44+ and CD24+/CD44+ subphenotypes of SUM149-PT cell line produced clusters of cells but no tumorspheres (*cf.* description in section 8.2.1 on page 124). These clusters are made of small round cells which sometimes aggregate more densely.

Regarding the future of the subphenotypes of SUM159-PT cell line, they form at T1 and T2 tumorspheres in which cells are densely grouped. Morphological differences were noted between T1 tumorspheres derived from the CD24-/CD44+ subphenotype and those derived from the CD24+/CD44+ subphenotype. Thus, tumorspheres issued from CD24-/CD44+ have a very defined area and tumorspheres from CD24+/CD44+ exhibit a significantly less defined area outline and are surrounded by refractive cells.

At T2, the tumorspheres derived from CD24-/CD44+ and CD24+/CD44+ no longer exhibit significant morphological differences. They are still densely aggregated but their boundaries are blurrier and covered with small, round and highly refractive cells (*cf.* fig. 8.5 on the facing page).

Subphenotypic characteristics of the original plated SUM159-PT subpopulations cells have no influence on the clonogenic property across tumorspheres generations (T1 and T2) (*cf.* fig. 8.5 on the next page subfig. C). These observations are quite inconsistent with previous studies which published that CD24-/CD44+ subphenotype derived from breast origin retains higher tumor-initiating capability [114, 246]. Nevertheless, subphenotype plays an important role in the cells ability to proliferate, and thus, derived tumorspheres to grow. Indeed, tumorspheres derived from CD24-/CD44+ subpopulation are less prone to grow compared to the one derived from CD24+/CD44+ subpopulation which respectively measure in average $6419.0 \pm 862.9 \mu\text{m}$ and $9642 \pm 1052 \mu\text{m}$ at T1, $4143.0 \pm 714.8 \mu\text{m}$ and $11\,400 \pm 1414 \mu\text{m}$ at T2 (*cf.* fig. 8.5 on the facing page subfig. D).

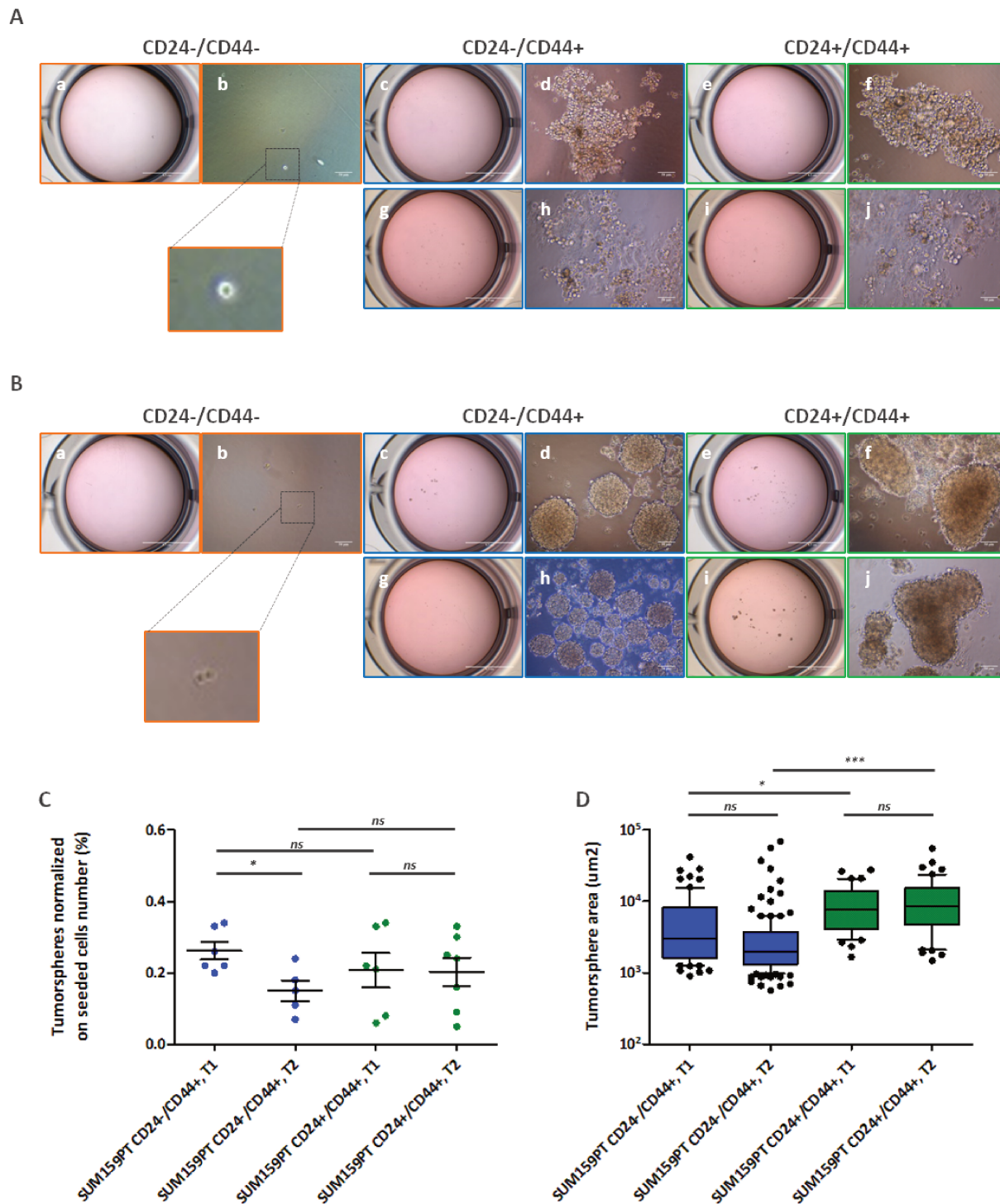


Figure 8.5 : Clonogenic abilities of first and secondary re-plating of sorted SUM149-PT and SUM159-PT cells subphenotypes after respectively 7 and 14 days in 3D culture conditions

A. Representative calibrated pictures of first and secondary tumorspheres assay performed on re-plated sorted SUM149-PT CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ cells subphenotypes. **B.** and on re-plated sorted SUM159-PT CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ cells subphenotypes. (a, c, e, g, i) Representative calibrated pictures of growing tumorspheres through a binocular loop (x6.3 magnification) (picture pixel size= 1pixel/6.23 µm and white scale equal 0.5 cm) (b, d, f, h, j) and through a transmission microscope (x10 magnification) (picture pixel size=1pixel/0.35 µm and white scale equal 50 µm). (a, b, c, d, e, f) Tumorspheres developed from cells subphenotypes first plating in 3D conditions. (g, h, i, j) and from secondary plating. **C.** Grouped data of the normalization of tumorspheres number on seeded SUM159-PT cells number. Calculation is based on tumorspheres up to 40 µm in diameter and each manip replica on the graph is represented thanks to a bullet colored point. **D.** Grouped data of tumorspheres area derived from SUM159-PT 3D cultured cells. Each tumorsphere area is plotted on the graph as a bullet black point and colored whiskers regroup 10 to 90 percentile.

Statistical analyses were performed using paired 2-tailed Student-t test, 99% of confidence interval and error bars represent SEM. In blue : CD24-/CD44+ cells, in orange : CD24-/CD44- cells, in green : CD24+/CD44+ cells sorted phenotype

8.3.2 The subphenotype CD24-/CD44-

For the CD24-/CD44- subphenotype of SUM149-PT and SUM159-PT lines, we observe at T1 few and isolated small round, highly refractive cells as presented in fig. 8.5 on page 129. Their low number suggests that they did not multiply. Perhaps due to the NSA culture medium lacking insulin and hydrocortisone, the proliferative capacity of the CD24-/CD44- cells and their metabolism is hindered [9]. Indeed, the same cultured cell subpopulation in 2D but with a medium containing both growth factors had proliferated (*cf.* fig. 5.2 on page 52 subfig. B and C). These isolated cells were still present in the medium after 7 days and were capable of withstanding anoikis signals throughout that duration [111] and this without proliferating. Maybe, they went in quiescent or “poised” state [54, 156]. This study could not be completed because of the small number of cells. Nevertheless, we plan both to confirm the quiescent state (G0 phase of the cell cycle) by a co-staining of CD24-/CD44- cells grown in 2D with BrdU agent and to characterize their behavior over cell cycling with Fucci technology [46, 89, 201, 268].

This work highlighted that CD24-/CD44+ and CD24+/CD44+ populations have both strong clonogenic and proliferative capabilities in 3D culture conditions. Thus, these two subphenotypes have the potential to initiate tumors *in vitro*. Moreover, overtime subculturing in non-adherent conditions (T2) provides evidence for the presence of self-renewing initiating tumorsphere cells in both CD24-/CD44+ and CD24+/CD44+ subpopulations [49, 71, 117, 325]. Nevertheless, the CD24-/CD44- subphenotype survives in the single cell status and this for 7 days, but does not generate tumorsphere.

Influence of 3D spatial relations on heterogeneity and plastic cell capabilities

«Les petites herbes ne s'avouent pas vaincues par les monts enneigés. Chaque année, elles renaissent au vent printanier.»

- He Yifu -

In previous studies, we found out that the organization of cells in tumorspheres was influenced by the phenotypic state of the original cells. Furthermore, we noticed that cells from SUM149-PT and SUM159-PT cell lines did not behave the same way in 3D culture conditions. The SUM149-PT cells formed shapeless heap while cells of SUM159-PT cells aggregated into polarized clusters. So, we wondered if the clonogenic potential was dependent on the level of expression of certain markers (CD24 and CD44) within cell aggregates.

9.1 Differential expression of phenotypes in 2D and 3D

Our first task was to compare the average expression level of CD24 and CD44 markers in 2D and 3D culture conditions for both SUM149-PT and SUM159-PT cell lines. Thus, we monitored the expression of these markers in 2D and 3D culture as described previously in section 5.2 on page 53.

2D cells were grown in standard culture medium (*cf.* appendix D.2 on page 249) following the recommendations described in section 5.3.1 on page 57. While 3D cells have been cultured in the NSA medium at concentration of 10 000 or 50 000 single cells/ml following the protocol described in appendix D.20 on page 283 (*cf.* fig. 9.1 on the next page).

The CD24 and CD44 markers were measured by flow cytometry after 7 days of culture (*cf.* fig. 9.2 on page 133). Their expression level was analyzed qualitatively and quantitatively using the same analytical procedures previously used to analyze the expression of 2D cells and detailed in chapter 5 on page 49 and more particularly described in section 5.2.3 on page 53.

The profiles of expression of CD24 and CD44 markers vary drastically depending on culture conditions (plane or 3D and with different culture media) and these changes concern both lines of interest (*cf.* fig. 9.3 on page 134). Furthermore, the proportions of the different subpopulations (CD24 and CD44-low-high) have drastic variations, such as :

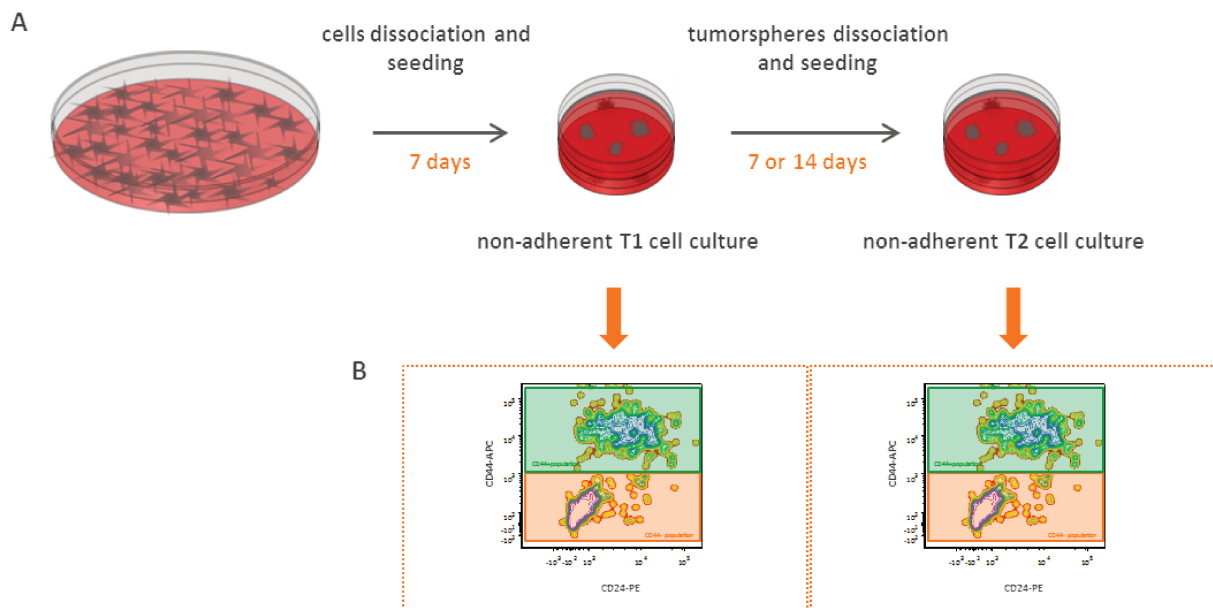


Figure 9.1 : Workflow of the markers characterization of tumorspheres assay.

A. Experimental workflow to assess CD24 and CD44 markers characterization of tumorspheres derived from SUM149-PT and SUM159-PT cell lines and their respective subphenotypes. B. Flow cytometry gating strategy of tumorsphere cells within CD24 and CD44 space of expression. (T1 : first plating and T2 : secondary plating)

- In 2D, CD44-high cell subpopulation is prevalent and accounts for 99% of the population.
- In 3D, the proportion of the CD44-low subpopulation increases (between 50 and 90% for SUM149-PT line and between 10 and 45% for SUM159-PT).

Besides, the average expression level of these markers is influenced by the culture conditions. Thus, there is for SUM149-PT cell line :

- An increase in the average level of expression of CD24 to CD44-high subpopulation
- A decrease in the expression level of the CD44 marker for the subpopulation CD44-low
- An increase in the average level of expression of CD24 expression marker for CD44-low subpopulation

As for the SUM159-PT cell line, we noticed :

- An increased level of expression of the CD44 marker to the CD44-high subpopulation
- A decrease in the expression level of the CD44 marker for CD44-low subpopulation

The dispersion for the CD44-high population :

- Increases for the CD44 marker for both cell lines.
- Decreases for the CD24 marker for SUM159-PT line.

Also, we noticed a significant difference in the expression level of CD24 marker in the cell aggregates of SUM149-PT and SUM159-PT cell lines. Shapeless cell aggregates from SUM149-PT cell line are composed of a CD44-high pool expressing CD24 marker while tumorspheres

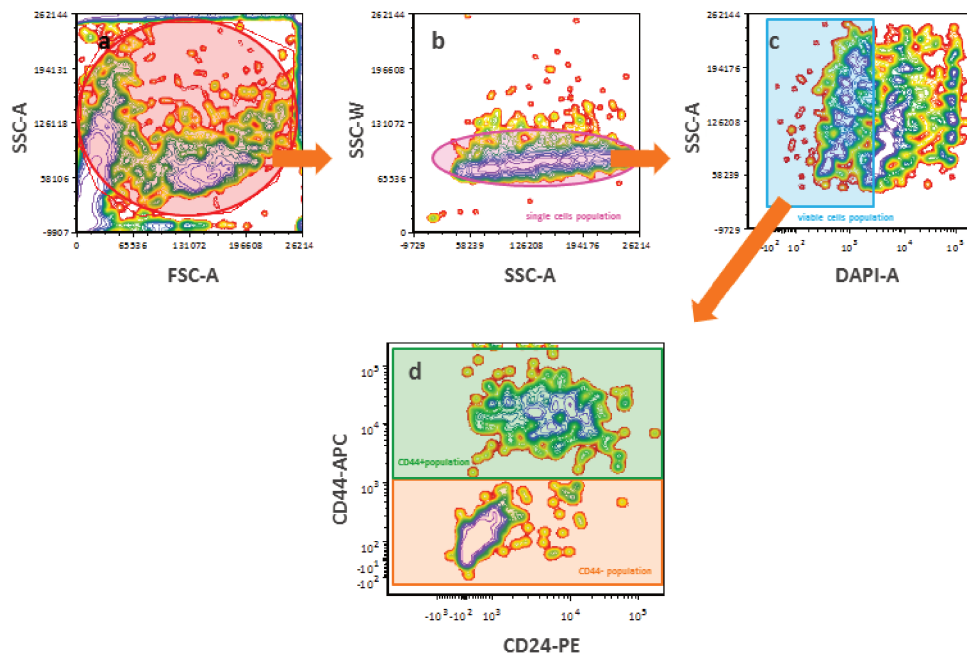


Figure 9.2 : Flow cytometry data analysis gating strategy workflow to characterize tumorspheres expression for CD24 and CD44 markers.

(a) Cytochrome contour plot of light scattered cells characteristics. (b) Cytochrome contour plot of Side Scatter characteristics. Pink selection stands for single cells population. (c) Cytochrome contour plot of cells viability parameter. Light blue selection stands for viable cells population. (d) Cytochrome contour plot for CD24 and CD44 expression markers.

from the SUM159-PT cell line exhibit a CD44-high population that does not express or only slightly the CD24 marker. These observations must be compared with the morphology of the cell clusters. Indeed, cell clusters from SUM149-PT cell line, shapeless, are unable to form typical tumorspheres. While the ones from SUM159-PT cell line densely aggregate and form well outlined tumorspheres. Thus, it appears that the ability to aggregate seems to be anti-correlated with the expression level of the CD24 marker.

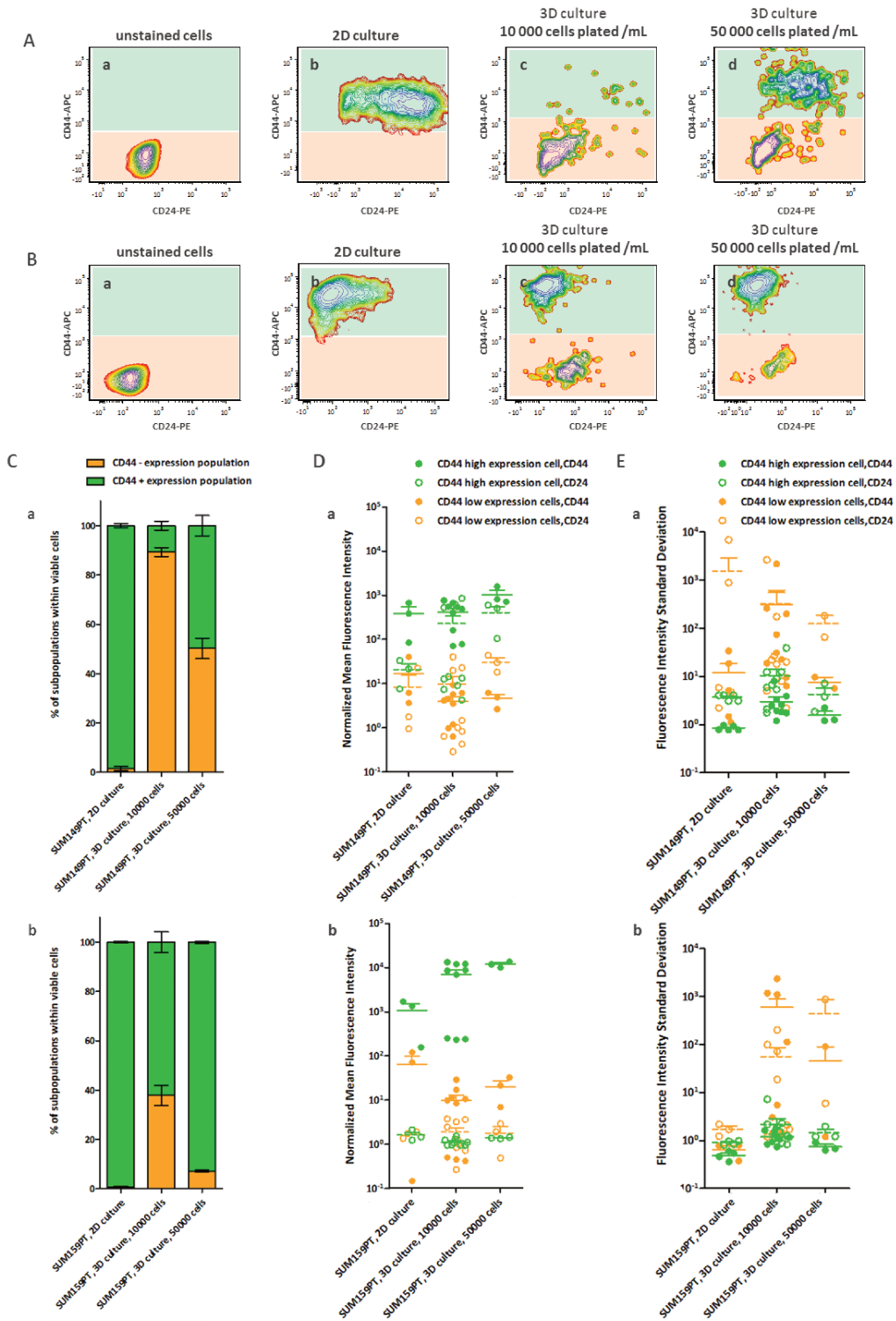


Figure 9.3 : Influence of culture conditions (2D and 3D culture conditions) on CD24 and CD44 expression profile of SUM149-PT and SUM159-PT cells

A. Flow cytometry contour plots of growing tumorspheres viable cells derived from SUM149-PT cells cultured either in 2D or 3D culture conditions. B. and derived from SUM159-PT cells. C. Percentage of CD44 negative and CD44 positive expressing tumorsphere cells derived from (a) SUM149-PT and (b) SUM159-PT cells cultured in 2D and in 3D conditions. D. CD24 and CD44 normalized mean fluorescence intensity of (a) SUM149-PT and (b) SUM159-PT tumorspheres viable cells. E. CD24 and CD44 fluorescence intensity standard deviation of (a) SUM149-PT and (b) SUM159-PT tumorspheres viable cells. Each replicative data are plotted on the graph as a bullet point (Error bar represents SEM and horizontal bar the arithmetic mean of CD24 and CD44 fluorescence intensity data)

9.2 Influence of the quantity of plated cells on phenotypes

As we saw in the previous chapter, changing the number of plated cells in the initial state is correlated with changing the proliferative capacity of the cells. So we wanted to know if a particular phenotype held proliferative functions.

Therefore, we phenotypically characterized in 3D the clusters formed by plated cells (10 000 cells/ml or 50 000 cells/ml plated in the initial state). After 7 days of culture, the clusters were analyzed by flow cytometry for CD24 and CD44 markers. Then to confirm that the generation of tumorspheres kept these properties, cell aggregates were dissociated and the single cell suspension placed in culture at 10 000 cells/ml or 50 000 cells/ml.

In both cell lines, the expression pattern of CD24 and CD44 markers varies depending on the proliferative potential developed by the cell aggregates.

The proportions of CD24 and CD44 phenotypes vary according to concentrations of cells plated at T1 :

- For SUM149-PT cell line, the proportion of CD44-low is 90% of the population for the condition 10.000 cells/ml while it is 50% of the population for the condition 50.000 cells/ml.
- For SUM159-PT cell line, the proportion of CD44-low is 40% for the condition of 10.000 cells/ml and 10% for the condition 50.000 cells/ml.

Thus, the proportion of CD24-/CD44- cells decreases when the concentration of cells plated in the initial state increases and the proliferative capacity increases. This property is true again for the T2 generation.

- For SUM149-PT cells, the proportion of CD44-low cells is 55% for the condition 10.000 cells/ml and it drops to 20% for condition 50.000 cells/ml.
- For SUM159-PT cells, the proportion of CD44-low cells is 30% for the condition 10.000 cells/ml and is 20% for condition 50.000 cells/ml.

We remark that for SUM159-PT cell line with the condition 50.000 cells, the pool of CD24-/CD44- cell subpopulation increases between T1 and T2.

This finding correlates with the decrease in the area of tumorspheres (*cf.* fig. 9.5 on page 138) shrinking from $9047 \pm 1284 \mu\text{m}$ to $2134 \pm 273 \mu\text{m}$. Thus, when the pool of CD24-/CD44- decreases there is an increase of proliferative potential and vice versa, when the pool of CD24-/CD44- increases the proliferative power decreases. The proliferative capacity is indeed maintained by the CD44-high pool and this, whatever the level of expression of the CD24 marker. This observation is consistent with previous studies which highlighted that CD44 positive compartment should retain most proliferative abilities [102]

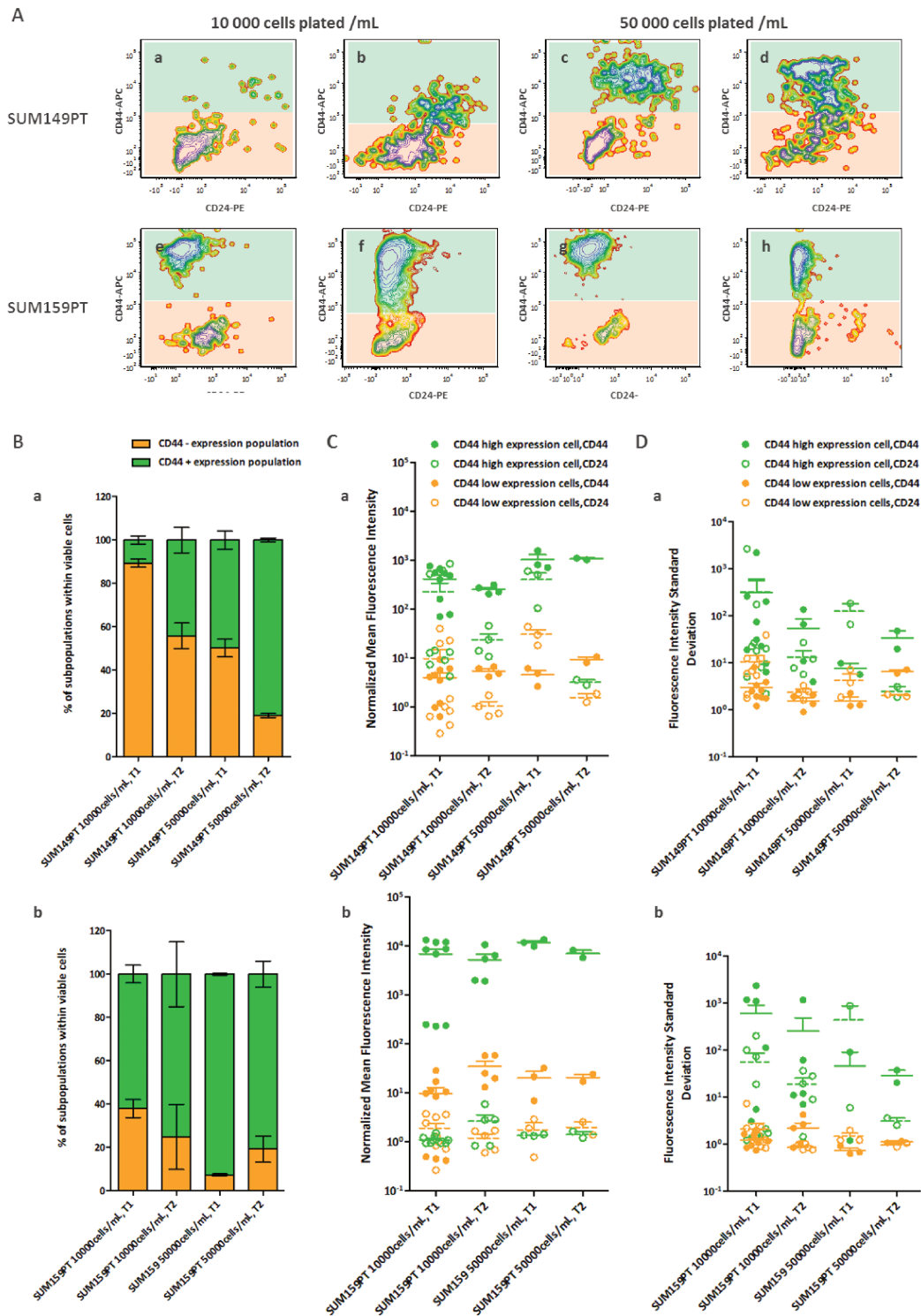


Figure 9.4 : Comparison of the CD24/CD44 expression profile of cells issued from either first or secondary re-plating of SUM149-PT and SUM159-PT cells after respectively 7 and 14 days in 3D culture conditions

A. Flow cytometry contour plots of growing tumorspheres viable cells derived from 3D culture of first and secondary plating SUM149-PT and SUM159-PT. (a, b, c, d) Flow cytometry contour plots of tumorspheres viable cells derived from SUM149-PT cells (e, f, g, h) and from SUM159-PT cells (a, c, e, g) Flow cytometry contour plots of first plating tumorspheres viable cells (b, d, f, h) and of secondary plating tumorspheres viable cells. B. Percentage of CD44 negative and CD44 positive expressing tumorsphere cells derived from (a) SUM149-PT and (b) SUM159-PT cells 3D culture. C. CD24 and CD44 normalized mean fluorescence intensity of (a) SUM149-PT and (b) SUM159-PT tumorspheres viable cells. D. CD24 and CD44 fluorescence intensity standard deviation of (a) SUM149-PT and (b) SUM159-PT tumorspheres viable cells.

Each replicative data are plotted on the graph as a bullet point (Error bar represents SEM and horizontal bar represents the arithmetic mean of CD24 and CD44 fluorescence intensity data).

9.3 Phenotypic plasticity in 3D culture conditions

In our previous 2D experiments, we have demonstrated that CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- subphenotypes of SUM149-PT and SUM159-PT cell lines were of great plastic capacity and produced back diversified populations consisting of the three original phenotypes. Our first intention was to check if this characteristic was also valid in 3D.

9.3.1 The methodology that has been followed

This study was conducted to every sub-phenotypic subpopulations : CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+. Isolated cell phenotypes (*cf.* section 6.1.1 on page 72) were put back into NSA medium culture at a concentration of 10 000 cells/ml. Cell aggregates of these CD24-/CD44+ and CD24+/CD44+ subpopulation were analyzed by flow cytometry at T1 and T2 for the CD24 and CD44 markers. This experiment was performed two times each with 2 to 3 replica.

9.3.2 CD24-/CD44+ and CD24+/CD44+ and re-diversification in 3D

At first glance, it is obvious that both CD24-/CD44+ and CD24+/CD44+ subpopulations exhibit plastic capacities whichever the SUM149-PT or SUM159-PT cell lines they derived from.

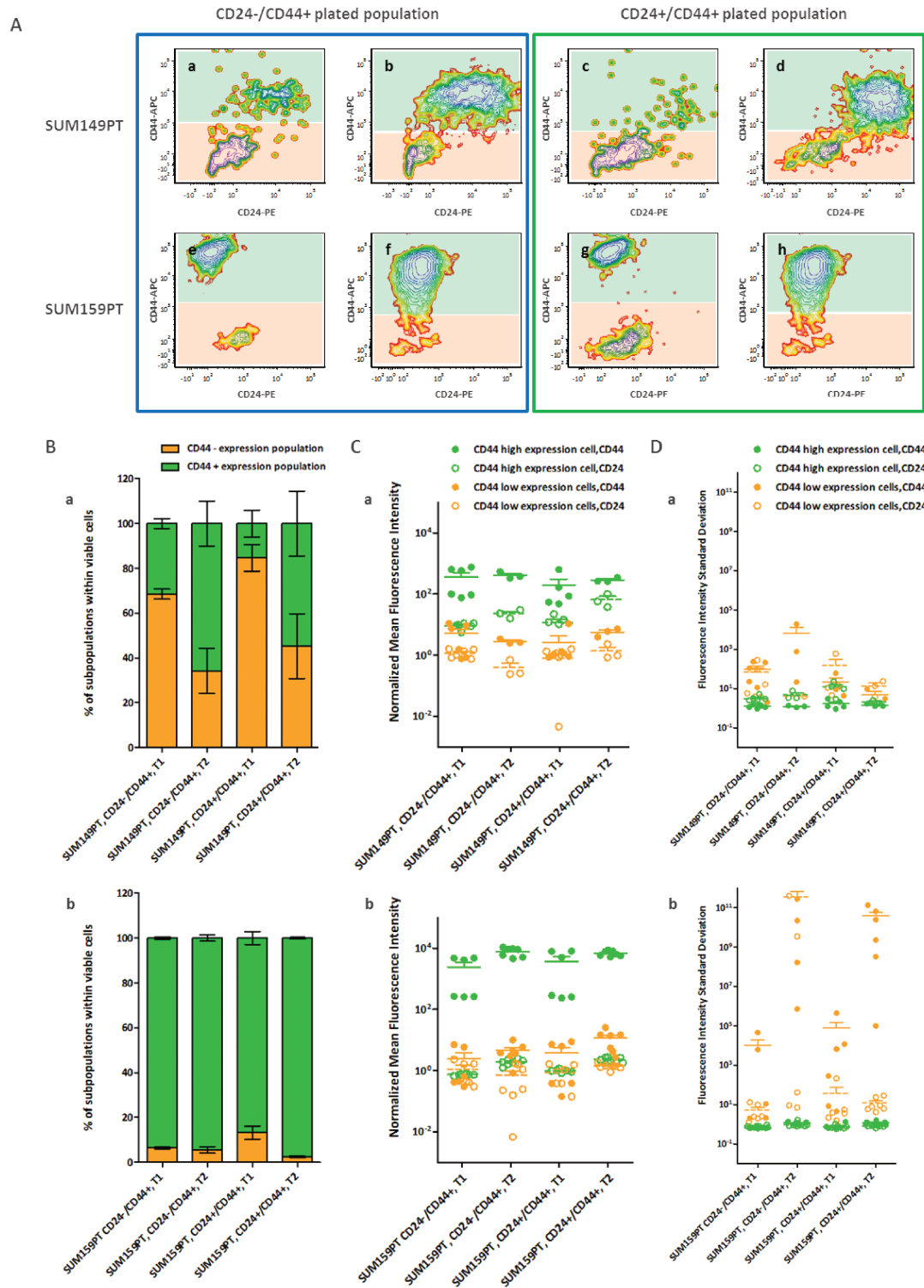


Figure 9.5 : Comparison of the CD24 and CD44 expression profile of cells issued from either first or secondary re-plated sorted SUM149-PT and SUM159-PT cells subphenotypes after respectively 7 and 14 days in 3D culture conditions

A. Flow cytometry contour plots of growing tumorspheres viable cells derived from 3D culture of first and secondary plating sorted SUM149-PT and SUM159-PT cells subphenotypes. (a, b, c, d) derived from SUM149-PT cells subphenotypes (e, f, g, h) and from SUM159-PT cells subphenotypes (a, c, e, g) Flow cytometry contour plots of first plating tumorspheres viable cells (b, d, f, h) and of secondary plating tumorspheres viable cells. B. Percentage of CD44 negative and CD44 positive expressing tumorsphere cells derived from (a) SUM149-PT and (b) SUM159-PT subphenotypes cells 3D culture. C. CD24 and CD44 normalized mean fluorescence intensity of tumorspheres derived from (a) SUM149-PT and (b) SUM159-PT viable cells. D. presents a grouped column scatter graphing CD24 and CD44 fluorescence intensity standard deviation of tumorspheres viable cells. Each replicative data are plotted on the graph as a bullet point and statistical analyses were performed using 2-tailed Student-t test (Error bar represents SEM). Blue color framing : CD24-/CD44+ cells sorted phenotype, green color framing : CD24+/CD44+ cells sorted phenotype

The cell phenotype CD24-/CD44+ transit to other phenotypic states

Regarding SUM149-PT cell line, subpopulation CD24-/CD44+ (derived from 2D culture) produces back two distinct populations, CD24+/CD44+ and CD24-/CD44-, while the same CD24-/CD44+ pool derived from SUM159-PT generates the following populations : CD24-/CD44+ and CD24-/CD44- (*cf.* fig. 9.5 on page 138).

The cell phenotype CD24+/CD44+ transit to other phenotypic states

This isolated subpopulation of SUM149-PT gives two distinct populations : a CD24+/CD44+ population and a CD24-/CD44- population of high importance. The phenotype resulting from the SUM159-PT cell line transitions to two distinct phenotypic states CD24-/CD44+ and CD24-/CD44-.

These populations, as described in T1, change somewhat in the second generation in their dynamics of expression of the CD24 and CD44 markers. Subpopulations T1 (CD24-/CD44+ and CD24+/CD44+) derived from the SUM149-PT cell line increase their dispersion of expression on the CD24 marker for CD44-high pool. However, the subpopulation T1 derived from the SUM159-PT cell line increases their dispersion on the CD44 marker and expresses a new intermediate subpopulation CD24-/CD44med.

These observations validate that the change of culture environment (moving from 2D to 3D) does not affect the plastic qualities of phenotype CD24-/CD44+. However, environmental changes influence the level of expression of CD24 and CD44 markers and thereby alter the expressed phenotypes and their proportions. In this context we note both the almost disappearance of the CD24-/CD44+ phenotype described in SUM149-PT cell line grown in 2D and the almost disappearance of CD24+/CD44+ phenotype characterized in SUM159-PT line grown in 2D.

9.3.3 CD24-/CD44+ et CD24+/CD44+ subpopulations have different proliferative capacity

The expression profiles of CD24 and CD44 markers exhibit the great particularity of being identical to the T1 generation for both CD24-/CD44+ and CD24+/CD44+ subphenotypes. It is the same for T2. However, changes happen on the proportions CD44-low and CD44-high cells pool. Thus, we note that the aggregates derived from the CD24-/CD44+ subpopulation exhibit a smaller CD44-low pool than those derived from the CD24+/CD44+ subpopulation. This observation holds true for the two cell lines. At T1 :

- For SUM149-PT, the proportion of the CD44-low pool deriving from the subpopulation CD24-/CD44+ is 70% and is equal to 90% for the one deriving from CD24+/CD44+ subpopulation.
- For SUM159-PT, the proportion of the CD44-low pool deriving from the subpopulation CD24-/CD44+ is 5% and is equal to 15% for the one deriving from CD24+/CD44+ subpopulation.

This difference in proportion of CD44-low pool is also true at the T2 generation.

If we follow the logic previously developed on the relationship between the proportion of CD44-high pool and proliferative capabilities present within the cell clumps, then it follows that the subpopulation of CD24-/CD44+ should exhibit proliferative capabilities superior to those of phenotypes CD24+/CD44+ [239, 282]. However, previous studies regarding tumorspheres areas measurements have shown that tumorspheres derived from CD24-/CD44+ subpopulation are smaller than the ones derived from CD24+/CD44+ (*cf.* section 8.3.1 on page 128). In order to

clearly determine what is or what are the cellular phenotypes that exhibit proliferative capacities within tumorspheres, it would be necessary to continue the characterization of cellular functions by the bias of lineage tracing experiments coupled with cell cycle and apoptosis analysis [241].

In the light of this work, we can draw the following conclusions :

- All cell mass has an equilibrium of CD24-/CD44- cells and CD24-/CD44+ or CD24+/CD44+ cells.
- CD24-/CD44+ and CD24+/CD44+ cells characterized in 2D are capable of plasticity in a 3D environment and return to a dynamic phenotypic steady state. This equilibrium is characteristic of the tumor origin and thus has dynamics of expression of CD24 and CD44 markers similar to those described in 2D conditions.
- The subphenotype CD24-/CD44-, while not proliferating by himself in this 3D culture medium (NSA medium), is always re-expressed by subphenotypes in transition in this environment. So this raises a question about the role played by this subphenotype. In 2D culture, we observed that only repopulations triggered by the rediversification of this phenotype reached an equilibrium in the expression of CD24 and CD44 markers equivalent to that described in the non-sorted population. Thus we assume that the CD24-/CD44- phenotype could play a key role in phenotypes homeostasis.
- When the level of expression of CD24 marker is important, then the cell clusters are shapeless. However, when the level of expression of CD24 is low, the cells aggregates very strongly and form well outlined tumorspheres.
- At the T2 generation, a CD24-/CD44med population appears for the SUM159-PT cell line. Its appearance is correlated with a change in morphology of tumorspheres which becomes less outlined. Thus, one can hypothesize that this particular subpopulation presents invasive characteristics.

Part IV

Strategies proposal to search social rules that govern cell phenotypic changes

*«L'effacement volontaire de soi, et le respect profond d'autrui, constituent les deux vertus grâce auxquelles le Sage espère profondément donner la paix au monde»
- Liou K'ia-Hway -*

Through previous parts, we highlighted an essential feature that is shared by all functional units of my system. Indeed, each cell, regardless of its initial phenotypic state, is characterized by its plasticity that is to say its inherent ability to switch from a phenotype to another one. The perspective I have taken in previous experiments focused on the overall behavior followed by both homogeneous and heterogeneous cell populations. More precisely, we monitored the behavioral path followed by two unsorted cancer cell lines and three subphenotypes sorted from these same cell lines. And such analyses showed that :

- A heterogeneous population for specific phenotypes is not stable over time in the proportion occupied by the phenotypes. Instead, I have observed nearly cyclical fluctuations that it would be very interesting to further characterize.
- A homogeneous population (obtained by selection within a heterogeneous population) invariably returns towards a heterogeneous population which re-expresses the unsorted cell phenotypes.
- Some phenotypes are faster than others in transiting from one phenotype to another.
- Phenotype appears to transit from a phenotype to another according to a defined and in a way which is characteristic for each cell phenotype.
- Only certain populations defined for a specific phenotype reform after a while a population whose heterogeneity characteristics are approaching the dynamic balance achieved by the original heterogeneous population.
- The environment (2D or 3D) in which the cells evolve, gives structure to cell diversity and influences the expression of their phenotypes.

So far the perspectives we have adopted enabled to carry out very long dynamic scale studies of cell populations, yet these studies have been performed as static analysis at the population level. From these observations we have the desire to establish a dynamic study at the single cell level of cells operating within a cell population. From this innovative viewpoint, many questions arose due to the observations followed on cell populations studies. Do individual cells fluctuate in the same manner as the general population? What are the characteristics of single-cells transitions of cells cultivated within a large population?

It seems obvious that there are perpetual single cell transitions between different phenotypic

states. Under the conditions of study, it seems like each cell is equally likely to transit to another phenotype. However, are these transitions synchronized and according to what parameters, what phenotypes...? What is the causality of these cells state transitions? Are there any triggers or restrictors of those state changes?

Through these questions the key issue of my thesis emerges and may be summarized as follows. Are cell phenotype transitions due to chance, to the autonomous decision-taking of each cell or determined by neighboring cells? Therefore, for the purpose of answering these questions of cell autonomy and following single-cell behavior changes over the long term in real-time that I just started developing a tool capable of tracking these phenotypic transitions. I named it the PhenotypeTracker.

This next part of my thesis will present the development of a molecular tool box able to refer in real time the phenotypic state of each cultured cells. I will begin by elaborating on the various units of the PhenotypeTracker molecular platform and by characterizing the main strategic decisions that have been taken over its development. I will continue with approving of the various units constituting this molecular tool and the improvement we wish to bring in the foreseeable future to ensure that this tool becomes fully operational to meet the expected needs.

Chapter 10

Long term follow-up at the single cells level of cell state transitions

“C’est souvent qu’il faut accepter de prendre le risque de faire quelques vagues pour trouver sa voie. Un bateau qui reste au port n’est jamais en danger, mais il ne va nulle part.”

- *Sagesse Populaire* -

In the two previous parts, we highlighted phenotypes transitions affecting all CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- subpopulations. Isolated, each of these subpopulations produces back a heterogeneous population including the three CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- phenotypes. These phenotypes are different from each other by the expression level of CD24 and CD44 markers. So the phenotypic space is characterized by a gradient of expression on the CD24 marker and another gradient for CD44 marker (*cf.* fig. 10.1 on the following page subfig. A).

10.1 Some criteria to model cell state transitions

10.1.1 On the way of self-renewing phenotypes

Each of the three identified phenotypes exhibits unique repopulation characteristics as previously described in part II on page 47. Each one displays specific functional capabilities. Especially some phenotypes (CD24-/CD44+ and CD24+/CD44+) trigger tumorspheres whose morphological characteristics differ depending on the original phenotype. The major issue is to identify different phenotypes that make up a tumor and understand their functions. Some studies suggest that some solid cancer (breast, prostate, liver, lung) are born from cancer stem cells [117, 156]. The highly sought out phenotype is therefore one from which emerges all phenotypes present within a tumor and which is capable of self-renewal.

Because of plastic abilities of all CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- phenotypes, we are still undecided on a key aspect : the self-renewal capacity of the cells. The key historical experiment to demonstrate self-renewal is the “tumor assay” [44, 241, 331] over generations. The idea is that if there is a T1 stem cell then there will be a tumorsphere. Stem cell will self-renew, will be present in the T1 tumorsphere and after re-plating restore a T2 tumorsphere, will self-renew

indefinitely. This line of thought does not take into account the phenotypic transitions since phenotype can appear at T3, following a transition from one state to another but not from any self-renewal abilities.

The CD24-/CD44- phenotype, representing less than 1% of the population of SUM149-PT and SUM159-PT cell lines, holds our attention because it has more plastic capacity :

- A differentiation potential to return to the initial equilibrium
- Anaoikis resistance

Moreover, Dey et al. identified in normal mammary glands the presence CD24-/CD44- phenotype as containing a distinct side population (SP) able to effluete Hoechst 33342 [71, 255].

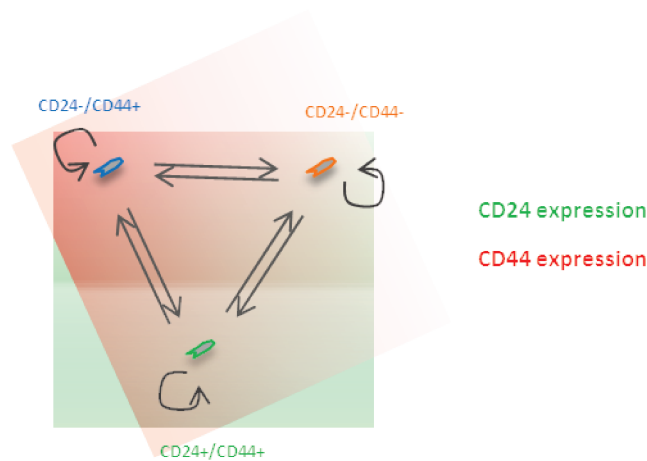


Figure 10.1 : *CD24 and CD44 phenotypes space within which cells evolve and transit from one phenotype to another one.*

10.1.2 Rhythm of cells division and phenotypic transitions

To model the future of a population, we must know the phenotype division rate constants. Measuring the division ratio is based on experiments in which we calculated the rate of doubling the number of cells in a population during a specific period. In a heterogeneous model, we isolated the different phenotypes, we re-cultured them, and measured the time it took for a homogeneous population to double in size (*cf.* fig. 10.2 on the facing page subfig. B).

However, in the population we study, the phenotypes are plastic and we found out that the cells interconverted themselves very quickly. Thus a homogeneous system at $t = 0$ was again very heterogeneous at $t + 3$. This extreme speed is particularly visible for CD24-/CD44- phenotype as described in section 6.3 on page 92. We have previously estimated that the average time of cell division in heterogeneous populations (CD24-/CD44+, CD24+/CD44+, CD24-/CD44-) of SUM149-PT and SUM159-PT lines ranged between 24 and 36 hours (i.e. at most three divisions in three days). So we asked ourselves when were scheduled phenotypic transitions with regard to the cell cycle :

- Are they coupled to cell divisions ?
- Or do they take place independently of cell division, as in the *C. elegans* model where the cell identity change allows switching from the Y rectal cell to PDA motoneuron [148, 260] ?

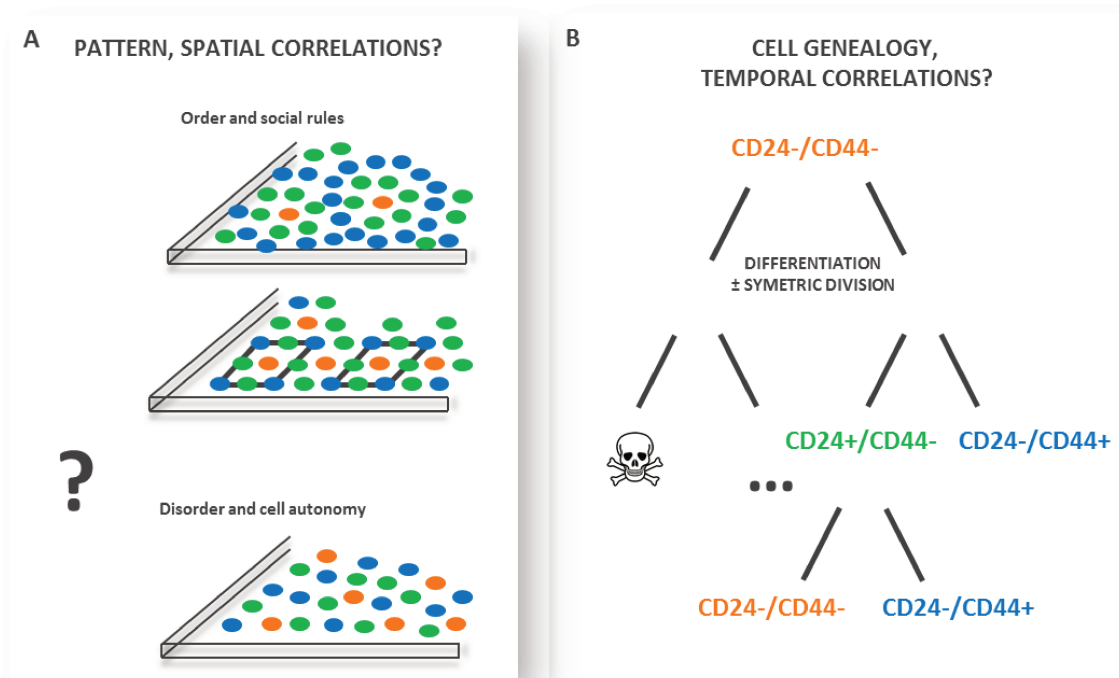


Figure 10.2 : Overview of working assumptions with respect to the search for correlations regarding the expression of cell phenotypes over time.

A. Thought to the search for spatial patterns in cells phenotypes expression. B. and for temporal patterns within cell phenotypes expression characteristics.

To answer this question, we must compute the division time not on the scale of the population or subpopulation, but at the single cell level (for each identified phenotype) and monitor the phenotypic transitions.

10.1.3 Influence of memory on the probabilities to transit from one phenotype to another

The phenotypic steady state (proportion and average expression level) of a population at $t = 0$ depends on the average phenotype of the population expressed in earlier generations as described in section 7.1 on page 103. Thus, at cell level the probability of transiting from a phenotypic state to another depends on the phenotypes previously expressed by its ancestors (*cf.* fig. 10.1 on page 146). How dependent are each cell to its past? And how many following generations can an ancestral phenotype influence? If the degree of dependency reaches a very high level, then there probably are phenotypic expression patterns (which impose themselves and are repeated over generations).

10.1.4 Expression of phenotypes and spatial organization

Each cell, following successive divisions, is the origin of a family of cells. Will its descendants mix with descendants of other families or will they remain clustered between members of the same family? In any case, each cell will interact with its neighbors and potentially communicate. At a time t , several cell generation mix and multiple families are present. Is a social order emerging from or within these cancer cells and will have an influence on some phenotypic changes? So when a cell, in the middle of others, must decide whether to transit to another phenotype, is it influenced by the phenotype of its neighbors? If it is, what is the critical distance of influence? (*cf.* fig. 10.2 on page 147 subfig. B)

The only way to answer all these questions is to work at the unicellular level to follow the phenotypic transitions and link them to asymmetric cell division (*cf.* fig. 10.1 on page 146 subfig. B). That is why we will discuss in the next paragraph the development of a molecular biology tool able to track in real-time the phenotypic transitions in the 2D space.

10.2 The characteristics of the molecular tool

The PhenotypeTracker should permit us to measure and monitor the level of endogenous expression of one or two markers (CD24 and CD44) at the single cell level. This monitoring must be done over several generations. For this, the endogenous expression of the marker will be revealed by a fluorescent marker and phenotypic transitions at " $t + 1$ " will therefore be determined by a decrease or an increase of the level of fluorescence measured at " $t = 0$ " (*cf.* orange and gray curve shown in fig. 10.3 on the facing page subfig. B).

To track several generations of phenotypic transitions (change in fluorescence level of the expression reporter of CD24 and CD44 markers) the PhenotypeTracker molecular platform is to be incorporated into the genome of host cells. The transduction of the molecular platform will be provided by second generation lentiviruses. This technology was chosen due to its high ability to transduce cells, whichever the cell type, even cells known as difficult as stem cells and progenitor, or cells in quiescent phase (neurons) [189, 220].

So we designed the PhenotypeTracker in the form of three expression blocks (*cf.* fig. 10.4 on page 150) :

- Two blocks of expression retrieve phenotypic transitions (green frame and red frame)

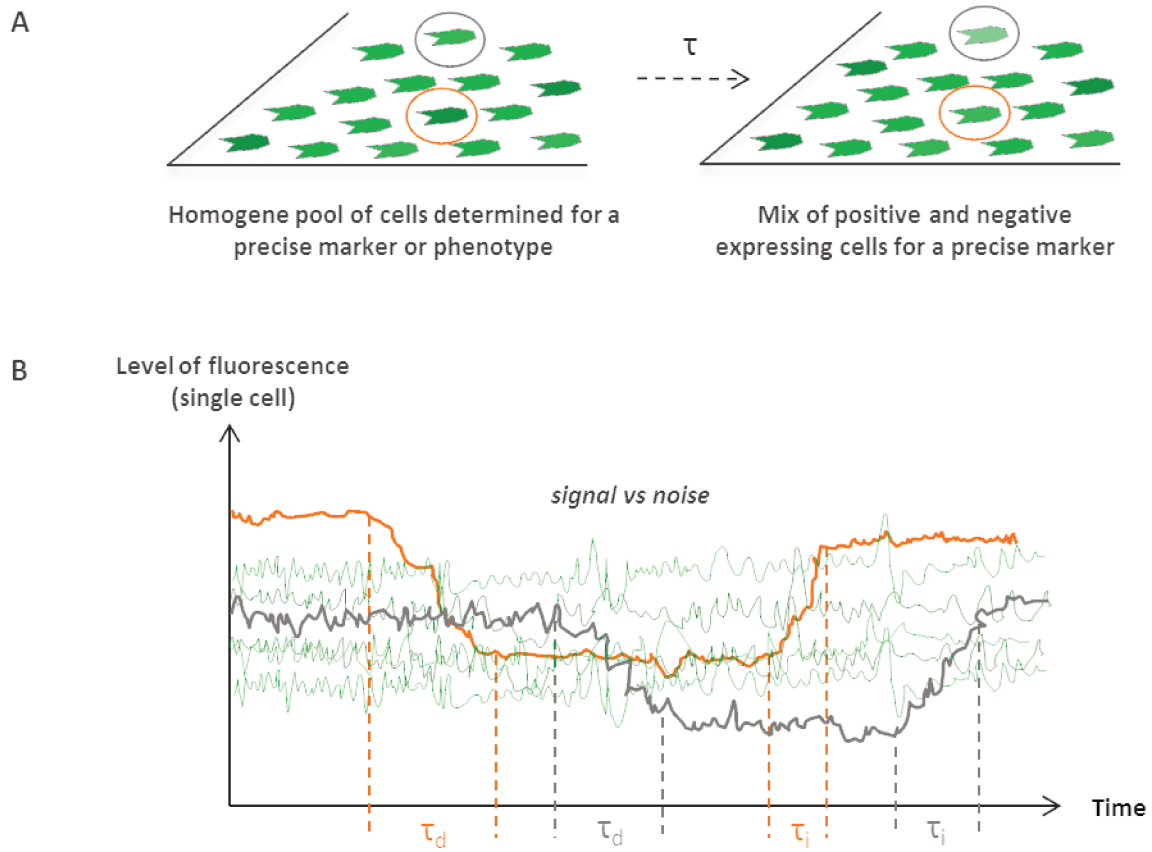


Figure 10.3 : Schematic representation of a real-time phenotype change tracking at the single cell level

A. Schematic picture of cells reporting a phenotype state through fluorescence signal and tracking at the single cell level of phenotypes changes. B. Schematic chronogram of single cell fluorescence expression level.

- A transduction reporter block and fluorescence normalizer (blue frame)

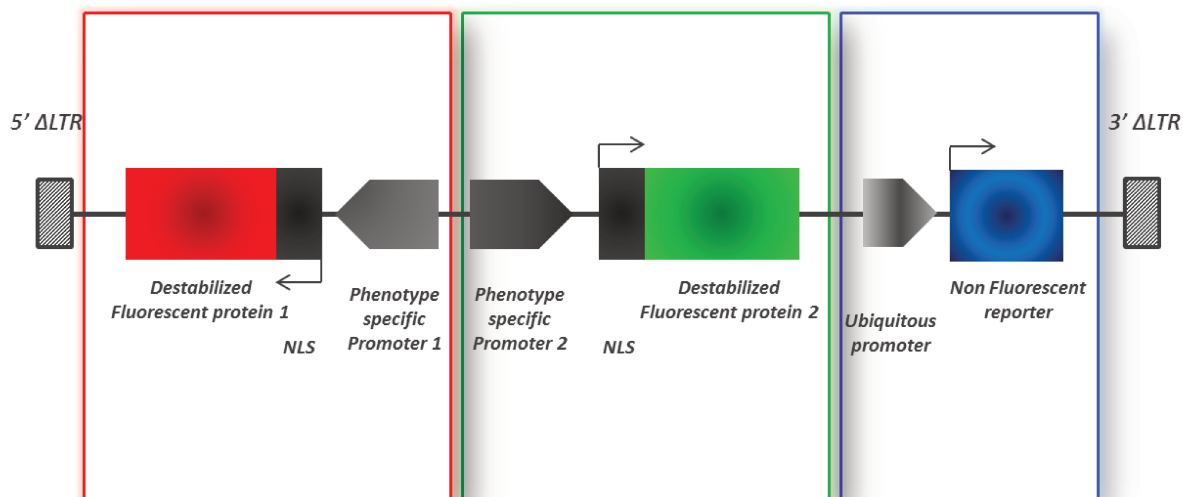


Figure 10.4 : Schematic illustration of the PhenotypeTracker lentiviral platform.

10.2.1 Phenotypes fluorescent reporter, a combination of two expressing endogeneous promoters

The expression level of each marker of interest is reported by a destabilized fluorescent protein whose short fluorescence half-life (between 2 and 3 hours) which allows following rapid phenotypic transitions [180]. Each expression reporter block of phenotypic state markers (in red and green on the figure (*cf.* fig. 10.4) expresses a fluorescent protein. The spectra (excitation and emission) of these two reporter proteins differ as much as possible to limit the cross-spectrum interference.

In the meanwhile, to achieve both the time and spatial tracking of phenotypes changes, cells will be monitored through fluorescent microscopy. The major difficulty is to identify to which cell belongs the cellular border in order to display the measured fluorescence to the right cell.

Thereby, to limit the identification bias, each of the fluorescent protein is preceded by a nuclear localization signal (NLS) whose ubiquitous protein sequence described for the first time in the large T Simian Virus 40 early-promoter (SV40) gene is : Pro-Lys-Lys-Lys-Arg-Lys-Val [153, 172, 354]. Thus, the matured fluorescent protein should be concentrated in the nuclei of cells and our identification of phenotypes should be based on the quantification of nuclear fluorescence. However, this approach has a minor limitation in respect to the turn-over of fluorescent proteins and so the identification of phenotypic transitions (phenotypes and state change speed). Indeed, once in the nucleus, the half-life of proteins depends on the nuclear ubiquitylation signal (and particularly the proteasome activation signals) and the abundance of several key agents in the proteasome. On the other hand, the export ability of ubiquitylated protein into the cytosol appears to be a key process of the degradation of certain proteins present in the nucleus [213]. However, to our knowledge, the half-life of fluorescent proteins affected by this nuclear localization feature remains unpublished.

Each of these two described elements (NLS signal and fluorescent proteins) are driven by a promoter specific for each of the monitored phenotypes (CD24 and CD44). Blocks, as described in green and red (*cf.* fig. 10.4) are placed in opposite directions to each other in order to limit the risk of serial activation of promoters (if placed in same direction) which could disrupt the

specificity of expression sought.

10.2.2 A reporter of molecular platform delivery

The blue block pattern (*cf.* fig. 10.4 on page 150) is important to :

- Titer lentiviral production.
- Validate the effectiveness of lentiviral transduction within the cell lines of interest.
- Normalize the fluorescence resulting of the expression of the markers and limit fluctuations in the fluorescence level related to the number of integration of the platform and chromatin context of the integrated area [74]. The normalization of fluorescence fluctuations observed for every measurement (made at $t + x$) limit the variations bias due to the integration of the molecular platform (number and location of integrations).

Thus, the expression bloc consists of a non-fluorescent reporter protein driven by a promoter having a strong ubiquitous expression.

Finally, we designed a molecular platform whose main objective is to report the endogenous expression of two specific markers of expression of a phenotypic state. Each of the above-described elements are easily adapted (change of reporter proteins and promoters) by targeted digestion with restriction enzyme or homologous recombination. This flexibility has been particularly refined to facilitate the implementation of this molecular platform to other cell lines and other endogenous markers.

Chapter 11

Validation of individual functional units forming the PhenotypeTracker molecular platform

«La vertu est la demeure du Sage, le devoir, sa voie, la politesse, son vêtement, la prudence, son flambeau et la sincérité, son sceau» - Tchouang-Tseu (IVème siècle avant Jésus-Christ)

As a beginning, we will confirm that the Δ LNGFR reporter gene will be a great candidate to monitor SUM149-PT and SUM159-PT lentivirus transduction efficiency. Then, you will appreciate the expression localization of the different subunits which compose the PhenotypeTracker platform thanks to fluorescent imaging analysis. Then we will go forward to the long-term monitoring of Δ LNGFR transgene expression and slightly focus on the efficiency of SUM149-PT and SUM159-PT subphenotypes transduction. Finally, the validation of CD44 human promoter will be discussed thanks to the tracking of its level of expression on specific control human cell lines.

I just want to take few lines to remind you that all molecular construction works you will encounter through this chapter are described at length within the table that contains all the constructions that were designed in the context of this thesis (*cf.* appendix F on page 309). Moreover, you probably are interested in knowing the strategies that was implemented to build these several molecular platform. You can follow the link to the appendix E on page 285 where these following strategies are fully detailed.

11.1 Δ LNGFR as reporter of the lentiviral transduction of target cells

LNGFR gene is mainly expressed within neuronal tissues [162]. However, colleagues have previously reported its expression within some epithelial tissues as the normal prostate epithelia [73] and mesenchymal tissues such as for instance inflammatory muscular tissue [128]. Up to my knowledge no papers mentioned its expression within neither normal nor cancerous mammary epithelia. Nonetheless, LNGFR protein and its interacting proteins are playing an important

role in regulating cells oncogenic processes [73] and due to that latest argument I considered Δ LNGFR expression within my cancerous cells of interest.

Thus, I found very important to validate that the genetic marker that we want to overexpress is not endogenously expressed within SUM149-PT and SUM159-PT mammary cancerous derived cell lines. Indeed, a too strong expression for that specific endogenous gene marker would not, for example, enable to efficiently sort positively transduced cells.

11.1.1 Δ LNGFR gene is poorly endogenously expressed within SUM149-PT and SUM159-PT cell lines

Thus, I lingered in quantitatively studying Δ LNGFR endogenous expression within SUM149-PT and SUM159-PT cell lines (*cf.* fig. 11.1).

A quantitative flow cytometry analysis was set up and living SUM149-PT and SUM159-PT cells were stained with Δ LNGFR primary antibodies coupled with Allophycocyanin (APC) labeled secondary antibodies according to the established protocols in appendix D.17 on page 274. After any cells doublet withdrawal, viable cells selected on the basis of DAPI negative signal were plotted into LNGFR-APC flow cytometry histograms. Then fluorescence emission intensity from these latest labeled cells was confronted with that emitted from controlled unstained and APC labeled secondary antibodies stained cells (*cf.* fig. 11.1).

This figure (*cf.* fig. 11.1) is a representative one emerging from more than 3 independent replicates that have been implemented.

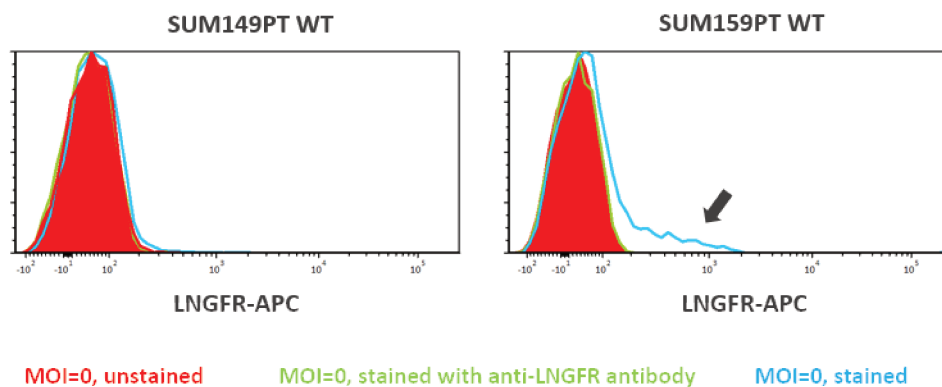


Figure 11.1 : *Validation of the non endogenous expression of LNGFR protein within SUM149-PT and SUM159-PT cell lines.*

SUM149-PT cell line does not endogenously express Δ LNGFR gene (*cf.* fig. 11.1). However, it is interesting to note that within SUM159-PT cell line around 20% of SUM159-PT cells are actively expressing LNGFR proteins with a fluorescence intensity 10 times higher compared to unstained cells (*cf.* fig. 11.1). This latter observation did not affect the cloning strategy that was previously described (*cf.* fig. 10.4 on page 150). However, on the basis of this information, sorting procedures of SUM159-PT transduced cells were performed in carefully selecting cells with highest LNGFR-APC fluorescence intensity, *i.e.* more than the maximum endogenously expressed.

As a conclusion of these experiments the use of Δ LNGFR as a positive lentivirus transduction reporter was validated and we took that advantage to go further into the validation of the transduction reporter sub-unit.

11.1.2 Sur-expression of Δ LNGFR transgene within HEK293T cells

The design of the PhenotypeTracker platform implies that the efficiency of the transduction monitoring highly depends on spatial, temporal and strength of Δ LNGFR transgene expression. In this way one of the major requirement can be formulated as the LNGFR expression do not fluctuate as a function of cell phenotypes transitions and more generally according to cell fate decision. Therefore, I considered a range of house-keeping gene promoters able to drive LNGFR transgene and whose characteristics can be summarized by these points :

- Strong expression abilities
- Over-time stable expression
- Ubiquitous expression cells tropism
- Uniform expression among positively transduced cells
- No differential expression between stem-like and more differentiated cells

Numerous house-keeping genes and their associated promoters have been discovered [79]. Many of these constitutive promoters as the SV40, CytoMegalovirus immediate-early promoter (CMV), human Elongation Factor 1 α (EF1 α), ubiquitin C, or PhosphoGlycerate Kinase 1 derived from mouse strain (PGK) promoters have been extensively cloned to overexpress ectopic genes for both *in vitro* and *in vivo* studies. Moreover, synthetically designed promoters have enriched the ranks of strong expression promoters as the promoter Chicken β -Actin promoter coupled to the CMV early enhancer (CAG) [152, 229] and other CMV-derived promoters [272]. Within this huge number of possibilities, I selected those whose expression was compatible with stem-like conditioning cellular environment [249]. And this is thanks to the help of comparative studies performed on established cell lines [249], ES cells, iPS cells [52, 136, 230] or even hematopoietic stem cells that we got interested in two putative candidate promoters : the human EF1 α and the mouse PGK promoters [99, 175].

Thus the two constructs driving transgene Δ LNGFR have been devised with the two promoters below :

- phEF1 α
- pmPGK

The lentiviral productions were produced according to the protocols previously described from the 3 plasmids [18, 60], including :

- The envelope which encodes the VSV-G protein as a result of its broad tropism infectivity (plasmid constructed by Bob Weinberg).
- The packaging which encodes the GAG protein, POL, Rev, Tat (8.74 plasmid constructed by Didier Trono).
- The expression plasmids of interest.

Lentiviral productions were then assayed by assessing the average level of transgene expression expressed in HEK293T cells by flow cytometry and Mean Fluorescence Intensity (MFI) characteristic measurement of each transduction. Once the number of lentiviral particles per μ l was evaluated, HEK293T cells were transduced with either phEF1 α - Δ LNGFR or pmPGK- Δ LNGFR at different MOI (*cf.* fig. 11.2 on page 157 subfig. A).

The Multiplicity of Infection (MOI) indicates the mean number of vector particles per cell. In HEK293T cells, MOI=1 corresponds to the integration of one molecular platform per cell assuming lentiviral particles can integrate cells in every step of cell cycle [219]. The HEK293T cells were transduced at a MOI equal to 1, 3 or 10 and the MFI of positively transduced cells was evaluated at 21 days post lentiviral transduction (*cf.* fig. 11.2 on the next page subfig. B).

These measurements were performed once for each MOI and for each molecular platforms.

The fluorescence histograms measuring the level of transgene expression Δ LNGFR show that both constructs of interest are expressed in more than 80% of the plated cells (*cf.* fig. 11.2 on the facing page subfig. B). The level of transgene expression (of positive cells) is not influenced by the MOI of 1 and 3. However, we observe an increase in MFI of 0.2 log for a MOI of 10.

Accordingly, we quantified the number of actual integration of the constructions in the genome of the cells positively expressing the transgene Δ LNGFR. To do this, we sorted the cells of interest (positive Δ LNGFR) by flow cytometry and had their DNA extracted. In collaboration with Gentner et al., we quantified the number of construction integration for each condition. So we set up a quantitative PCR using probes specifically recognizing the LTR sequences carried by the integrated lentiviral constructs and all raw data are normalized to the TELO genomic locus [175]. The results, reported by Gentner et al., demonstrated that the positive cells issued from MOI 1 and 3 integrated one lentiviral construction only while those with MOI 10 had integrated two lentiviral constructions (data not shown). Thus, the MFI measured by flow cytometry, representing the level of Δ LNGFR transgene expression, is concordant with the number of measured lentiviral integrations.

However, the expression efficiency of these selected promoters still remains highly dependent on target cell type and large variations in reporter gene expression level have been reported [249]. In this framework we decided to study their expression abilities within SUM149-PT and SUM159-PT cells.

11.1.3 Promoters, dependencies of over-expression of transgenes in SUM149-PT and SUM159-PT cell lines

The expression of promoters varies according to the cellular origin (species, organ, cancer state, differentiation stage) [187, 218], we wanted to find the promoter capable of driving the expression of Δ LNGFR transgene in both cell lines of interest. Also, we transduced at different MOI (5, 10 or 20) each of SUM149-PT and SUM159-PT cell lines with :

- On one hand, the pEF1 α - Δ LNGFR construction
- On another hand, the pmPGK- Δ LNGFR construction

The percentage of cells positively expressing Δ LNGFR and MFI were assessed by flow cytometry at 10 days post-transduction. These analyzes were carried out at two independent times (*cf.* fig. 11.2 on the next page subfig. C).

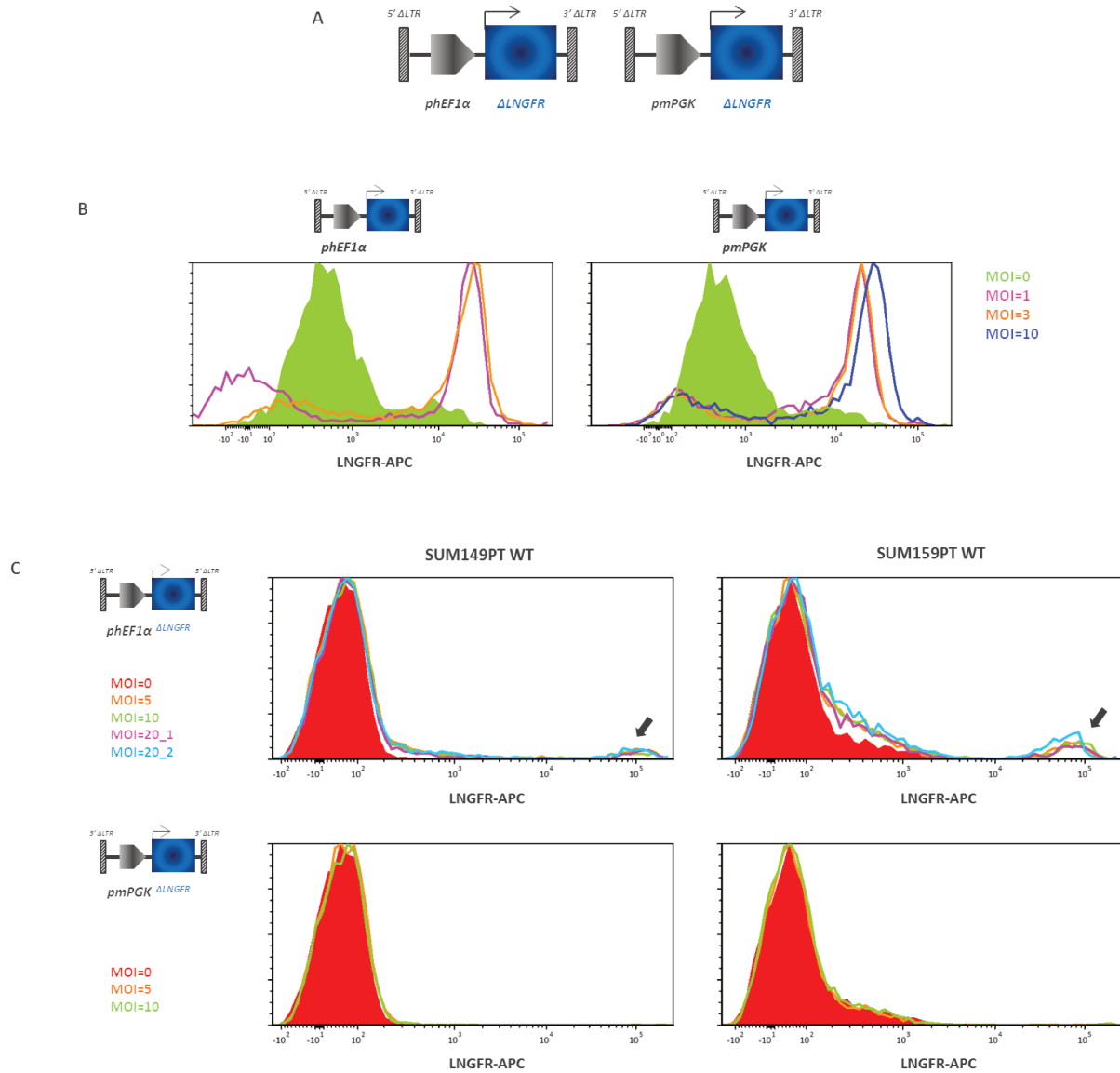


Figure 11.2 : Choice of promoter driving Δ LNGFR reporter whose expression is stable regardless of phenotypes' expression over time.

A. Schematic view of Δ LNGFR ubiquitous expression PhenotypeTracker platforms. B. Flow cytometry histograms of Δ LNGFR expression on HEK293T transduced cells in function of the transduction of the PhenotypeTracker platforms and their MOI. C. and on SUM149-PT and SUM159-PT transduced cells.

PGK promoter is not expressed neither in SUM149-PT nor in SUM159-PT cell lines. In opposition to previous promoter, $\text{phEF1}\alpha$ clearly over-expresses the transgene in around 6% of cells from SUM149-PT cell line and approximatively 8% of SUM159-PT cells. Finally, the MFI of cells positively expressing ΔLNGFR does not vary depending on the MOI. In transgenic mice McBurney et al. have already put in evidence that mouse PGK promoter level of expression varies from one adult tissue to another and that within tissues important cell-to-cell variabilities prevail ; suggesting that its activation is highly dependent on cells' metabolism requirements [206]. Following these observations, we chose as both transduction reporter and fluorescence normalizer of the PhenotypeTracker (blue frame of fig. 10.4 on page 150) the promoter $\text{phEF1}\alpha$. Thanks to confocal microscope observations, we were able to validate that there is no significant inter-cellular differences in the intensity of expression of the LNGFR marker (data not shown). This investigation, at the single-cell level, was validated regardless of the number of integration of the lentiviral platform (*cf.* fig. 11.3) and perfectly match flow cytometry measurements (*cf.* fig. 11.2 on page 157). The program to quantify the fluorescence signal is available within the supplementary data (*cf.* appendix G on page 317).

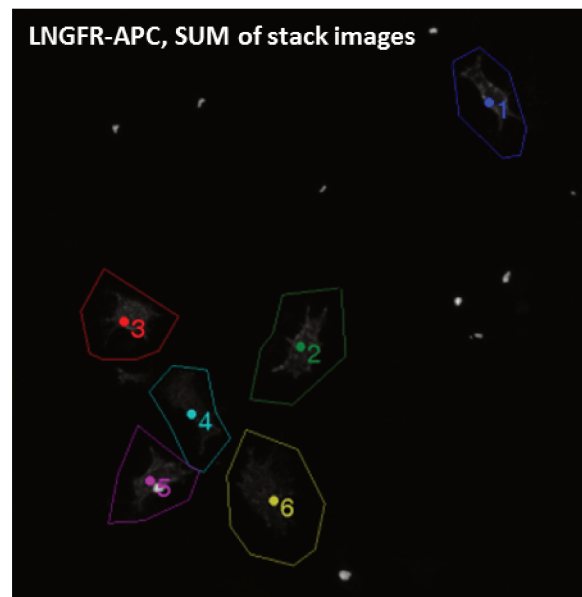


Figure 11.3 : *Confocal image analysis strategy developed by Isabelle Bonnet (UPMC, Institut Curie) to measure transduced single cell LNGFR expression level.*

11.2 Cell localization of transgenes encoded through the PhenotypeTracker molecular platform

The transduction reporter being chosen, we wanted to study the localization of expression of each of the blocks (green and blue expression block on the diagram of PhenotypeTracker (*cf.* fig. 10.4 on page 150)). Thus, we have enriched the $\text{phEF1}\alpha$ - ΔLNGFR construction with a block (in green) overexpressing fluorescent protein eGFP flanked in C-ter with a Nuclear Localization Sequence (NLS) [123, 309]. This fluorescent transgene is driven by the ubiquitous CMV early promoter to locate each of the component in the cell.

HEK293T cells were transduced with MOI=3 through a construction (all blocks constructed in sens orientation) pCMV-NLS-eGFP- $\text{phEF1}\alpha$ - ΔLNGFR (*cf.* fig. 11.4 on page 160 subfig. A). 4 days post-transduction, cells were stained with ΔLNGFR primary antibody followed with

secondary antibody labeled with APC fluorescent dye (*cf.* table F.5 on page 316). And then cells were fixed for further confocal microscopy acquisitions as described in supplementary protocols (*cf.* appendix D.17 on page 274).

The two frames (green and blue) form a promoter and a transgene enabling expression for each of the two transgenes. The transgene Δ LNNGFR is mostly localized at the cytoplasmic membrane. It is also found in lower density in the form of aggregates in the cytosol. The nucleus is devoid of it (*cf.* fig. 11.4 on the following page subfig. Bb).

EGFP fluorescent protein flanked by the NLS sequence is found in the cytoplasm and with a higher density in the nucleus (*cf.* fig. 11.4 on the next page subfig. B and C). This dual localization of the fluorescent protein leads us to discuss two strategic points relating to :

- The phenotypes discrimination
- The monitoring of phenotypic states transitions

Due to the cytosol localization of the eGFP protein, it will be difficult to identify the phenotype of a cell based solely on the measured intensity within the nucleus. We have therefore to take into account the expression of the transgene throughout the cell. This implies that a method to subtly segment cells will be incremented to limit the bias of phenotypes identification. The presence of a large amount of protein in the cytosol may be due to several factors :

- A high level of traduction of the proteins.
- An important export of the nucleus protein to the cytosol for degradation by cytoplasmic proteasome. In this case, there will be a faster turnover of the fluorescent proteins, reporters of phenotypic expression. This will highlight more precisely the phenotypes transitions and measure the time required to switch to a different phenotypic state.

To quantify the rate and speed of NLS-eGFP protein export from the nucleus to the cytoplasm, we should inhibit all agents in the export process. If there is a decrease in the amount of fluorescent proteins in the cytosol, this means that the export rate from the nucleus into the cytosol is very high.

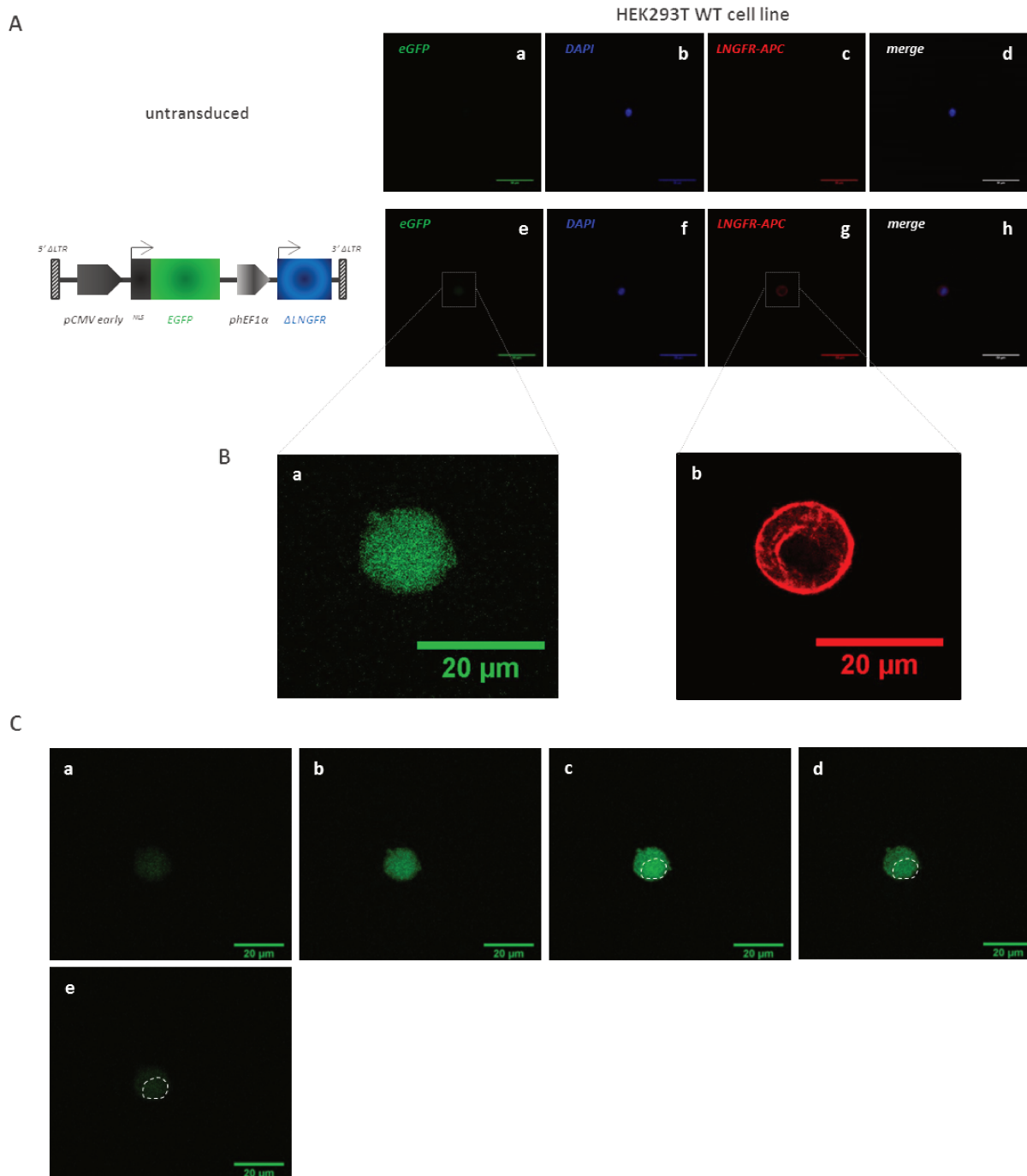


Figure 11.4 : Validation of NLS-Fluorescent protein expression on HEK293T cell line.
A. Representative calibrated fluorescent pictures obtained from untransduced and eGFP (in green) ubiquitous Phenotypes Tracker platform transduced HEK293T cells and stained for 2 markers DAPI (in blue) and LNGFR-APC (in red) (x20 magnification confocal microscope and picture pixel size= 1pixel/0.166 053 2 μm). White and colored scale bar equal 50um. B. Zoom of pictures (e) and (g) from A. Colored scale bar equal 20um. C. Representative Z stack of calibrated fluorescent picture taken from eGFP (in green) ubiquitous Phenotypes Tracker platform HEK293T transduced cells. Colored scale bar equal 20um.

11.3 Transgene expression bias within SUM149-PT and SUM159-PT subphenotypes

During our first experiments of transduction of SUM149-PT and SUM159-PT cell lines, we noticed that the percentage of cells positively expressing the molecular construct was particularly low (around 8%) compared to HEK293T (around 80%) (*cf.* fig. 11.2 on page 157). This remark take place in the context of the same molecular platform have been transduced in these three cell lines : phCD44-NLS-eGFP-phEF1 α - Δ LNGFR.

The expression level of the Δ LNGFR transgene was controlled at different time points by flow cytometry and sorting procedures were adapted from previously described procedures in section 5.1.2 on page 51.

17 days after transduction the SUM149-PT cells positively expressing the transduction marker (Δ LNGFR) represented only 1% of the analyzed living cells. These cells were purified to 99.9% and placed in culture for 45 days. The percentage of cells positive for Δ LNGFR was only 18%. 77% of negative cells had appeared.

Thus, after each enrichment in positive cells, cells which were not expressing the integrated transgene appeared. This may be explained as follows :

- One or more subphenotypes do not express the lentiviral construct
- Δ LNGFR positive cells are capable of phenotypic transitions

This reflection has led us to look for potential bias in subphenotype inactivation. For each of our analyzes, we therefore coupled the quantification of the level of Δ LNGFR transgene expression with the characterization of subpopulations according to their level of expression of the markers CD24 and CD44.

Both SUM149-PT et SUM159-PT cell lines have been transduced with phCD44-NLS-eGFP-phEF1 α - Δ LNGFR molecular platform at a MOI=10. And then monitoring have been performed as described earlier following the protocols you can find in section 5.2 on page 53.

74 days post transduction, only 1.44% were positively expressing Δ LNGFR transgene. Both negative and positive expressing cells were stained for CD24 and CD44 markers. Surprisingly, positively Δ LNGFR expressing cells were CD24+/CD44+. While the negative cells for Δ LNGFR transgene expression exhibit dynamic of expression for CD24 and CD44 markers, similar to that of non-transduced cells. Following the two successive sorting procedures of cells positively expressing the Δ LNGFR marker, there is resurgence of negative cell for the same marker (*cf.* fig. 11.6 on page 164). The analysis of these two populations in the space of CD24 and CD44 markers has validated the characterization made at 74 days post transduction :

- Δ LNGFR positive cells correspond to a cell subpopulation expressing the CD24+/CD44+ phenotype tropism.
- Δ LNGFR negative cells exhibit for CD24 and CD44 an expression map quite equivalent to that of Wild Type (WT) cells. However, double negative population is more represented there.

However, after cell sorting I and II, the expression maps of SUM149-PT Δ LNGFR repopulating cells, in CD24 and CD44 space, were quite different compared to SUM149-PT untransduced cells as illustrated in chapter 5 on page 49. The CD24-/CD44+ population is much more dispersed. The percentage of the subpopulation is also higher (against 1% for SUM149-PT WT cells).

So transducing cells with lentiviral particles of VSV-g tropism changes the behavior of cells in the expression of CD24 and CD44 space.

While the proposed tool is functional because :

- It reports the expression of Δ LNGFR a witness both of lentiviral transduction and transcriptional state of cell phenotypes (*cf.* fig. 11.5 on the next page).
- It demonstrates the expression of the fluorescent protein derived from the upstream promoter, which proves that the bi-block construct (in blue in green) in sens works.

However, the tool is only valid for a particular subpopulation of CD24+/CD44+. This does not meet the objective to follow the transitions between the different phenotypic states (CD24-/CD44 +, CD24+/CD44+ and CD24-/CD44-) in real time. This is unexpected because, the promoter phEF1 α (expressing Δ LNGFR) is described as strong, ubiquitous. It should therefore not be turned off automatically in some phenotypes (CD24-/CD44-, CD24-/CD44+, a majority of the population CD24+/CD44+) in a pluri-clonal context because it is less prone to silencing and should provide more stable long term expression [249, 288, 289].

A massive de-expression of the reporter and transduction normalizer means that there is probably a drastic chromatin remodeling and the appearance of a local structure inhibiting the transgene transcription Δ LNGFR [66]. Especially as, the number of actors (histone acetyltransferase, deacetylase, methyltransferase, kinase) involved in this process of remodeling chromatin and their activity is altered during the cancer arising process [179, 194, 351].

Hopefully, transcriptomic studies that we have launched for the three phenotypes CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- should inform us on the level of expression of agents involved in chromatin remodeling based on phenotypes (*cf.* fig. A.15 on page 222).

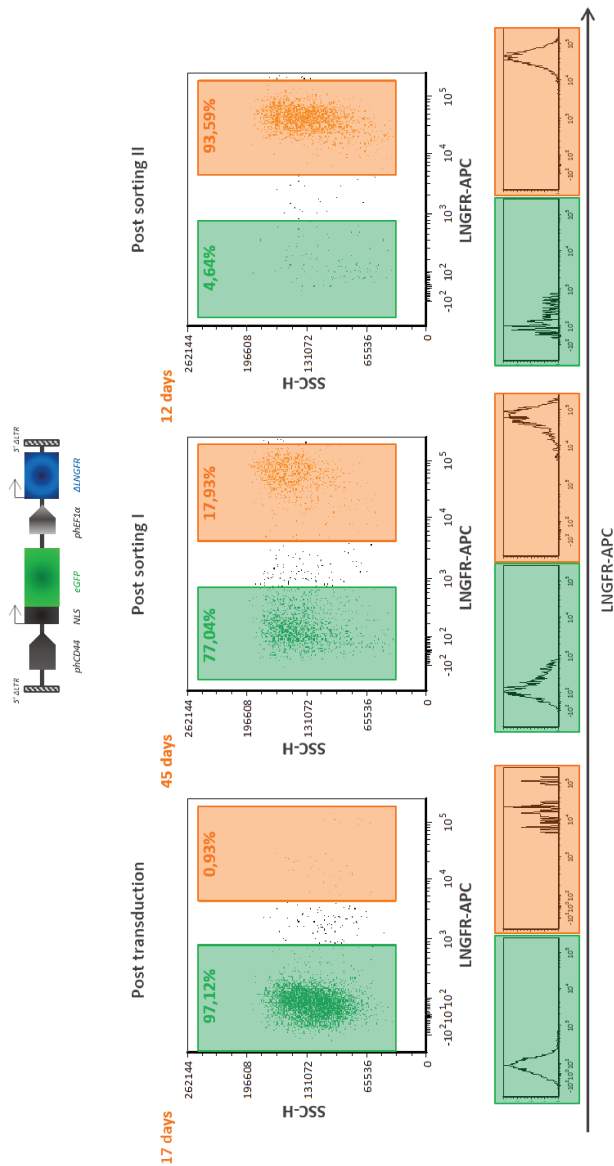


Figure 11.5 : Monitoring cell sorting procedures of Δ LINGFR expression within CD24 and CD44 phenotypes of SUM149-PT cell line transduced with CD44 expression PhenotypeTracker platform.

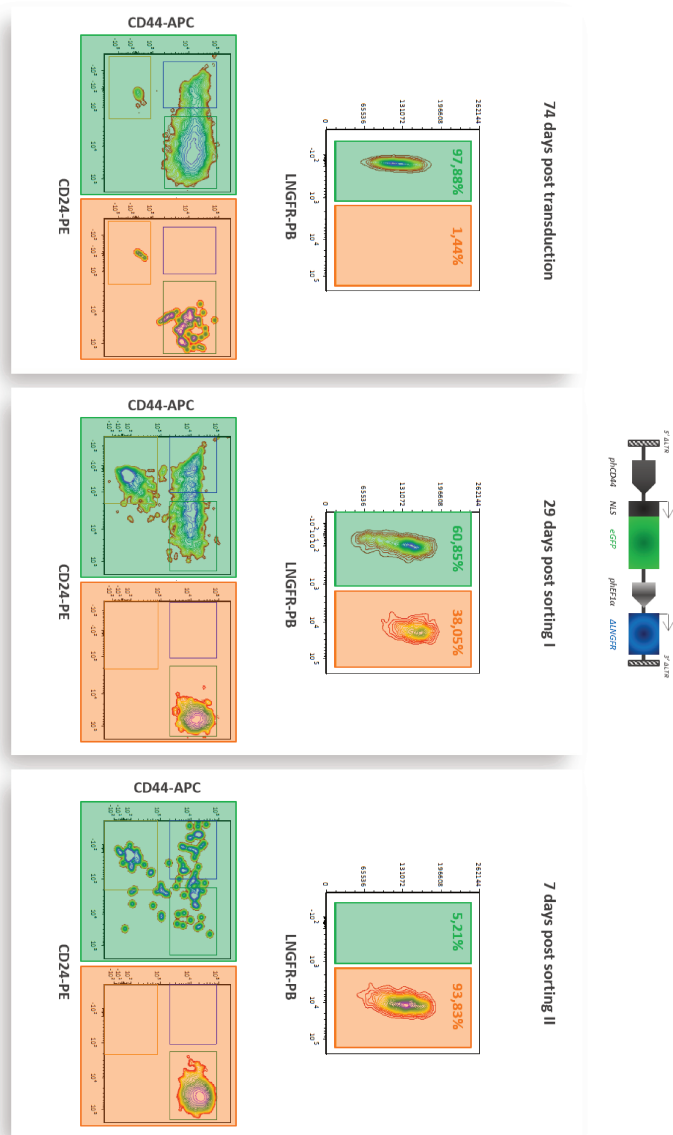


Figure 11.6 : *Monitoring the cell sorting procedures of ΔLINGFR expression pattern within SUM149-PT cell line transduced with CD44 expression PhenotypeTracker platform.*

11.4 Validation of some human driving genes promoters

Monitoring of phenotypic transitions (activation and inactivation) of both CD24 and CD44 markers was based on the exogenous input from constitutive promoters. These later should be able both to follow the fluctuations of endogenous expression of markers of interest and to validate the good expression of regulatory sequences (described as promoter sequences for expression of the gene of interest).

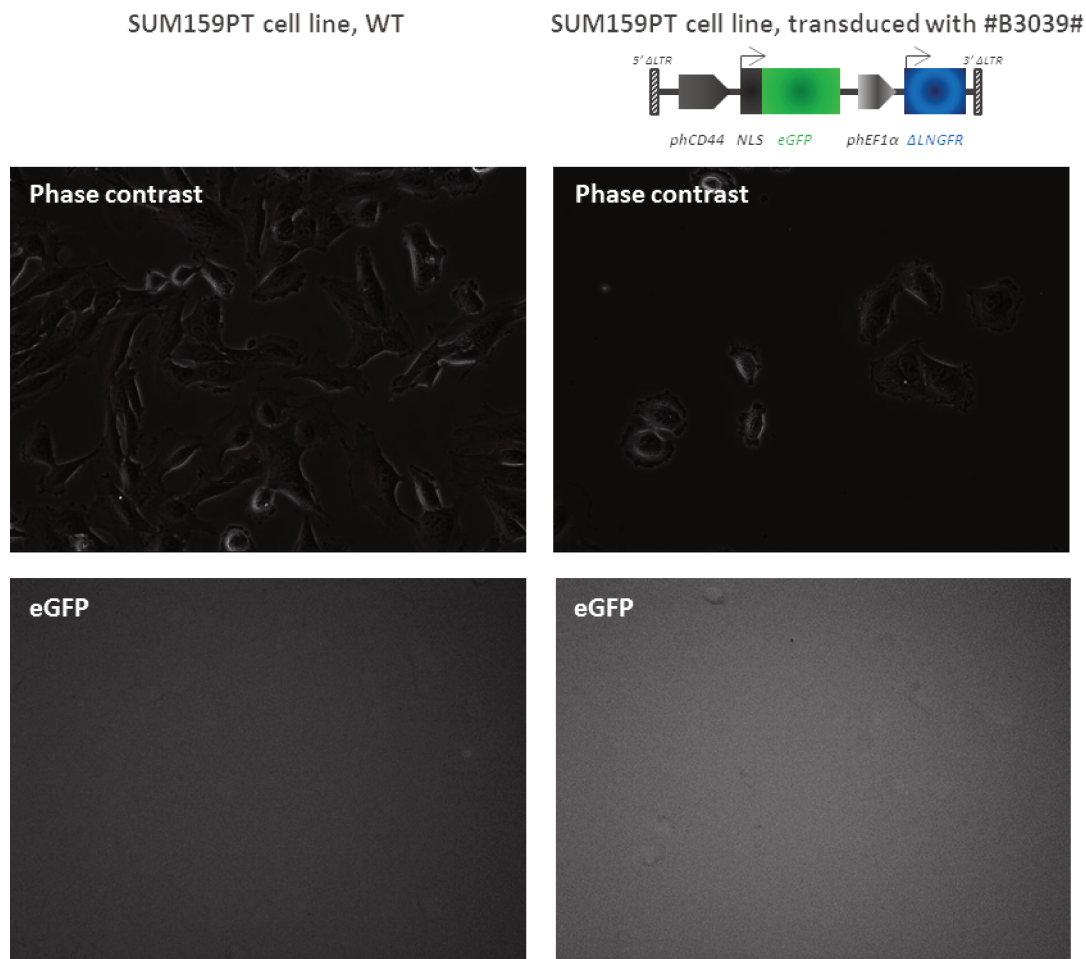
11.4.1 Regulatory sequences of human CD44 gene

The human CD44 gene promoter, isolated by the team of Godard et al. (plasmid deposited within Addgene® plasmids bank, reference : #19122) carries many regulatory elements such as the pro-oncogenic factor p53, the EGF regulatory element [355] and the site of binding Nuclear Factor Kappa-light-chain-enhancer of activated B cells (NF κ -B) [292]. The phCD44 promoter, amplified by Polymerase Chain Reaction (PCR) from genomic DNA (of the human normal breast cell line : HMLE) and 2021 base pairs upstream of translation initiation site, have been cloned in place of the pCMV promoter initially tested.

The SUM159-PT cells positively transduced by viral particles containing the sequence phCD44-NLS-eGFP-phEF1 α - Δ LNGFR at MOI = 10 were purified. Positive Δ LNGFR cells were observed by fluorescence microscopy 24 hours post-sorting for 24 hours (one image per 30mn) (and performed thanks to the Nikon Center microscopy platform of the Curie Institute). Positive Δ LNGFR cells, corresponding to a subpopulation CD24+/CD44+ (*cf.* fig. 11.6 on page 164), are supposed to strongly express the CD44 marker. However, nowhere in the field we observe fluorescence signals for eGFP.

This highly low enhanced Green Fluorescent Protein (eGFP) signal intensity could be a matter of human CD44 promoter inefficiency or hugely low transcriptional activity. Indeed, frequently highly specific promoters are inefficient activators of transcription as it is regarding the von Willebrand factor promoter [146] or even the promoter of the sucrose-isomaltase gene [341]. This concern is one of the major obstacle that the field of gene therapy is facing nowadays and which is to combine at the same time temporal and spatial specificity of gene target transcription and high efficiency of gene re-expression.

In order to enhance the activity of a weak promoter without losing cell-type and tissue-specificity, one strategy has been proposed by Nettelbeck et al. in developing a self-enhancing promoter through the establishment of a positive feedback loop transcription activation [222]. This strategy consists in coupling a cell type-specific promoter to drive the simultaneous expression of the desired reporter gene with a strong artificial transcriptional activator. This activator sequence tally with a chimeric protein which combine the DNA-binding domain of LexA29 gene and the strong herpes simplex virus VP16 transcriptional activation domain to stimulate the transcription through LexA binding sites introduced into the promoter. However, under this positive feedback conditioning, no one knows if gene transcription signal is still following endogenous expression variations. This key remark would condition the sensitivity of the PhenotypeTracker platform in reporting and mimicking endogenous CD24 and CD44 expression.



**Figure 11.7 : Fluorescence microscopy of SUM159-PT cell line after 72h of CD44 expression
PhenotypeTracker platform**

11.4.2 Regulatory sequences of human CD24 gene

The endogenous promoter region (present in chromosome 6Q21) of the human CD24 marker has been extensively studied to determine the Single Nucleotide Polymorphism (SNP), the regulatory sequences and external factors (as for example hypoxia criteria) which modify sensitivity to transcription signaling. The regulatory sequences of the human CD24 marker has not been determined with consensus. And moreover, the different regulatory elements described (Specificity Protein 1 (SP-1) binding domain, Hypoxia-Inducible Factors (HIF)1- α binding domain, E-box binding TWIST2 area) significantly influence the level of expression of reporter protein [186, 312, 330].

Experiments are in progress by transient transfection of MDA-MB-468 cells (breast cancer cell line which highly expresses CD24 markers) to determine the promoter sequence which would mimic the best the endogenous expression of the CD24 marker. The choice of this specific human derived cell line is based on transcriptomic data kindly provided by Dr François Radvanyi (UMR144, Curie Institute) and Dr David Gentien (Genomics platform, Curie Institute) [254]

The main objective of this work was to build a molecular tool capable of simultaneously tracking the endogenous expression variations of two markers (CD24 and CD44). The strategy is to use the endogenous regulatory sequences of the two markers of interest as a promoter element which would trigger the expression of two short half-life fluorescent proteins. Thus, this tool consists in three independent units :

- Two units (red and green blocks) are used to monitor the level of expression of the two markers by combining their expression and to determine in real-time the phenotypic state of each cell.
- A unit (blue block) plays at the same time the roles of transduction marker, the chromatin state normalizer and copies-to-host-genome normalizer.

The design of the construct allows us to track simultaneously the expression of the transduction reporter (blue block) and the expression of a fluorescent marker driven by another promoter. However, we have identified an important bias of expression of the construct for a subpopulation of CD24+/CD44+ phenotype. These observations suggest that significant remodeling of the chromatin must take place during phenotypic transitions, significantly altering the transcriptional state of each of the phenotypes described in the space of CD24 and CD44 expression (CD24-/CD44+, CD24+/CD44+ and CD24-/CD44-).

On the other hand, the human promoter sequence of the tested CD44 gene is not capable of driving the expression of a reporter protein according to the level of endogenous expression of the CD44 marker. Clearly, the isolated sequence by our colleagues does not hold all the necessary regulatory sequences to mimic the endogenous expression of the CD44 gene. Thus, monitoring an endogenous marker using an exogenous construct does not seem to be the best method.

To meet the challenge it would be interesting to adapt other technology :

- Either genetic by specifically (or even surgically) targeting the constitutive promoter and by grafting in C-ter of Open Reading Frame (ORF) a fluorescent protein with linker sequence of P2A/T2A type. This procedure would slightly disrupt the mRNAs splicing and maturation process. Various technologies of genome editing are possible such as Zinc Finger Nucleases, TALENs, or Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas system [77, 133, 190, 252, 321]. The biggest challenge will be to determine the best method to bring the molecular engineering system within each target cells. This will firstly require optimizing vector tropism to efficiently transduce cells diversity (using for example human foamy virus [129], murine leukemia virus [85]...) and secondly necessitate to non-defectively trigger the expression of a signal normalizer .
- Or non-genetic using quantum dots technology on which would be grafted epitopes capable of specifically recognizing the CD24 or CD44 markers. Quantum dots are nanocrystals whose crown composed of limits the toxicity of the semiconductor. By adjusting the core sizes of quantum dots specifically targeting for CD24 and CD44 markers (and add in excess in the cell culture medium) we could monitor cells staining to determine their phenotype and track in real-time their phenotypic transitions [294, 349].
- Besides, a technology emerged in recent years to follow non-phenotypic transitions without an invasive manner. Indeed, Raman spectroscopy is an established laser-based technology to track changes of biochemical structures or metabolism within cells as markers of cell state transitions. By relying on the major characteristic that each Raman spectra are specific for each cell type and provide additional information on cell viability, differentiation status, and tumorigenicity. Thus, in our case this technology would bring the opportunity

to dynamically follow cells status changes without labeling issues, and with the ability to increment researches towards *in vivo* follow-up [33, 140, 240].

Chapter 12

Order or disorder? Resulting of a spatial dependence of cells

*«Dans la nature, tout a toujours une raison. Si tu comprends cette raison, tu n'as plus besoin de l'expérience.»
- Léonard de Vinci -*

Previous chapter “the issue of social relationships among cells” has shown that the average phenotype of population was influenced by the phenotype of the ancestral population. And we discussed the existence of possible temporal patterns (*cf.* fig. 10.2 on page 147). Thus, the probability of transiting at $t = 0$ from one state to another varies in function of the phenotypic state of the ancestral cell at $t = -1$ (*cf.* fig. 10.1 on page 146). Furthermore, we found that there was always a state of equilibrium of both the phenotypic proportions and the average level of expression of markers.

Our task now is to focus on a single cell and confirm whether its neighboring cells influence its phenotypic behavior (expressed phenotype, phenotypic transition, and speed of changes).

The first step was to find a tool to monitor the phenotypic transitions at the cellular level. A first draft of this tool was released in the previous chapter. In this chapter we will propose a methodology to analyze and quantify the social rules governing both the organization of the cells and the phenotypic transitions.

This way of addressing social relations and the resulting consequences (with respect to the expression of the cells) was inspired by an observation of 55 years old. Indeed, in 1960, Harry Rubin showed that some cells affected by the Rous Sarcoma Virus (RSV), and therefore presenting all the potential for developing a tumor, do not develop tumors under certain conditions [264, 265]. In contact with the surrounding non-infected cells, the infected cells do not develop tumors. These experiments merely demonstrate the influence that may be exercised by the cellular environment on both the phenotypic and the behavioral expression of a cell. Thus, do cells become malignant (or at least do express cancerous characteristics), as described by Douglas Hanahan et Robert Weinberg (*cf.* section 2.1 on page 13), as a function of the resilience of the spatial environment they are part of?

12.1 Statistical methodology to determine cells autonomy or cells dependence

In the context of cancer cells, we would like to know if the cellular phenotypes are expressed in a particularly unyielding differentiation tree (both spatially and temporally) or whether the expression phenotype is under the control of a local demand to meet an average phenotypic balance within a given space. That is why we wondered about the process of cell decision. During the decision-making process, are the cells independent or are they under the influence of the community?

- If they are independent, then the phenotypes should be haphazardly organized in space.
- If the spatial structure is ordered, we should find patterns in the spatial organization (*cf.* fig. 12.1 on the next page).

The spatial organization of a population of two phenotypes, a blue (-1) one and a yellow ($+1$) one was simulated so that a pixel of the image corresponds to a cell. In this map, we measured the average phenotype of a Region Of Interest (ROI) whose surface (and therefore the number of cells) varies.

In the simulation, the black circles frame 5 pixels, *i.e.* 5 cells, blue circles 10 cells, pink circles 30 cells, and red circles 100 cells. The average phenotype was computed for each ROI and each time the entire population is crisscrossed, thereby allowing to plot the distribution of the phenotypes for each ROI. The variance of the distribution means phenotypes ROI population is thus our criterion measure of cell behavior.

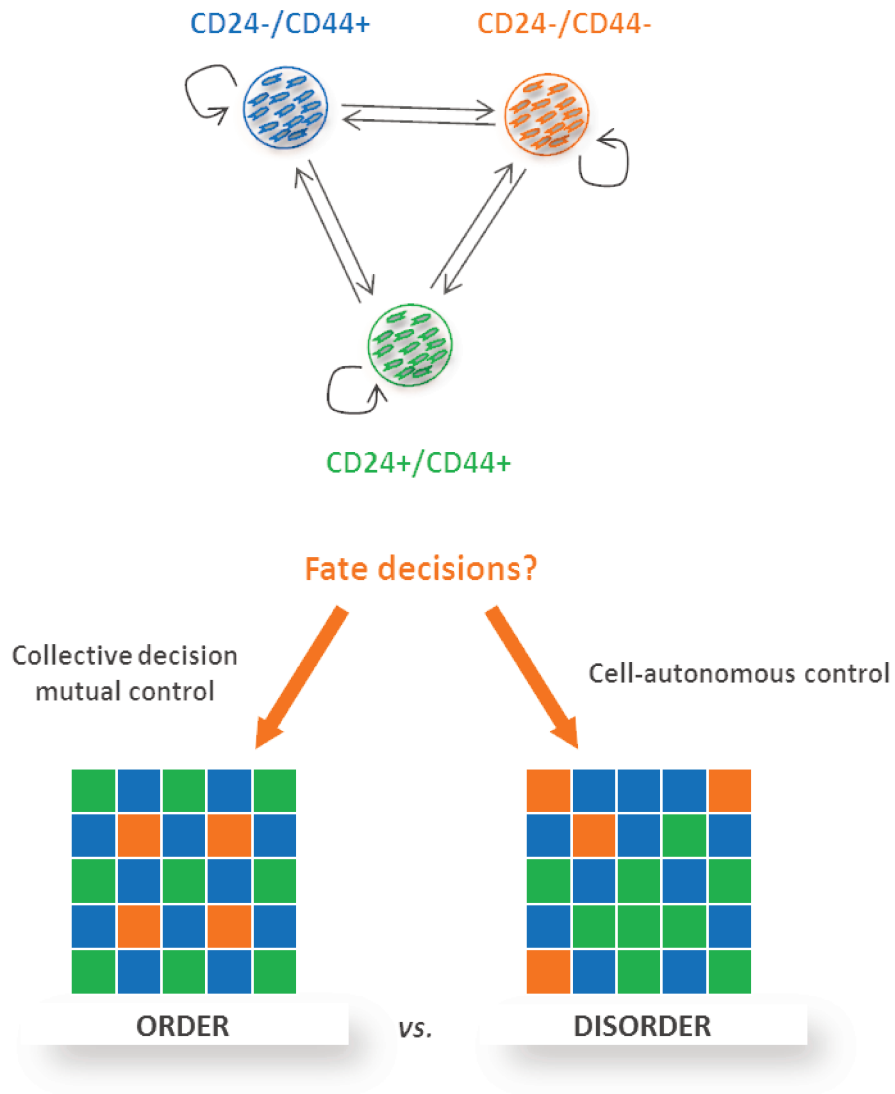


Figure 12.1 : Questioning about spatial correlations between cells fate decisions and cells phenotypes patterning.

In blue : $CD24-/CD44+$ cells sorted phenotype, in orange : $CD24-/CD44-$ cells sorted phenotype, in green : $CD24+/CD44+$ cells sorted phenotype

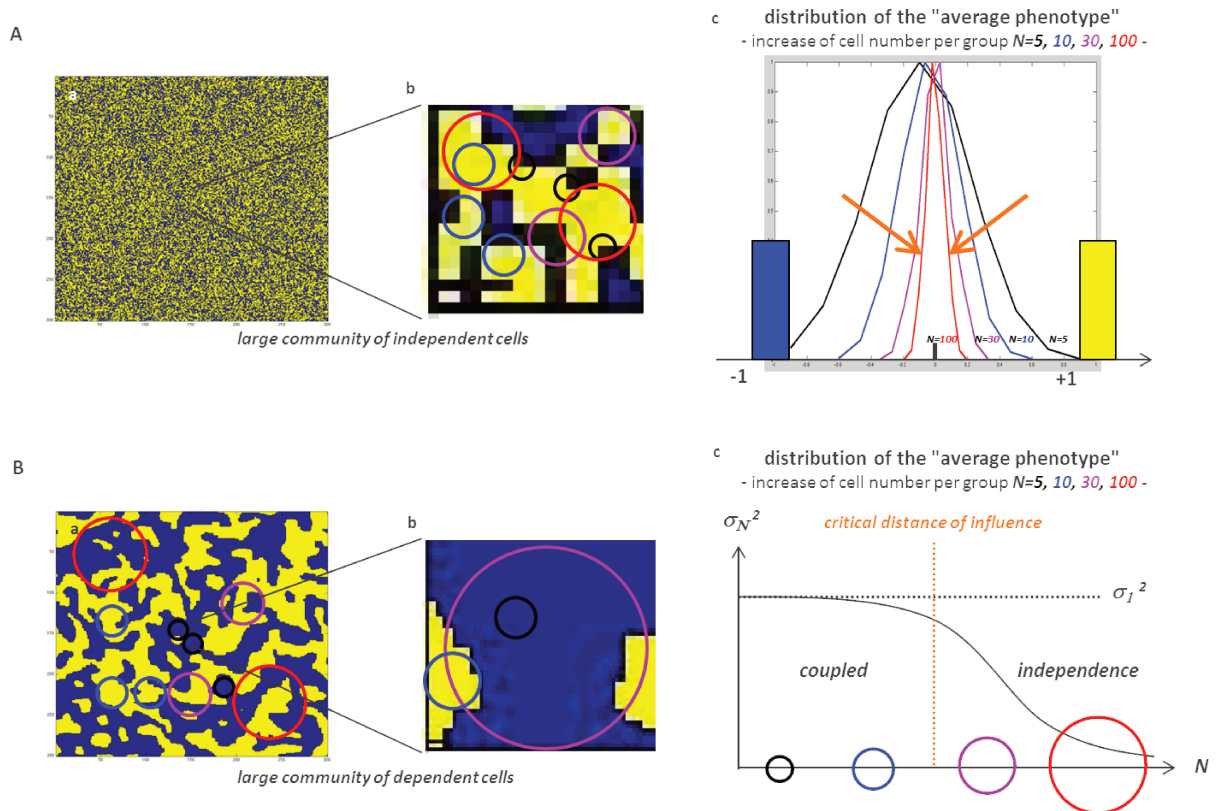


Figure 12.2 : Variance of the distribution of the "average phenotype" calculated per ROI and in function of cells degree of dependence

A. Simulations of large community of independent cells . B. and of coupled cells. Respectively (a) large scale simulation of cells phenotypic expression, (b) zoom of (a), (c) distribution of the "average phenotype" in function of cells number included per determined ROI.

12.1.1 If cells are autonomous

When the behavior of independent cells in the expression of their phenotype (coupling = 0), the phenotypes expression maps at the population level consist of yellow and blue dots are homogeneously and randomly dispersed on the whole area.

The distribution of average phenotypes is according to a normal distribution whose mode is zero. The variance of the distribution follows the law of variances independence and thus decreases linearly while the surface of the ROI increases (*cf.* fig. 12.2 on page 172 subfig. A).

12.1.2 If cells are dependent of the others

When the behavior of cells is dependent from their neighbors (*i.e.* positive coupling), cells of the same phenotype clustered and thus, blue and yellow areas appeared on the maps of expression phenotypes. The law of variances independence does not apply. Thus, the variance is stable until the critical ROI from which it then decreases as the variance of independent law (*cf.* fig. 12.2 on page 172 subfig. B). More cells are coupled then critical ROI is higher.

Thus, in studying the behavior of the variance with regard to the ROI, we may conclude on whether the cells are independent or not of their neighborhood.

12.2 A first approach thanks to homogeneous systems

Since the Phenotype Tracker was not yet functional, we were not able to dynamically monitor the distribution of phenotypic states over time. However, we have made a first approach from sorted phenotypes CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- and re-cultured for 5 days (*cf.* fig. 12.4 on page 175). Then we wanted to compare the dispersion spectrum of the expression level of the cell population for CD24 and CD44 markers at local and global scales (*cf.* fig. 12.3).

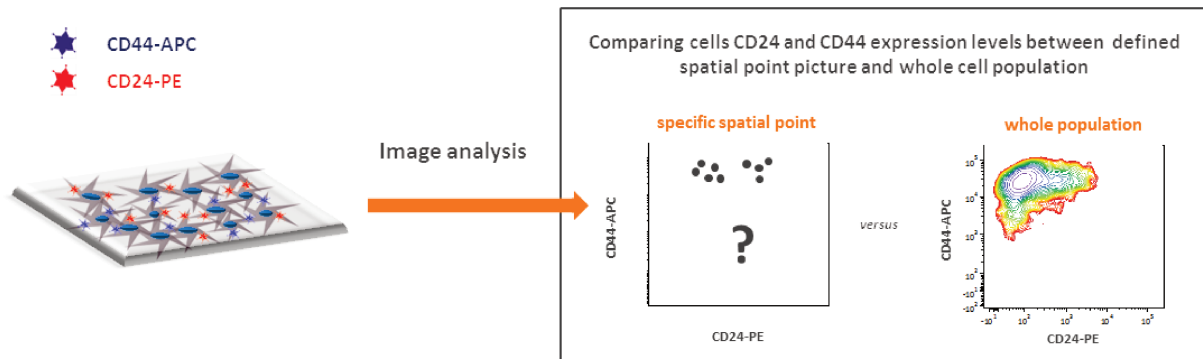


Figure 12.3 : *Analytical method implemented to demonstrate a bias with respect to CD24 and CD44 expression cells distribution between the whole cell population and geographically defined areas*

12.2.1 Method that assesses a local spatial organization within cells phenotypes

Regarding our work on the cell population of interest, each of the subphenotype CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- of both SUM149-PT and SUM159-PT cell lines was isolated as previously described in section 5.1.2 on page 51. Cells were plated on non-coated glass slides so as not to influence behaviors of phenotypic repopulation [122, 181, 344] and grown under standard conditions described in appendix D.2 on page 249. After 5 days of culture, two analyzes were performed for each condition (*cf.* fig. 12.4 on page 175) :

- Flow cytometry analysis
- Analysis by immuno-fluorescence microscopy of cells grown on slide

During the flow cytometry analysis, cells cultured on glass slide were trypsinized and stained for CD24 and CD44 markers according to the protocols described in appendix D.17 on page 274. For the analysis of the spatial organization by confocal microscopy, the cells grown on glass slides were stained for the same markers (CD24 and CD44) using the protocol described in appendix D.19 on page 280. And two to three samples per culture condition were carried out and 5 pictures per slide were randomly imaged by high resolution confocal microscopy (picture pixel size = 1 pixel/0.166 μm).

To illustrate our method of analysis, we will make a parallel between the distribution of CD24-/CD44+, CD24+/CD44+ et CD24-/CD44- subphenotypes of the population of interest and the geographical distribution of various ethnic groups in a large city like Paris [13]. Analyses from the census of the French population in 1999 helped to draw a map which corresponds to the installation of the people according to their ethnic origin (European, African, Asian and American). For example, it is interesting to note that African ethnic groups are particularly represented in the 12th, 13th, 18th, 19th and 20th districts and few represented in the 8th and 16th districts. Regarding Asian ethnic groups, they settled in very determined neighborhoods within 13th, 3rd and 10th districts (*cf.* fig. 12.5 on page 176). These implantation zones thus correspond to center of attraction or repulsion with respect to the settling of people based on their ethnic groups. From these maps, it seems possible to extrapolate the probability that a new inhabitant of Korean origin settles in the 6th district is lower than in the 13th district.

In our cell model, it is quite possible that the average phenotype in a given space influences the phenotypic expression of a cell, which would lead to a clustering of phenotypes (as shown in the map above). Therefore, we compared the dispersion of phenotypes expression within a rectangular ROI (340 μm x 340 μm) to the overall expression (*cf.* fig. 12.3 on page 173).

12.2.2 Preliminary results regarding the expression of the phenotypes depending on spatial localization

As explained previously, we compare the dispersion of phenotypes expressed locally (using confocal imaging technique) and globally (using flow cytometry method). These preliminary experiments were carried from three subphenotypes : CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- that were isolated and re-cultured for five days.

If we take a qualitative interest in the repopulation of phenotype CD24-/CD44-, we note that cell population re-expressed almost 80% of CD44+ cells at the same levels as those previously described in section 6.3 on page 92. However, locally we observe a clustering of expression phenotypes (*cf.* fig. 12.6 on page 179 subfig. C). So for all the cells of the studied field, the latter express little or even not at all the CD24 marker and expressed in only a few locations the CD44 marker. Thus, it seems that the CD24-/CD44- and CD24-/CD44med phenotypes are clustering without being represented by both of the CD24-/CD44+ and CD24+/CD44+ phenotypes (which are nonetheless well represented phenotypes in the general population).

Similar observations have been observed at the spatial organization level of the cells derived from the repopulation of CD24+/CD44+ and CD24-/CD44+ phenotypes (*cf.* fig. 12.6 on page 179 subfig. A and B).

In collaboration with Isabelle Bonnet (Curie Institute, UPMC and UMR168, Paris) and Romain Cendre (CIC-IT, Nancy) we worked on the development of a software to perform an automatic sorting/analysis of our datasets of confocal images (*cf.* fig. 12.7 on page 180).

Our set of data currently consists in 2D microscopy images of different fields of view of SUM149-PT and SUM159-PT cells for several stages and for different initial culture conditions as previously introduced. Each field of view is a stack of four images corresponding to the four confocal microscope channels : a pseudo-DIC channel and three fluorescence channels (DAPI for nucleus detection, Cy5 for CD44-APC and eGFP for CD24-PE).

The first stage of the software is purely based on the extraction of the signal contained in the images. It aims to detect all the cells present within the field of view and, for each cell detected, to quantify the expression level of the two proteins, based on the level of the corresponding fluorescence signals. Once this quantification step is completed, each cell will be represented by the level of expression of the two proteins of interest. Thus, the second stage will consist in analyzing these populations. The questions we want to address are :

- Can we automatically sort cells according to their phenotype (*i.e.* the triplet of fluorescence signal) ?
- Is there cluster of data (meaning cluster of cells expressing the same level of markers within the same area) ?
- And are there spatial correlations among the different cell populations ?

Here are the outlines of the procedure we used :

- **Detection of the cells.** We use both the DAPI and pseudo-DIC pictures. From the DAPI channel, we make a binary image by applying an automatic threshold. We then apply some morphological operations to clean the obtained binarized image that mainly consist in filling of the interior gaps and removing isolated pixels. At that step, only the nuclei that are too close, for example, those of daughter cells just after cytokinesis cannot be properly distinguished. From the DIC channel, we roughly locate each cell edges, by applying an entropic filter. We thus intersect these two informations to properly segment the cells. Briefly, we used marker-based watershed segmentation. It consists in marking a group of connected pixels (nuclei) inside each object (cell) to be segmented.

- **Fluorescence signal.** Once the cells are correctly segmented, we defined a list of pixels of interest that are those located along the cell edges. For all of these pixels, we extract the fluorescence signal for the three color channels.

The results obtained from images exhibiting isolated cells were particularly promising (*cf.* fig. 12.7 on page 180). However, the clear detection of cell membranes in a monolayer did not efficiently succeed. So we are still working on the methodology to clearly determine cells edges and various methods we developed are presented in appendix B on page 233.

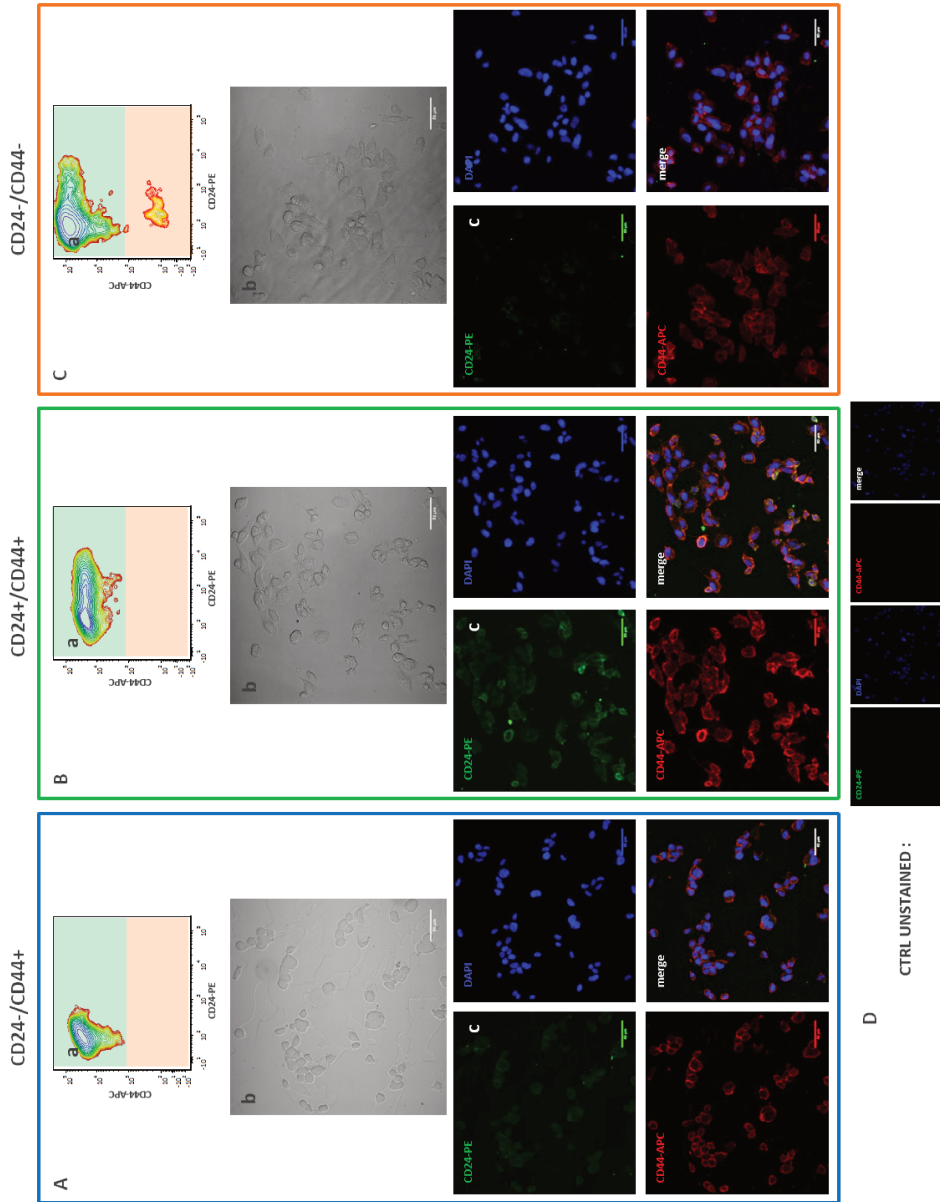


Figure 12.6 : Spatial organization assessment of repopulating sorted SUM159-PT subphenotypes after 5 days of culture.

A. Preliminary results of repopulating CD24-/CD44+ sorted cells. B. of repopulating CD24+/CD44+ sorted cells. C. of repopulating CD24-/CD44- sorted cells. Respectively (a) Flow cytometry cells analysis of CD24 and CD44 markers quantification of repopulating cells 6 days after sorting procedure, (b) Calibrated pseudo-DIC picture, (c) and calibrated fluorescent pictures for three markers : DAPI (in blue) CD24-PE (in green) and CD44-APC (in red) (picture pixel size= 1pixel/0.166 μm).

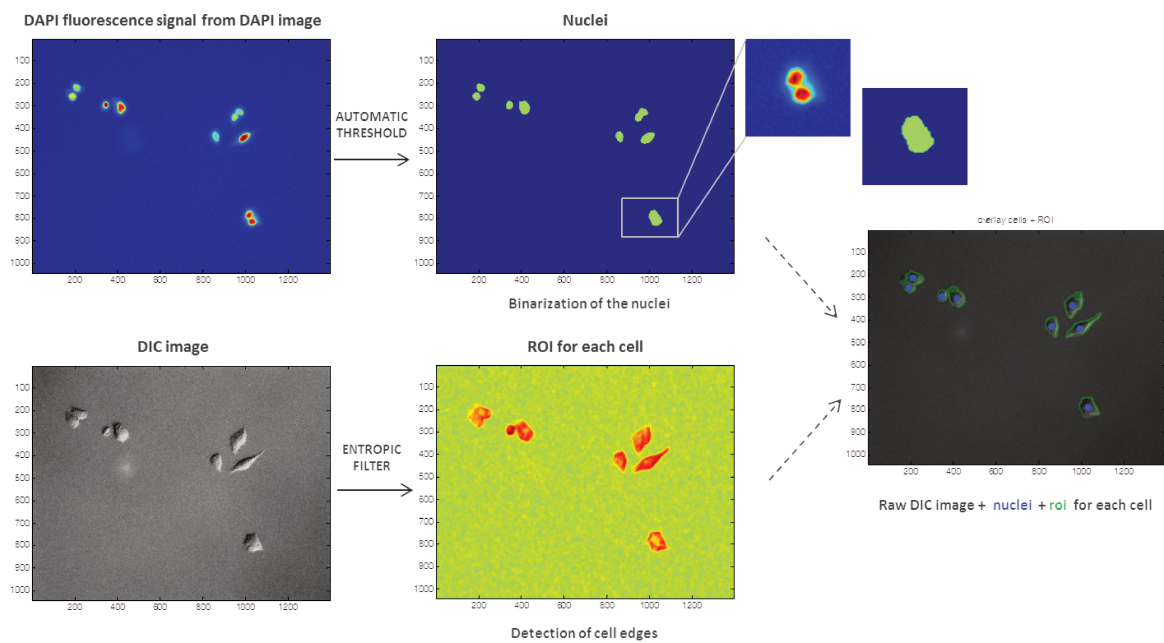


Figure 12.7 : *Confocal image analysis workflow developed by Isabelle Bonnet (UPMC, Institut Curie) and Romain Cendre (CIC-IT Nancy) to detect single cells edges.*

12.3 Conceptualizing and modeling of interactions among cells

- This work has been conducted under the close collaboration of Nicolas Grorod, Louise Magalhaes, Veronica Marek, Thomas Palomares, Konstantin Shmelkov and Thomas Vignon as part of their collective scientific project [PSC2013] -

Following these initial observations, we wished to conceptualize and modeling the overall operation of a cell population, the strength of the interactions and mechanisms of correlations between cells. So initially we developed a cellular automaton capable of modeling the interactions that are taking place between cells using mathematical and physical methods. Then the addition of some IT tools (mathematics and physics) have allowed us to study the scales correlation characteristics between cells as defined in section 12.1 on page 170.

This section is at the crossroad of experimentation and theory (*cf.* fig. 12.8) :

- The experiment allows us to obtain spatial distribution maps of the cells.
- The development of different observables from theoretical cell maps that allow us to follow the reverse way, that is to say, to search for specific characteristics of cell correlations of neighborhood interactions.

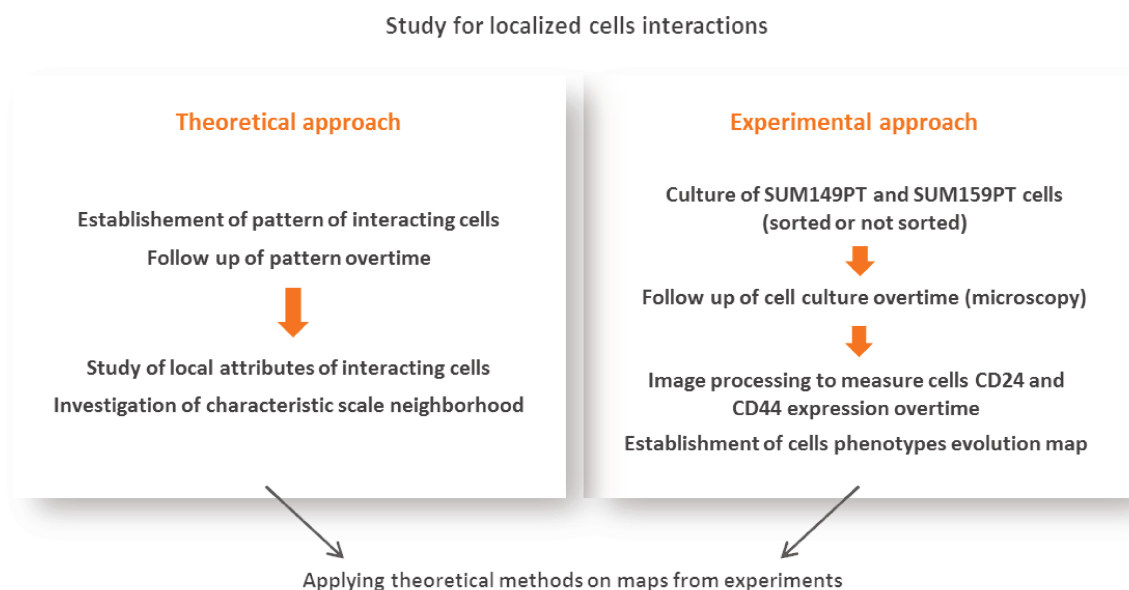


Figure 12.8 : *Interconnections between theoretical and experimental approaches to examine the spatial interactions among various cell phenotypes.*

12.3.1 The basic features of cellular automaton

In order to model the evolution of phenotypic state of the cells, we have established a cellular automaton. The basic characteristic is that the non-dividing cells are connected to each other by a local interaction, *i.e.* they change their phenotypic state depending on the phenotypic state of their surroundings. Some criteria based on biological observations during our experiments, were used to determine the characteristics of the cellular automaton :

- Balanced proportions of different cell types on average over the map, with the possibility to choose the value of this balance.
- Independence of this equilibrium with respect to initial conditions (Further characteristic is in contradiction with the descriptions made with respect to our research system (*cf.* chapter 6 on page 71), nevertheless this was initially stated to simplify the implementation of the automaton).
- Dynamic equilibrium meaning that the phenotype state of the cells is not fixed and therefore still evolving.

To ensure the feasibility of this model, we have also attached some biological limits :

- The internal processes of the cells, at the molecular level in particular, are inaccessible to modeling due to a complexity beyond the scope of the modeling. We strive to remove complexity to highlight the general characteristics of the interactions, the initial finding of the study being the return to equilibrium proportions across the population (*cf.* chapter 6 on page 71).
- Certain physical parameters are considered fixed such as the number of cells, the minimum distance between cells, the neighborhood radius of each cell and the location of cells in the area studied.

The cellular automaton is based on the following :

- Three types of cells corresponding in practice to the three cellular phenotypic states (CD24+/CD44+, CD24-/CD44+ and CD24-/CD44-) and respectively represented by three colors (blue, green and yellow) (*cf.* fig. 12.9 on the next page subfig. A).
- The infinite plane cell medium (compared to the scale of a cell) was modeled by a finite rectangle with periodic boundary conditions (the left and right edges and the top and bottom edges being connected).
- Time is discretized.
- Each cell has a set of probabilities to express the phenotype **blue**, **yellow**, **green** such $pB + pY + pG = 1$ and pB , pY and pG are between 0 and 1.
- Each cell has a neighborhood comprising any cell at a distance less than a fixed constant (*cf.* fig. 12.9 on the facing page subfig. C).
- At $t = 0$, a global state of the system is selected.
- At every moment, the phenotypic state of each cell evolves following a transition function taking as parameters :
 - The probabilities of the cell phenotypes
 - The proportions of each phenotype in the vicinity of the cell

and a phenotype state is then drawn randomly according to that law of probability of phenotype evolution.

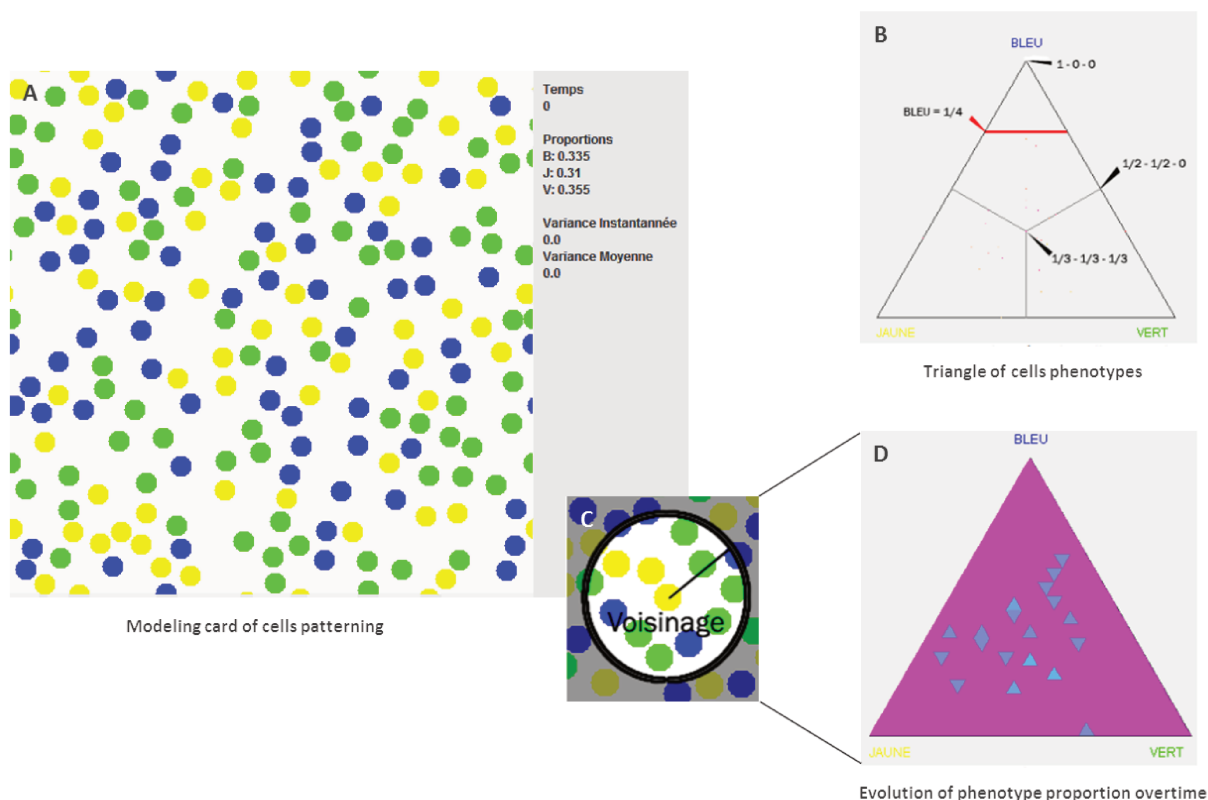


Figure 12.9 : Automaton modeling the behavior of 3 cells phenotypes and cells decision-taking as an interacting system.

A. Cell automaton which simulates changes of interacting cells and their types of differentiation through time. B. Triangle of phenotypes space used to plot the average phenotype within the ROI. C. Example of a selected ROI to monitor overtime changes of cells phenotypes and to follow spatial patterns. D. Triangle of space phenotypes to follow overtime the evolution of phenotypic ratios calculated within selected ROI

- This informatics tool was developed by a group of youths Ecole Polytechnique students (Nicolas Grorod, Louise Magalhaes, Veronica Marek, Thomas Palomares, Konstantin Shmelkov and Thomas Vignon) as part of their 2nd year collective scientific project -

And each map were then cut into macro-cells of different scales (and therefore having a certain number of cells) as previously developed in section 12.1 on page 170. Within each macro-cell (ROI), the average phenotype, computed from the ratio of each of the three phenotypes (blue, green and yellow) was then determined.

Furthermore, some observables have been set up to locally and globally monitor (ROI) the influence of the laws of phenotypic transitions on phenotypes proportions.

- Display of each phenotype proportion (overall and per each ROI)
- Representation of the evolution of the average proportions calculated for each ROI within the space of proportions

12.3.2 Features and simulation of cell states transitions

The important point at this stage was to work on the functions of phenotypic evolution over time, *i.e.* to establish a model on how likely a cell with a phenotype (P1) at t would evolve to another phenotype (P2) at $t + 1$ so that :

$$\begin{pmatrix} p_B \\ p_Y \\ p_G \end{pmatrix} (t+1) = f \left(\begin{pmatrix} n_B \\ n_Y \\ n_G \end{pmatrix} \right), \epsilon \begin{pmatrix} p_B \\ p_Y \\ p_G \end{pmatrix} (t)$$

With p_B , p_Y and p_G , the transition probabilities to the phenotypes “blue”, “yellow” or “green”; and n_B , n_Y and n_G the number of cells “blue”, “green” or “yellow” present in the vicinity of the cell (ROI).

Two methods have been tested.

The probability matrix

This first model was inspired by the neuronal interactions modeling system described by Yaner Bar-Yam and named “the Hopfield networks” [21]. As part of this neural system, each neuron has an orientation (1 or -1) and interacts with all the other neurons of the system. The evolution of the neural network as a whole (at time “ t ”) is thus modeled by linear laws depending only on the state of the system at time “ $t - 1$ ” and where the phenotypic state of each cell influences the phenotypic state of all the other cells of the system.

Under this model, we extended this function to change to a 2D system, for which the cells can to express 3 different phenotypic states.

Thus, the law of evolution as part of our biological system can be expressed as :

$$\begin{pmatrix} p_B \\ p_Y \\ p_G \end{pmatrix} (t+1) = \begin{pmatrix} p_B \\ p_Y \\ p_G \end{pmatrix} (t) + \begin{pmatrix} n_B \\ n_Y \\ n_G \end{pmatrix} \begin{pmatrix} a, b, c \\ d, e, f \\ g, h, j \end{pmatrix}$$

In the context of this function of evolution, the essential parameters are the coefficients of the transition matrix (“a” to “j”).

First and to simplify the model, we considered the matrix of phenotypic transition to be symmetric. This means that the influence of phenotypic state “blue” on “yellow”, for example, is similar to that exerted by the phenotype “yellow” on “blue”. In other words, in this simplified model, no phenotype exerts a greater influence than another phenotype in conducting phenotypic transitions, whose aim is to re-balance the local average phenotype expressed. This approximation will certainly not match with certain biological phenomena that were previously described (*cf.* chapter 6 on page 71). Indeed, specific cell phenotypes dramatically influence the behavior of the global cell population and certainly of the adjacent cells and consequently of the local behavior of the cellular tissue. A striking example to illustrate this observation corresponds to the healing process and how leader cells play a key role in the initiation of cell relocation process of the other surrounding cells [83, 204, 284].

The purpose of this methodology is to obtain at steady state a global system whose proportions remain stable over time, without freezing completely. By choosing high coefficient values, we obtained too great a divergence in the proportions of the system. Thus, in order to maintain a certain inertia in the system, that is to say that the terms :

$$\begin{pmatrix} p_B \\ p_Y \\ p_G \end{pmatrix} (t)$$

will not negligible compared with the transition matrix for the calculation of :

$$\begin{pmatrix} p_B \\ p_Y \\ p_G \end{pmatrix} (t + 1)$$

we decided to take the values of these coefficients in the magnitude of a hundredth. It was then possible to obtain systems leading to an equilibrium of overall proportions.

Finally, we wanted to characterize the effect of the coefficients of the matrix of phenotype transition on the overall equilibrium of the system. In other words, is it possible to change the balance of phenotype proportions of the overall system by changing the coefficients of the transition matrix ?

By changing the diagonal coefficients “ a, e, j ”, the overall equilibrium is deeply changing while keeping a state of equilibrium of phenotypic proportions. This provides an equilibrium, a balance for a particular phenotypic state (**blue** for example) within the system.

By modifying the non-diagonal coefficients “ b, c, f ”, the overall equilibrium can be adjusted, but above all the system can be stabilized or boosted locally.

Despite the many advantages of the matrix, this model is difficult to interpret biologically. Indeed, the coefficients of phenotypic transitions do not have their own physical reality, you cannot deduct their effect on the system by trial and error through simulations. This major limitation has led us to develop another law of evolution, in accordance with the statistical physics theories, and therefore more easily interpretable and adjustable.

Potential energy and statistical physics

The energy approach is based on statistical physics tools. Indeed, in so far as there is a balance of the proportions of the medium while maintaining a dynamic evolution, it is appropriate to associate with a cell energy depending on their proximity and their status. The probability of getting the partner state is then provided by the Boltzmann weights, and we can define overall balance from local interactions. The probability of obtaining a energy state \mathcal{E} is :

$$p(\mathcal{E}) = \frac{\exp(-\alpha\mathcal{E})}{\mathcal{Z}}$$

Thus, α is a constant that we can evolve to alter the equilibrium state of the system at the local level and,

$$\mathcal{E} = \sum i \in_{(B,G,Y)} p_i.$$

We considered that the cell in order to decide its future phenotypic state, takes into account an average medium information. This means that the changing probabilities of expression of a phenotype depends on the proportions of the phenotypes (B, G, and Y) in its vicinity (ROI) and not on established interactions with neighboring cells. For each cells of the studied ROI (characterized by its neighborhood, meaning the proportion of the several phenotypes), an amount of energy is defined to describe the desired states (repulsion or attraction for the expression of a phenotype).

To each energy is associated a behavioral effect of cell medium that we wish to observe on the map of proportions. Potentials used are one-dimensional, *i.e.* an electric potential $\mathcal{V}(x_B, x_G, x_Y)$, and can be written as :

$$\mathcal{V} = \sum i \in_{(B,G,Y)} \mathcal{V}(x_i).$$

To calculate a probability associated with the phenotype of (i), we will compute the energy while considering that the neighborhood contains an additional neighbor phenotype (i) :

$$p_B(n_B, n_G, n_Y) = \frac{\exp(-\alpha \mathcal{E})\left(\frac{n_{B+1}}{N}, \frac{n_G}{N}, \frac{n_Y}{N}\right)}{\mathcal{E}}$$

$$p_G(n_B, n_G, n_Y) = \frac{\exp(-\alpha \mathcal{E})\left(\frac{n_B}{N}, \frac{n_{G+1}}{N}, \frac{n_Y}{N}\right)}{\mathcal{E}}$$

$$p_Y(n_B, n_G, n_Y) = \frac{\exp(-\alpha \mathcal{E})\left(\frac{n_B}{N}, \frac{n_G}{N}, \frac{n_{Y+1}}{N}\right)}{\mathcal{E}}$$

where $N = n_B + n_G + n_Y + 1$

In addition, a potential to limit the medium to standardize is added to each phenotype. Thus, each phenotype will tend to converge to a point of equilibrium while maintaining the phenotypic diversity.

Secondly, we have added in each transition of a phenotype memory expressed by introducing a potential that tends to slow change at time of transition (t) to ($t + 1$)

Finally, we tried to translate the phenomenon by which a cell, whatever the phenotype it expresses, is more likely to express a phenotype at the time (t) if probability at ($t - 1$) is also important. This in order to take account of continuity in the development of probability, and translate somehow a form of pseudo-determinism in the behavior of a cell.

As indicated above, α parameter must be chosen carefully as it determines the influence of the energy system and thus on the transition probabilities. α must remain low enough to keep both the dynamic nature of the development and the independence with regard to the initial

conditions, but high enough to ensure convergence of phenotypic proportions across the map toward a stable equilibrium.

The choice of α depends on both the size of the neighborhood and the energy used. On the other hand, the probability of expression of a phenotype is done by calculating the energy of the cell so that it adopts this phenotype. If the cell takes into account 100 neighbors, its phenotype change will be less significant in this population if it considers that a neighborhood of 10 cells. Thus, the model's ability to meet the objectives of independence of initial conditions and convergence to a dynamic equilibrium depends on the neighborhood and by extension of α .

Two limitations to this model still remain to be addressed :

- The validation of the model is subject to more parameters and is therefore less robust than the matrix model.
- Observables introduced earlier are less efficient and require mastering both parameters and theory to analyze the outcomes of the experiment maps.

12.3.3 Looking for spatial correlation, influence of cells coupling and referential

The objective is to be able to trace back to the key parameters of inter-cellular interactions of a same map, such as the characteristic vicinity scale. From our simulations, which by their construction have a characteristic scale of cellular interactions (δ), we would validate our methodology by finding the critical distance of integrated influence (δ).

In order to obtain local information, simulations map was divided into different sub-ROI, where we followed the local evolution of the phenotype proportions. We sought through this methodology to highlight the influence of local interactions on the overall behavior. Thus, we have defined a variance corresponding to the dispersion of macro-cells proportions around the equilibrium point of global proportions.

Thus, at the scale of an ROI, the first records of variances, by changing the scale of analysis, allowed us to demonstrate a change in variance when the scale of analysis crossed the characteristic scale of neighborhood interaction (*cf.* fig. 12.10 on the next page).

$$\mathcal{V} = \frac{1}{|\mathcal{M}|} \sum_{\mathcal{M} \in \mathcal{C}} (x_B^{\mathcal{M}} - x_B)^2 + (x_Y^{\mathcal{M}} - x_Y)^2 + (x_G^{\mathcal{M}} - x_G)^2$$

where (x_B, x_Y, x_G) are the average proportions in \mathcal{C} and $(x_B^{\mathcal{M}}, x_Y^{\mathcal{M}}, x_G^{\mathcal{M}})$ are the average proportion in \mathcal{M} .

The second step was to compare the variances (dispersion of ROI proportions around the equilibrium point of the overall population) from numerous ROI.

An analogy with the behavior of an element which has two separate state (-1 to +1) would allow us to better understand the study of variances that we want to implement in 2D with cells that may be three distinct phenotypic states. So in the next simulation (*cf.* fig. 12.11 on page 189) we present the phenotype expressed over time and depending on the strength of the coupling (*i.e.* influence a phenotype expressed at instant “ t ” exerted on the expression of the phenotype at instant “ $t + 1$ ”). Thus, the dispersion of phenotypic state around its mean value is very low when considered time scale is short and grow for long time scale. Finally, if the phenotypic state is independent of the previous ones, we would observe an increasing variance in proportion to the number of state considered.

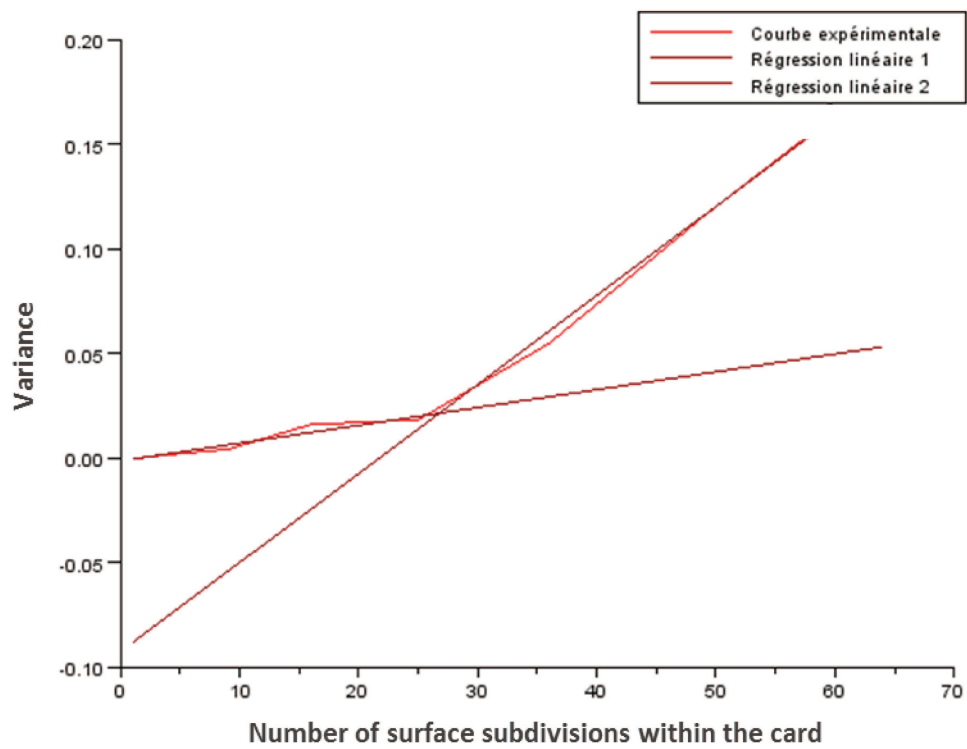


Figure 12.10 : Study of the change of variance regime depending on the scale of analysis.

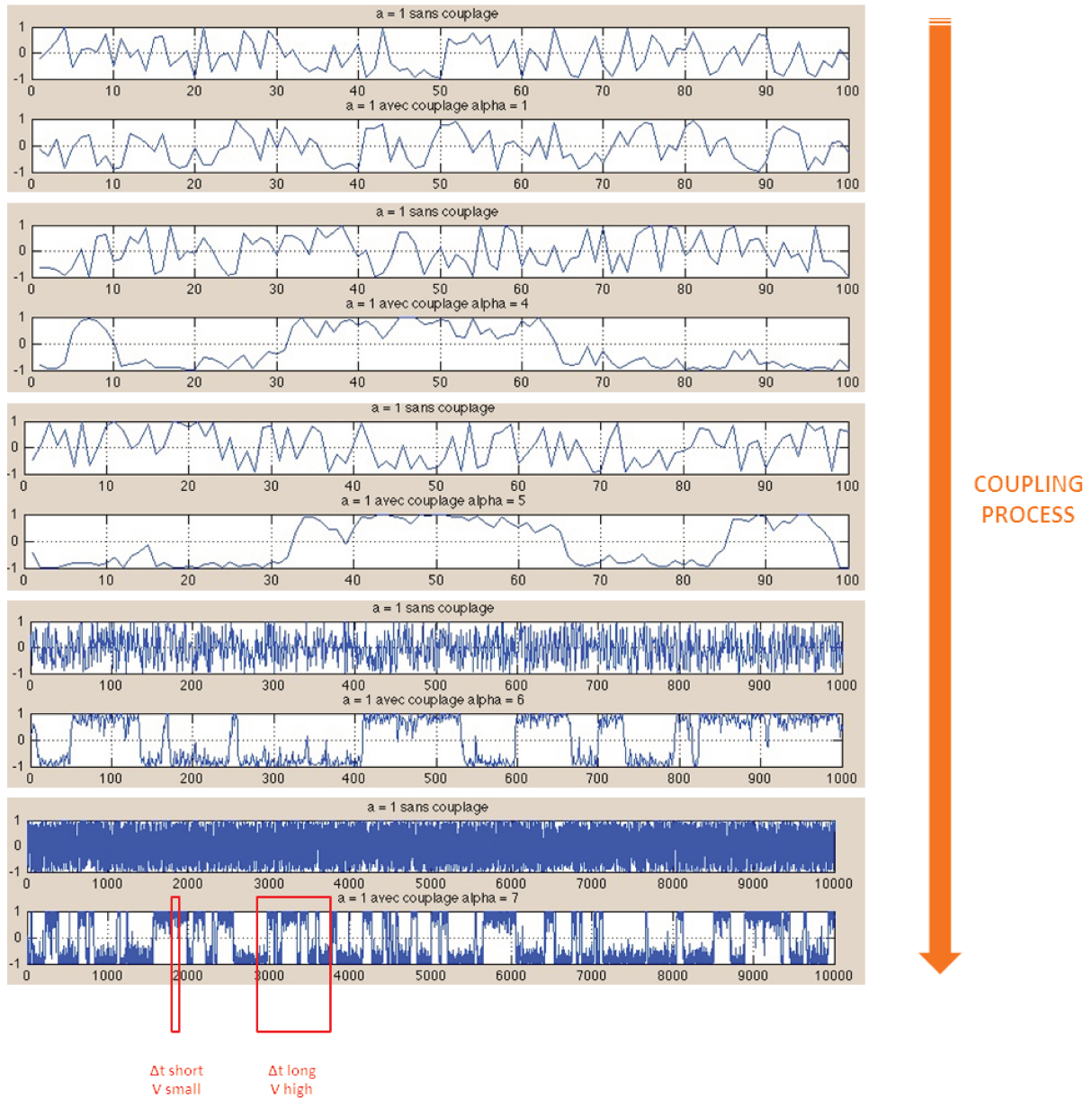


Figure 12.11 : Simulations monitoring the expression pattern of two phenotypes (0 and 1) which phenotypic expression may or may not be coupled to previously expressed phenotypes.

As for the 2D system described above in section 12.3.1 on page 181), studies were conducted an area of near 200 cells with corresponding displacement and made from maps whose potential attractiveness (\mathcal{G}) could be more or less important (*cf.* fig. 12.12). And the analyzes carried out highlighted a peak of variance when the vicinity of analysis is the same as the vicinity of simulation.

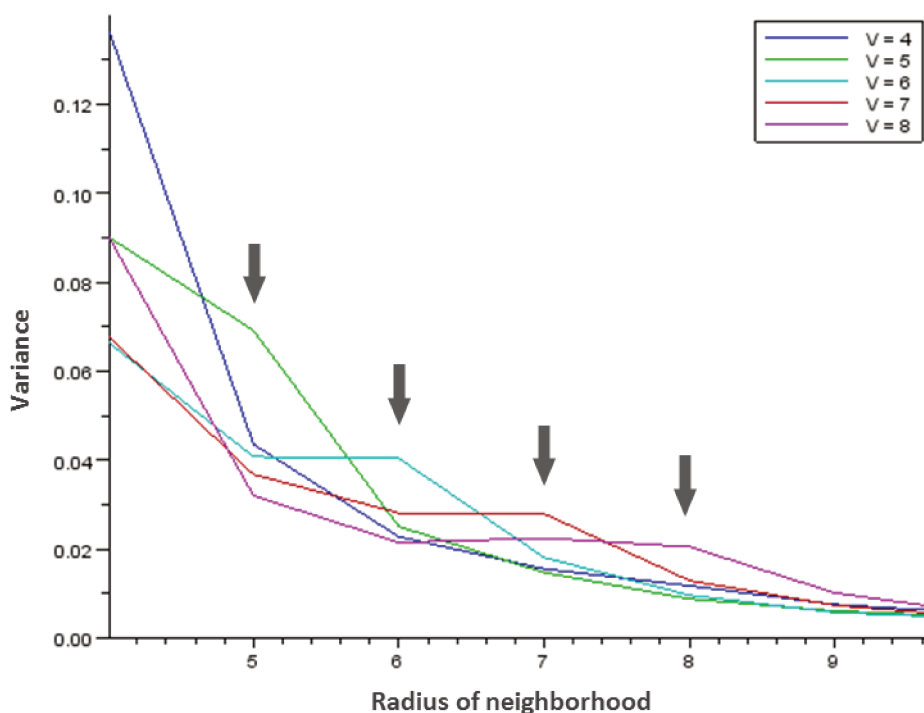


Figure 12.12 : *Simulations monitoring the expression pattern of two phenotypes (0 and 1) which phenotypic expression may or may not be coupled to previously expressed phenotypes.*

This first approach of simulation and research of distance of cells influence, even if promising, needs to be more fully characterized. For example, it would best to determine whether the method of comparison of variances is sensitive to the analysis step size and the number of characteristic cells of ROI. On the other hand, the presented modeling methodologies took into account a number of approximations contradictory to the state of the biological system studied (stable cell number over time and non-displacement of the latter, independently of the balance overall overlooked the initial conditions of the system ...).

The possibility of applying the previous method on small scale (hundred cells) and on fixed maps (no temporal evolution) remains a considerable advantage for experiments. However, the maps obtained to date from the experiment did not reach that level because they represent at the most fifty cells. We therefore did not have the opportunity to compare our theoretical results on some real maps. However, this theoretical analysis project should continue in parallel with the image analysis obtained previously (*cf.* fig. 12.6 on page 179).

This chapter closes our first investigations related to the search for a spatial organization of cell phenotypes in relation to each other. And by extension we seek to measure the degree of independence or cell dependency for transition of a phenotype (P1) to another phenotype (P2) depending on the phenotypic state of the cells present around the cell of interest. While

searching for the degree of correlation between the expression of a phenotype and that of the surrounding cells, we found that the behavior of the dispersion of macro-cells proportions around the equilibrium point of global proportions changes depending on the degree of dependence of the cells. Our spatial correlation search strategy is thus based on a comparison of the balance of local and phenotype proportions that obtained on a global scale.

Initially, we wanted to confirm a difference between the expression of phenotypes (the dispersion of expression of the markers CD24 and CD44) locally and globally expressed. Our preliminary results (obtained by flow cytometry and confocal imaging) gave an indication of the clustering representativeness of the local expression differences which are particularly noteworthy for the cells derived from the repopulation of CD24-/CD44- phenotype.

To determine the critical distance of influence of the different cellular phenotypes on neighboring cells and extrapolate social rules governing the phenotype behavior of cells locally and globally, we created a cellular automaton able to change cell phenotypes overtime and depending on defined functions. Thus, it is possible to implement the interactions and roles to vary degrees of interactions between cells according to the show and the size of the neighborhood phenotypes. At equilibrium, obtaining 2D map representing the spatial organization of the cells against each other allowed the establishment of a methodology point research critical distance influence. Based on change variance behavior of local phenotype proportions with respect to those generally expressed. These preliminary works, both theoretical and experimental, are the basis of reflection for the continuation of this work in the laboratory.

General conclusion

“Il faut toujours se réserver le droit de rire le lendemain de ses idées de la veille.”
- Napoléon Bonaparte -

Dear Reader,

These last pages have presented, one thing leading to another, several experiences, methodologies of analysis and statistics, and advanced some new concepts. Thus, they gradually have built a promising base for future research studies. In parallel, we have taken the opportunity to somewhat frolic within the reflection meanders of that current project. However, time has come to conclude ! And these final words will bring little updates with respect to :

- What this PhD journey brought about cells plasticity and overtime dynamics of phenotypes expression within SUM149-PT and SUM159-PT cell populations.
- The future projects which are supported by these latest investigations.

And it will be the noteworthy points we will tackle throughout the next three sections.

Main contributions of this PhD thesis

To summarize the main points that were developed in this work I would start by recalling that this PhD took its essence through the questioning of cell dependence to its neighboring cells in the context of cells plasticity.

Characterization of the heterogeneity of SUM149-PT and SUM159-PT cell lines regarding CD24 and CD44 markers

The first goal of this thesis has been to characterize both qualitatively and quantitatively the level of heterogeneity within two breast cancer cell lines respectively named SUM149-PT and SUM159-PT. CD24 and CD44 markers, largely used over literature to describe cells differentiation processes and phenotypes characteristics within breast cells, have played a leading role to track cells subphenotypes. Among cell lines, three cells subpopulations have been distinguished on the basis of the level of expression of these two surface markers, as follows : CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ (*cf.* fig. 5.1 on page 50). CD24-/CD44+ and CD24+/CD44+ subpopulations have been previously portrayed and studied by colleagues. Nonetheless, CD24-/CD44- phenotype had, as far as we know, not yet been disclosed in breast cancer. So this phenotype state is made conspicuous from the other phenotype for its small size and low granularity ; these latest

physical qualifying terms could corroborate with cellular fragments. Based on flow cytometry data, CD24-/CD44- subpopulation has been isolated and put back into standard culture condition. Their ability to adhere in less than one day to the culture dish, and their capacity to divide and proliferate have validated our hypothesis that such small representative population (0.5 to 1% of the overall cell line) matched with a new cell sub-type (*cf.* fig. 5.2 on page 52).

It is interesting to note that disclosed phenotype and most preferably the CD24-/CD44- subpopulation have previously been identified within normal breast cells. In that context, this specific population is not caused to any culture artifacts and does fit for specified (*i.e.* in the context of this study, in cancer) cell populations [71, 277, 301]. It has been argued that there is a strong functional and behavioral link between the phenotypes described in both a healthy and a cancerous context. This supports that we should intensely combine these two environmental contexts at the behavioral, organizational and functional levels, without forgetting to mention markers. This would help both in identifying cell subpopulations and in understanding their role within the organization of a tissue or an organ.

Initially, we monitored the overall behavior of the cell populations SUM149-PT and SUM159-PT. The level of expression of CD24 and CD44 markers was measured for up to 130 days through flow cytometry. Each of the cell lines exhibit the three phenotypes listed above although subpopulations proportions are not the same depending on the cell lines (*cf.* section 5.3 on page 55). Thus, the SUM149-PT line is particularly rich in subpopulation CD24+/CD44+ while SUM159-PT line is in turn largely represented by the subpopulation CD24-/CD44+ (and account respectively around 80–85% of total population). The expression profile of the cell population based on CD24 and CD44 markers change over time. However, this study allowed us to observe that the profile fluctuated over time at the following levels : phenotypic proportion, mean and dispersion of expression markers, according to a specific temporality and tendency to recur at regular intervals for each cell lines as described in section 5.3.2 on page 57 and in section 5.3.3 on page 63.

Following this monitoring, we determined the behavior of three extreme subphenotypes of both SUM149-PT and SUM159-PT cell lines (CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+). For that purpose, we purified each of them from the total cell populations to perform a long-term monitoring of their behavior.

Phenotypes plasticity and recovery of heterogeneity from homogeneous populations

The previous study highlighted on the scale of entire cell lines the cell plasticity that exists within cell populations. Sorted subphenotypes showed from 3 days of culture onward, that each homogeneous cell subpopulations re-express phenotypic diversity. And this in such a manner that each of the studied phenotype (CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+) is again represented.

Interestingly, none of the subphenotype behave in the same way in the process of repopulation. And more specifically, the dynamic of equilibrium achieved through the process of plasticity is characteristic of each of the initial subphenotype (*cf.* section 6.2 on page 76 and section 6.3 on page 92). Therefore, the balance of phenotypic expression achieved by the repopulation is limited by the attractive potential of the initial phenotype from which it originated. For example, cell population derived from CD24-/CD44+ phenotype remains attracted by the potential well of the initial CD24-/CD44+ phenotype. As a result, the average expression of markers of repopulation approximates the average CD24-/CD44+ phenotype. Moreover, it should be noted that only the phenotype CD24-/CD44- (especially the subpopulation derived from SUM159-PT cell line) is capable of reforming a population whose marker expression profile CD24 and CD44 is similar to that characteristic of the non-sorted initial population.

Previous studies have shown that the equilibrium state of the repopulation, *i.e.* the recovery of a dynamic heterogeneity, is dependent on the original phenotype. In addition, we then exposed that the equilibrium state of the population “ $P = 0$ ” is influenced by the ancestral phenotypes “ $P - 2$ ” and “ $P - 1$ ” (*cf.* section 7.1 on page 103). So, there is an imprint of the phenotype origin (of previous generations) on the dynamics of the phenotypic expression of the repopulations.

These studies have not only raised the question of how the equilibrium of the markers in cell populations is subjected to the plasticity process. However, they have also informed us that the three identified subpopulations behave distinctly in the process of repopulation. Thus, only the CD24-/CD44- phenotype is plastic so that all derivatives cell repopulation are close to the unsorted population (regarding the dynamics of expression of the studied markers). Does the CD24-/CD44- subpopulation keep the phenotypic equilibrium (homeostasis of the markers expression) of cell populations? Or even does this subpopulation preserve or trigger the social relations established by and/or between cellular phenotypes?

Singular tumorigenic abilities of cell lines subphenotypes

Previous experiments have been conducted in 2D and we had a look on cells behavior in 3D. In this new environment, we monitored on one hand, the ability of tumorigenic cell populations and subpopulations CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+. On another hand, we assessed the plastic capacity of the three identified phenotypes. That is why, the number of tumorspheres and their sizes were measured while the cells phenotypes of tumorspheres were determined by flow cytometry according to the CD24 and CD44 markers.

After these experiments, only the cells from the cell line SUM159-PT showed tumorigenic capacity, *i.e.* the ability to develop dense tumorspheres with diameter equal to or greater than 80 μm for one single cell, while cells derived from the SUM149-PT cell line formed shapeless mass of highly refractive cells. In more detail, the three cell subpopulations (CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+) cultured in 3D conditions do not behave the same way. Thus, the phenotype CD24-/CD44- of both cell lines, although resistant to the anoikis environmental contextual pressure, is unable to form tumorspheres and has very limited capacity multiplications. CD24-/CD44+ and CD24+/CD44+ subpopulations, in turn behave like the original cell line. Nevertheless, some differences with respect to the morphology of tumorspheres were identified. The tumorspheres derived from CD24-/CD44+ are clearly bounded by a peripheral corolla of cells compared with those from CD24+/CD44+ population.

The dynamics of expression of cell phenotypes within the spheres tumor vary significantly from cell populations cultured in 2D. The difference is keen for phenotypes proportions. For example, the CD24-/CD44- subpopulation is enriched from 20 and 50x in 3D cultures compared to 2D conditions (*cf.* section 9.1 on page 131). However, the plastic potential of CD24-/CD44+ and CD24+/CD44+ cellular phenotypes proved to be just as developed as those presented in 2D condition (*cf.* section 9.3 on page 137). We noted that the repopulation derived from both CD24-/CD44+ and CD24+/CD44+ cells phenotype exhibited substantially the same dynamic expression of CD24 and CD44 markers in equilibrium and this whatever the stage of the analysis.

Morphological and phenotype characteristics are not the only ones to be influenced by cell lines and the original phenotype. Indeed, the concentration of single individualized and initially plated cells, influences the proliferating capacity of cells within tumorspheres (*cf.* section 8.2.2 on page 126 and section 9.2 on page 135).

Monitoring cell-state transition thanks to the lentiviral “PhenotypeTracker” platform

The transitions between two cellular phenotypic states require time tracking to define key characteristics of these events :

- The transition speed based on the initial and final states
- The coordination of expression of CD24 and CD44 markers across the genealogical history of the cells
- Or the spatial location of the phenotypes with respect to each other

Thus, we considered building a genetic platform to simultaneously track the expression of CD24 and CD44 phenotypic markers. Thus, follow up of the phenotypes transitions relies on the monitoring of the expression level of short half-life (up to 3 to 6 hours) fluorescent proteins triggered by endogenous promoters (*cf.* chapter 10 on page 145).

The PhenotypeTracker is composed of a third unit to both validate transduction efficiency and normalize the number of integrated constructs (from one cell to another). The choice of both transgene and promoter for this third unit was then used both to understand the sensitivity of the promoters at the cellular context and to develop a methodology for quantifying the number of genomic integrations (*cf.* section 11.1 on page 153).

In parallel, the validation of specific units monitoring the different markers was confronted with some limits, linked to the expression of endogenous sequence. Indeed, transfection and transduction studies of the PhenotypeTracker platform reporting CD44 expression, revealed that phCD44 sequence (previously isolated by colleagues [104]) was not able to drive signals of eGFP reporter protein. Similar works are ongoing to validate the promoter sequence of the human CD24 surface marker (*cf.* section 11.4 on page 165).

Thereafter, a peculiar attention was given to the differential and effective transduction of cell phenotypes. Indeed, the validation of positively transduced cells (via a lentiviral platform carrying a VSV-g tropism) have highlighted that only one quarter of CD24+/CD44+ subpopulation was actively expressing the carried transgene (*cf.* section 11.3 on page 161). While other cells, from previously studied subpopulations, do not expressed the transgene, yet driven through a promoter with ubiquitous features. These comments have questioned both the possible bias within the infectious potentialities of the different subpopulations and the differential chromatin environment of cells subphenotypes inducing transgene dis-regulation.

Methodology to measure the level of reliance regarding the neighborhood

Our final objective consisted in measuring the degree of cell-to-cell dependence or independence. The issued assumption is that the expressed phenotype of each cell at time “*t*” is dependent on the phenotypic states of the surrounding cells, and consequently of the averaged phenotype of the cellular environment. So the last part of this thesis helped clear brush the subject through a set of preliminary practices and theoretical work. We presented a statistical method to assess a local spatial organization within cells phenotype. This later, based on studying the behavior of the variance in average phenotype of the studied areas which vary in size (*cf.* section 12.1 on page 170). This study is at the crossroads of both experimental and theoretical studies.

A first prototypic approach was performed to comprehend the existence of spatial phenotype patterns. We compared the dynamics of phenotypic expression within cell repopulations (*i.e.* map of expression of CD24 and CD44 markers) at two different scales : the global scale (by flow cytometry) and locally (by fluorescent imaging). Qualitative observations have highlighted differences of the

pattern of phenotype expression between these two scales. Indeed, clustering of cells sharing the same level of expression characterize the spatial organization of the cells relative to each other (*cf.* section 12.2 on page 173). Moreover, some phenotypes were differentially expressed (and even sometimes not depicted). These early findings led us to precisely study the phenotype locally expressed. To achieve this, we have worked on the development of a method to automatically analyze 2D microscopic images at the single cell level (both segmentation of cells edges or nuclei and measurement of CD24/CD44 level of expression) and without losing information regarding the spatial organization of each cells (*cf.* appendix B on page 233). Nuclei segmentation, based on the level of expression of DAPI marker (fluorescent stain that highly fix adenin- and thymine-rich regions of the DNA), specifically individualize cell nuclei to determine the number of cells per field. Moreover, several methods to segment cellular membranes have been implemented. Unfortunately, this work did not lead to fine analysis of DIC images in order to measure the level of expression of CD24 and CD44 markers per each cells.

Moreover, we aimed to determine whether the phenotypic expression of the cells is dependent upon the cell environment, which each cell is characterized by a specific phenotypic state. This entails a detailed analysis of cell interactions at the single-cell level, between small groups of cells and at the scale of the total population. More specifically we wish to describe and determine the social rules underlying cell phenotype expression at the single cell level. To address this challenge, this required to compare the fate of cells in culture with the one whose behavior would have been determined by pre-established and modeled social rules. Thereby in parallel with the development of the work described above, we have initiated the developing of a program to model the behavior of collaborative cells with the help of six students from the Ecole Polytechnique (*cf.* section 12.3 on page 181). This first cellular automaton models changes of the phenotype of non-dividing blue, green and yellow cells (this major limitation has been initially implemented both to facilitate the development of the tool and to abstain from both space constraint and cell movement) which depends on the degree of interaction within cells. Two approaches have been developed and tested to model the functions of the phenotypic transitions (*cf.* section 12.3.2 on page 183) :

- The probability matrix
- The potential energy

In addition, some statistical tools have been designed and tested on the theoretical maps obtained from the model portrayed above.

At the end of this crossing through my thesis we will turn our thoughts with a pragmatic manner on the contribution of these data. Then we will focus on the understanding of tumors life with a peculiar gaze fixed to the clinic and therapeutics.

The first afterthought relates to cell plasticity. According to the outline presented in chapter 6 on page 71 the three highlighted cell phenotypes are able to switch from a state to another over the population lifespan (*cf.* fig. 6.1 on page 72). We determined that those transitions occur over a very short period. Indeed, in less than 3 days, the CD24-/CD44- subpopulation re-expressed the positive state of markers CD24 and CD44 (*cf.* fig. 6.15 on page 93). This trans-phenotypic plasticity has major implications for the envisioned therapeutics to limit tumor growth or even reduce it. Indeed, the main therapeutics considered today target the cancer stem cells. Scientific literature describes those cells as the nodal point of tumor growth [3, 145, 276]. However, previously published results and the ones observed in this thesis show that differentiated cells are able to

de-differentiate themselves and revert back to stem cell state [48, 89, 113]. Therefore, targeting and killing cancerous stem cells will not give permanent improvements, the treatment will have to continue for an indefinite time. The system will never cease to revert back to an equilibrium—certainly different from the initial state regarding the expression of cells markers and the cells behaviors—however it seems likely that all phenotypes will be re-expressed over time, including the stem cells phenotype.

I invite you to continue this ultimate discussion through which we will reach the end of our voyage whose theme will be : the research project we are currently envisaging for both the near future and the more or less distant expectations.

Research topics that we are currently exploring as a result of this thesis work

The PhD thesis that you just explored is the beginning of a greater research work. So in the wake of this work, we have right now implemented a certain number of research axis and we will use this subsequent section to describe the efforts we are currently implementing. Our first goal is to question the balance of a cell population with respect to the expression of its phenotype in the context of cells plasticity. Thus, two research topics are now favored :

- Determining the genetic profile of expression that cells were taken to return to population's steady state through whole transcriptomic analysis performed at different stages of cells repopulation.
- Modelizing the process of re-equilibrium according to the initial cell phenotypes.

Finally, the second project that we are following aims to study spatial organization of single-cells phenotypes in real time. Thus, the cornerstone for this project is to specifically target markers promoter within the context of cells genome thanks to a fluorescent-reporter genomic platform in order to monitor in real-time changes in markers expression. To complete, we are currently establishing an image analysis methodology both to effectively delineate cells boundaries and to correctly affect CD24 and CD44 expression to each single cells. This will help to determine cells phenotype and their change overtime.

Whole transcriptomic analysis of sorted CD24/CD44 subpopulations and overtime phenotypic re-equilibrium of cell populations

Within this work we have characterized two breast cancer cell lines (respectively named SUM149-PT and SUM159-PT) according to two surface markers : CD24 and CD44. Each of those cell lines is characterized by its typical pattern of expression within the space of CD24 and CD44 markers. Thus, three cell subpopulations have been distinguished on their level of expression of the above-mentioned markers and could be isolated : CD24-/CD44-, CD24-/CD44+ and CD24+/CD44+ populations. Sorted and placed back into culture, these three subpopulations have displayed different behavior regarding their ability to reform a heterogeneous cell population.

Our first intention consists in characterizing these specified cell subpopulations through global gene expression profiling. This whole genome transcriptomic study would help us in ascertaining qualifiers genes or even better molecular network of these three different cell subpopulations [195]. However, many cellular functions are shared between different cell types, and identifying those features that best discriminate between phenotypes still remains a challenge.

With this aim in mind these three cell subpopulations were isolated by flow cytometry cell sorting procedures, on three separate occasions, by following the workflow as illustrated in fig. A.15 on

page 222. In collaboration with the Genomics Platform of the Translational Research Department from the Curie Institute (David Gentien and Audrey Rapinat) RNAs of these cells subpopulations were extracted with the micro-RNeasy® kit from Qiagen, while changing some buffers to retrieve microRNAs and lncRNAs for further analysis. During RNA extraction procedures, a DNase treatment was performed for each sample in such a way minimizing potential contamination in DeoxyriboNucleic Acid (DNA) that would deceive thereafter the quantification of transcripts. In addition, high standards quality controls were carried out to ensure the purity and wholeness of extracted RiboNucleic Acid (RNA) :

- Measurement of the A230/260 ratio thanks to the Nanodrop® (Agilent) which indicates the presence of organic contaminants, such as phenol, TRIzol, chaotropic salts and other aromatic compounds. In a pure sample, the A260/230 ratios should be close to 2.0 and samples with 260/230 ratios below 1.8 are considered to have a significant amount of these contaminants that will interfere with downstream applications as the reverse transcription procedure. Our data shows variable values of A230/260 ratio and clearly indicate discrepancies of RNA organic contamination among samples extractions.
- Measurement of the A260/280 ratio with the Nanodrop® (Agilent) to determine protein contamination of nucleic acid samples. For pure RNA A260/280 ratios should be around 2.1. A lower ratio indicates the sample is protein contaminated and this contamination may have an effect on downstream applications that use nucleic acid samples. Measurements, ranging from 1.7 and 2.2, ensure high level of quality of all RNAs samples.
- Quantification of RNAs degradation was performed thanks to agarose gel electrophoresis. Each sample present clear peaks of 28S and 48S ribosomal RNAs and no RNA smear have been observed ensuring pure integrity of the extracted RNAs.

We further agreed that purified RNAs were upstanding, slightly contaminated with proteins and organic compounds, nevertheless in a suitable extent for downstream quantifications applications. Concomitantly purified RNAs concentration of each sample was quantified by Nanodrop® (Agilent). And all those results are available among table A.1 on page 223. Henceforth, we are considering quantifying the amount of RNA transcript within the various subpopulations through the Nanostring technology.

We have previously described in chapter 6 on page 71 that each of the sorted cells phenotypes - *i.e.* according to their level of expression of CD24 and CD44 markers - were able to give back to cells phenotypes which differ in their level of expression of CD24 and CD44 markers. As well cell plasticity leads to shift from a homogeneous to a heterogeneous cell population in a very short time frame. Nonetheless, over the experiments of cells repopulation we could appreciate that the way and the time taken to reform a heterogeneous cell population were varying in accordance to the starting cell phenotype. For the purpose to further describe the path that each phenotypes was following to reform a diverse cells population, we examine which genes could be the drivers of these phenotypic transitions. That is why we have extracted cells RNAs undergoing the repopulation process after 5 days following above-cited RNAs extraction procedures. Through these analyzes we want to determine :

- The genes which are differentially expressed of relevant activated biological pathways among the three repopulating cells phenotypes in the early days of the transition processes. Each specific gene would help us in characterizing the typical drivers of the different phenotypic transitions.

- Specific genes and pathways which are universally expressed during the phenotypes repopulating process. This would enable us to discriminate potential triggering genes of cells plasticity.

Thus, we may appreciate genes which are committed along the transitions paths pursued by different cell populations in order to return to the equilibrium within the phenotypic space. In the future this knowledge would likely allow the manipulation of characteristic phenotypic transitions to lead cell populations to balanced pre-determined phenotypic proportions and expressions.

Through cell plasticity phenomenon a homogeneous population inescapably evolves to a heterogeneous population. And slowly a balance between the diverse cell phenotype (proportions, level of expression of the markers...) is being established in a restricted period. Nevertheless, it is much to observed that the equilibrium which is set up particularly dependent on the phenotype of initial homogeneous population. Whatsoever cells from the first generation of cell sorting as well as the cells derived from a secondary sorting, the resulting heterogeneous cell populations perpetuate through their behavior the memory of their origin. The resultant question that we raised could be summarized as : How is maintained the memory of the cell of origin across generation of an evolving population ? Are there specific and persistent markers of the cell of origin which are expressed through cell generations. That is why we first started this work by extracting RNAs from :

- Freshly sorted cells subphenotypes.
- Cells subphenotypes derived from first-sorting strategy and cultivated for 57 days long.
- Cells subphenotypes derived from secondary-sorting procedure (at day 32 post first-sorting) and put back into culture for 25 days long.

RNA extractions were monitored as described in section 12.3.3 on page 198 and results regarding their purity, integrity and concentration are available on table A.1 on page 223. Nevertheless, we expect from this study to pinpoint a number of genes involved in the memory of the past. These discoveries will help us to investigate whether these genes are temporally and spatially expressed in a particular cell phenotypes. At last, we may surely edit the expression pattern of these genes and consider the resulting consequences for the re-balancing of the population.

Real time follow-up of cells phenotype transitions

Early analysis of phenotype transitions has been performed on several time points. This approach displays a major limitation regarding both the kinetic study of the phenotypic transitions expressed over time and cell generations. And above all, we cannot follow the spatial organization of the phenotypes and its evolution throughout the decisions taken at the cellular level and the collective level of cell groups.

This requires knowledge of the phenotype of each cell at every instant and that without damaging the potential cell interactions and their ability to transit from one phenotype to another. For that purpose, we wished to firstly follow the phenotypic transitions in real time in a 2D cells culture.

We initiated to integrate within the genome of the cells (and this whatever the phenotype expressed by the cell at the time “ $t = 0$ ” transduction) a molecular construct able to report in real time the endogenous variations of expression of CD24 and CD44 genes. If this project has not yet been fully satisfactory, nevertheless some new tracks are under development to change the methods, technologies and specific features as previously developed in section 11.4.2 on page 166.

This ambitious project to follow-up cell state changes requires discriminating cells boundaries at the single cell level, each cells even if they are aggregated to each others in a dense cell monolayer. We therefore undertook this work in collaboration with Dr Isabelle Bonnet (Institut Curie, UMR168) and then with Romain Cendre (CIC-IT Nancy). We have tested different image analysis methods to initially discriminate cell edges and in a second time to determine the expression of CD24 and CD44 markers at the single cell level (*cf.* appendix B on page 233). Algorithm to soften noise of the source images provided optimal results and areas of overexposure have been homogenized. And then several algorithms to discriminate cell boundaries have been incremented (thresholding, filtering, spreading, deviation around the mean methodologies). Nevertheless, only the algorithms based on edge detection were close to our expectations without giving full satisfaction.

To improve both image quality and automatic cell detection, we are currently working on better algorithms of images homogenization, background subtraction and cell edges detection thanks to detection of both crashes and cell structure.

Besides these technological considerations, we expect of these tools they enable us to monitor phenotype transitions in real time and to study their behavior at different scales. For example, we will be able to determine the intermediate levels of phenotype expressions ; because indeed we believe that the transitions of phenotype expressions are rather part of a gradual than a binary process. Moreover, we would interest us in detail to :

- The dynamics of rediversification within cell populations to better determine the phenotypes which drive that process.
- How does take place the equilibrium of expressed phenotypes within the rediversified populations.
- The characteristics of periodicity that single out rediversified populations derived from CD24-/CD44+, CD24+/CD44+ and CD24-/CD44- homogeneous subpopulations in order to compare them with the ones described for each of non-sorted cell lines (*cf.* section 5.4.2 on page 68).

Modeling phenotypes plasticity and dynamic steady-state

The question underlying the general problem can be summarized as follows : “does the behavior of an individual affect the behavior of a population?”. And contrariwise, “does the global behavior of a population affect the behavior of an individual?”. A sociological study of the populations’ organization is actually needed at different levels, from the individual cell to the whole population. To begin with, we created a method to study the behavior at the uni-cellular level and at the population level based on flow cytometry analysis of markers CD24 and CD44. Then we monitored the changes of expression of both markers at several time points at the cellular and population level. Through recurrent analysis over time, we estimated the changes in the overall cell population in CD24 and CD44 phenotype space.

Throughout the analyses described in this thesis, we manually delimited the cells of interest. We suppose that this process might introduce bias regarding our study of the behavior of cell subpopulations, especially fo CD24-/CD44- subphenotype. We certainly wish to remove human bias when gating cells to limit bias to dispersion and mean of markers CD24 and CD44. We suggest automating cell population gating. We started on this work by testing several software programs such as the FlowCAP project developed by Aghaeepour et al. [4], a transfer learning approach presented by Lee et al. [177] and an automatic clustering approach with Density-Based Merging by Walther et al. [329]. Nonetheless, we noticed that cell populations whose proportions

are below 8%, such as CD24-/CD44- phenotype, are not properly detected. Therefore, we are considering improving those tools further regarding the characterization of under-represented cell subpopulations. Furthermore, we also would like to automate bi-parametric detection and the derived statistical analysis. For that purpose, we looked into few methods developed by our colleagues Pyne et al. and Wallin et al. [248, 328].

The study of phenotypic transitions was conducted through three independent experiments for which three to six intra-experiment replicas were included. For technical reasons the harvesting of raw data of inter-experiment replicas was not always carried the same time points. Nevertheless, we wanted to ensure that the fluctuations observed after the pooling of all results were relevant. Therefore, we wanted to normalize the flow cytometry data measured over time and to confirm the relevance of the expression trajectory of the cell population in the space of CD24 and CD44 markers. We presented the following process on section 5.2.3 on page 53.

- Normalization of noise originating from flow cytometry, reduction and control of noise coming from equipment and staining techniques
- Normalization of noise originating from the biological sample, reduction and control of cell auto-fluorescence

However, given the numerous publications, flow cytometry data normalization is not an easy task [86, 116, 127, 232]. This step is critical before using flow cytometry data, highlighting the evolutionary behavior of the expression level of the markers, and preventing bias in variations. Furthermore, the monitoring of phenotypes requires computation of statistical characteristics such as mean, dispersion of fluorescence map and proportion of subphenotypes because they only apply to the most represented cells. Under-represented cell phenotype are mostly ignored even though some like CD24-/CD44- might have a significant role. Pr. Pyne et al. [130] developed a methodology we consider following. Its steps are monitoring the time course profile of Sca-1 expression marker, monitoring phenotypic transitions and repopulation, comparing the re-balancing of populations with a reference cell line. The overall methodology can be applied to our case. Nonetheless, strive to improve it on the bi-parametric aspect [130, 248, 328].

Finally, we focused on modeling the re-popularization and re-balancing of phenotype subpopulations as described in chapter 7 on page 103. Based on past work, we would like to propose a model that takes into account the variability of probability laws that characterize phenotype transition, especially due to phenotype memory. In the meantime, we will continue on the spatial modeling of phenotype transitions.

Future prospects for this research project

In previous parts we have turned our attention to former work and research still in progress. Nonetheless, we will take advantage of this next section informations to discuss the long-term projects that we wish to implement in the wake of our previous work. Thus, these elements will enable us to better consider the overall framework of reflection within which our research project takes place.

This latter has to be with collective independence or dependence of decision-taking regarding cells neighborhood. And therefore this following section will go through these various aspects :

- Observation over the long-term and in real-time phenotype transitions at the individual and the collective scale.
- Validation or invalidation of the assumption that cells are dependent on their adjacent neighborhood with respect to phenotype transitions.

- Search for spatial patterns of cells phenotype organization.
- Browse for more cell lines (initially with a breast origin) which intrinsically present a context of cell heterogeneity.

Do cancer cells spatially organize themselves ?

Determining whether the cellular phenotype is expressed in terms of a spatially defined pattern is the ultimate issue and not the least of this work. And more specifically this questioning will be the completion of this research project. Answering to this precise controversy will be the result of all previous efforts, both in terms of :

- Techniques implementation in terms of monitoring of phenotype transitions, of live microscopy, and of automation of image analysis procedures (*cf.* appendix B on page 233).
- Analyzes of the experimental outcome (quest for spatial expression patterns, exploration of potential correlations between phenotype transitions and influence of the cellular environment, relationship between temporality of phenotype expression and the phenotype history of cell progenies).
- Modeling of cellular plasticity events and re-balancing of dynamic biological systems.

And so the main issue that concerns us can therefore be formulated as follows : what impact would social rules play regarding cell-decision making (phenotypic plasticity and transitions, population dynamics, division rates...)?

The hypothesis we formulate is that the cells transit of a phenotype to another according to the influence of its neighbors. That is why we have initiated this work by reflecting on an approach that would enable us to measure the extent to dependence of cells next to each other. The methodology implemented has been described over chapter 12 on page 169 and mainly rests on the computation of the phenotype average of a geographically preset area (surface and geographical coordinates). Thus, the average phenotype of several geographic areas is determined and every averaged value is compared on the basis of surface area. Based on these data, we shall be subsequently able to determine the critical distance thanks to which one cell exerts influence on its neighboring cells.

We expect from such information much more than just calculated distance of influence that a cell phenotype may have on other neighboring cells. In continuation of this rationale, this expected value will give us information regarding the nature of cell communication. Thus, a distance of influence of just a few cells will set us on the trail of direct communication way that cells used to communicate between partners. While a critical distance of influence of fifty to a hundred cells would correspond to the use of excreted molecules into the external environment as a manner to efficiently propagate a message.

On the way of other cell lines which are phenotypically heterogeneous

This work took a large interest in two breast cancer lines so-named as SUM149-PT and SUM159-PT. These cell models have been particularly interesting by virtue of their heterogeneity and distinctive plasticity with regard to the expression of CD24 and CD44. We suspect that these features not only qualify these two outlined cell lines and may be deferred to a number of additional cell lines. Therefore, we have conducted a study to identify the level of heterogeneity for miscellaneous markers as (Cytokeratin (KRT)5/KRT6, KRT8/18, KRT14, CD44) a small quarantine breast cell lines agglomerated in the form of a TMA and kindly transmitted by Thierry

Dubois et al.. You will see an overview of cell lines for which the staining has enabled us to quantify the level of heterogeneity within table A.3 on page 228 and we are considering to continue this study by including the CD24 marker. So each cell line, included in a block of paraffin, was stained for the inventoried markers according to the conventionally used methodology for clinical diagnostic and being developed by Nicolas André and Didier Meseure, who belong to the Biopathology Department of the Curie Institute. As a first step, we determined the gradient of six classes describing the level of expression of the markers ; and thus the standard was as follows :

- 0 corresponding to **no expression**
- 0-1 “ ” to **light** degree of expression
- 1 “ ” to **middle** expression
- 2 “ ” **almost high** expression
- 3 “ ” **high** degree of expression
- 4 “ ” **very great** expression

Then, we visually semi-quantified the percentage of cells that corresponds to each class of level of expression. Estimated data regarding the expression of KRT5/KRT6, KRT8/18, KRT14, and CD44 markers are presented as heat maps within the supplementary data (*cf.* fig. A.17 on page 229, fig. A.18 on page 230, fig. A.19 on page 231 and fig. A.20 on page 232). As indicated (by a yellow, orange or red star) in the figures, about one third of the cell lines exhibit an heterogeneous nature regarding the level of expression of the markers. And as previously conducted for SUM149-PT and SUM159-PT cell lines, it would be possible to :

- Determine the periodic expression of each of these cell lines for the paired markers CD24/CD44, KRT5/KRT8 et KRT14/KRT18 (*cf.* chapter 5 on page 49).
- Study the phenotypes plasticity (*cf.* chapter 6 on page 71).

The screen that we realized is the preamble to broader research studies regarding the characterization of both cell fate and social rules governing the spatio-temporal organization of cell populations. It would be possible to ascertain whether there are correlations between the phenotypic transitions (periodicity, speed ...) and social rules that regulate the phenotypic balance both at the local and the global scales of the different cell lines and the pathological conditions. Thus, one could establish a repository serving as a tool for high throughput screening of under development molecules to determine the effectiveness of these molecules to change social rules.

Thus, the work that is currently submitted has launched numerous issues which still remain unresolved. Nonetheless, we have been careful in anticipating further necessary requirements to the successful pursuit of this research project. Thus, a certain number of future tools and technologies have been roughly trimmed and some collaborations have already been approached. The best is yet to come and I wish favorable luck and good continuation to all those who will take this project over !

Part V

Appendix

Appendix **A**

Supplementary data

«L'esprit d'observation procure un bien plus grand avantage que celui de satisfaire une curiosité raisonnable : en nous apprenant combien en général les hommes sont injustes, il nous prépare à l'injustice. Ainsi, lorsqu'elle nous atteint, nous en sommes moins blessés, et la considérant comme une infirmité de l'espèce plutôt que comme le tort d'un individu, nous sommes moins exposés au malheur de haïr.»
- Pierre-Marc-Gaston de Levis - *Maximes et préceptes* (1808)

This chapter regroups

A.1 Additional data for part II to part V

A.1.1 Supplementary data associated to part “natural history of cells diversity”

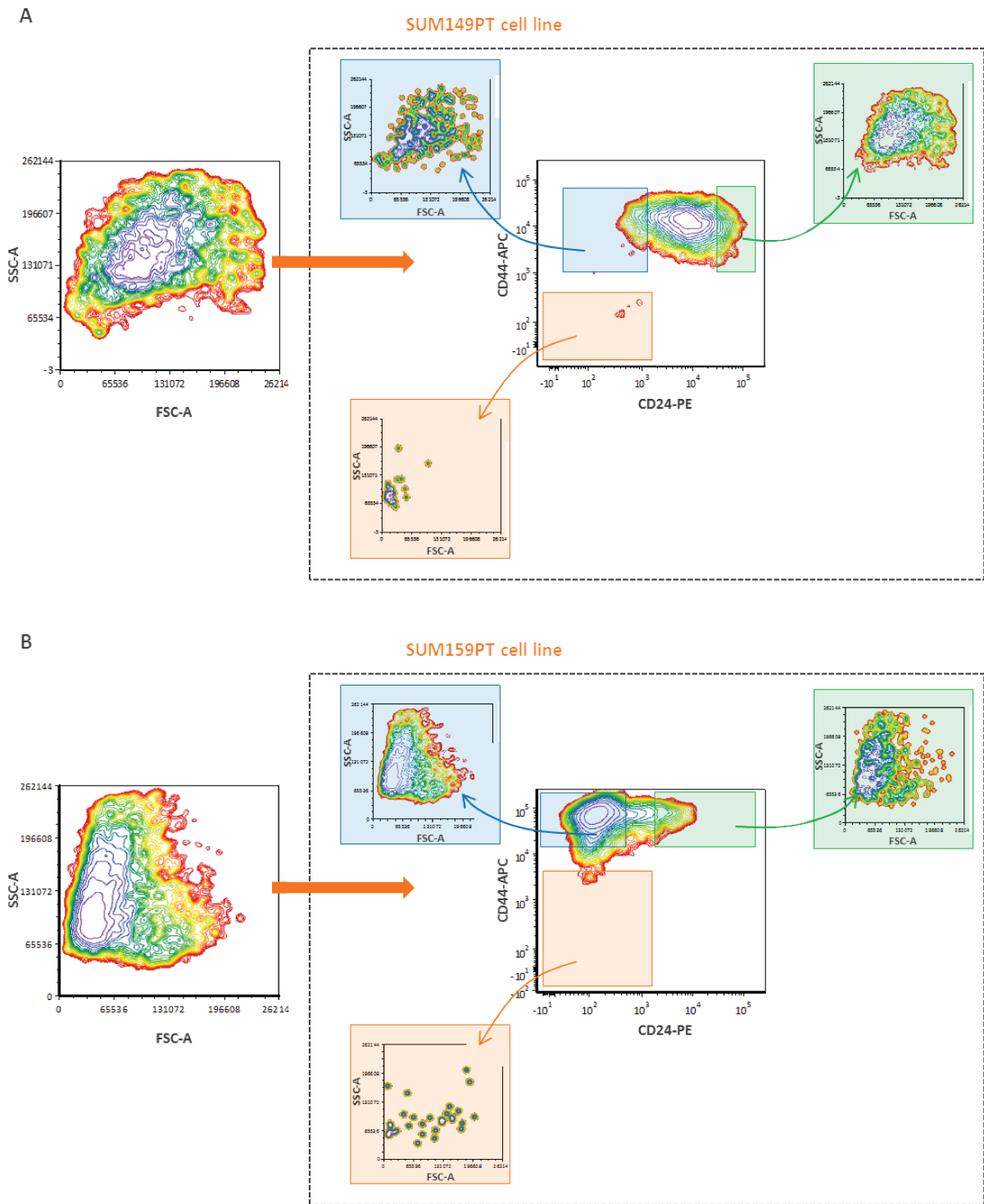


Figure A.1 : FSC and SSC characterization of SUM149-PT and SUM159-PT sorted subphenotypes.

A. FSC and SSC flow cytometry contour plots of SUM149-PT unsorted cell line and its sorted subphenotypes. B. and of SUM159-PT unsorted cell line and its sorted subphenotypes.

In blue : CD24-/CD44+ phenotype, in orange : CD24-/CD44- phenotype, in green : CD24+/CD44+ phenotype

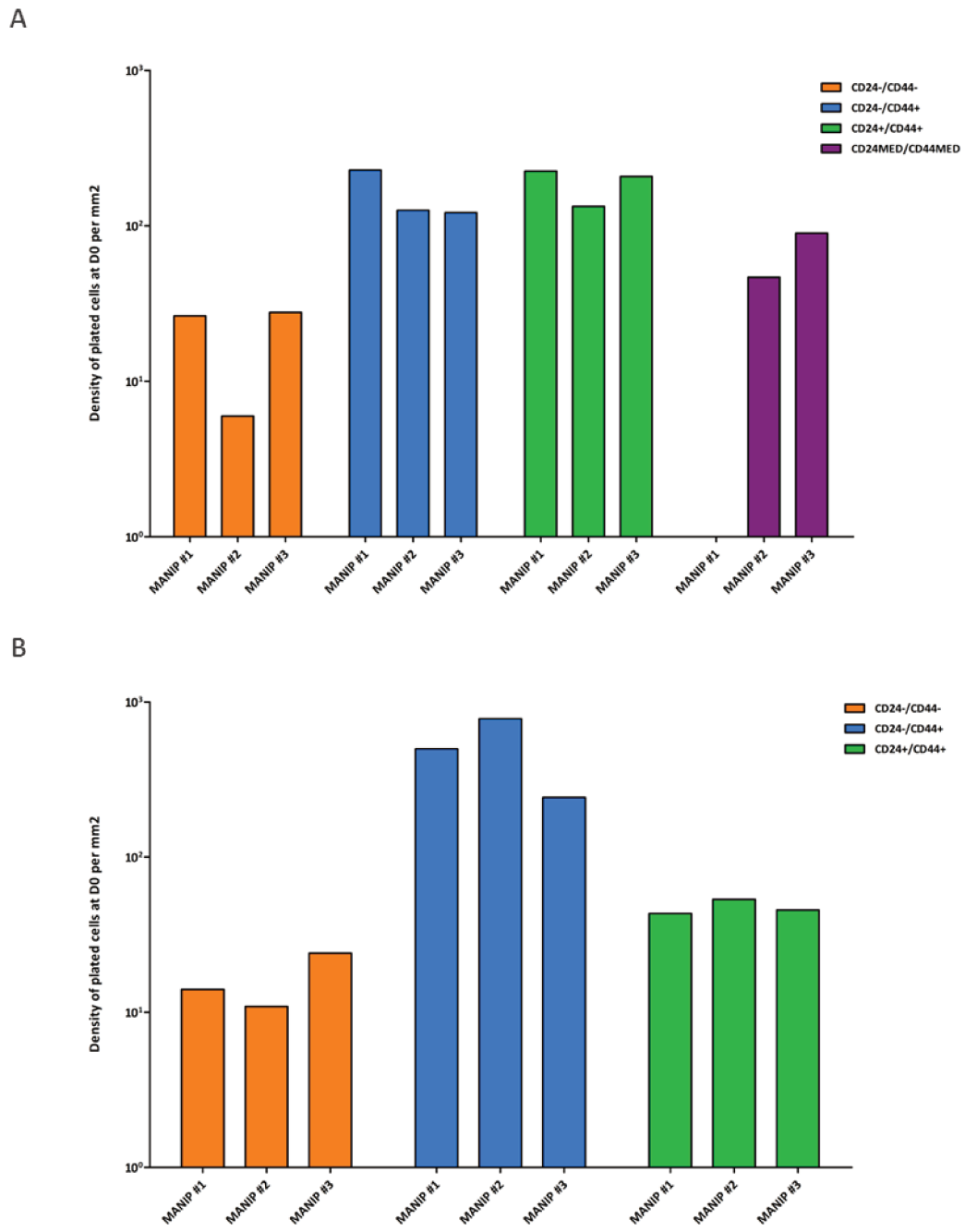
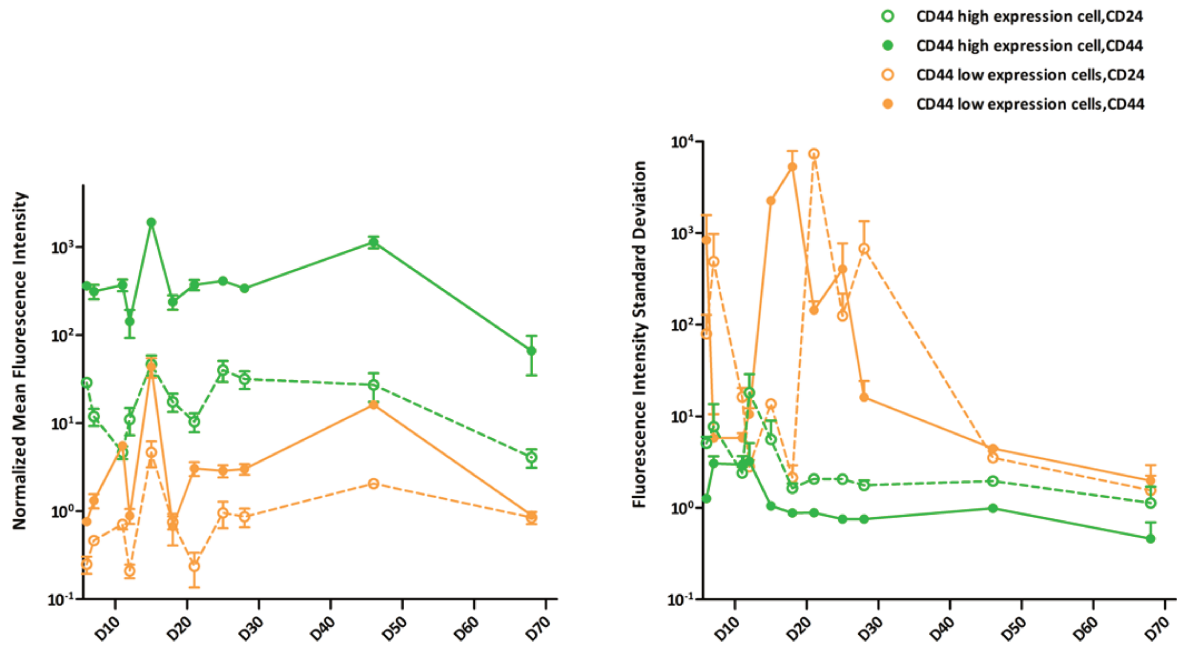


Figure A.2 : Density plating of cells sorted subphenotypes per assay

A. Column graph to compare density plating of SUM149-PT sorted cells subphenotypes for each replicative sorting experiments. B. and of SUM159-PT sorted cells subphenotypes.



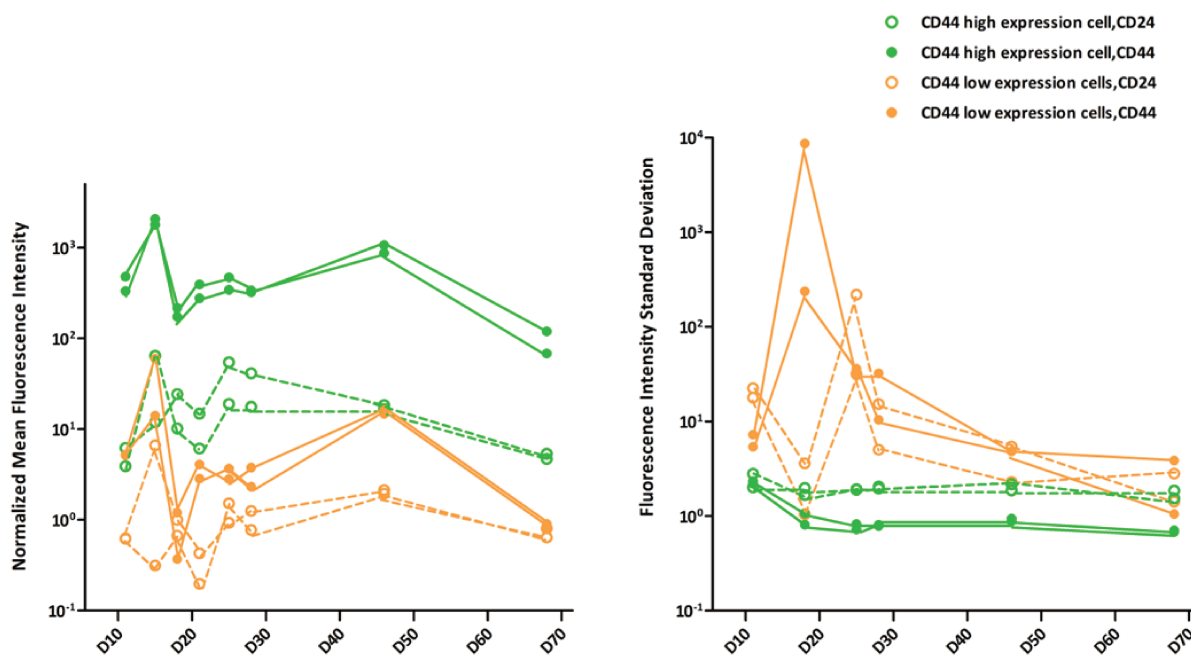


Figure A.4 : Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM149-PT CD24-/CD44- subphenotype cells.

A. Grouped column scatters graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean of intra-experiment replica of SUM149-PT repopulating viable cells. B. and of CD24 and CD44 standard deviation fluorescence intensity.

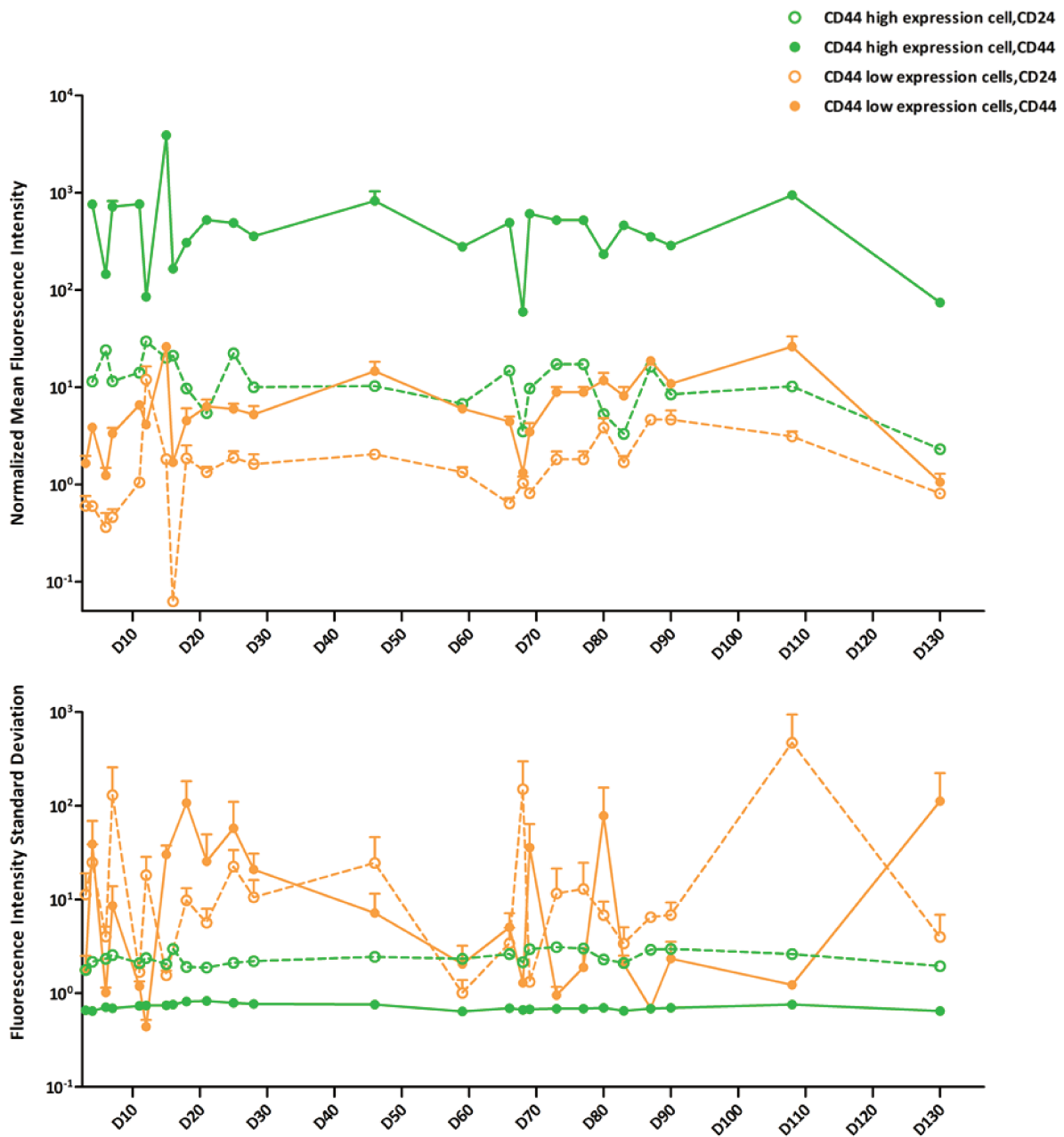


Figure A.5 : Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM149-PT CD24-/CD44+ subphenotype cells.

A. Grouped column scatter graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean fluorescence intensity of SUM149-PT repopulating viable cells. B and of CD24 and CD44 fluorescence intensity standard deviation. Arithmetic pseudo-geometrical mean of both normalized pseudo-geometrical mean and standard deviation of fluorescence intensity are plotted on the graph as bullet points (Error bar represents SEM).

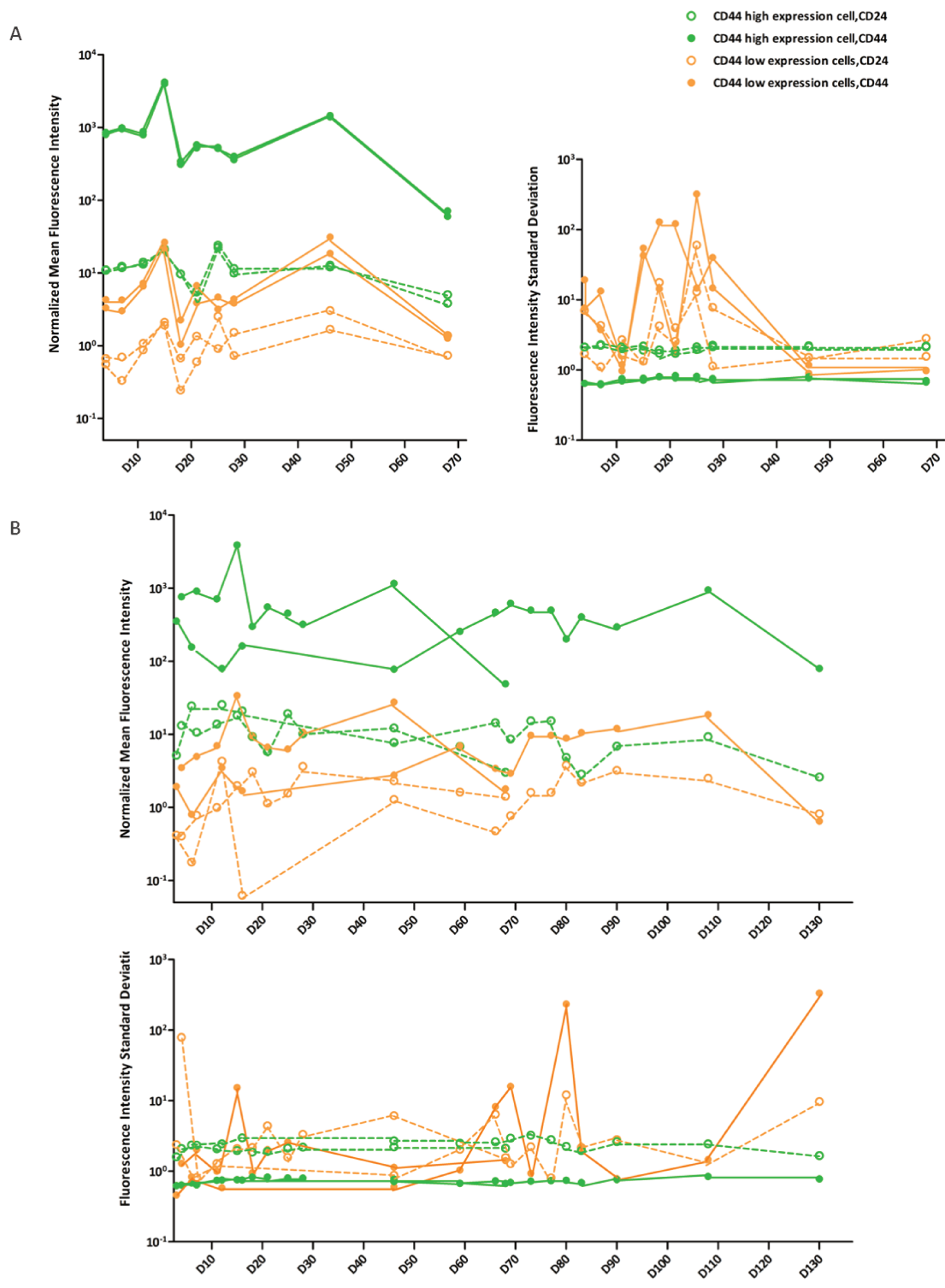


Figure A.6 : Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM149-PT CD24-/CD44+ subphenotype cells.

A. Grouped column scatters graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean and standard deviation fluorescence intensity of intra-experiment replica of SUM149-PT repopulating viable cells. B. and of inter-experiment replica.

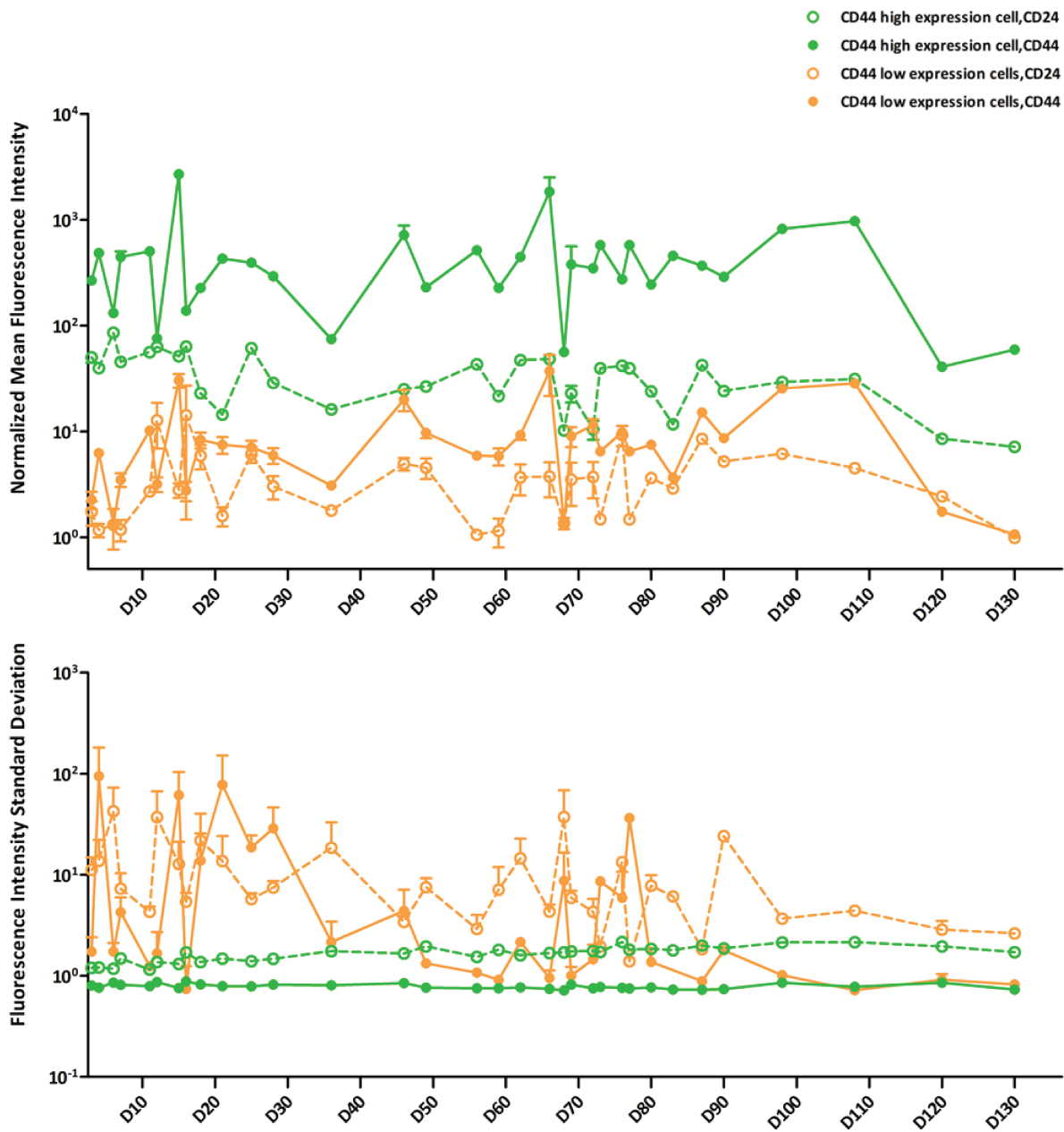


Figure A.7 : Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM149-PT CD24⁺/CD44⁺ subphenotype cells.

A. Grouped column scatter graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean fluorescence intensity of SUM149-PT repopulating viable cells. B and of CD24 and CD44 fluorescence intensity standard deviation. Arithmetic pseudo-geometrical mean of both normalized pseudo-geometrical mean and standard deviation of fluorescence intensity are plotted on the graph as bullet points (Error bar represents SEM).

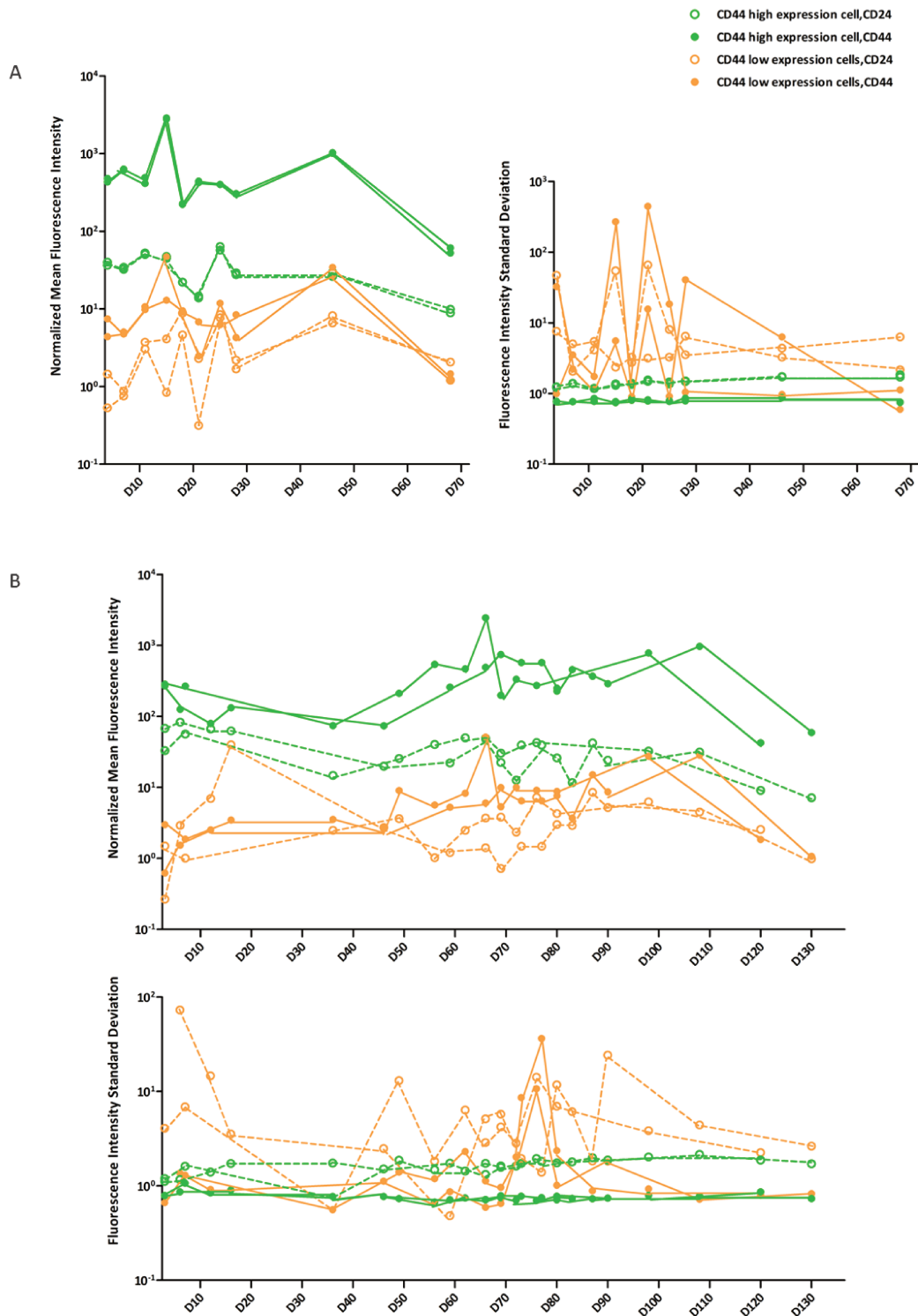


Figure A.8 : Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM149-PT CD24+/CD44+ subphenotype cells.

A. Grouped column scatters graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean and standard deviation fluorescence intensity of intra-experiment replica of SUM149-PT repopulating viable cells. B. and of inter-experiment replica.

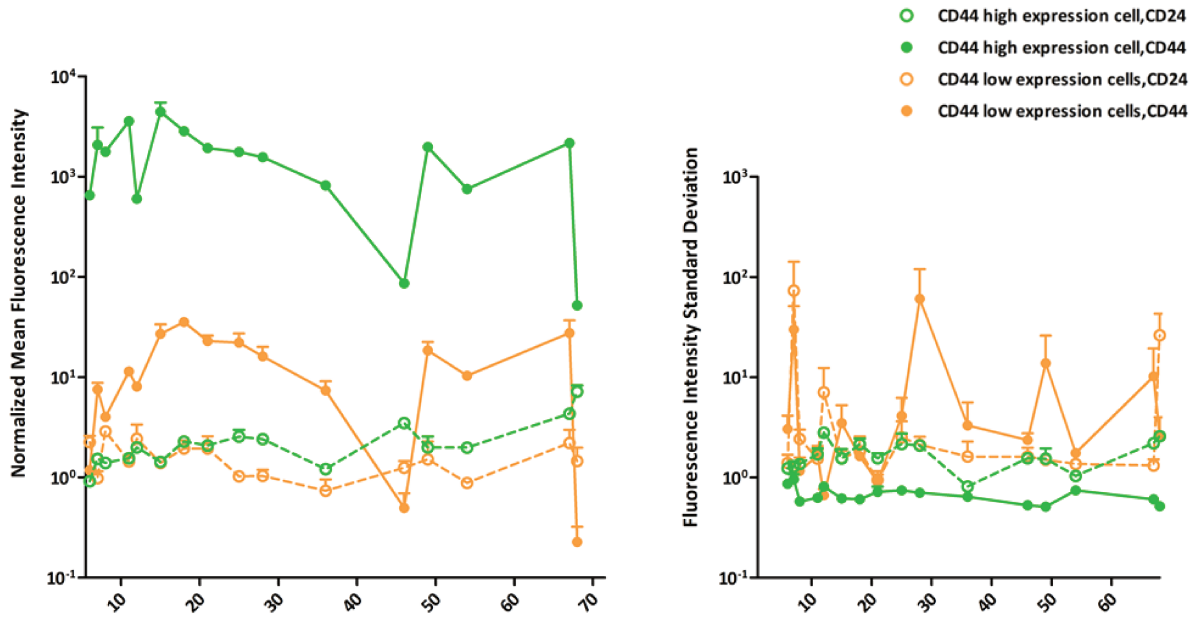


Figure A.9 : Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM159-PT CD24-/CD44- subphenotype cells.

A. Grouped column scatter graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean fluorescence intensity of SUM149-PT repopulating viable cells. B and of CD24 and CD44 fluorescence intensity standard deviation. Arithmetic pseudo-geometrical mean of both normalized pseudo-geometrical mean and standard deviation of fluorescence intensity are plotted on the graph as bullet points (Error bar represents SEM).

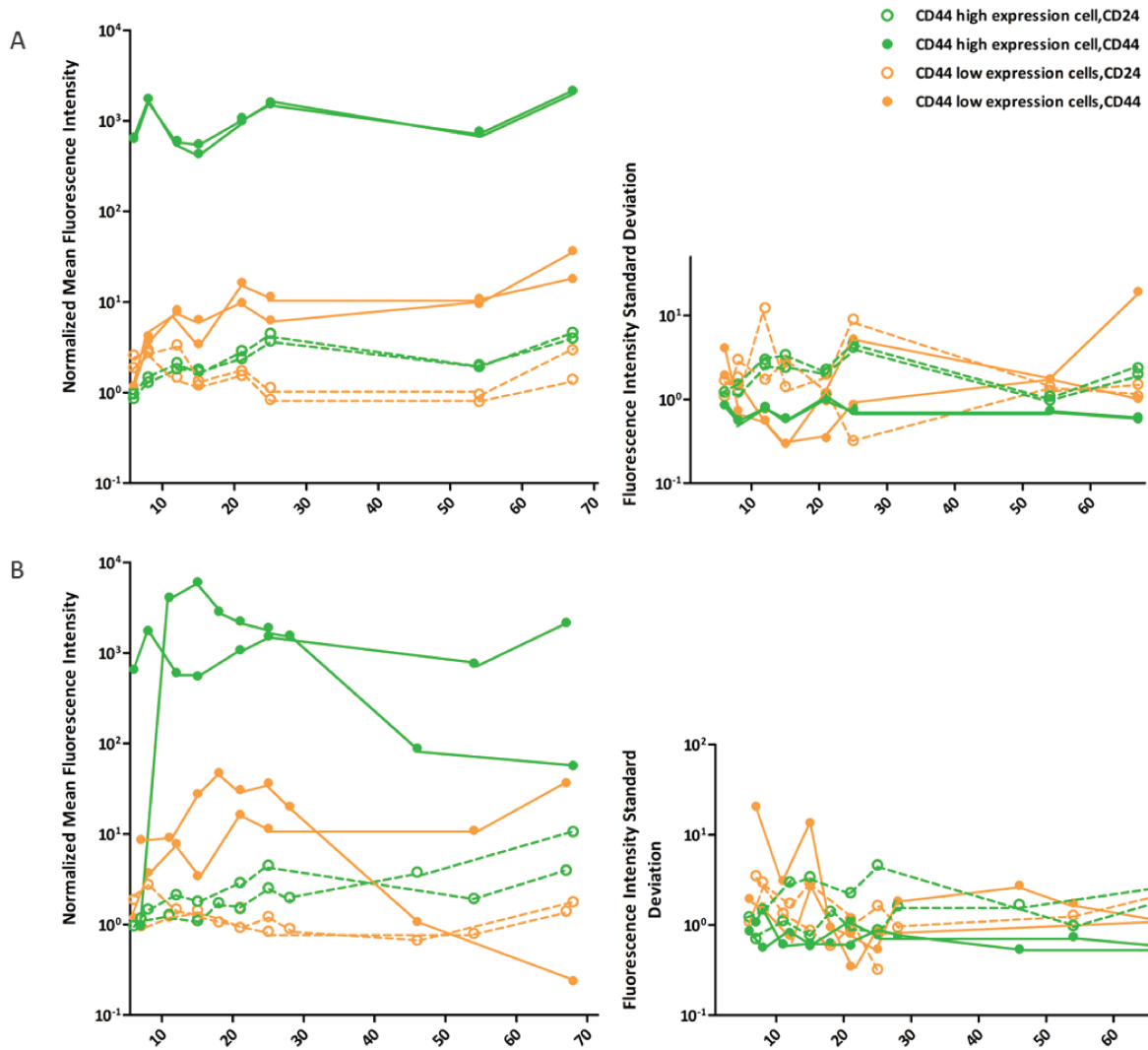


Figure A.10 : Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM159-PT CD24-/CD44- subphenotype cells.

A. Grouped column scatters graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean and standard deviation fluorescence intensity of intra-experiment replica of SUM149-PT repopulating viable cells. B. and of inter-experiment replica.

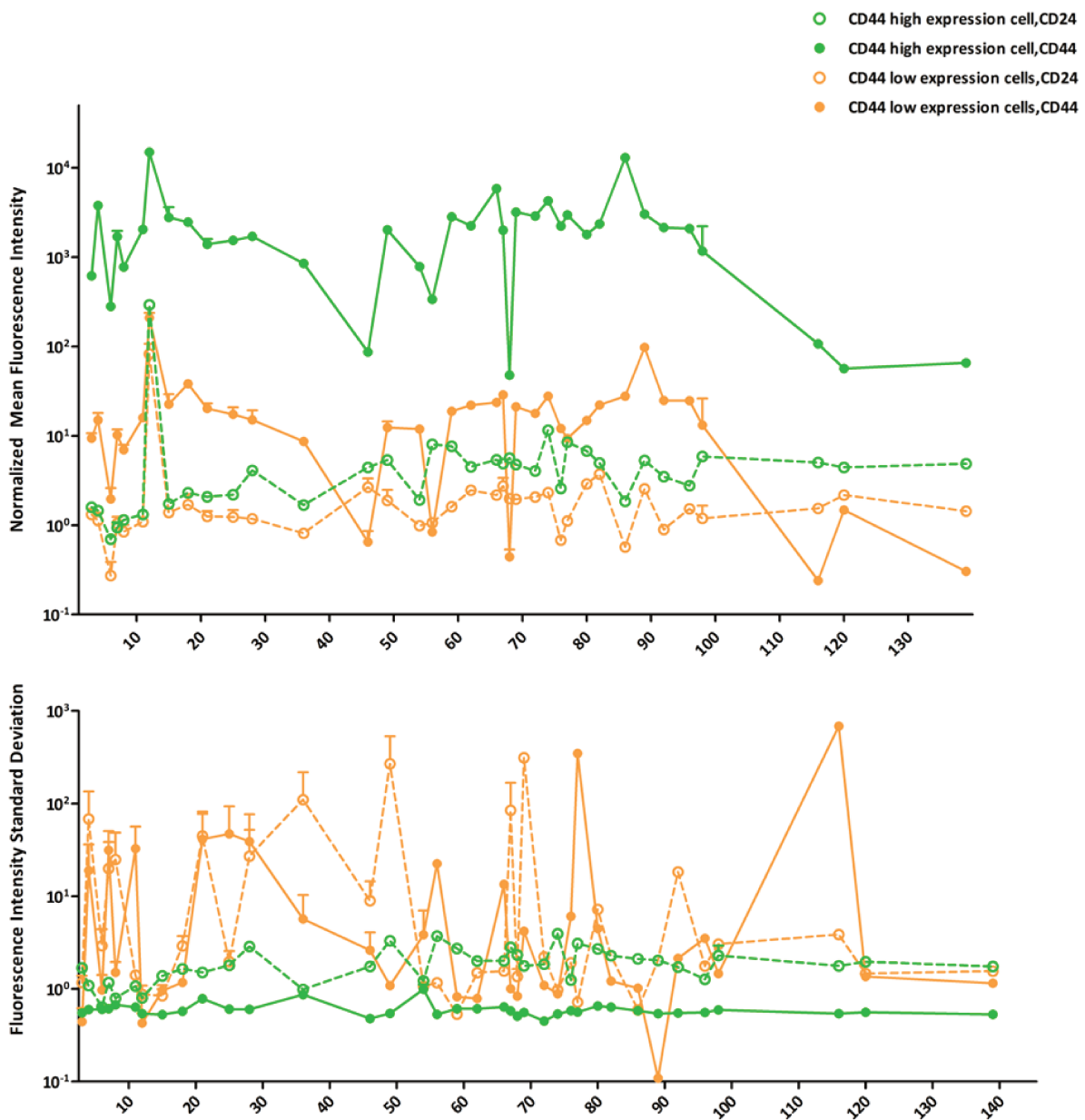


Figure A.11 : Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM159-PT CD24-/CD44+ subphenotype cells

A. Grouped column scatter graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean fluorescence intensity of SUM149-PT repopulating viable cells. B and of CD24 and CD44 fluorescence intensity standard deviation. Arithmetic pseudo-geometrical mean of both normalized pseudo-geometrical mean and standard deviation of fluorescence intensity are plotted on the graph as bullet points (Error bar represents SEM).

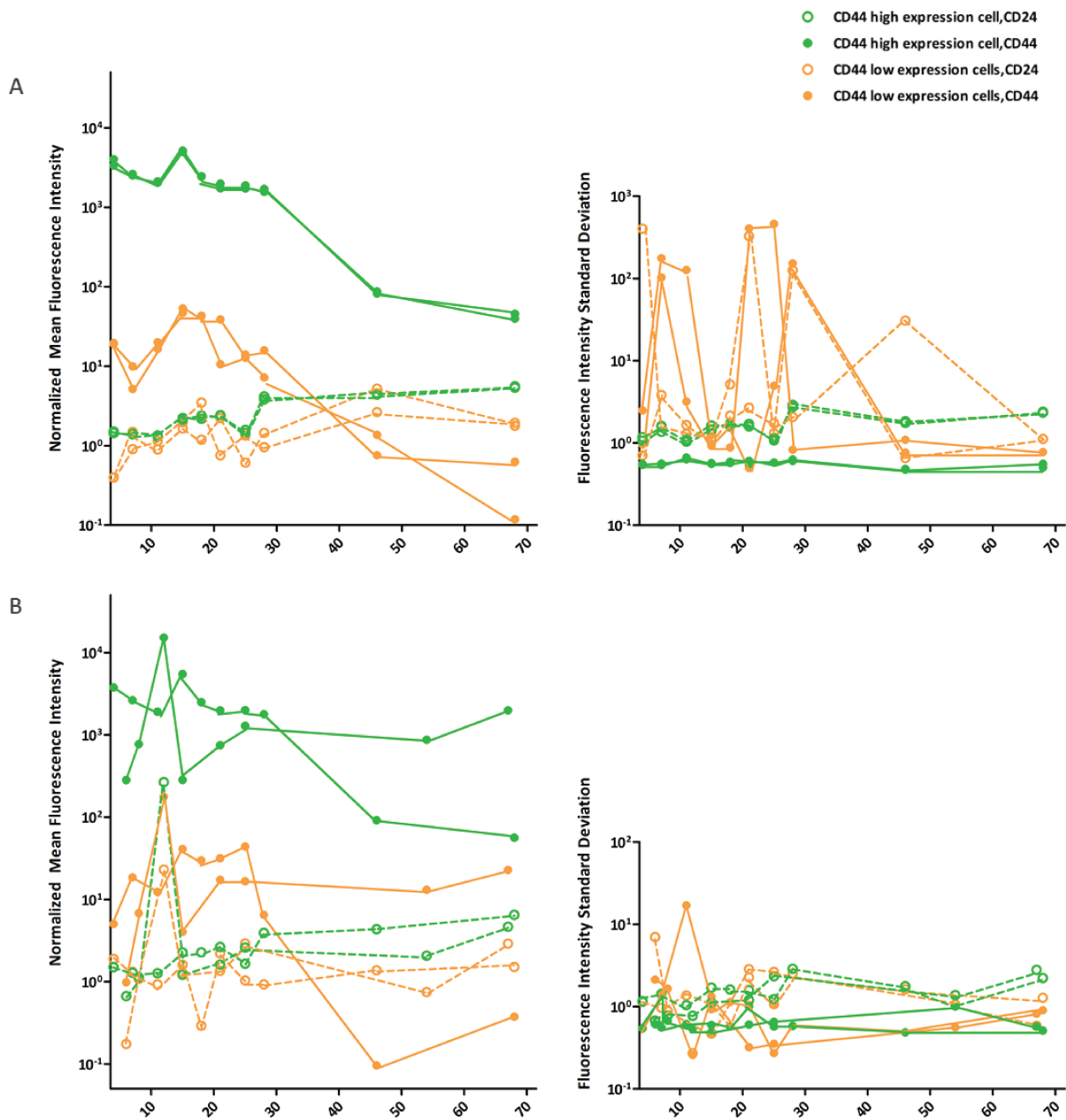


Figure A.12 : Long-term monitoring of CD24/CD44 expression profile of repopulating sorted SUM159-PT CD24-/CD44+ subphenotype cells

A. Grouped column scatters graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean and standard deviation fluorescence intensity of intra-experiment replica of SUM149-PT repopulating viable cells. B. and of inter-experiment replica.

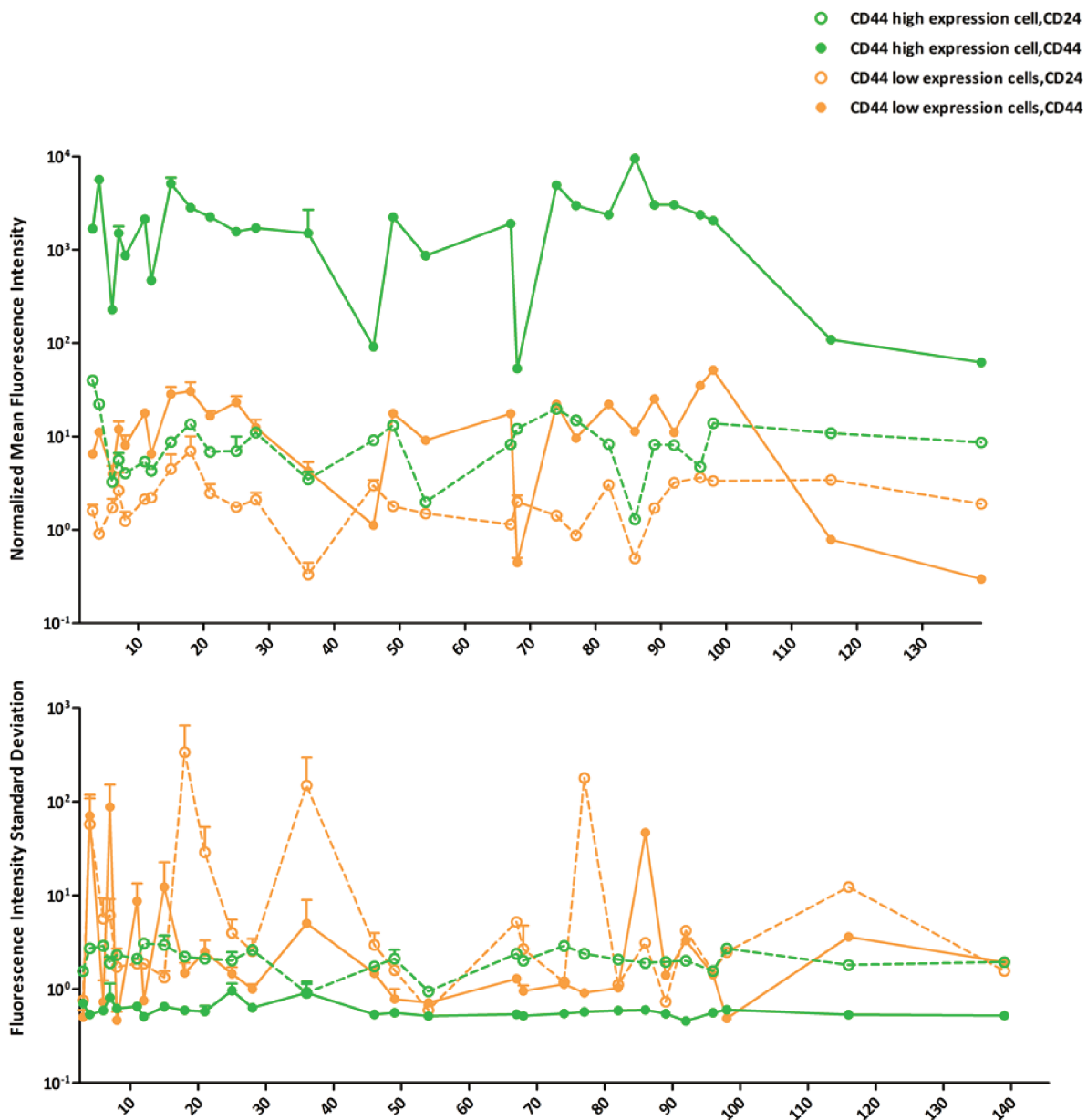


Figure A.13 : Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM159-PT CD24⁺/CD44⁺ subphenotype cells

A. Grouped column scatter graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean fluorescence intensity of SUM149-PT repopulating viable cells. B and of CD24 and CD44 fluorescence intensity standard deviation. Arithmetic pseudo-geometrical mean of both normalized pseudo-geometrical mean and standard deviation of fluorescence intensity are plotted on the graph as bullet points (Error bar represents SEM).

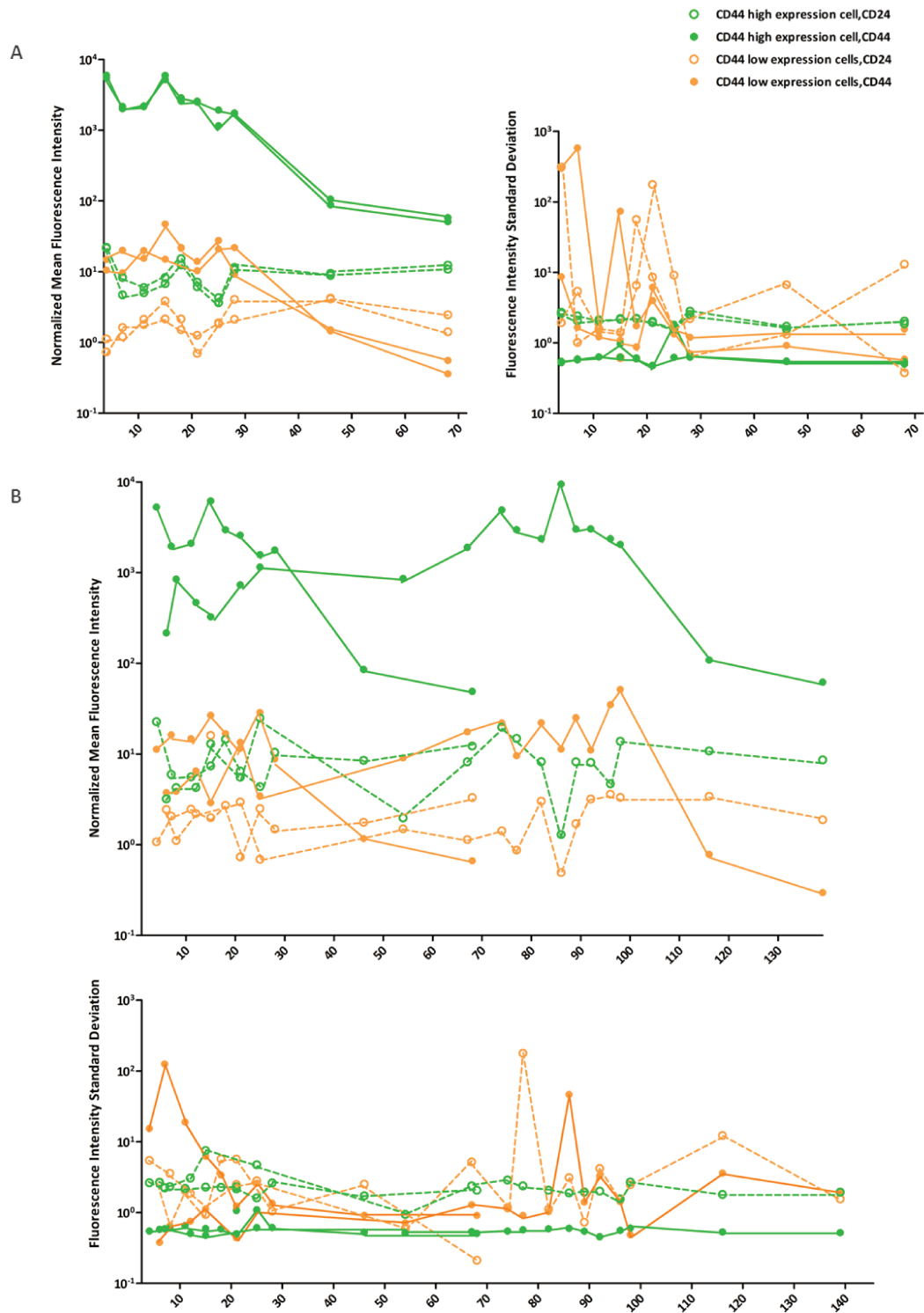


Figure A.14 : Long-term monitoring of CD24/CD44 expression profile of pooled repopulating sorted SUM159-PT CD24+/CD44+ subphenotype cells
A. Grouped column scatters graphing long term following of CD24 and CD44 normalized pseudo-geometrical mean and standard deviation fluorescence intensity of intra-experiment replica of SUM149-PT repopulating viable cells. B. and of inter-experiment replica.

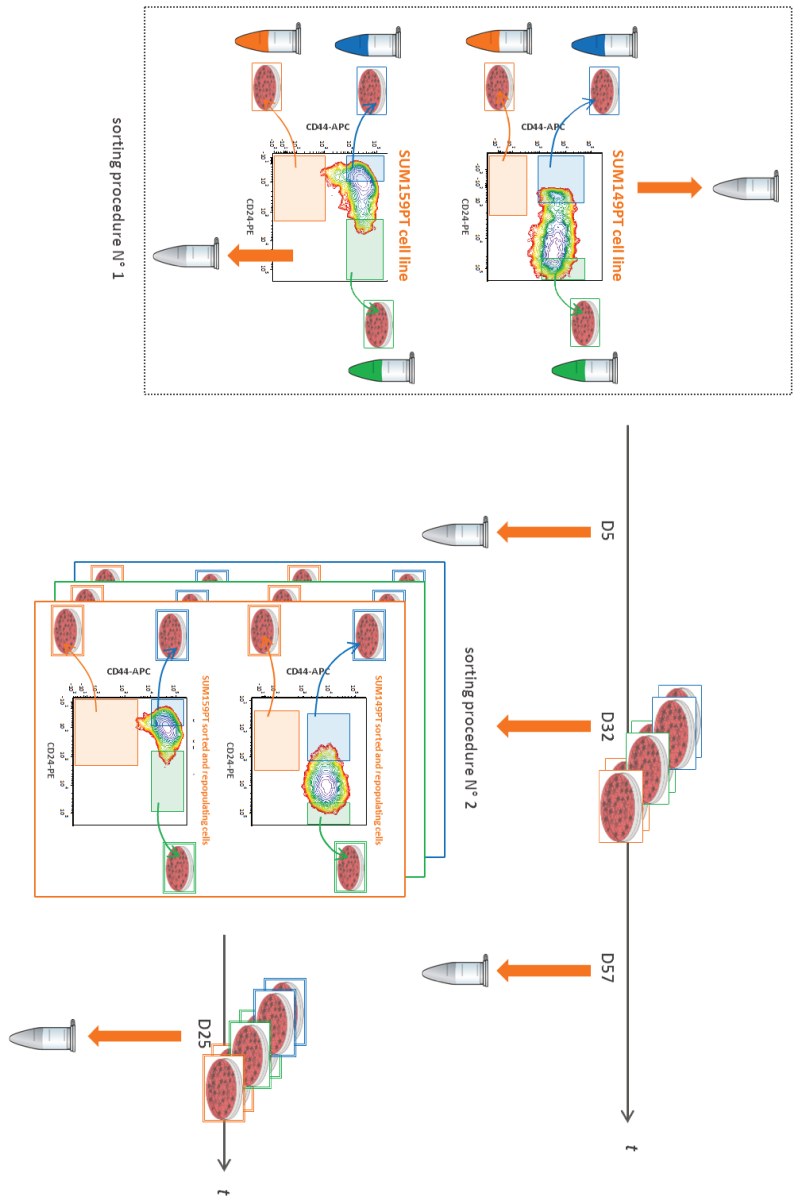


Figure A.15 : *Experiment workflow followed to perform transcriptomic analysis on SUM149-PT and SUM159-PT cell lines, cells sorted subphenotypes and their derivative repopulating cells*

name cell line	sorted or not sorted cells	number of sorting	days of culture after sorting	name of cells subpopulation	number of cells per tube	concentration (pg/ μ l) calculated by Agilent	amount of RNA (ng) calculated by Agilent	ratio 260/280	ratio 260/230	ratio 28S/18S	RIN	Conclusion about RNA extraction quality
					50000							
SUM159PT	non sorted cells				50000							
SUM159PT	non sorted cells				50000							
SUM149PT	non sorted cells				50000							
SUM149PT	non sorted cells				50000							
SUM159PT	sorted cells	first sorting	D0	CD24-CD44-	10466	77	0,92			1,5	8,4	OK
SUM159PT	sorted cells	first sorting	D0	CD24-CD44+	20000	382	4,58			1	8,5	OK
SUM159PT	sorted cells	first sorting	D0	CD24-CD44+	20000	176	2,11			2,4	6,1	OK
SUM149PT	sorted cells	first sorting	D0	CD24-CD44-	1700	8 316	99,79			1	7,5	OK
SUM149PT	sorted cells	first sorting	D0	CD24-CD44+	20000	377	4,52			1,5	8,7	OK
SUM149PT	sorted cells	first sorting	D0	CD24+CD44+	20000	605	7,26			1,6	8,6	OK
SUM159PT	sorted cells	first sorting	D0	CD24-CD44-	13000	166	1,99			1,9	7,1	OK
SUM159PT	sorted cells	first sorting	D0	CD24-CD44+	20000	322	3,86			2,2	8,8	OK
SUM159PT	sorted cells	first sorting	D0	CD24+CD44+	20000	827	9,92			2	8,7	OK
SUM149PT	sorted cells	first sorting	D0	CD24-CD44-	20000	76	0,84			0	1	~
SUM149PT	sorted cells	first sorting	D0	CD24-CD44+	20000	357	4,28			1,9	7,7	OK
SUM149PT	sorted cells	first sorting	D0	CD24+CD44+	20000	280	3,36			1,5	8,6	OK
SUM159PT	sorted cells	first sorting	D0	CD24-CD44-	20000	109	1,31			1,9	7,3	OK
SUM159PT	sorted cells	first sorting	D0	CD24-CD44+	20000	329	3,95			2,2	7,9	OK
SUM159PT	sorted cells	first sorting	D0	CD24-CD44+	20000	447	5,36			2,3	8,2	OK
SUM149PT	sorted cells	first sorting	D0	CD24-CD44-	30000	113	1,36			1,3	4,6	OK
SUM149PT	sorted cells	first sorting	D0	CD24-CD44+	31800	3 474	41,69			1,3	9,1	OK
SUM149PT	sorted cells	first sorting	D0	CD24+CD44+	32331	3 387	40,64			1,4	9,1	OK
SUM159PT	sorted cells	first sorting	D0	CD24-CD44-	30118	51	0,61			1,1	5,1	OK
SUM159PT	sorted cells	first sorting	D0	CD24-CD44+	30530	583	7			1,9	9,4	OK
SUM159PT	sorted cells	first sorting	D0	CD24+CD44+	30159	4 423	53,08			1,5	8,9	OK
SUM149PT	sorted cells	first sorting	D5	CD24-CD44-	less than 50000	37	0,44			1,6	4,2	~
SUM149PT	sorted cells	first sorting	D5	CD24-CD44+	less than 50000	18 624	223,49			0,9	7,4	OK
SUM149PT	sorted cells	first sorting	D5	CD24+CD44+	less than 50000	11 069	132,83			1,1	7,4	OK
SUM159PT	sorted cells	first sorting	D5	CD24-CD44-	less than 50000	9 896	118,75			1,4	8,1	OK
SUM159PT	sorted cells	first sorting	D5	CD24-CD44+	less than 50000	16 949	203,39			0,9	7,7	OK
SUM159PT	sorted cells	first sorting	D5	CD24+CD44+	less than 50000	4 097	49,16			1,8	7,2	OK
SUM149PT	sorted cells	first sorting	D5	CD24-CD44-	less than 50000	38	0,46			0	2,4	~
SUM149PT	sorted cells	first sorting	D5	CD24-CD44+	less than 50000	8 647	103,76			0,4	6,7	OK
SUM149PT	sorted cells	first sorting	D5	CD24+CD44+	less than 50000	9 746	116,95			1,4	8	OK
SUM159PT	sorted cells	first sorting	D5	CD24-CD44-	less than 50000	149	1,64			1,9	10	OK
SUM159PT	sorted cells	first sorting	D5	CD24-CD44+	less than 50000	5 075	60,9			0,7	8,2	OK
SUM159PT	sorted cells	first sorting	D5	CD24-CD44+	less than 50000	26 348	316,18			0,3	5,9	OK
SUM149PT	sorted cells	first sorting	D5	CD24-CD44-	less than 50000	47	0,56			0	1	~
SUM149PT	sorted cells	first sorting	D5	CD24-CD44+	less than 50000	32 028	384,34	2,02	1,26	1,1	9	OK
SUM149PT	sorted cells	first sorting	D5	CD24+CD44+	less than 50000	20 908	250,9	2	1,01	1,3	9	OK
SUM159PT	sorted cells	first sorting	D5	CD24-CD44-	less than 50000	3 187	38,24			1,5	9	OK
SUM159PT	sorted cells	first sorting	D5	CD24-CD44+	less than 50000	40 184	482,21	1,97	1,85	1,3	8,8	OK
SUM159PT	sorted cells	first sorting	D5	CD24+CD44+	less than 50000	35 972	431,66	1,98	0,92	1,5	8,5	OK
SUM149PT	sorted cells	first sorting	D57	CD24-CD44-	less than 50000	3 180	38,16			1,3	8,6	OK
SUM149PT	sorted cells	first sorting	D57	CD24-CD44+	less than 50000	12 892	154,7	2,2	1,2	1,3	8,9	OK
SUM149PT	sorted cells	first sorting	D57	CD24MECD44M	less than 50000	16 352	196,22	2,1	1,54	1,3	8,9	OK
SUM149PT	sorted cells	first sorting	D57	CD24+CD44+	less than 50000	16 044	192,53	1,9	1,47	1,2	8,6	OK
SUM159PT	sorted cells	first sorting	D57	CD24-CD44-	less than 50000	16 975	203,7			1,4	5,7	OK
SUM159PT	sorted cells	first sorting	D57	CD24-CD44+	less than 50000	29 852	358,22	1,9	1,69	1,4	8,7	OK
SUM159PT	sorted cells	first sorting	D57	CD24+CD44+	less than 50000	29 936	329,3	1,5	0,57	0,2	2,6	~
SUM159PT	sorted cells	second sorting from CD24+/CD44+	D25	CD24-CD44-	less than 50000	11 443	137,32			0,1	5,8	OK
SUM159PT	sorted cells	second sorting from CD24+/CD44+	D25	CD24-CD44+	less than 50000	47 615	571,38	1,9	1,67	1,4	9	OK
SUM159PT	sorted cells	second sorting from CD24+/CD44+	D25	CD24+CD44+	less than 50000	54 500	654	1,9	1,32	1,4	9	OK
SUM159PT	sorted cells	second sorting from CD24-/CD44+	D25	CD24-CD44-	less than 50000	13 606	149,67			1,4	6,9	OK
SUM159PT	sorted cells	second sorting from CD24-/CD44+	D25	CD24-CD44+	less than 50000	28 653	315,18	1,8	1,33	0,9	6,2	OK
SUM159PT	sorted cells	second sorting from CD24-/CD44+	D25	CD24+CD44+	less than 50000	23 560	259,16	1,9	1,75	1	6,9	OK
SUM159PT	sorted cells	second sorting from CD24-/CD44-	D25	CD24-CD44-	less than 50000	23 603	283,24			1,2	7,6	OK
SUM159PT	sorted cells	second sorting from CD24-/CD44-	D25	CD24-CD44+	less than 50000	36 510	438,12	1,8	1,74	1,3	8,7	OK
SUM159PT	sorted cells	second sorting from CD24-/CD44-	D25	CD24+CD44+	less than 50000	36 725	440,7	1,9	0,82	1,4	8,7	OK
SUM149PT	sorted cells	second sorting from CD24-/CD44-	D25	CD24-CD44+	less than 50000	19 085	229,02	1,7	1,11	1,5	9	OK
SUM149PT	sorted cells	second sorting from CD24-/CD44-	D25	CD24+CD44+	less than 50000	17 250	189,75	1,8	1,32	1,3	8,8	OK

Table A.1 : Data list of cell lines and repopulating cells samples whose transcriptomic analysis was performed. - Part I

name cell line	sorted or not sorted cells	number of sorting	days of culture after sorting	name of cells subpopulation	number of cells per tube	concentration (pg/ μ l) calculated by Agilent	amount of RNA (ng) calculated by Agilent	ratio 260/280	ratio 260/230	ratio 285/185	RIN	Conclusion about RNA extraction quality
SUM149PT	sorted cells	second sorting from CD24-/CD44+	D25	CD24-CD44-	less than 50000	10 390	124,68			1	8	OK
SUM149PT	sorted cells	second sorting from CD24-/CD44+	D25	CD24-CD44+	less than 50000	31 575	378,9	1,7	1,45	1,2	9	OK
SUM149PT	sorted cells	second sorting from CD24-/CD44+	D25	CD24+CD44+	less than 50000	23 195	278,34	1,8	1,45	1,2	9,1	OK
SUM149PT	sorted cells	second sorting from CD24MED/CD44MED	D25	CD24-CD44-	less than 50000	10 839	130,07			1,1	7,6	OK
SUM149PT	sorted cells	second sorting from CD24MED/CD44MED	D25	CD24-CD44+	less than 50000	25 080	300,96	1,7	0,56	1,2	9,1	OK
SUM149PT	sorted cells	second sorting from CD24MED/CD44MED	D25	CD24+CD44+	less than 50000	25 140	301,68	1,7	1,62	1,5	8,8	OK
SUM149PT	sorted cells	second sorting from CD24+/CD44+	D25	CD24-CD44+	less than 50000	18 585	223,02	1,7	0,44	1,3	9,2	OK
SUM149PT	sorted cells	second sorting from CD24+/CD44+	D25	CD24+CD44+	less than 50000	23 615	283,38	1,8	0,28	1,2	9,1	OK

Table A.2 : Data list of cell lines and repopulating cells samples whose transcriptomic analysis was performed. - Part II

A.1.2 Supplementary data associated to part “strategies proposal to search social rules that govern cell phenotypic changes”

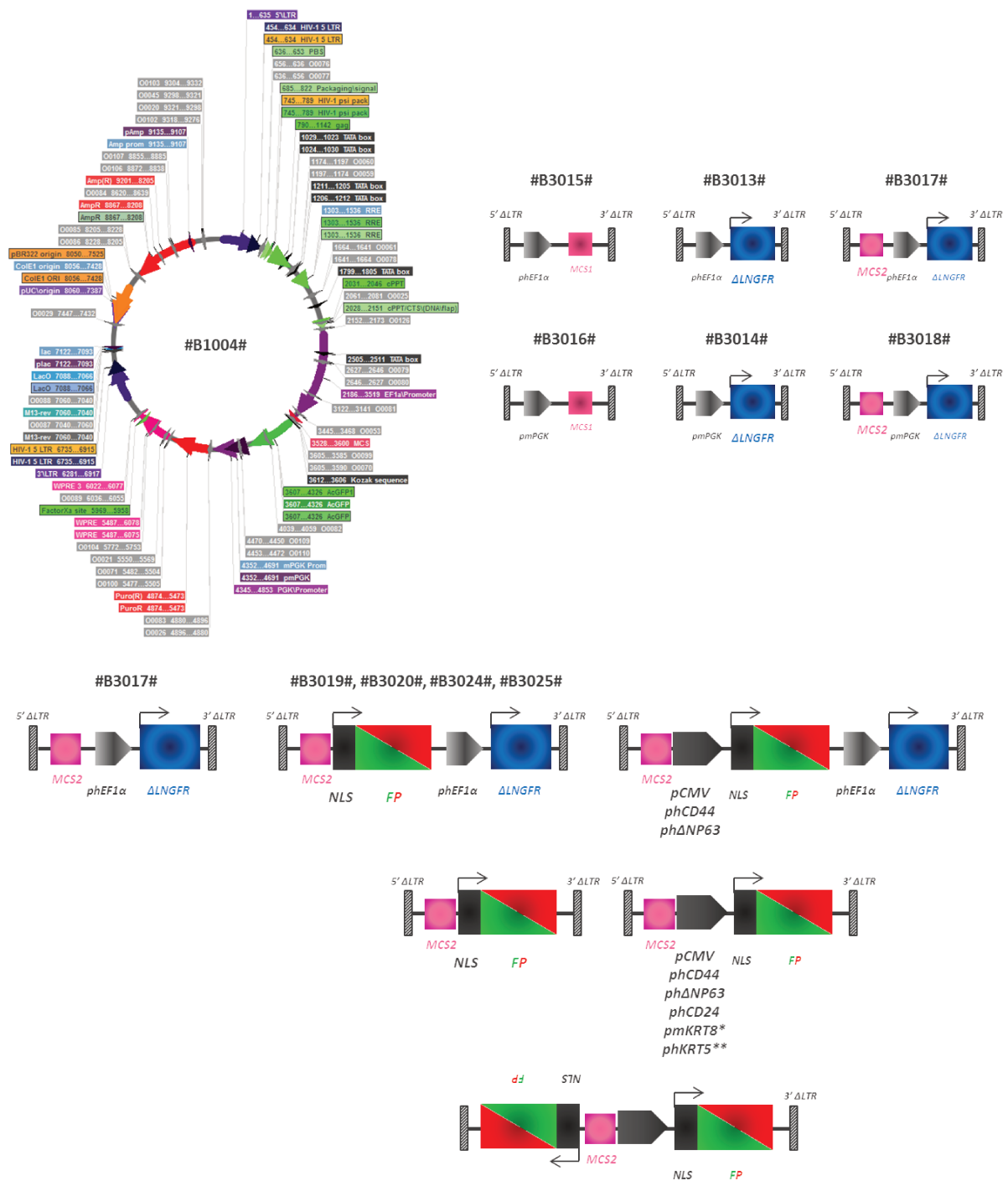


Figure A.16 : Schematic representations of PhenotypeTracker lentiviral platforms.

A.2 Looking for heterogeneity and screening of breast cancer cell lines.

- In collaboration with Dr Didier Meseure and M. André Nicolas, Biopathology Departement, Institut Curie -

NAME	Tissue	Cell type	Disease	Tumor/Stage	Status
HCC937	mammary gland	Lymphoblast, Epithelial	TNM stage IIb, grade 3, primary ductal carcinoma	Yes	?
B720	mammary gland	Epithelial	carcinoma	Yes	?
MDAMB468	mammary gland/breast; derived from metastatic site; pleural effusion	?	adenocarcinoma	Yes	?
Site-3	mammary gland/breast; derived from metastatic site; pleural effusion	?	adenocarcinoma	Yes	?
MDAMB331	mammary gland/breast; derived from metastatic site; pleural effusion	Epithelial	adenocarcinoma	Yes	?
A851	SKIN/epidermis	?	epidermoid carcinoma	Yes	?
MCF10A	mammary gland/breast	Epithelial	fibrocystic disease	No	?
B7549	mammary gland/breast	Epithelial	ductal carcinoma	?	?
MDAMB436	mammary gland/breast; derived from metastatic site; pleural effusion	?	adenocarcinoma	No	?
MCF124	mammary gland/breast	Epithelial; spontaneous	fibrocystic disease; focal areas of intraductal hyperplasia	No	?
HS578T	mammary gland/breast	Immortalization	carcinoma	No	?
HCC1148	breast; mammary gland/duct	Epithelial	TNM stage IIIA, grade 3, primary ductal carcinoma	?	?
HCC1187	mammary gland/breast	Epithelial	TNM stage IIIA, grade 3, primary ductal carcinoma	?	?
HCC70	breast; mammary gland/duct	Epithelial	TNM stage IIIA, grade 3, primary ductal carcinoma	?	?
18985	mammary gland/breast	Epithelial	?	?	?
MDAMB157	mammary gland; breast; mediulla	?	medullary carcinoma	Yes	?
MCF7	mammary gland; breast; derived from metastatic site; pleural effusion	Epithelial	adenocarcinoma	?	?
PMD42	mammary gland/breast	?	carcinoma	?	?
MDAMB53	mammary gland/breast; derived from metastatic site; pericardial effusion	?	metastatic carcinoma	No	?
ZR75-1	mammary gland/breast/duct; derived from metastatic site; ascites	Epithelial	ductal carcinoma	Yes	?
HCC1954	mammary gland; breast/duct	Epithelial	TNM stage IIIA, grade 3, ductal carcinoma	?	?
T47D	mammary gland; derived from metastatic site; pleural effusion	Epithelial	Infiltrating ductal carcinoma of the breast	?	?
HCC38	mammary gland; breast/duct	Epithelial	TNM stage IIb, grade 3, primary ductal carcinoma	?	?
HCC359	mammary gland; breast	Epithelial	TNM stage IV, grade 3, primary metastatic carcinoma	?	?
MDAMB361	mammary gland/breast; derived from metastatic site; brain	?	adenocarcinoma	?	?
HCC20	breast; mammary gland/duct	Epithelial	TNM stage IIIA, grade 3, primary ductal carcinoma	?	?
HCC428	mammary gland/breast; derived from metastatic site; adenocarcinoma and pleural effusion	Epithelial	TNM stage IV, grade 4, adenocarcinoma	?	?
MDAMB34V	mammary gland; breast/duct; Derived from metastatic site; pleural effusion	?	ductal carcinoma	?	?
HCC959	mammary gland; breast/duct	epithelial, lymphoblast	TNM stage IIIA, grade 3, primary ductal carcinoma	?	?
B7474	mammary gland; breast/duct	?	ductal carcinoma	Yes	?
MDAMB231, MBI	?	?	?	?	?
MCF10A	mammary gland; breast	Epithelial	fibrocystic disease	No	?
MDAMB15	mammary gland/breast; derived from metastatic site; pleural effusion	?	adenocarcinoma	No	?
MCF124					
MCF147	mammary gland; breast	Epithelial	Normal; reduction mammoplasty from a nulliparous patient with fibrocystic breast disease that contained focal areas of intraductal hyperplasia	?	?
SUM149P1	mammary gland; breast	Epithelial; Luminal	Primary Inflammatory Breast cancer	No	?
SUM159P7	mammary gland; breast	Epithelial; Luminal	Anaplastic Carcinoma; primary tumor	Yes	ER negative, PR negative

Table A.3 : Data list of breast cancer cell lines screened through immunostaining procedures to semi-quantitatively study the level of expression of 6 markers of interest

A.2.1 Breast cancer cell line screening for KRT5/KRT6 markers

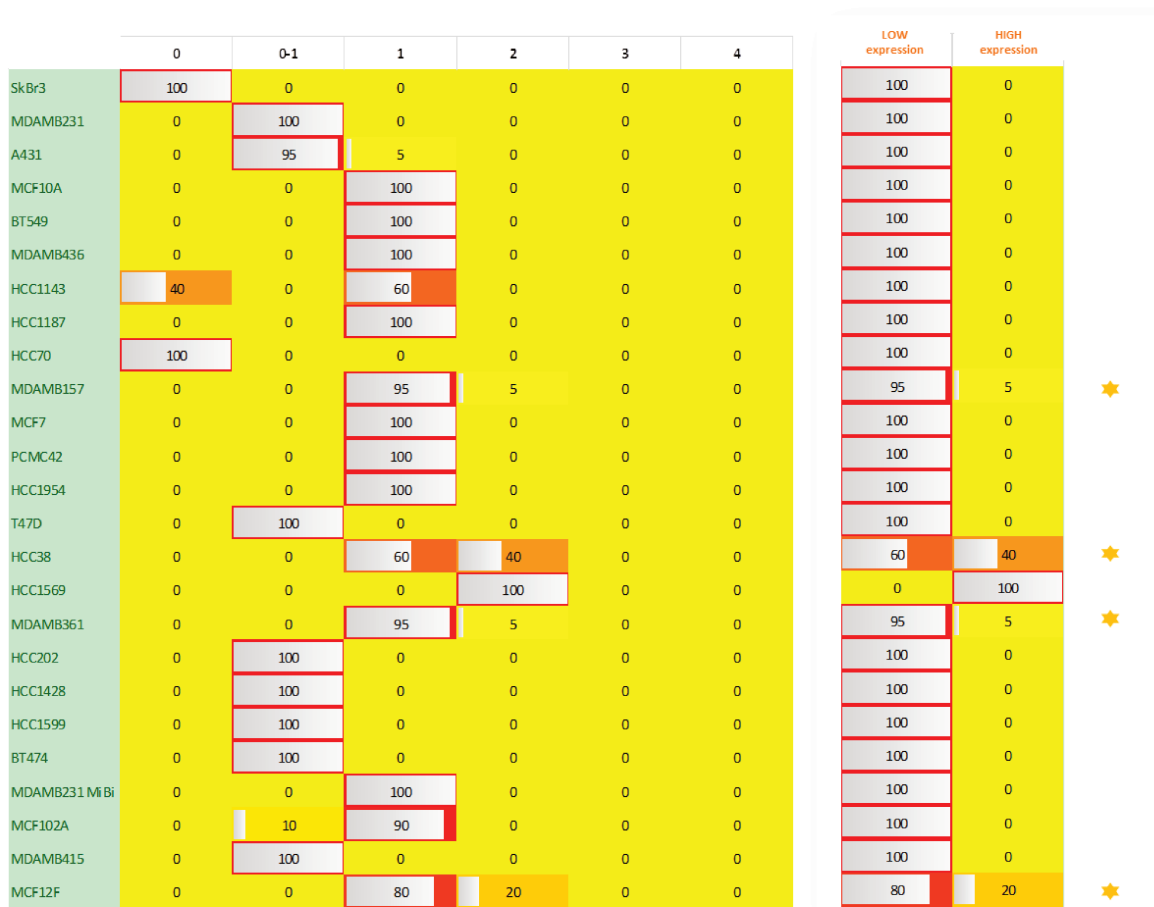


Figure A.17 : Large scale screening of breast cancer cell lines for their expression of KRT5 and KRT6 biomarkers

A.2.2 Breast cancer cell line screening for KRT8/KRT18 markers

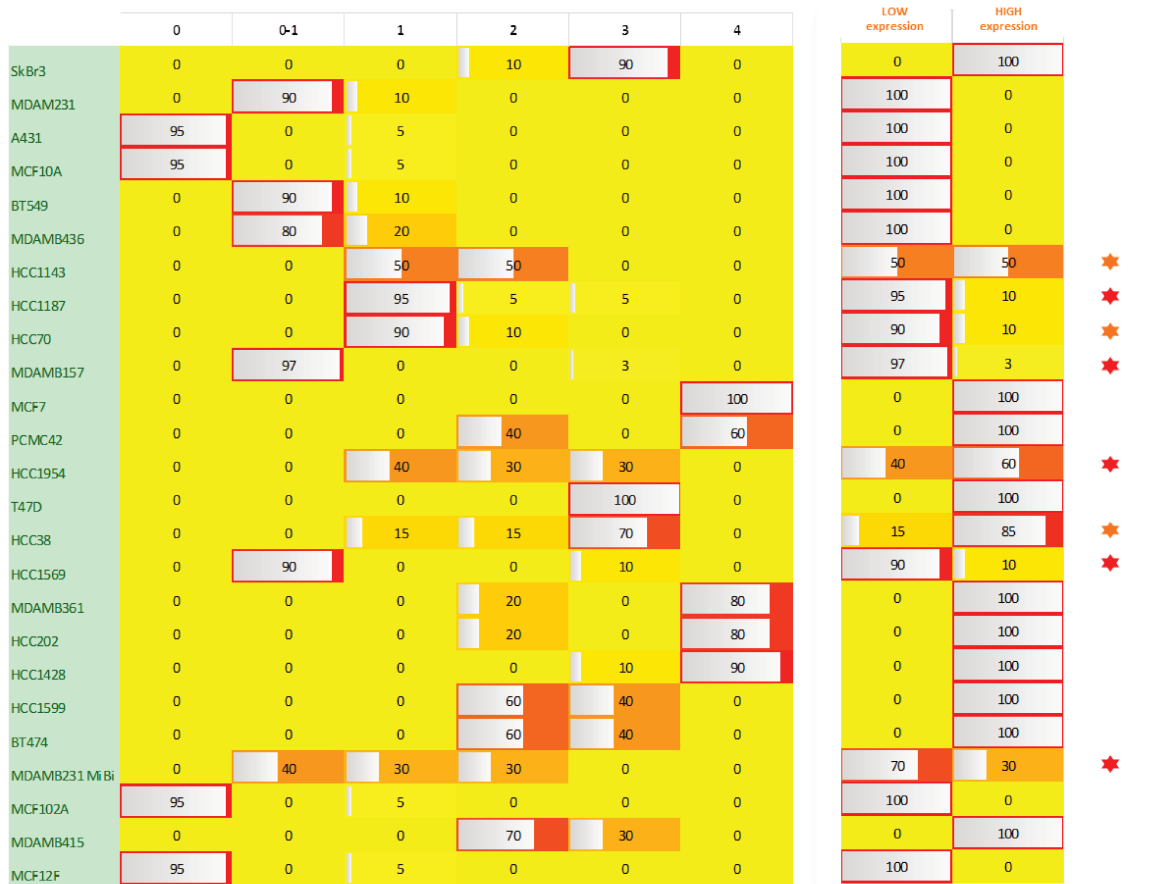


Figure A.18 : Large scale screening of breast cancer cell lines for their expression of KRT8 and KRT18 biomarkers

A.2.3 Breast cancer cell line screening for KRT14 marker

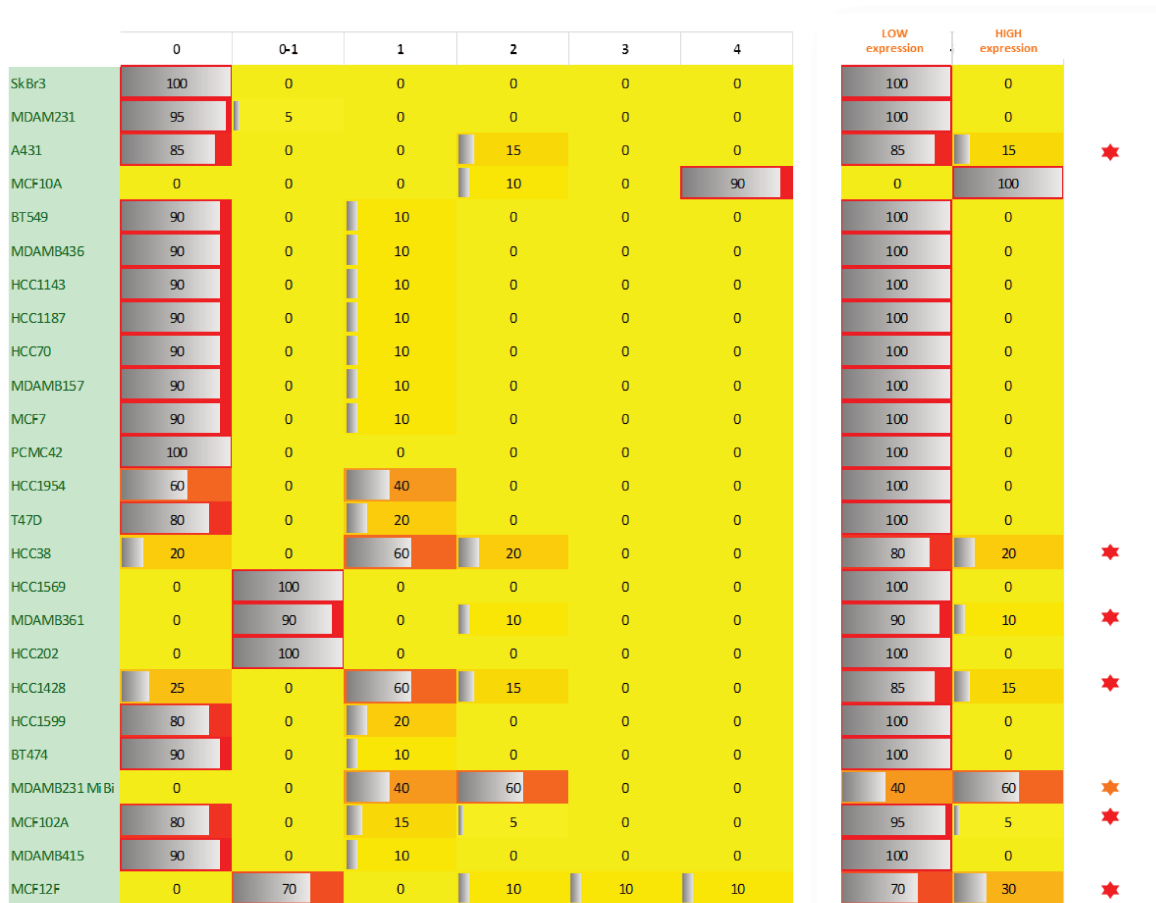


Figure A.19 : Large scale screening of breast cancer cell lines for their expression of KRT14 biomarkers

A.2.4 Breast cancer cell line screening for CD44 marker

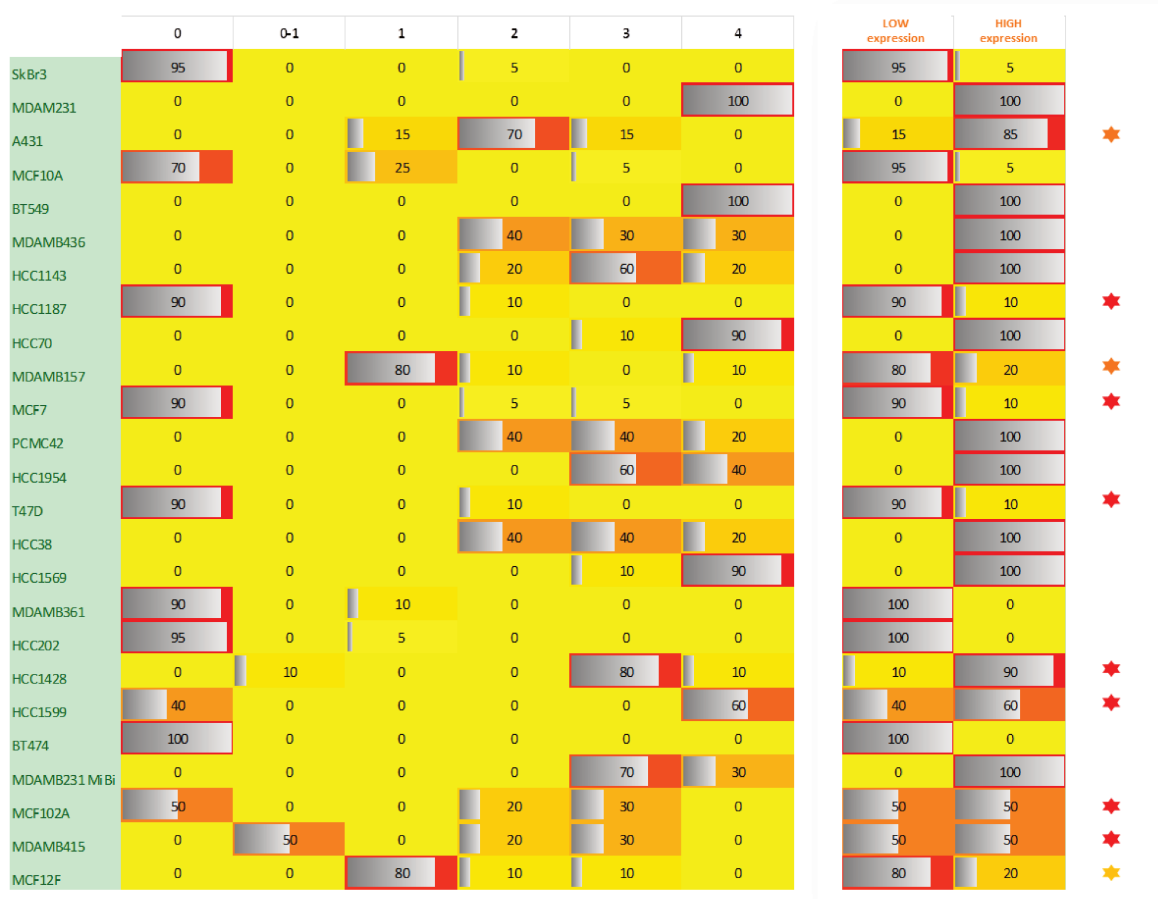


Figure A.20 : Large scale screening of breast cancer cell lines for their expression of CD44 biomarkers

Appendix **B**

Methodologies developed to segment cells edges from confocal images

« Les images vraies, les images pures de vérité trouvent asile dans l'écriture, dans la compassion de solitude de celui qui écrit. »

- Christian Bobin - L'inespérée

- This work regarding the segmentation of cells nuclei and edges have been performed under the close collaboration with Romain Cendre from the CIC-IT Nancy Center -

Our work on cells segmentation is based upon existing data taken through a confocal microscope (Zeiss LSM NLO710). These data are substantial : there exists almost 300 datasets, comprised of various images taken at different frames and modalities, all files contained within a .lsm format. Each set of pictures will involve the following modalities (*cf.* fig. B.1 on the next page subfig. A) :

- Polarized light, to observe cells edges
- Nucleus staining with DAPI
- CD24-PE staining
- And CD44-APC staining

The purpose of this adjacent work is to develop a segmentation technique capable of isolating each cell of the overall dataset we have already acquired. These data consist of almost 300 different datasets.

These acquired data enable the extraction of quantitative information related to two markers CD24 and CD44. This extraction of information necessitates to isolate each cell present within all datasets, more commonly known as segmentation. This segmentation can be performed in various ways including manual one, however we would like to avoid it for obvious reasons related to the amount of data.

If we only consider data linked to our analysis modalities, namely to recall : the polarized light and DAPI stained pictures, we could isolate some obvious problems. The images associated

with DAPI staining are a top quality and we will allow to easily isolate cell's nucleus. In contrast, polarized data, which remains to cells boundaries, are much more problematic. Here is a short list related to roaming problems :

- Noise of the background is quite important.
- Picture background is globally non-homogenous.
- There is no high level of contrast between cells signal and background signal.

B.1 Automatic or manual cells segmentation, principle of image segmentation

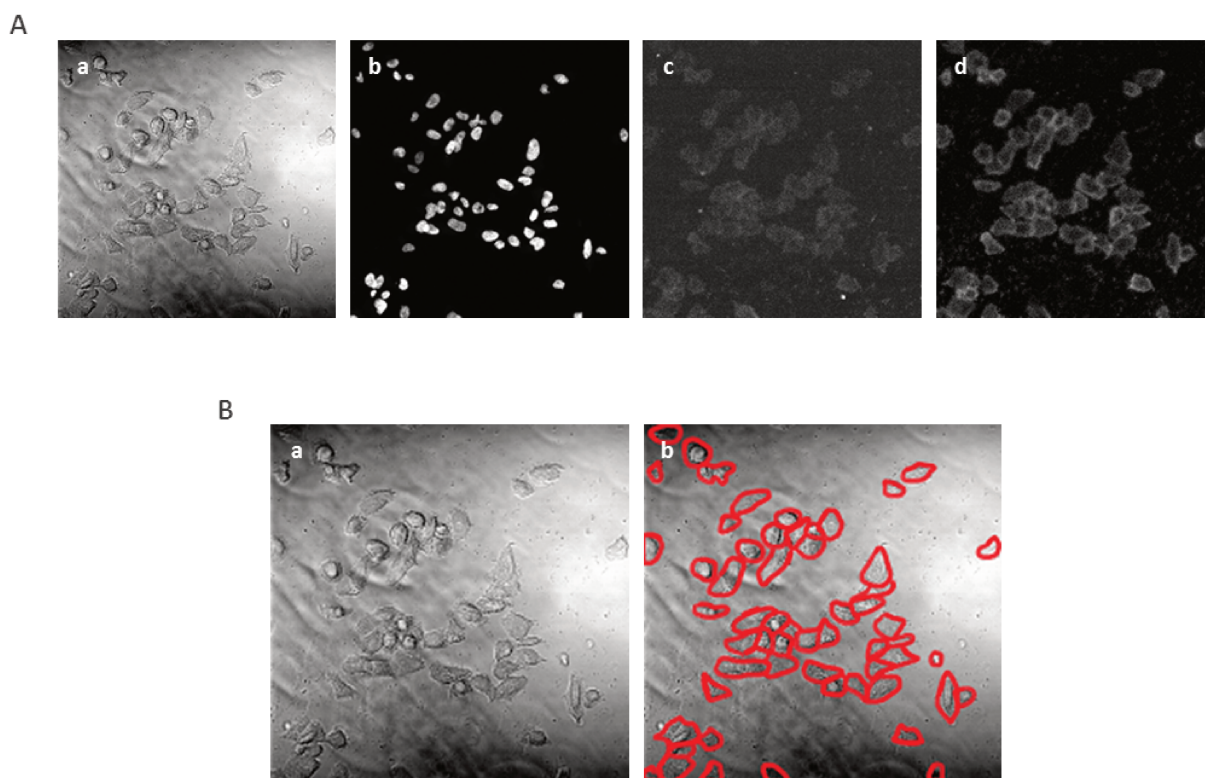


Figure B.1 : Principle of single cell segmentation.

A. Non calibrated pictures from confocal microscopy (Zeiss, LSM.NLO710) (a) DIC image, (b) DAPI, (c) CD24 and (d) CD44 imaged pictures. B. Example of expected cells' segmentation (a) from original image and (b) segmented picture manually obtained.

In computer science imaging, segmentation consists of grouping in a same area, pixels which share a link. This connection may be either from the strength of the different pixels or from strength of the shapes they may serve in the image. In our case, the purpose of this tool is to detect and delineate individual cells which compose a microscopic picture (*cf.* fig. B.1 subfig. B).

The interests of such method are manifold. One of them is, obviously, a significant time saving compared to manual segmentation. As the number of available data is significant (approximately 1500 images), the gain is much more consistent. The second point as to deal with the reproducibility of the analysis. Indeed, an algorithm, as opposed to man, has a constant behavior over time and avoids unintentional human error or linked to the human ability to interpret information.

B.2 Segmentation of cells nuclei

This segmentation is the first phase of our work, the nuclei of data are very useful in our situation. Indeed, apart from the quality of these images, the nuclei appear more independent than cells, less intertwined. These nuclei may be used to identify the number of cells and the body that are associated with them. These nuclei may also be the starting point for segmentation algorithms.

The DAPI images, although noisy have a lot of contrast. It is easy to isolate each of them with a human eye. Besides, we have 3D information that we will use to optimize our detection (*cf.* fig. B.2 subfig. A). The first observation that we can establish is the important distinction between the bottom and our cell nuclei forming two classes of data (the almost black bottom and the shape around the white values). So we easily obtain a separation between these two classes by Otsu thresholding (*cf.* fig. B.2 subfig. Ba).

The resulting binary image has slight inconsistencies (noise, holes ...) we can fill with closing operations (fill our areas cores) and opening (noise suppression) (*cf.* fig. B.2 subfig. Ab). This step is done, we can proceed to a label, that is to say, any pixel group belonging to the same group. This label is used by custom connectivity 4 or 8 (corresponding to the consideration of neighbors). In our case we will use 3D information, for this we will use a 26 connectivity, allowing us to take into account the neighboring contiguous sections (*cf.* fig. B.2 subfig. Bb).

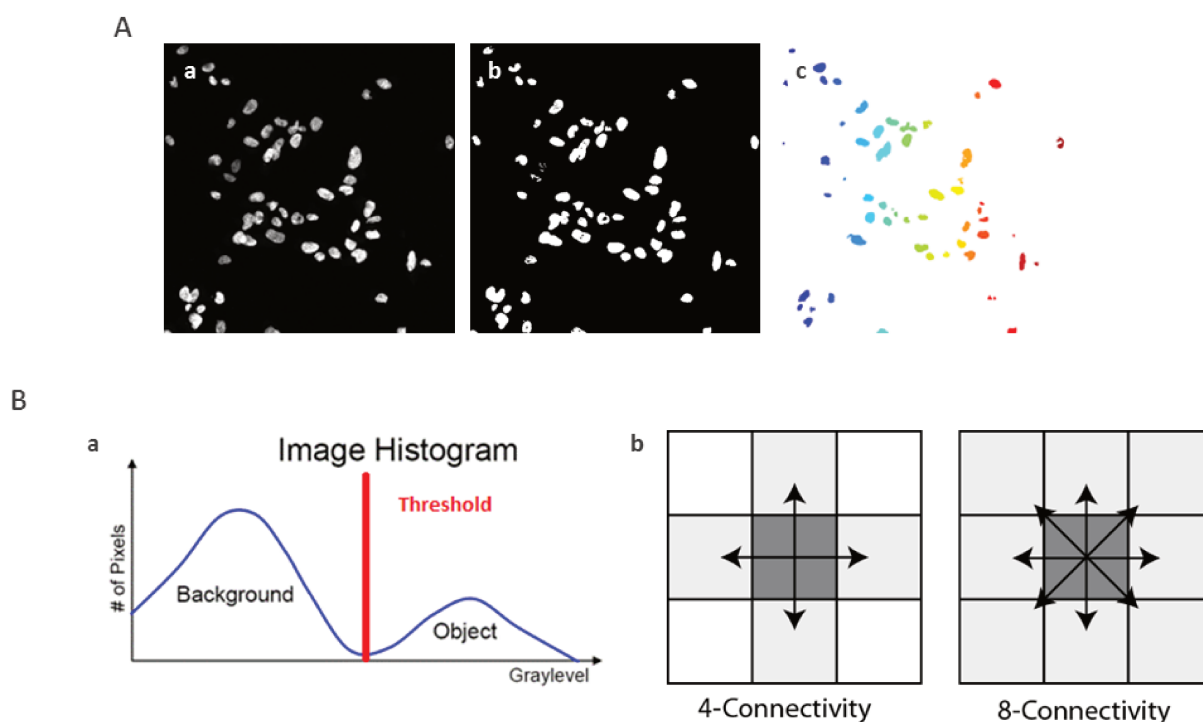


Figure B.2 : Principle of nuclei segmentation from DAPI pictures

A. Segmentation of cells nuclei from (a) Image of origin (b) Image obtained from the Otsu binarization and (c) from the 3D labelization. B. Principles of nuclei detection (a) Otsu threshold (b) Schematic explanation of 4- and 8-degree of pixels connectivity

The challenge seems successful for this first step ; we get a result consistent with our expectations. Although all nuclei are not fully detected, each of them seems to be at least isolated (*cf.* fig. B.2 subfig. Ac).

B.3 Segmentation of cells edges

We will strive in this part to segment the various elements in our possession while being aware of the problems mentioned above. We will discuss the second part from different angles corresponding to various attempts we have made to realize this work.

B.3.1 Thresholding methodology

In this section we will experiment with a manual and automatic value approach, even though automatic threshold is effective for images with sharp contrast and consistent values, which our data seems to lack.

The quite mediocre results below demonstrate the failure to achieve the desired result with such technique (*cf.* fig. B.3). Supposing that an “ideal” threshold can be found manually or automatically, the image would be affected by a partial decrease in brightness on a certain area of our images.



Figure B.3 : Segmentation of cells edges through threshold methods
(a) DIC Image of origin (b) Segmented image obtained with a low-threshold and (c) with a high-threshold

We will focus in the next section to elements which allow us to correct our source values.

B.3.2 Filtering methodology

The primary benefit of the polarized light (pseudo-DIC) is “to highlight the outline of our cells”. Thus the use of filters dedicated to edge detection would highlight the outlines of our cells. We will implement in MATLAB™ the major existing techniques to date : Prewitt, Sobel and Canny. The data obtained seem rather conclusive for these alternatives (*cf.* fig. B.4). We get a rough detection our cells, diffuse and fragmented. We could use the results from these methods as material for subsequent detection processes.

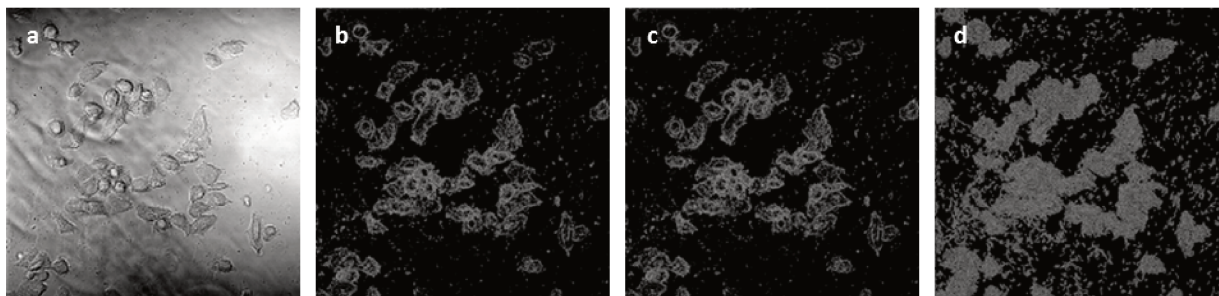


Figure B.4 : Segmentation of cells edges through filters methods
(a) DIC Image of origin (b) Segmented image obtained with a filter of Prewitt (c) of Sobel and (d) of Canny

B.3.3 Spreading methodology

This third used technique is based on region growth. It is often used on outlined shapes. We therefore used the propagation zone provided by the segmentation of our nuclei.

The expected spread does not happen and the region stays stable or even decreases in regards to the nuclei masks provided (*cf.* fig. B.5). This method is also very time and CPU consuming. Due to this, we can dismiss this method.

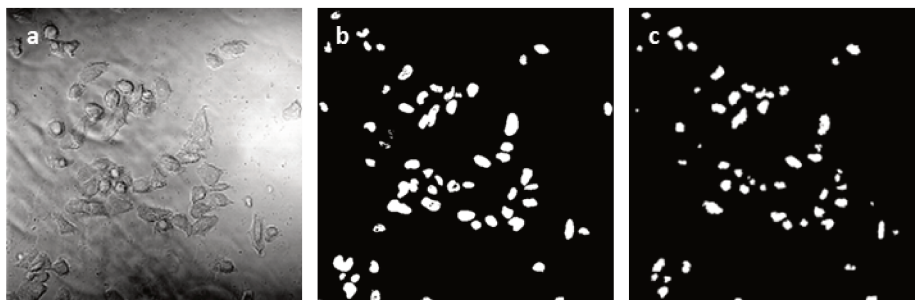


Figure B.5 : Segmentation of cells edges through spreading methods.
 (a) DIC Image of origin (b) Mask used for spreading method (c) Segmented image obtained with spreading method

B.3.4 Deviation around the mean methodology

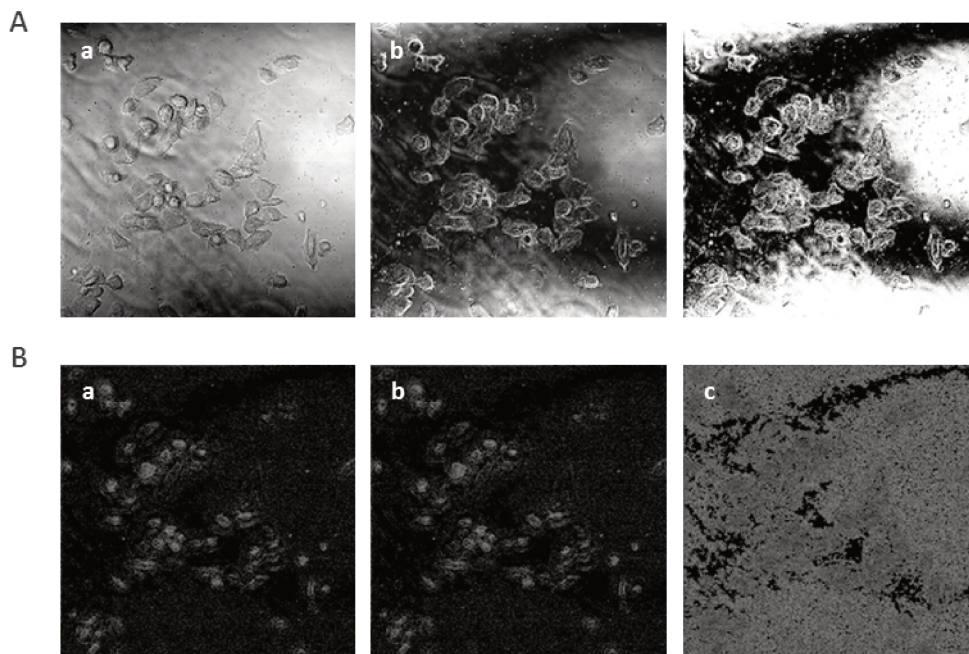


Figure B.6 : Segmentation of cells edges through deviation around the mean method.
 A. Segmentation of cells edges from (a) Image of origin (b) Background subtraction (c) Looking of a threshold. B. Detection of cells edges within subtracted picture (a) with a filter of Prewitt (b) of Sobel and (c) of Canny

We started from the premise that only cells generate information on our image. Thus, theoretically, it would be sufficient to perform the computation of a value representative of this noise. We calculated the average of our pictures that we subtracted from each pixel of the image

and determined an absolute value to remove this theoretical noise.

However, the result is below all expectations, and this mainly due to lack of homogeneity of the source image (*cf.* fig. B.6 on page 237). This result totally undermines this methodology as evidenced by the images.

B.3.5 Correction of background lack of homogeneity methodology

We note that a major limitation of our segmentation can be characterized by an in-homogeneity of the noise. To compensate for the background, we must first assess this noise in order to balance the values at each pixel. The first technique used is the use of a filter medium therefore sufficient to assess a representative point value from the bottom of our image (convolution filter with unitary coefficients). This technique was abandoned in favor of a noise detection technique proposed by MATLAB™ based on the opening and closing functions of MATLAB™ (often used on binary images for morphological processing).

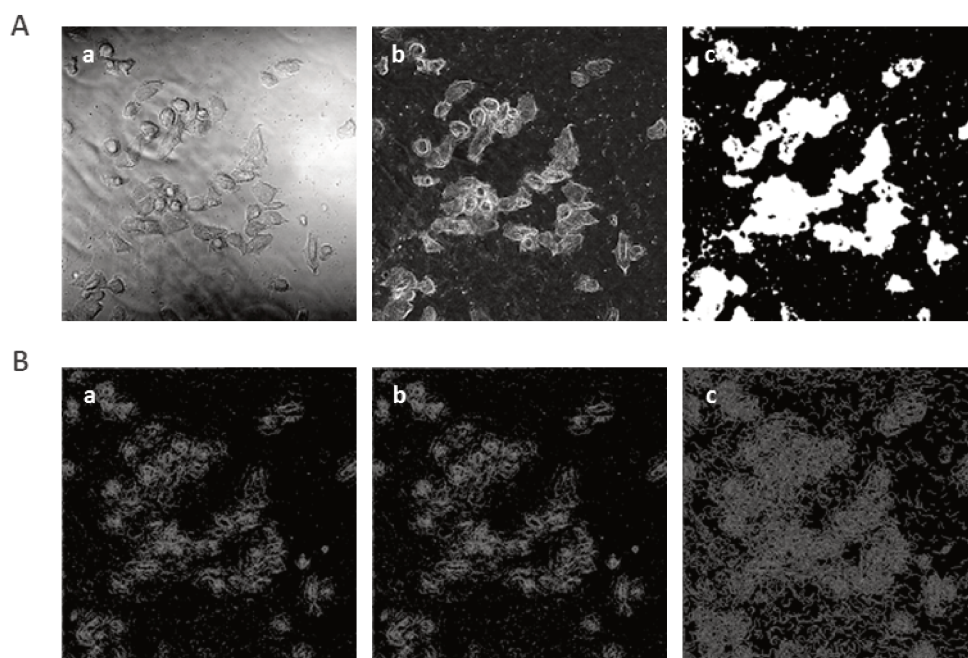


Figure B.7 : Segmentation of cells edges through correction of picture background non-homogeneity.

A. Segmentation of cells edges from (a) Image of origin (b) Background subtraction (c) Looking of a threshold. B. Detection of cells edges within subtracted picture (a) with a filter of Prewitt (b) of Sobel and (c) of Canny

The results of this operation are positive, areas of overexposure are homogenized as we wanted (*cf.* fig. B.7). The thresholding and edge detection operations bring a real gain compared to our first attempt without correction of the source images. We still have to isolate each cell, which a region growth technique could help us with.

B.4 Future perspective of cells segmentation

To conclude, all techniques relying on the raw data were a failure, only the algorithms based on edge detection were close to our expectations. In contrast, the image correction seems to provide optimal results. So we tried to isolate the data from our cells, an ineffective process in our situation,

but also to compensate the exposure of the image by removing the noise, this methodology seems promising.

B.4.1 Image homogenization

The currently applied technique seems indeed to bring good results based on simplistic resolution elements. Advanced algorithms could provide better uniformity of lighting and thereby segmentation.

B.4.2 Background subtraction

This methodology could be considered to refine the extraction of various regions, provided that the previous step provides results of high quality.

B.4.3 Edges detection and thresholding

Both of these techniques are generally possible if we improve the quality of images with the two processes mentioned above. However, on our raw data only the detection seems really appropriate, other advanced techniques might be effective. The technique of an adaptive thresholding, whose threshold value would vary locally, was considered but should not provide any additional benefit compared to the homogenization technique mentioned above.

Spreading and detection of crashes

Yet another attempt would be to maintain the growth principle in a more permissive manner, while managing potential collisions between different propagation points (nuclei). This technique could be implemented after thresholding on homogenized images.

Cells structure

To conclude on this outlook, it could be interesting to focus on the granularity of cell areas and attempt an extraction from such parameters

Validation of SUM149-PT and SUM159-PT cell lines, an epic adventure

«L'Homme et sa sécurité doivent constituer la première préoccupation de toute aventure technologique.»
- Albert Einstein -

17% to 36% of cell cultures in use are cross-contaminated either by intraspecies contamination (unrelated cells from the same species) or interspecies contamination (cells from another species) [23, 196]. As an example, ECV304 cell line, which was widely considered a model cell line for endothelial cells, has also been reported to be a subclone of the T24 human bladder carcinoma cell line [169].

Quality control for cell line cross-contamination is not regularly practiced in most laboratories, despite the obvious importance and frequency of the problem. This may be due to several factors, including :

- Failure to appreciate the occurrence of cross-contamination because it is often not visually recognizable.
- Unsuitability of detection methods for extensive cell-culture screening.
- Lack of awareness of cross-contamination due to insufficient reporting.

However, in recent years, scientific publications, journals and research institutions require that the biological material used during research to be certified according to its origin. To overcome this problem, several methods are available to establish cell authentication and provenance, as for example :

- Human Leukocyte Antigen (HLA) typing (as used in patient-donor tissue typing)
- Karyotyping (visualization of chromosomal pattern and architecture) or specific mutation sequencing
- Isoenzyme typing
- Perhaps the most sophisticated, DNA fingerprinting : Short Tandem Repeats (STR) profiling [182, 352]

STR profiling is now a well-established and well-regarded technique for unambiguously characterizing a number of different loci in the human genome. Depending on Connexin et al. point of view, STR profiling represents the optimal candidate at this time and it has been frequently utilized in forensics, population genetics, genetic genealogy, and cell line origin validation [17, 39, 317]. It is now advisable that cells should be put to this authoritative reference standard test before the beginning of a project and at different interval during that time in order to effectively eliminate scientific misrepresentation due of cell line cross-contamination [59, 202, 256].

C.1 Screening for SUM149-PT and SUM159-PT safety, batch number 1

First batches of SUM149-PT and SUM159-PT cell lines, received on November 2012, were a kind gift from collaborators at the “Institut de Recherche en Cancérologie de Montpellier” (France).

In March 2014, as we wanted to validate the origin of these current cells, we looked at their STR data to do so. However, STR characterization for SUM149-PT and SUM159-PT cell lines were neither available through diverse cell line collections (American Type Culture Collection (ATCC), German Collection of Microorganisms and Cell Cultures (DSMZ), European Collection of Cell Cultures (ECACC), RIKEN Bioresource Center Cell Bank (RIKEN)) nor within Asterand™, the custodian of SUM derivative cell lines.

So we searched in literature a couple of mutations that could be used to discriminate both SUM149-PT and SUM159-PT from other cell lines, and thereby validating that we were working for over a year on the right cells. Three mutations were retained and searched by pyrosequencing within the genome of these two cell lines [20, 80, 135] :

- Harvey Rat Sarcoma Viral Oncogene Homolog (HRAS), point mutation 81 T>C
- HRAS, point mutation 35 G>A
- BRCA1, point mutation 2288delT

Thus, each of SUM149-PT and SUM159-PT cell lines present some genetic features as described in the attached table :

	SUM149-PT	SUM159-PT
HRAS 81 T>C	Mutated	Mutated
HRAS 35 G>A	Wild Type	Mutated
BRCA1 2288delT	Mutated	Mutated

Thus, in a first time, sequencing of previous markers has reported that :

- SUM149-PT cell line from batch 1 did not carry correct genetic characteristics regarding HRAS 35G>A and BRCA1 2288delT mutations (*cf.* fig. C.1 on the next page subfig. Aa).
- SUM159-PT cell line from batch 1 present genetic characteristic as reported by the literature (*cf.* fig. C.1 on the facing page subfig. Ba).

In a second time, the level of CD24, CD44 and CD326 markers expressed by each cell populations clustering has been measured by flow cytometry and compared to the data based on literature [113]. The three cellular phenotypes are thus represented by [89] :

- CD24+/CD44+/CD326- and CD44+/CD24+/CD326+ populations, which are highly enriched in luminal cells.
- CD24-/CD44+/CD326- population which is enriched in basal cells.
- CD24-/CD44+/CD326+ which represents the stem-like population.

For SUM149-PT line batch 1, we find more stem-like cells and less luminal cells. And for SUM159-PT cells, we found less basal cells than the standard.

On the other hand, a CD24-/CD44- cell population was present in from 55% for SUM149-PT and 25% for SUM159-PT whereas this population was not reported in the graph of flow cytometry previously published [89, 113]. Following these results (both genetic and phenotypic) and having no knowledge regarding the culture conditions within the previous laboratory, we concluded that SUM149-PT and SUM159-PT cell lines from batch 1 did not carry the characteristics of the cell lines derived by Stephen Ethier. Thus, we had to give up the previously accomplished work from such two cell lines of batch 1 and search for a new batch of conforming cells.

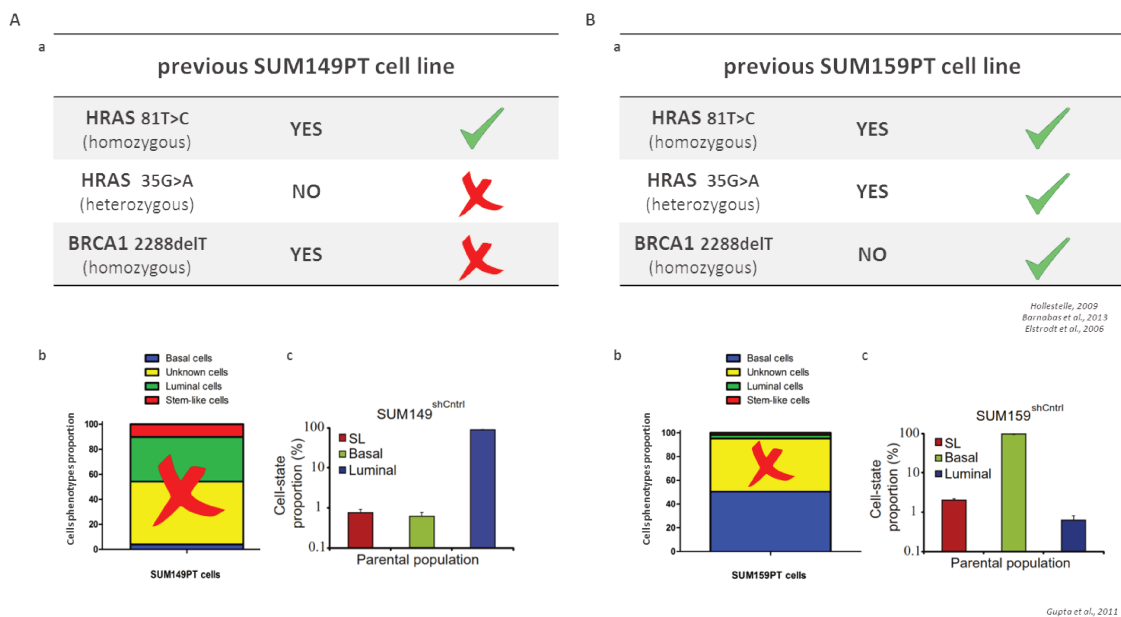


Figure C.1 : Sequencing of *HRAS* and *BRCA1* specific mutations on SUM149-PT (A) and SUM159-PT (B) cell lines, batch 1.

(a) Characterization of genetic mutations. (b) Phenotype characterization regarding the level of expression of CD24, CD44 and CD326 markers of cells from batch 1. (c) from the literature [113]

C.2 Authentication of SUM149-PT and SUM159-PT, batch number 2

A batch number 2 of SUM149-PT and SUM159-PT was delivered to us from Asterand (official custodian lines derived by Stephen Ethier et al.) in October 2014. Relying on the recently characterized STR profiling, we validated the origin and the integrity of these two cell lines. Thereby, all the data presented in this manuscript is the result of work exclusively carried out with cell lines from batch number 2.

Meanwhile, mycoplasma tests were performed once every three weeks on mother's cells and cell lines derived from manipulations.

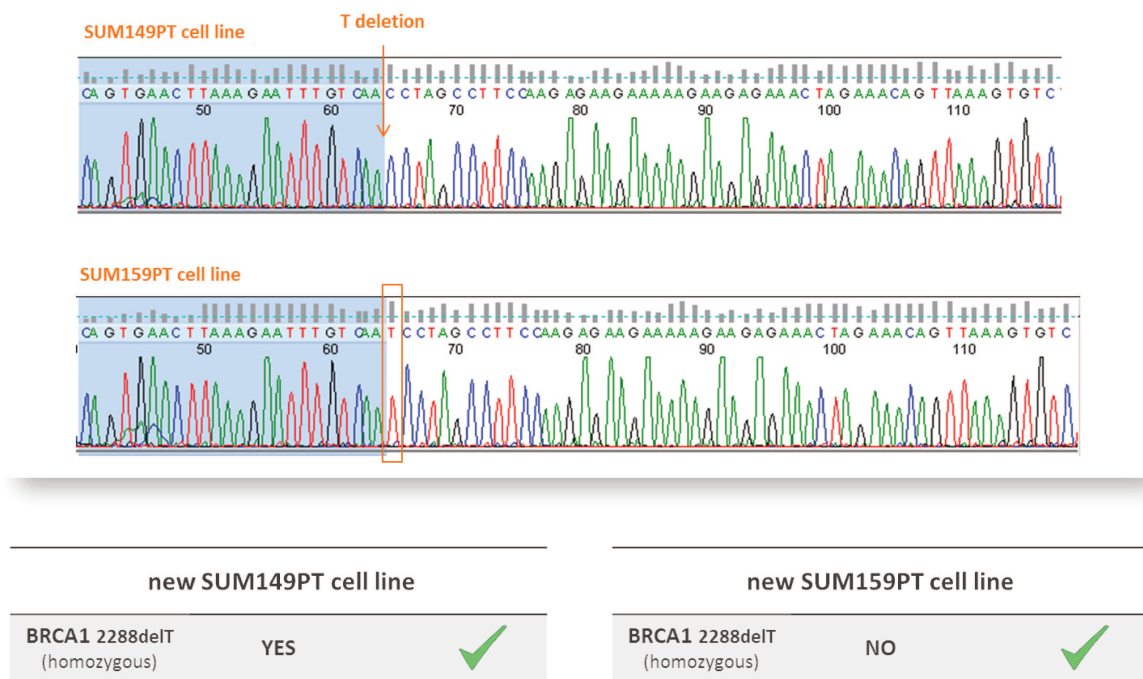


Figure C.2 : Sequencing of BRCA1 specific mutation on SUM149-PT and SUM159-PT cell lines, batch 2.

A

SUM-149PT	ID: 28064A1	
Marker	Allele Size	Expected Alleles Asterand #: 28064A1
Amelogenin	X	X
CSF1PO	12	12
D13S317	12	12
D16S539	11	11
D5S818	11	11
D7S820	11	11
TH01	9.3	9.3
TPOX	9	9
vWA	16, 18	16, 18

B

SUM-159PT	ID: 28066A1	
Marker	Allele Size	Expected Alleles Asterand #: 28066A1
Amelogenin	X	X
CSF1PO	11	11
D13S317	12	12
D16S539	11	11
D5S818	11	11
D7S820	10	10
TH01	6, 7	6, 7
TPOX	8, 11	8, 11
vWA	16	16

Figure C.3 : STR screening of SUM149-PT and SUM159-PT cell lines, batch 2.

Appendix **D**

Protocols developed and used throughout the experimental work

D.1 Protocol – Cell culture – HEK293T

⚠ PAY ATTENTION TO WIPE EVERYTHING GOING UNDER THE PSM WITH 70% ETHANOL OR SURFANIOS AND CLEAN CAREFULLY THE HOOD AND THE TUBULURES WITH SURFA'SAFE FOR AROUND 10 MIN

Materials

Technical material

- Flasks for cells culture (TPP)
- Sterile tips WITH filters (StarLab)
- Gloves
- Sterile distilled Water (Gibco)
- Ethanol 100% and 70% or 0.025% Surfanios solution
- Cryotubes of 1.8 ml
- Mister Frosty, -20 °C with isopropanol inside
- DMSO
- Liquid nitrogen (cisterns are at the -1 floor)

Reagents

- DMEM + GlutaMAX, 4 °C, protected from light (Gibco, 61965-026)
- Heat-inactivated Fetal Bovin Serum, -20 °C (Gibco)
- Sodium pyruvate, 4 °C (Gibco, 11360-039)
- MEM NEAA, 4 °C (Gibco, 11140-050)

- Trypan Blue Stain 0.4%, RT (Sigma Aldrich) Working solution : 0.04% in PBS, 4 °C
- Calcium-magnesium-free Phosphate Buffer Saline pH 7.4 1X (Gibco)
- TrypLe Express - EDTA 0.02%, 4 °C (Gibco)
- Cell specific culture medium
- PenStrep, -20 °C (Invitrogen, 15070-063)

Solution

Cell culture specific medium for 500 ml

DMEM + GlutaMAX		434.5 ml
Fetal Bovine Serum		50 ml
Sodium Pyruvate	100 mM, 100X	5 ml
MEM NEAA	100X	5.5 ml
PenStrep	5000 units of penicillin and 5000 µg of streptomycin per ml, 100X	5 ml

D.1.1 Thawing of frozen cells

1. Wipe the frozen vial with 70% ethanol or with 0.025% surfanios solution
2. Quickly thaw the vial in a 37 °C water bath. Allow to thaw until a last crystal piece is left
3. Wipe the outside of the vial again with 70% ethanol and place it into a PSM
4. Use a 15 ml conical tube to resuspend the contents of the vial in 9 ml of pre-warmed to 37 °C growth medium. Make sure to wash out any remaining cells in the tube
5. Gently pipette the cell suspension a few times to create a homogenous solution
6. Take 100 µl of cell suspension and mix it with 100 µl of 0.04% Trypan Blue Stain. Proceed with assessing cell count and viability
7. Centrifuge the tube at 1000 rpm for 10 min at room temperature
8. Aspirate off the supernatant and resuspend the cells in appropriate volume of warmed, fresh medium in order to achieve a density of 1.0×10^{10} cells/ml.

⚠ SEED AT LEAST 1.5×10^6 CELLS IN A T25CM² FLASK AND 5×10^6 CELLS IN A T75CM²
FINAL TOTAL VOLUME FOR A T25CM² FLASK SHOULD BE 10 ML AND 25 ML FOR A T75CM²

9. Gently rock the culture flasks to evenly distribute the cells and place them into a 37 °C incubator at 5% CO₂
10. Completely change the culture medium the day after culture initiation

D.1.2 Culture and subculture of the cells

Maintaining the culture

1. Warm the appropriate amount of medium to 37 °C in a sterile container.
2. Remove the old medium, wash the cells with around 6 ml of PBS and replace it with the warmed, fresh medium and return the flask to the incubator at 37 °C, 5% CO₂

 CHANGE THE MEDIUM EVERY TWO DAYS

Cells subculturing


1. Discard the culture medium from the culture dish or flask
2. Lift up the cells with TrypLE Express using a sterile pipette tip (1 ml for a 3 cm dish or 2 ml for a T25/10cm dish or 5 ml for a T75)
3. Incubate at 37 °C, 5% CO₂ for 3–5 min
4. Check the detachment of the cells under an inverted phase-contrast microscope. The cells have to be in suspension.
5. Add culture medium with a sterile pipette tip (2 ml for a 3 cm dish or 6 ml for a T25/10cm dish or 10 ml for a T75) to inactivate the TrypLE express action
6. Gently pipette up and down to form a single cell suspension, transfer in a 15 ml sterile Falcon tube
7. Rinse the original dish or flask with 1–3 ml of culture medium and transfer in the same 15 ml sterile Falcon tubes
8. Take 100 µl of cell suspension and mix it with 100 µl of 0.04% Trypan Blue Stain. Perform a cell number and viability count
9. Centrifuge at 1000 rpm for 10 min at room temperature
10. Aspirate off the supernatant and resuspend the cells in appropriate volume of warmed, fresh medium in order to achieve a density of 1.0×10^6 cells/ml.

11. Distribute cells into new flasks

 SEED AT LEAST 1.5×10^6 CELLS IN A T25CM² FLASK AND 5×10^6 CELLS IN A T75CM².

FINAL TOTAL VOLUME FOR A T25CM² FLASK SHOULD BE 10 ML AND 25 ML FOR A T75CM²

12. Gently rock the culture flasks to evenly distribute the cells and place them into a 37 °C incubator at 5% CO₂
13. Completely change the culture medium the day after culture initiation

 DON'T FORGET TO WRITE ON THE FLASK, THE NAME OF THE CELL LINE, THE DATE, THE NUMBER OF PASSAGES, AND YOUR NAME.

PERFORM CELLS PASSAGES BEFORE CELLS REACH 90–95% CONFLUENCE. FOR HEK293T CELLS IT SHOULD BE EVERY 2–3 DAYS.

D.1.3 Freezing and cryopreservation of cells, DMSO 10%

1. Discard the culture medium from the culture dish or flask
2. Lift up the cells with TrypLE Express using a sterile pipette tip (1 ml for a 3 cm dish or 2 ml for a T25/10cm dish or 5 ml for a T75)
3. Incubate at 37 °C, 5% CO₂ for 3–5 min
4. Check the detachment of the cells under an inverted phase-contrast microscope. The cells have to be in suspension.
5. Add culture medium with a sterile pipette tip (2 ml for a 3 cm dish or 6 ml for a T25/10cm dish or 10 ml for a T75) to inactivate the TrypLE express action
6. Gently pipette up and down to form a single cell suspension, transfer in a 15 ml sterile Falcon tube
7. Rinse the original dish or flask with 1 to 3 ml of culture medium and transfer in the same 15 ml sterile Falcon tubes
8. Take 100 µl of cell suspension and mix it with 100 µl of 0.04% Trypan Blue Stain. Perform a cell number and viability count
9. Centrifuge at 1000 rpm for 6–8 min at room temperature
10. Aspirate off the supernatant and resuspend the cells in appropriate volume of warmed, fresh medium in order to achieve a density of 5.0×10^6 cells/ml.
11. Distribute 5×10^6 cells per cryotubes
12. Add 100 µl of DMSO into each cryotube. And **very quickly** rock the cryotubes and put them into the Mr Frosty. Place it at –80 °C.
13. Let the cryotubes into the –80 °C for 7 days and transfer them afterwards into the liquid nitrogen cisterns.

 **DON'T FORGET TO WRITE ON THE NOTEBOOK THE EMPLACEMENT OF YOUR TUBES**

D.2 Protocol – Cell culture – SUM149 PT and SUM159 PT 5%

From Asterand@Biotech

⚠ PAY ATTENTION TO WIPE EVERYTHING GOING UNDER THE PSM WITH 70% ETHANOL AND CLEAN CAREFULLY THE HOOD AND THE TUBULURES WITH SURFA'SAFE FOR AROUND 10 MIN

Materials

Technical Material

1. Flasks for cells culture (TPP)
2. Sterile tips WITH filters (StarLab)
3. Gloves
4. Sterile distilled Water (Gibco)
5. Ethanol 100% and 70%
6. Cryotubes of 1.8 ml
7. Mister Frosty, -20°C with isopropanol inside
8. DMSO
9. Liquid nitrogen (cisterns are at the -1 floor)

Reagents

1. Ham's F-12 Medium + L-Glutamine, 4°C , protected from light (Gibco, 11765-062)
2. Heat-inactivated Fetal Bovin Serum, -20°C (Gibco)
3. HEPES, RT (Sigma_H3375) Working solution (filtered) : 1 M, 23.83 g/100ml in distilled water, 4°C
4. Insulin, 4°C , protected from light (Sigma_I9278)
5. Hydrocortisone, RT, protected from light (Sigma_H4001) Working solution (⚠ DO NOT FILTER) : 1 mg/ml, 40 mg/20ml of ethanol 100% and 20 ml of distilled water, -20°C , protected from light
6. Trypan Blue Stain 0.4%, RT (Sigma Aldrich) Working solution : 0.04% in PBS, 4°C
7. Calcium-magnesium-free Phosphate Buffer Saline pH 7.4 1X (Gibco)
8. TrypLe Express - EDTA 0.02%, 4°C (Gibco)

Solution


Cell specific culture medium for 500 ml

Ham's F12 medium		469 ml
Fetal Bovine Serum		25 ml
HEPES (1 M)	10 mM	5 ml
Insulin (1 mg/ml)	5	250 μ l
Hydrocortisone (1 mg/ml)	1	500 μ l

 DO NOT USE ANTIBIOTIC

D.2.1 Thawing of frozen cells

1. Wipe the frozen vial with 70% ethanol
2. Quickly thaw the vial in a 37 °C water bath. Allow to thaw until a last crystal piece is left
3. Wipe the outside of the vial again with 70% ethanol and place it into a PSM
4. Use a 15 ml conical tube to resuspend the contents of the vial in 9 ml of pre-warmed to 37 °C growth medium. Make sure to wash out any remaining cells in the tube
5. Gently pipette the cell suspension a few times to create a homogenous solution
6. Take 100 μ l of cell suspension and mix it with 100 μ l of 0.04% Trypan Blue Stain. Proceed with assessing cell count and viability
7. Centrifuge the tube at 1000 rpm for 6–8 min at room temperature
8. Aspirate off the supernatant and resuspend the cells in appropriate volume of warmed, fresh medium in order to achieve a density of 1.0×10^6 cells/ml.

 SEED AT LEAST 5×10^5 CELLS IN A T25CM² FLASK AND 5×10^6 CELLS IN A T75CM² FINAL TOTAL VOLUME FOR A T25CM² FLASK SHOULD BE 10 ML AND 25 ML FOR A T75CM²

9. Gently rock the culture flasks to evenly distribute the cells and place them into a 37 °C incubator at 5% CO₂
10. Completely change the culture medium the day after culture initiation

D.2.2 Culture and subculture of the cells**Maintaining the culture**

1. Warm the appropriate amount of medium to 37 °C in a sterile container.
2. Remove the old medium, wash the cells with around 6 ml of PBS and replace it with the warmed, fresh medium and return the flask to the incubator at 37 °C, 5% CO₂

 CHANGE THE MEDIUM EVERY 2 DAYS

Cells subculturing

1. Discard the culture medium from the culture dish or flask
2. Lift up the cells with TrypLE Express using a sterile pipette tip (1 ml for a 3 cm dish or 2 ml for a T25/10cm dish or 5 ml for a T75)
3. Incubate at 37 °C, 5% CO₂ for 3–5 min
4. Check the detachment of the cells under an inverted phase-contrast microscope. The cells have to be in suspension.
5. Add culture medium with a sterile pipette tip (2 ml for a 3 cm dish or 6 ml for a T25/10cm dish or 10 ml for a T75) to inactivate the TrypLE express action
6. Gently pipette up and down to form a single cell suspension, transfer in a 15 ml sterile Falcon tube
7. Rinse the original dish or flask with 1–3 ml of culture medium and transfer in the same 15 ml sterile Falcon tubes
8. Take 100 µl of cell suspension and mix it with 100 µl of 0.04% Trypan Blue Stain. Perform a cell number and viability count
9. Centrifuge at 1000 rpm for 6–8 min at room temperature
10. Aspirate off the supernatant and resuspend the cells in appropriate volume of warmed, fresh medium in order to achieve a density of 1.0×10^6 cells/ml.
11. Distribute cells into new flasks
**⚠ SEED AT LEAST 5.10×10^5 CELLS IN A T₂₅CM² FLASK AND 1×10^6 CELLS IN A T₇₅CM²
FINAL TOTAL VOLUME FOR A T₂₅CM² FLASK SHOULD BE 10 ML AND 25 ML FOR A T₇₅CM²**
12. Gently rock the culture flasks to evenly distribute the cells and place them into a 37 °C incubator at 5% CO₂
13. Completely change the culture medium the day after culture initiation

⚠ DON'T FORGET TO WRITE ON THE FLASK, THE NAME OF THE CELL LINE, THE DATE, THE NUMBER OF PASSAGES, AND YOUR NAME
PERFORM CELLS PASSAGES BEFORE CELLS REACH 80–85% CONFLUENCE. FOR SUM149-PT IT SHOULD BE EVERY 3–4 DAYS AND FOR SUM159-PT IT SHOULD BE EVERY 3 DAYS

D.2.3 Freezing and cryopreservation of cells, DMSO 10%

1. Discard the culture medium from the culture dish or flask
2. Lift up the cells with TrypLE Express using a sterile pipette tip (1 ml for a 3 cm dish or 2 ml for a T25/10cm dish or 5 ml for a T75)
3. Incubate at 37 °C, 5% CO₂ for 3–5 min

4. Check the detachment of the cells under an inverted phase-contrast microscope. The cells have to be in suspension.
5. Add culture medium with a sterile pipette tip (2 ml for a 3 cm dish or 6 ml for a T25/10cm dish or 10 ml for a T75) to inactivate the TrypLE express action
6. Gently pipette up and down to form a single cell suspension, transfer in a 15 ml sterile Falcon tube
7. Rinse the original dish or flask with 1–3 ml of culture medium and transfer in the same 15 ml sterile Falcon tubes
8. Take 100 μ l of cell suspension and mix it with 100 μ l of 0.04% Trypan Blue Stain. Perform a cell number and viability count
9. Centrifuge at 1000 rpm for 6–8 min at room temperature
10. Aspirate off the supernatant and resuspend the cells in appropriate volume of warmed, fresh medium in order to achieve a density of 1.0×10^6 cells/ml.
11. Distribute 1.0×10^6 cells per cryotubes
12. Add 100 μ l of DMSO into each cryotube. And **very quickly** rock the cryotubes and put them into the Mr Frosty. Place it at $-20\text{ }^{\circ}\text{C}$
13. Let the cryotubes into the $-2\text{ }^{\circ}\text{C}$ for 7 days and transfer them afterwards into the liquid nitrogen cisterns.

 DON'T FORGET TO WRITE ON THE NOTEBOOK THE EMPLACEMENT OF YOUR TUBES

D.3 Protocol – Cloning – Bacteria Transformation DG1, Stellar, or TOP10

From Eurogentec or Invitrogen

Materials

Technical material

- Water bath at 42 °C
- LB Agar plate with antibiotic (100 µg/ml of Ampicillin or Kanamycin)
- Incubator at 37 °C
- Shaking incubator at 37 °C, 225 rpm
- Camping stove
- Timer!

Reagents

- One shot of electro or chemically competent E.Coli (Invitrogen or Eurogentec)
- SOC medium (Invitrogen)


Preparation of the transformation procedure

1. Equilibrate a water bath at 42 °C
2. Warm the vial of SOC medium to 37 °C

Caution

 YOU ARE WORKING WITH DNA! PAY ATTENTION TO WORK PROPERLY TO AVOID ANY DNA CONTAMINATION. WORK WITH A CAMPING STOVE AT ANY TIME OF THE PROCEDURE.

D.3.1 Transformation procedure

1. Thaw on ice one vial of bacteria for each transformation
2. Add 10 ng of the cold DNA into a vial of DG1 cells and mix gently.
 DO NOT MIX BY PIPETTING UP AND DOWN
3. Incubate the vials on ice for 30 min
4. Heat-shock the cells for 30–42 s at 42 °C without shaking
5. Remove the vials from the 42 °C bath and place them on ice for 2 min
6. Aseptically add 250 µl of pre-warmed SOC medium to each vial
7. Cap the vials tightly and shake horizontally at 37 °C for 1 h at 225 rpm in a shaking incubator
8. Spread 50–200 µl from each transformation on a selective plate. Invert the selective plates and incubate overnight at 37 °C

9. Store the remaining transformation mix at 4 °C for at least **3 months long**
10. Select colonies and analyze by plasmid isolation, PCR or sequencing

D.3.2 Rapid transformation procedure

1. Thaw on ice one vial of bacteria for each transformation
2. Add 10 ng of the DNA into the vial of bacteria and mix gently. Do not mix by pipetting up and down
3. Incubate the vials on ice for 30 min
4. Heat-shock the cells for 30–42 s at 42 °C without shaking
5. Remove the vials from the 42 °C bath and place them on ice for 2 min
6. Immediately spread 50–200 µl from each transformation on a selective plate. Invert the selective plates and incubate overnight at 37 °C
7. Store the remaining transformation mix at 4 °C for at least **3 months long**
8. Select colonies and analyze by plasmid isolation, PCR or sequencing

D.4 Protocol Cloning Blunt Quick Blunting

From NEB

Materials

Technical material

- Heat block at 70 °C

Reagents

- Quick Blunting Kit (NEB, E1201S)

D.4.1 Quick Blunting reaction

1. Set up of the reaction
 - Up to 10 ng purified DNA
 - 2.5 µl 10X blunting buffer
 - 2.5 µl 1 mM dNTP mix
 - 1 µl blunt enzyme mix
 - ddH₂O qsp 25 µl
2. Reactions containing restriction enzyme digested DNA are incubated at RT for 15 min. Reactions with sheared/nebulized DNA or PCR products are incubated at RT for 30 min.
3. Immediately inactivate enzyme in the blunting reaction by heating at 70 °C for 10 min
4. Continue to the dephosphorylation procedure
⇒ Following :
 - Protocol Cloning Dephosphorylation Antarctic Phosphatase
 - Protocol Cloning Dephosphorylation Alkaline Phosphatase Calf Intestinal
5. Store the rest at -20 °C

D.5 Protocol Cloning Dephosphorylation Antarctic Phosphatase

From New England Biolabs

Materials

Technical material

- Heat block at 37 °C
- Heat block at 65 °C

Reagents

- Alkaline Phosphatase (NEB, M0289S, 1000 units, 5000 units/ml)

D.5.1 Dephosphorylating DNA with Antarctic Phosphatase

1. Add 1/10 volume of 10X Antarctic Phosphatase reaction buffer to 1–5 µg of DNA cut with any restriction endonuclease in any buffer
2. Add 1 µl of Antarctic Phosphatase (5 units) and mix
3. Incubate 15 min at 37 °C for 5' extensions or blunt-ends, 60 min for 3' extensions
4. Heat inactivate for 5 min at 65 °C
5. Purify DNA by gel purification, spin column purification or phenol extraction
6. Store at –20 °C

D.6 Protocol Cloning Dephosphorylation Alkaline Phosphatase Calf Intestinal

From New England Biolabs

Materials

Technical material

- Heat block at 37 °C

Reagents

- Alkaline Phosphatase, Calf Intestinal (CIP) (NEB, M0290S, 1000 units, 10 000 units/ml)

Dephosphorylating DNA with CIP

1. Suspend DNA in 1X NEB Buffer 3 (0.5 µg/10µl)
2. Add 0.5 unit of CIP per µg vector DNA
3. Incubate 90 min at 37 °C
4. Purify DNA by gel purification, spin column purification or phenol extraction
5. Store at -20 °C

D.7 Protocol Cloning Gel & PCR Purification

by GENOMED, by affinity chromatography

Materials

Technical material

- Water bath at 50 °C

Reagents

- JetQuick DNA purification Kits, Gel extraction procedure (Genomed, 42-0050)

D.7.1 DNA purification from gel slice

1. Excise a minimal area of gel (up to 400 mg) containing the DNA fragment
2. Add Gel solubilization buffer (L1) to the excised gel in the a 2 ml tube
3. Place the tube with the gel slice and buffer L1 into a 50 °C water bath or heat block. Incubate the tube at 50 °C for 15 min. Invert the tube every 3 °C to mix
4. Place a column into a 2 ml receiver tube. Pipet the dissolved gel slice onto the column. Centrifuge the column at 12 000 *g* for 1 min. Discard the flowthrough and place the column into the 2 ml receiver tube
5. Add 500 µl wash buffer (L2) containing ethanol to the column. Centrifuge the column at 12 000 *g* for 1 min.
6. Empty the collection tube and redo step 5 one time again
7. Empty the collection tube and recentrifuge the column assembly for 1 min at 12 000 *g* for 1 min with the microcentrifuge lid open to allow evaporation of any residual ethanol
8. Carefully transfer minicolumn to a clean 1.5 ml microcentrifuge tube
9. Add 50 µl of TE buffer to the minicolumn. Incubate at room temperature for 1 min. Centrifuge at 12 000 *g* for 2 min
10. Repeat step 9 with the 50uL flowthrough from the 1.5 ml microcentrifuge tube. Incubate at room temperature for 1 min. Centrifuge at 12 000 *g* for 2 min
11. Discard minicolumn and transfer the flowthrough into a 1.8 ml cryotube. Store at -20 °C

D.8 Protocol Cloning Ligation In-Fusion

Materials

Technical material

- Heat block at 50 °C

Reagents

In-Fusion™ HD Cloning Kit (ClonTech, 639637)

In-Fusion cloning reaction

Set up of the reaction

- 2 µl of 5X In-Fusion HD enzyme premix
- 50–200 ng of linearized vector
- 10–200 ng of purified PCR fragment
- ddH₂O qsp 10 µl

Linearized Vector	<10 kbase = 50–100 ng	>10 kbase = 50–200 ng
Purified PCR fragment	0.5 kbase = 10–50 ng	0.5–10 kbase = 50–100 ng >10 kbase = 50–100 ng

The better quantity yield between linearized vector and purified PCR fragment is 100/50 versus 100/80.

For reactions with larger volumes of vector and PCR insert (>7 µl of vector+insert), double the amount of enzyme premix, and add ddH₂O for a total volume of 20 µl

1. Adjust the total reaction volume to 10 µl using deionized H₂O and mix the reaction
2. Incubate 15 min at 50 °C, then place on ice
3. Continue to the transformation procedure
⇒ following Protocol – Cloning – Bacteria Transformation DG1, Stellar, or TOP10
4. Store the rest at –20 °C

D.9 Protocol Cloning Ligation LR Clonase II

From Invitrogen, Life Technologies

Materials

Technical material

- Heat block at 25 °C
- Heat block at 37 °C

Reagents

- Gateway®LR Clonase II Enzyme mix (Invitrogen Life Technologies, 11791-020)

D.9.1 LR cloning reaction

1. Add the following components to a 1.5 ml microcentrifuge tube at RT and mix
 - Entry clone attL1-gene-attL2 (50–150 ng)
 - Destination vector attR1-ccdB-attR2 (150 ng)
 - TE buffer or TRIS buffer, pH 8.0 to pH 8.5 up to 8 µl

It's better to use a supercoiled attR plasmid and a linearized attL vector for the LR reaction.

⚠ NEVER FORGET TO TEST PENTR-GUS (KANAR) VECTOR AS A POSITIVE CONTROL FOR THE LR REACTION

2. Thaw on ice the LR Clonase II enzyme mix for about 2 min. Vortex the LR Clonase II enzyme mix briefly twice (2 s each time)
3. To each sample, add 2 µl of LR Clonase II enzyme mix to the reaction and mix well by vortexing briefly twice. Microcentrifuge briefly.

⚠ DON'T FORGET TO RETURN LR CLONASE II ENZYME MIX TO -20 °C OR -80 °C STORAGE. THE LR CLONASE II IS VERY SENSITIVE TO TEMPERATURE CHANGES.

4. Incubate reactions at 25 °C for 1 h
5. Add 1 µl of the proteinase K solution to each sample to terminate the reaction. Vortex briefly. Incubate samples at 37 °C for 10 min
6. Continue with the transformation procedure

⇒ Following Protocol – Cloning – Bacteria Transformation DG1, Stellar, or TOP10

7. Store the rest at -20 °C

D.10 Protocol Cloning Ligation Quick Ligase

From New England Biolabs

Materials

Reagents

- Quick ligation kit (NEB, M2200S, $-20\text{ }^{\circ}\text{C}$)

D.10.1 Quick ligation protocol

1. Combine 50 ng of vector with a 3-fold molar excess of insert. Adjust volume to 10 μl with ddH₂O
2. Add 10 μl of 2X Quick Ligation Reaction buffer and mix
3. Add 1 μl of Quick T4 DNA ligase and mix thoroughly
4. Centrifuge briefly and incubate at room temperature ($25\text{ }^{\circ}\text{C}$) for 5 min Incubation beyond this time provides no additional benefit. In fact, transformation efficiency starts to decrease after 2 hours and is reduced by up to 75% if the reaction is allowed to go overnight at $25\text{ }^{\circ}\text{C}$
5. Chill on ice and continue with the transformation procedure

⇒ Following Protocol – Cloning – Bacteria Transformation DG1, Stellar, or TOP10

6. Store the rest at $-20\text{ }^{\circ}\text{C}$

D.11 Protocol Cloning Oligos Annealing

Materials

Technical material

- Heat block at 95 °C

Reagents

- Annealing buffer, filtered (10 mM Tris pH 7.5, 1 mM EDTA, 50 mM NaCl)
 - 39.12 ml of ultrapure water
 - 400 µl of 1 M Tris pH 7.5
 - 80 µl of 0.5 M EDTA
 - 400 µl of 5 M NaCl
- Ultrapure RNase free water

D.11.1 Annealing reaction

1. Resuspend oligos to a stock concentration of 100µM in ultrapure RNase free water. Store oligos stocks at -20 °C
2. To a 1.5 ml tube add 10µL of the proper top and bottom strand oligos and then 80 µl of annealing buffer to obtain a 10 µmol/dm³ concentration of oligos
3. Bring heat block to 95 °C and inset 1.5 ml tubes with oligos to be annealed. Turn off the heat block and let the tube for 15 min. Afterward remove the heatblock from the heating unit and let tubes further cool down on the bench top for 45 min
4. Continue the cloning with the ligation procedure
⇒ Following Protocol Cloning Ligation Quick Ligase
5. Store the rest at -20 °C

D.12 Protocol Cloning Oligos Resuspension

Materials

Reagents

- Ultrapure RNase free water (Gibco)

D.12.1 Resuspension procedure

1. Resuspend oligos to a stock concentration of $100 \mu\text{mol}/\text{dm}^3$ in ultrapure RNase free water.
2. Store oligos stocks at $-20 \text{ }^\circ\text{C}$

D.13 Protocol Cloning Plasmid DNA Purification Maxiprep

From Macherel Nagel

Materials

Technical material

- Shaking incubator at 37 °C, 225 rpm
- Camping stove

Reagents

- NucleoBond Xtra Maxi Plus kit (Macherel Nagel, 740416.10)
- LB medium
- Antibiotic (100 µg/ml of Ampicillin or Kanamycin)

Caution

⚠ YOU ARE WORKING WITH DNA! PAY ATTENTION TO WORK PROPERLY TO AVOID ANY DNA CONTAMINATION.

Preheat the ELU buffer to 50 °C prior to elution to improve yields.

D.13.1 Low-copy plasmid purification

1. Prepare a large overnight culture. Inoculate a single colony clone or 100 µl of a miniprep preparation into 600 ml of LB medium+ATB. Grow the culture overnight at 37 °C at 225 rpm. ⚠ WORK WITH A CAMPING STOVE AT ANY TIME OF THE PROCEDURE
2. Harvest bacterial cells. Pellet the cells by centrifugation at 6000 g for 15 min at 4 °C and discard the supernatant completely. *At this step it is possible to freeze the pellet at -20 °C for 1 month if necessarily.*
3. Resuspend the pellet completely in 24 ml RES buffer+RNase A ⚠ (ATTENTION STORE THIS RES BUFFER AT 4 °C)
4. Add 24 ml of LYS buffer to the suspension. Mix gently by inverting the tube 5 times. Do not vortex in order not to release contaminating chromosomal DNA from cells debris. Incubate the mixture at RT for 5 min
5. Equilibrate a column with the inserted column filter with 25 ml of EQU buffer
6. Add 24 ml of NEU buffer to the suspension and immediately mix the lysate gently by inverting the tube 10–15 times.
⚠ DO NOT VORTEX.
Incubate crude lysate on ice for 5 min
7. Make sure to have a homogeneous suspension of the precipitate by inverting the tube 3 times directly before loading the column filter. Allow the column to empty by gravity flow
8. Wash the column filter with 15 ml of EQU buffer and allow the column to empty by gravity flow

9. Discard the column filter
10. Wash the column with 25 ml of WASH buffer and allow the column to empty by gravity flow
11. Elute the plasmid DNA with 15 ml ELU buffer. Collect the eluate in a 50 ml centrifuge tube
12. Add 10.5 ml room temperature isopropanol to precipitate the eluted plasmid DNA. Vortex well and let the mixture sit for 2–5 min. Centrifuge at 15 000 *g* for 90 min at 4 °C. Carefully discard the supernatant.
13. Add 5 ml endotoxin-free room temperature 70% ethanol to precipitate to the pellet. Centrifuge at 15 000 *g* for 30 min at room temperature. Carefully remove ethanol completely from the tube with a pipette tip and let dry for a couple of hours at room temperature
14. Dissolve the pellet in 500 µl of endotoxin free TE-EF or TRIS buffer and let it dissolve for 1 day at 4 °C
15. Gently pipette up and down and determine plasmid yield by UV spectrophotometry
16. Aliquot and place it at –20 °C for long term storage

D.13.2 High-copy plasmid purification

1. Prepare a large overnight culture. Inoculate a single colony clone or 100 µl of a miniprep preparation into 300 mL of LB medium+ATB. Grow the culture overnight at 37 °C at 225 rpm.
⚠ WORK WITH A CAMPING STOVE AT ANY TIME OF THE PROCEDURE
2. Harvest bacterial cells. Pellet the cells by centrifugation at 6000 *g* for 15 min at 4 °C and discard the supernatant completely. *At this step it is possible to freeze the pellet at –20 °C for 1 month if necessary.*
3. Resuspend the pellet completely in 12 ml RES buffer+RNase A ⚠ (ATTENTION STORE THIS RES BUFFER AT 4 °C)
4. Add 12 ml of LYS buffer to the suspension. Mix gently by inverting the tube 5 times.
⚠ DO NOT VORTEX IN ORDER NOT TO RELEASE CONTAMINATING CHROMOSOMAL DNA FROM CELLS DEBRIS.
Incubate the mixture at RT for 5 min
5. Equilibrate a column with the inserted column filter with 25 ml of EQU buffer
6. Add 12 ml of NEU buffer to the suspension and immediately mix the lysate gently by inverting the tube 10–15 times.
⚠ DO NOT VORTEX.
Incubate crude lysate on ice for 5 min
7. Make sure to have a homogeneous suspension of the precipitate by inverting the tube 3 times directly before loading the column filter. Allow the column to empty by gravity flow
8. Wash the column filter with 15 ml of EQU buffer and allow the column to empty by gravity flow

9. Discard the column filter
10. Wash the column with 25 ml of WASH buffer and allow the column to empty by gravity flow
11. Elute the plasmid DNA with 15 ml ELU buffer. Collect the eluate in a 50 ml centrifuge tube
12. Add 10.5 ml room temperature isopropanol to precipitate the eluted plasmid DNA. Vortex well and let the mixture sit for 2–5 min. Centrifuge at 15 000 *g* for 90 min at 4 °C. Carefully discard the supernatant.
13. Add 5 ml endotoxin-free room temperature 70% ethanol to precipitate to the pellet. Centrifuge at 15 000 *g* for 30 min at room temperature. Carefully remove ethanol completely from the tube with a pipette tip and let dry for a couple of hours at room temperature
14. Dissolve the pellet in 500 µl of endotoxine free TE-EF or TRIS buffer and let it dissolve for 1 day at 4 °C
15. Gently pipette up and down and determine plasmid yield by UV spectrophotometry
16. Aliquot and place it at –20 °C for long term storage

D.14 Protocol Cloning Plasmid DNA Purification Midiprep

From Macherel Nagel

Materials

Technical material

- Shaking incubator at 37 °C, 225 rpm
- Camping stove

Reagents

- NucleoBond Xtra Midi Plus kit (Macherel Nagel, 740412.10)
- LB medium
- Antibiotic (100 µg/ml of Ampicillin or Kanamycin)

Caution

⚠ YOU ARE WORKING WITH DNA! PAY ATTENTION TO WORK PROPERLY TO AVOID ANY DNA CONTAMINATION.



Preheat the ELU buffer to 50 °C prior to elution to improve yields.

D.14.1 Low-copy plasmid purification

1. Prepare a large overnight culture. Inoculate a single colony clone or 100 µl of a miniprep preparation into 200 ml of LB medium+ATB. Grow the culture overnight at 37 °C at 225 rpm.
⚠ WORK WITH A CAMPING STOVE AT ANY TIME OF THE PROCEDURE
2. Harvest bacterial cells. Pellet the cells by centrifugation at 6000 *g* for 15 min at 4 °C and discard the supernatant completely. *At this step it is possible to freeze the pellet at -20 °C for 1 month if necessary.*
3. Resuspend the pellet completely in 16mL RES buffer+**RNase A (attention store this RES buffer at 4°C)**
4. Add 16 ml of LYS buffer to the suspension. Mix gently by inverting the tube 5 times. Do not vortex in order not to release contaminating chromosomal DNA from cells debris. Incubate the mixture at RT for 5 min
5. Equilibrate a column with the inserted column filter with 12 ml of EQU buffer
6. Add 16 ml of NEU buffer to the suspension and immediately mix the lysate gently by inverting the tube 10–15 times.
⚠ DO NOT VORTEX.
Incubate crude lysate on ice for 5 min
7. Make sure to have a homogeneous suspension of the precipitate by inverting the tube 3 times directly before loading the column filter. Allow the column to empty by gravity flow

8. Wash the column filter with 5 ml of EQU buffer and allow the column to empty by gravity flow
9. Discard the column filter
10. Wash the column with 8 ml of WASH buffer and allow the column to empty by gravity flow
11. Elute the plasmid DNA with 5 ml ELU buffer. Collect the eluate in a 50 ml centrifuge tube
12. Add 3.5 ml room temperature isopropanol to precipitate the eluted plasmid DAN. Vortex well and let the mixture sit for 2–5 min. Centrifuge at 15 000 *g* for 90 min at 4 °C. Carefully discard the supernatant.
13. Add 2 ml endotoxin-free room temperature 70% ethanol to precipitate to the pellet. Centrifuge at 15 000 *g* for 30 min at room temperature. Carefully remove ethanol completely from the tube with a pipette tip and let dry for a couple of hours at room temperature
14. Dissolve the pellet in 200 µl of endotoxine free TE-EF or TRIS buffer and let it dissolve for 1 day at 4 °C
15. Gently pipette up and down and determine plasmid yield by UV spectrophotometry
16. Aliquot and place it at –20 °C for long term storage

D.14.2 High-copy plasmid purification

1. Prepare a large overnight culture. Inoculate a single colony clone or 100 µl of a miniprep preparation into 100 ml of LB medium+ATB. Grow the culture overnight at 37 °C at 225 rpm.
 **WORK WITH A CAMPING STOVE AT ANY TIME OF THE PROCEDURE**
2. Harvest bacterial cells. Pellet the cells by centrifugation at 6000 *g* for 15 min at 4 °C and discard the supernatant completely. *At this step it is possible to freeze the pellet at –20 °C for 1 month if necessarily.*
3. Resuspend the pellet completely in 8 ml RES buffer+**RNase A (attention store this RES buffer at 4 °C)**
4. Add 8mL of LYS buffer to the suspension. Mix gently by inverting the tube 5 times. Do not vortex in order not to release contaminating chromosomal DNA from cells debris. Incubate the mixture at RT for 5 min
5. Equilibrate a column with the inserted column filter with 1 ml of EQU buffer
6. Add 8 ml of NEU buffer to the suspension and immediately mix the lysate gently by inverting the tube 10–15 times.
 **DO NOT VORTEX.**
Incubate crude lysate on ice for 5 min
7. Make sure to have a homogeneous suspension of the precipitate by inverting the tube 3 times directly before loading the column filter. Allow the column to empty by gravity flow
8. Wash the column filter with 5 ml of EQU buffer and allow the column to empty by gravity flow

9. Discard the column filter
10. Wash the column with 8 ml of WASH buffer and allow the column to empty by gravity flow
11. Elute the plasmid DNA with 5 ml ELU buffer. Collect the eluate in a 50 ml centrifuge tube
12. Add 3.5 ml room temperature isopropanol to precipitate the eluted plasmid DNA. Vortex well and let the mixture sit for 2–5 min. Centrifuge at 15 000 *g* for 90 min at 4 °C. Carefully discard the supernatant.
13. Add 2 ml endotoxin-free room temperature 70% ethanol to precipitate to the pellet. Centrifuge at 15 000 *g* for 30 min at room temperature. Carefully remove ethanol completely from the tube with a pipette tip and let dry for a couple of hours at room temperature
14. Dissolve the pellet in 200 μ l of endotoxine free TE-EF or TRIS buffer and let it dissolve for 1 day at 4 °C
15. Gently pipette up and down and determine plasmid yield by UV spectrophotometry
16. Aliquot and place it at -20 °C for long term storage

D.15 Protocol Cloning Plasmid DNA Purification Miniprep

From Qiagen

Materials

Technical material

- Shaking incubator at 37 °C, 225 rpm
- Camping stove

Reagents


- QIAprep Spin Miniprep kit (Qiagen, 27106)
- LB medium
- Antibiotic (100 µg/ml of Ampicillin or Kanamycin)

Caution

 YOU ARE WORKING WITH DNA! PAY ATTENTION TO WORK PROPERLY TO AVOID ANY DNA CONTAMINATION.


Preheat the buffer EB to 50 °C prior to elution to improve yields

D.15.1 Low-copy plasmid purification

1. Prepare a small overnight culture. Inoculate a single colony clone into 10 ml of LB medium+ATB. Grow the culture overnight at 37 °C at 225 rpm.
 WORK WITH A CAMPING STOVE AT ANY TIME OF THE PROCEDURE
2. Harvest bacterial cells. Pellet the cells by centrifugation at 6000 *g* for 15 min at 4 °C and discard the supernatant completely. At this step it is possible to freeze the pellet at -20 °C for 1 month if necessarily.
3. Resuspend the pellet completely in 250 µl buffer P1+RNase A (**attention store this buffer P1 at 4 °C**) and transfer to a microcentrifuge tube. No cell clumps should be visible after resuspension of the pellet
4. Add 250 µl of buffer P2 to the suspension. Mix gently by inverting the tube 4 to 6 times. Do not vortex in order not to release contaminating genomic DNA from cells debris. If necessary, continue inverting the tube until the solution becomes viscous and slightly clear. Do not allow the lysis reaction to proceed for more than 5 min
5. Add 350 µl buffer N3 and invert the tube immediately but gently for 4 to 6 times. To avoid localized precipitation, mix the solution gently but thoroughly, immediately after addition of buffer N3. The solution should become cloudy
6. Centrifuge for 10 min at 17 000 *g* in a table-top microcentrifuge. A compact white pellet will form
7. Apply the supernatants from step6 to the QIAprep spin column by pipetting
8. Centrifuge for 60 s at 17 000 *g*. Discard the flow-through

9. Wash QIAprep spin column by adding 0.75 ml buffer PE and centrifuging for 60 s
10. Discard the flow-through, and centrifuge for an additional 1 min to remove residual wash buffer
11. Place the QIAprep column in a clean 1.5 ml microcentrifuge tube. To elute DNA, add 50 μ l buffer EB (10 M Tris-Cl, pH 8.5) to the center of each QIAprep spin column, let stand for 1 min and centrifuge for 1 min

D.15.2 High-copy plasmid purification

1. Prepare a small overnight culture. Inoculate a single colony clone into 5 ml of LB medium+ATB. Grow the culture overnight at 37 °C at 225 rpm.
 **WORK WITH A CAMPING STOVE AT ANY TIME OF THE PROCEDURE**
2. Harvest bacterial cells. Pellet the cells by centrifugation at 6000 *g* for 15 min at 4 °C and discard the supernatant completely. *At this step it is possible to freeze the pellet at -20 °C for 1 month if necessary.*
3. Resuspend the pellet completely in 250 μ l buffer P1+**RNase A (attention store this buffer P1 at 4°C)** and transfer to a microcentrifuge tube. No cell clumps should be visible after resuspension of the pellet
4. Add 250 μ l of buffer P2 to the suspension. Mix gently by inverting the tube 4 to 6 times. Do not vortex in order not to release contaminating genomic DNA from cells debris. If necessary, continue inverting the tube until the solution becomes viscous and slightly clear. Do not allow the lysis reaction to proceed for more than 5 min
5. Add 350 μ l buffer N3 and invert the tube immediately but gently for 4 to 6 times. To avoid localized precipitation, mix the solution gently but thoroughly, immediately after addition of buffer N3. The solution should become cloudy
6. Centrifuge for 10 min at 17 000 *g* in a table-top microcentrifuge. A compact white pellet will form
7. Apply the supernatants from step 6 to the QIAprep spin column by pipetting
8. Centrifuge for 60 s at 17 000 *g*. Discard the flow-through
9. Wash QIAprep spin column by adding 0.75 ml buffer PE and centrifuging for 60 s
10. Discard the flow-through, and centrifuge for an additional 1 min to remove residual wash buffer
11. Place the QIAprep column in a clean 1.5 ml microcentrifuge tube. To elute DNA, add 50 μ l buffer EB (10 M Tris-Cl, pH 8.5) to the center of each QIAprep spin column, let stand for 1 min and centrifuge for 1 min

D.16 Protocol Cloning Sequencing BigDye® 1.1 and 3.1

From the sequencing platform, Curie Institute

Materials

Technical material

- Thermocycler

Reagents

- BigDye®Terminator v1.1 cycle, Sequencing RR-100 (AB, 4336768, $-20\text{ }^{\circ}\text{C}$)
- BigDye®Terminator v3.1 cycle, Sequencing RR-100 (AB, 4336911, $-20\text{ }^{\circ}\text{C}$)

D.16.1 Preparation of the PCR mix

DNA template to sequence		800 ng
Primer	at $1.6\text{ }\mu\text{mol}/\text{dm}^3$ concentration	2 μl
BigDye®v1.1 or v3.1		3 μl
Buffer for BigDye®		2 μl
ddH ₂ O	Until	20 μl

D.16.2 PCR protocol

Name of the protocol on the thermocycler : JPSEQ

Step	Temperature	Time
Initial denaturation	96 $^{\circ}\text{C}$	1:00
Denaturation X25	96 $^{\circ}\text{C}$	0:10
Annealing	50 $^{\circ}\text{C}$	0:05
Extension	60 $^{\circ}\text{C}$	4:00
Hold	4 $^{\circ}\text{C}$	∞

D.16.3 Administrative procedure

1. Put a piece of scotch on the top of a tubes bar with the following information
 - Name of the unit
 - Initial (First Name/Name) of the team's director
 - Initial (First Name/Name) of the customer
 - Number of the tubes from the bar
 - Phone number of the customer
 - Date
2. Drop off the tubes to the sequencing platform.
Curie Hospital, 4th floor. Put on a rack and bring it in the $-20\text{ }^{\circ}\text{C}$, 1st drawer up.

D.16.4 Analysis procedure

1. Open the sequence with the program Ape.exe or Serial Cloner.exe
2. Align the sequence with the designed sequence and look for any dissimilarity

D.17 Protocol FACS Cells Flow Cytometry – Hoechst 33258 or DAPI Staining

Materials

Reagents

- Calcium-magnesium-free Phosphate Buffer Saline pH 7.4 1X (Gibco)
- Heat-inactivated Fetal Bovin Serum, -20°C (Gibco)
- TrypLe Express - EDTA 0.02%, 4°C (Gibco)
- Cell specific culture medium

Solution

- FACS buffer : PBS with 2% FBS

⚠ PAY ATTENTION TO USE COLD FACS BUFFER (4°C)

Antibodies

- Anti-mouse Ig, k/Negative Control (FBS) Compensation Particles Set (BD Bioscience, Material Number : 552843)

⚠ PAY ATTENTION TO TITRATE ANTIBODIES

Dye for dead cells staining

Hoechst 33258 or DAPI, Room Temperature, protect from light *cell permeant*

- Stock solution : 10 mg/ml in **water**, 4°C , protect from light. **Stable** for 6 months. ⚠
HOECHST STAINS SHOUDN'T BE RESOLUBILIZED IN PBS! BUT EITHER WITH WATER, OR DIMETHYLFORMAMIDE OR DMSO
FOR LONG TERM STORAGE, STOCK SOLUTION CAN BE ALIQUOTED AND STORED AT $<-20^{\circ}\text{C}$
- Working solution : 10 $\mu\text{g}/\text{ml}$ in **PBS**, 4°C , protect from light. **Stable** for 6 months
Concentration for the staining : 1 $\mu\text{g}/\text{ml}$

Fluorescence Spectral Characteristics

- Excitation : 350 nm
- Emission : 461 nm

Caution

⚠ HOECHST STRAINS ARE KNOWN MUTAGENS AND SHOULD BE HANDLED WITH CARE !

D.17.1 Preparation of cells for flow cytometry

- Discard the culture medium from the culture dish or flask
- Lift up the cells with TrypLE Express using a sterile pipette tip (1 ml for a 3 cm dish or 2 ml for a T25/10cm dish or 5 ml for a T75)
- Incubate at 37 °C, 5% CO₂ for 3–5 min
- Prepare the appropriate number of 15 ml sterile Falcon tubes according to the cells conditions number. Check the detachment of the cells under an inverted phase-contrast microscope. The cells have to be in suspension.
- Add culture medium with a sterile pipette tip (2 ml for a 3 cm dish or 5 ml for a T25/10cm dish or 10 ml for a T75) to inactivate the TrypLE express action
- Gently pipette up and down to form a single cell suspension, transfer in a 15 ml sterile Falcon tube
- Rinse the original dish or flask with 1–3 ml of culture medium and transfer in the same 15 ml sterile Falcon tubes
- Perform a cell number and viability count
- Centrifuge at 1000 rpm for 10 min at room temperature

D.17.2 Direct immunofluorescence staining of cell surface antigens for flow cytometry

1. Aspirate supernatant and resuspend the pellet with the appropriate volume of FACS buffer in order to obtain a cell suspension at 5000 cells/ μ l


$$V_{FACSbuffer} = \frac{[Nbcells]}{5000}$$

2. Aliquot the cell suspension with the quantity of cells needed into as many sterile FACS test tubes needed. Don't forget the FMO and the unstained tubes
3. Centrifuge at 1000 rpm for 10 min at 4 °C
4. Transfer all the tubes on ice
⚠ FROM THIS STEP, CELLS SHOULD REMAIN ON ICE AT ALL THE TIMES CELL STAINING (EITHER CELL SURFACE OR INTRA-CELLULAR STAINING) SHOULD BE PERFORMED AT THIS POINT
5. Resuspend cells in 200 μ l of FACS buffer
6. Acquire data by flow cytometry. Analyze fixed cells within 24 hours. Just before the acquisition add 10 μ l/100 μ l of DAPI or Hoechst 33258 staining, vortex and acquire data.

D.17.3 Indirect immunofluorescence staining of cell surface antigens for flow cytometry

1. Aspirate supernatant and resuspend the pellet with the appropriate volume of FACS buffer in order to obtain a cell suspension at 5000 cells/ μ l

$$V_{FACSbuffer} = \frac{[Nbcells]}{5000}$$

2. Aliquot the cell suspension with the quantity of cells needed into as many sterile FACS test tubes needed. Don't forget the FMO and the unstained tubes
3. Centrifuge at 1000 rpm for 10 min at 4 °C
4. Transfer all the tubes on ice
 FROM THIS STEP, CELLS SHOULD REMAIN ON ICE AT ALL THE TIMES. CELL STAINING (EITHER CELL SURFACE OR INTRA-CELLULAR STAINING) SHOULD BE PERFORMED AT THIS POINT
5. Resuspend cells in 200 μ l of FACS buffer
6. Acquire data by flow cytometry. Analyze fixed cells within 24 hours. Just before the acquisition add 10 μ l/100 μ l of DAPI or Hoechst 33258 staining, vortex and acquire data.

D.17.4 Compensation particles settings for flow cytometry

1. Vortex BD™ CompBead thoroughly before
2. Add 100 μ l of FACS buffer to each sterile FACS tubes used for compensation settings
3. Add 1 full drop of the BD™ CompBeads Negative control and 1 drop of the the BD™ CompBeads Anti-mouse Ig,k beads to each sterile FACS tubes
4. Add the adequate quantity of antibody to the appropriate labeled tube. Make sure the antibody is deposited to the bead mixture, then vortex
5. Incubate 20 min at **RT** protected from exposure to direct light
6. Add 1–2 ml of FACS buffer to each tube and pellet by centrifugation at 200 *g* for 10 min, 4 °C
7. Discard supernatant from each tube by careful vacuum using a fine-tip Pasteur pipette
8. Resuspend bead pellet in each tube by adding 200 μ l of FACS buffer. Vortex thoroughly
9. Run each tube separately on the flow cytometer
- 10.

D.18 Protocol FACS Cells Sorting Antibodies Staining

Materials

Reagents

- Calcium-magnesium-free Phosphate Buffer Saline pH 7.4 1X (Gibco)
- Heat-inactivated Fetal Bovin Serum, -20°C (Gibco)
- TrypLe Express - EDTA 0.02%, 4°C (Gibco)
- Cell specific culture medium
- EDTA
- BSA 4°C (Invitrogen)

Solution


- FACS buffer : PBS with 2% FBS
- FACS resuspension buffer : PBS, 1% BSA, 2 mM EDTA

 PAY ATTENTION TO USE COLD FACS BUFFER AND FACS RESUSPENSION BUFFER (4°C)

Antibodies

- Anti-mouse Ig, k/Negative Control (FBS) Compensation Particles Set BD Bioscience
Material Number : 552843

 PAY ATTENTION TO TITRATE ANTIBODIES

 CELLS NEED TO STAY STERILE, PAY ATTENTION TO WORK PROPERLY UNDER THE PSM FOR ALL THE PROCEDURE LONG ! THIS PROCEDURE CAN ONLY BE PERFORMED USING CELL SURFACE ANTIGENS !

D.18.1 Preparation of cells for flow cytometry

1. Discard the culture medium from the culture dish or flask
2. Lift up the cells with TrypLE Express using a sterile pipette tip (10 ml for T175)
3. Incubate at 37°C , 10% CO_2 for 3–5 min
4. Prepare the appropriate number of 50 ml sterile Falcon tubes according to the cells conditions number. Check the detachment of the cells under an inverted phase-contrast microscope. The cells have to be in suspension.
5. Add culture medium with a sterile pipette tip (10 ml for T175) to inactivate the TrypLE express action
6. Gently pipette up and down to form a single cell suspension, transfer in a 50 ml sterile Falcon tube

7. Rinse the original dish or flask with 5 ml of culture medium and transfer in the same 50 ml sterile Falcon tubes
8. Perform a cell number and viability count
9. Centrifuge at 1000 rpm for 10 min at room temperature

D.18.2 Direct immunofluorescence staining of cells for flow cytometry

1. Aspirate supernatant and resuspend the pellet with the appropriate volume of FACS buffer in order to obtain a cell suspension at 5000 cells/ μ l

$$V_{FACSbuffer} = \frac{[Nbcells]}{5000}$$


2. Aliquot the cell suspension with the quantity of cells needed into as many sterile FACS test tubes needed. Don't forget the FMO and the unstained tubes
3. Centrifuge at 1000 rpm for 10 min at 4 °C
4. Transfer all the tubes on ice \triangle FROM THIS STEP, CELLS SHOULD REMAIN ON ICE AT ALL THE TIMES
5. Resuspend the cells with the appropriate volume of FACS buffer to obtain 500 μ l (as 100 μ l = volume of antibodies + volume of FACS buffer)
6. Add antibody at the recommended dilution. Mix well and incubate on ice for 30 min protected from exposure to direct light.
7. Wash cells with 2 ml of FACS buffer, centrifuge at 1000 rpm for 10 min at 4 °C and discard the resulting supernatant
8. Resuspend cells in FACS resuspension buffer to obtain a concentration of 20×10^6 cells/ml
9. Go to the FACS facility and filtrate the cells through the filter of the blue plug.
10. Acquire data by flow cytometry

D.18.3 Indirect immunofluorescence staining of cells for flow cytometry

1. Aspirate supernatant and resuspend the pellet with the appropriate volume of FACS buffer in order to obtain a cell suspension at 5000 cells/ μ l

$$V_{FACSbuffer} = \frac{[Nbcells]}{5000}$$

2. Aliquot the cell suspension with the quantity of cells needed into as many sterile FACS test tubes needed. Don't forget the FMO and the unstained tubes

3. Centrifuge at 1000 rpm for 10 min at 4 °C
4. Transfer all the tubes on ice  FROM THIS STEP, CELLS SHOULD REMAIN ON ICE AT ALL THE TIMES
5. Resuspend the cells with the appropriate volume of FACS buffer to obtain 500 μ l (as 100 μ l = volume of antibodies + volume of FACS buffer)
6. Add antibody at the recommended dilution. Mix well and incubate on ice for 30 min protected from exposure to direct light.
7. Wash cells with 2 ml of FACS buffer, centrifuge at 1000 rpm for 10 min at 4 °C and discard the resulting supernatant
8. Resuspend the cells with the appropriate volume of FACS buffer to obtain 500 μ l (as 100 μ l = volume of antibodies + volume of FACS buffer)
9. Add antibody at the recommended dilution. Mix well and incubate on ice for 30 min protected from exposure to direct light.
10. Wash cells with 2 ml of FACS buffer, centrifuge at 1000 rpm for 10 min at 4 °C and discard the resulting supernatant
11. Resuspend cells in FACS resuspension buffer to obtain a concentration of 20×10^6 cells/ml
12. Go to the FACS facility and filtrate the cells through the filter of the blue plug.
13. Acquire data by flow cytometry

D.18.4 Compensation particles settings for flow cytometry

1. Vortex BD™ CompBead thoroughly before use
2. Add 50 μ l of FACS buffer to each sterile FACS tubes used for compensation settings
3. Add 2 full drops of the BD™CompBeads Negative control and 2 drops of the BD™CompBeads Anti-mouse Ig,k beads to each sterile control FACS tubes
4. Add the adequate quantity of antibody to the appropriate labeled tube. Make sure the antibody is deposited to the bead mixture, then vortex
5. Incubate 20 min at **RT** protected from exposure to direct light
6. Add 1–2 ml of FACS buffer to each tube and pellet by centrifugation at 200 g for 10 min, 4 °C
7. Discard supernatant from each tube by careful vacuum using a fine-tip Pasteur pipette
8. Resuspend bead pellet in each tube by adding 200 μ l of FACS buffer. Vortex thoroughly
9. Run each tube separately on the flow cytometer

D.19 Protocol Immunofluorescence of Fixed Cells using PFA

Materials


Technical material

- Glass coverslip with cells cultured on it
- Parafilm
- A large plate as a support for the staining

Reagents

- Calcium-magnesium-free Phosphate Buffer Saline pH 7.4 1X (Gibco)
- Paraformaldehyde (PFA) (Sigma, 28793.292)
- Gelatin, 4 °C (Sigma, G7765)
- Triton-X-100 (EuroMedex, 2000C)
- Immu-Mount TM (ThermoScientific, 9990402)

Solution

- NaOH solution, 1 M, 4 °C
- PFA 4% solution, -20 °C  IT IS HIGHLY TOXIC, MANIPULATE UNDER THE HOOD
- Gelatin 0.25% solution, 4 °C
- Gelatin 0.125% solution, 4 °C
- Triton-X 0.25%

Preparation of solutions

PFA 4% solution

1. With a becher and under the chemical hood, weigh 4 g of PFA.
2. Add 50 ml of PBS 1X with a stirring rod and heat the preparation at 70 °C
3. Add 50 µl of NaOH solution, 1 M and wait 15 min. When the PFA solution becomes clear, it means that the PFA is dissolved
4. Adjust the pH to 7.4 thanks to HCl
5. And complete the solution with PBS, 1X until 100 ml
6. Aliquot and store for long-term the PFA 4% solution at -20 °C

Gelatin 0.25%

1. Scratch the gelatin from the bottle (4 °C) and put some into a Falcon 50 ml. Let it melt down at 50 °C in a bain-marie
2. Perform a dilution at 0.25% of gelatin into PBS 1X
 - 30 ml PBS + 166 µl gelatin
 - 15 ml PBS + 83 µl gelatin
 - 10 ml PBS + 55 µl gelatin

Gelatin 0.125%

1. 1) Scratch the gelatin from the bottle (4 °C) and put some into a Falcon 50 ml. Let it melt down at 50 °C in a bain-marie
2. 2) Perform a dilution at 0.125% of gelatin into PBS 1X
 - - 30 ml PBS + 83 µl gelatin
 - - 15 ml PBS + 55 µl gelatin
 - - 10 ml PBS + 27 µl gelatin

Dye and antibodies

⚠ PAY ATTENTION TO TITRATE ANTIBODIES.

IF DIFFERENT ANTIBODIES ARE USED, PAY ATTENTION TO TAKE DIFFERENT SPECIES WHEN YOU USE SECONDARY ANTIBODIES

Caution

⚠ PFA IS HIGHLY MUTAGEN, ONCOGEN AND TOXIC, MANIPULATE WITH CARE

Preparation of cells for immunofluorescence staining

1. Discard the culture medium from the culture dish or flask and take over the coverslip from the dish
2. Prepare a large plate with parafilm at the bottom ; put some sprayed sopalin around the plate. This preparation is performed to avoid those cells dry
3. Put down a drop of PBS as much as you have coverslips to stain. Transfer the coverslips to the plate and put on the drop of PBS. Lean smoothly on the coverslip in order that the drop of PBS covers the converslip.

⚠ PAY ATTENTION DO NOT LET DRY THE CELLS. ALSO IT IS RECOMMENDED TO USE 2 DIFFERENT PIPETTES WITH THE RIGHT AND LEFT HANDS.

Cells Surface markers staining for immunofluorescence

⚠ PAY ATTENTION DO CENTRIFUGE THE ANTIBODIES BEFORE THE OPENING OF THE TUBE BECAUSE PART OF THE SOLUTION STAYS INSIDE THE COVER

⚠ WORK ON COLD METAL PLATE ALL MANIP LONG!!!

1. Block the cells on the coverslip with 300 μ l of gelatin 0.25% for 30 min
2. Incubate the cells with the primary antibody diluted into gelatin 0.25%. Put on the coverslip 100 μ l of the solution and let it incubate for 1 h
3. Wash 3 times for 5 min long the coverslip with the solution of gelatin 0.125%
4. Incubate the cells with the secondary antibody diluted into gelatin 0.25%. Put on the coverslip 100 μ l of the solution and let it incubate for 45 min
5. Wash the coverslip 3 times for 5 min long with PBS 1X

Cells fixation for immunofluorescence

1. Fix the cells sucking up the PBS and put on the coverslip 300 μ l (for a coverslip of 18 mm of diameter) of PFA 4% for 15 min long at RT
2. Aspirate the PFA 4% and wash the coverslip 3 times with PBS 1X.
3. Attention do not let dry the cells. Also it is recommended to use 2 different pipettes with the right and left hands.
4. Afterwards it is possible to stain for nuclei with either DAPI or several Hoechst diluted into gelatin 0.25%. Afterwards, wash 3 times with PBS 1X
5. Mount the coverslip on a glass-slide with a drop of 10 μ l of Immu-Mount™
 - Pipette 10 μ l of Immu-Mount™ on the slide
 - Take carefully the coverslip with a pair of pliers and dry the side without cells on a paper
 - Put down the coverslip on the drop of Immu-Mount™, face with cells in contact with Immu-Mount™
 - Fix the slide with 4 drops of warmish

D.20 Protocol of tumorspheres assays

From Smart et al. [291]

⚠ PAY ATTENTION TO WIPE EVERYTHING GOING UNDER THE PSM WITH 70% ETHANOL AND CLEAN CAREFULLY THE HOOD AND THE TUBULURES WITH SURFA'SAFE FOR AROUND 10 MIN

Materials

Technical material

- 40 µm cell strainer (BD Falcon)
- Countess® automated cell counter (Invitrogen)
- Low-adherent 6-well plates (Nunc Thermofisher Scientific)

Reagents

- TrypLe Express - EDTA 0.02%, 4 °C (Gibco)
- DMEM/F12 (Invitrogen)
- Calcium-magnesium-free Phosphate Buffer Saline pH 7.4 1X (Gibco)
- EGF (Sigma)
- βFGF (R&D Systems)
- Heparin (Sigma)
- Human proliferation supplement (NeuroCult® ; Stem Cell Technologies)
- Bovine serum albumin (Sigma)
- Penicillin G-streptomycin solution (Gibco)

Solution

- NSA tumorspheres culture specific medium for 100 ml

medium DMEM/F12	1X	87.07 ml
rhEGFR	20 ng/mL	20 µl
rh FGF-β	10 ng/mL	10 µl
heparin	4 ng/mL	400 µl
human Neurocult®	10%	10 ml
BSA	0.15%	1.5 ml
Penstrep	1%	1 ml

Thawing of cells and culture of tumorspheres

1. Discard the culture medium from the culture flask
2. Lift up the cells with TrypLE Express using a sterile pipette tip (2 ml for a T25/10cm dish or 5 ml for a T75)
3. Incubate at 37 °C, 5% CO₂ for 3–5 min
4. Check the detachment of the cells under an inverted phase-contrast microscope. The cells have to be in suspension.
5. Add culture medium with a sterile pipette tip (6 ml for a T25/10cm dish or 10 ml for a T75) to inactivate the TrypLE express action
6. Gently pipette up and down to form a single cell suspension, transfer in a 15 ml sterile Falcon tube
7. Rinse the original dish or flask with 1–3 ml of culture medium and transfer in the same 15 ml sterile Falcon tubes
8. Centrifuge at 1000 rpm for 6–8 min at room temperature
9. Aspirate off the supernatant and resuspend the cells in appropriate volume of warmed calcium-magnesium-free PBS in order to rinse thoroughly cells from serum.
10. Achieve three times the last two phases
11. Resuspend cells with 10 ml of warmed calcium-magnesium-free PBS and gently pipette up and down to form a single cell suspension
12. Filter cell suspension through a 40 µm cell strainer and rinse it three time with 5 ml to recover the maximum of cells
13. Centrifuge at 1000 rpm for 6–8 min at room temperature
14. Aspirate off the supernatant and resuspend the cells in appropriate volume of warmed NSA tumorspheres culture specific medium to achieve the correct concentration of cells for the experiments.
15. Distribute cells into low-adherent 6-wells dishes.
16. Gently rock the culture flasks to evenly distribute the cells and place them into a 37 °C incubator at 5% CO₂

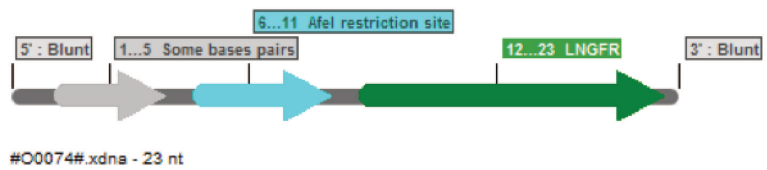
Appendix E

Design strategies developed to construct the molecular platforms

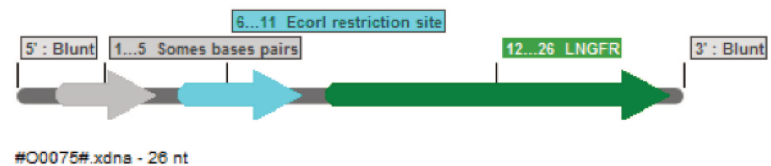
E.1 Design and construction of #B3013#

Construction #I0020#

PCR on #B2009#



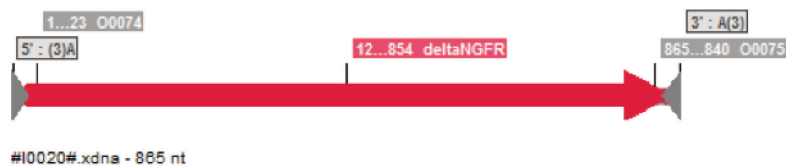
Oligo F : #O0074#



Oligo R : #O0075#

DNA matrix : #B2009# Size of the PCR product : 865 bp

NAME OF THE RESULT : #I0020-I1#



Digestion of #I0020-I1#

Restriction enzymes : AfeI (9) and EcoRI (856) Expected fragment sizes : 8bp and 11bp and 847bp Size of the fragment to keep : 847 bp

NAME OF THE RESULT :	#I0020#
----------------------	---------

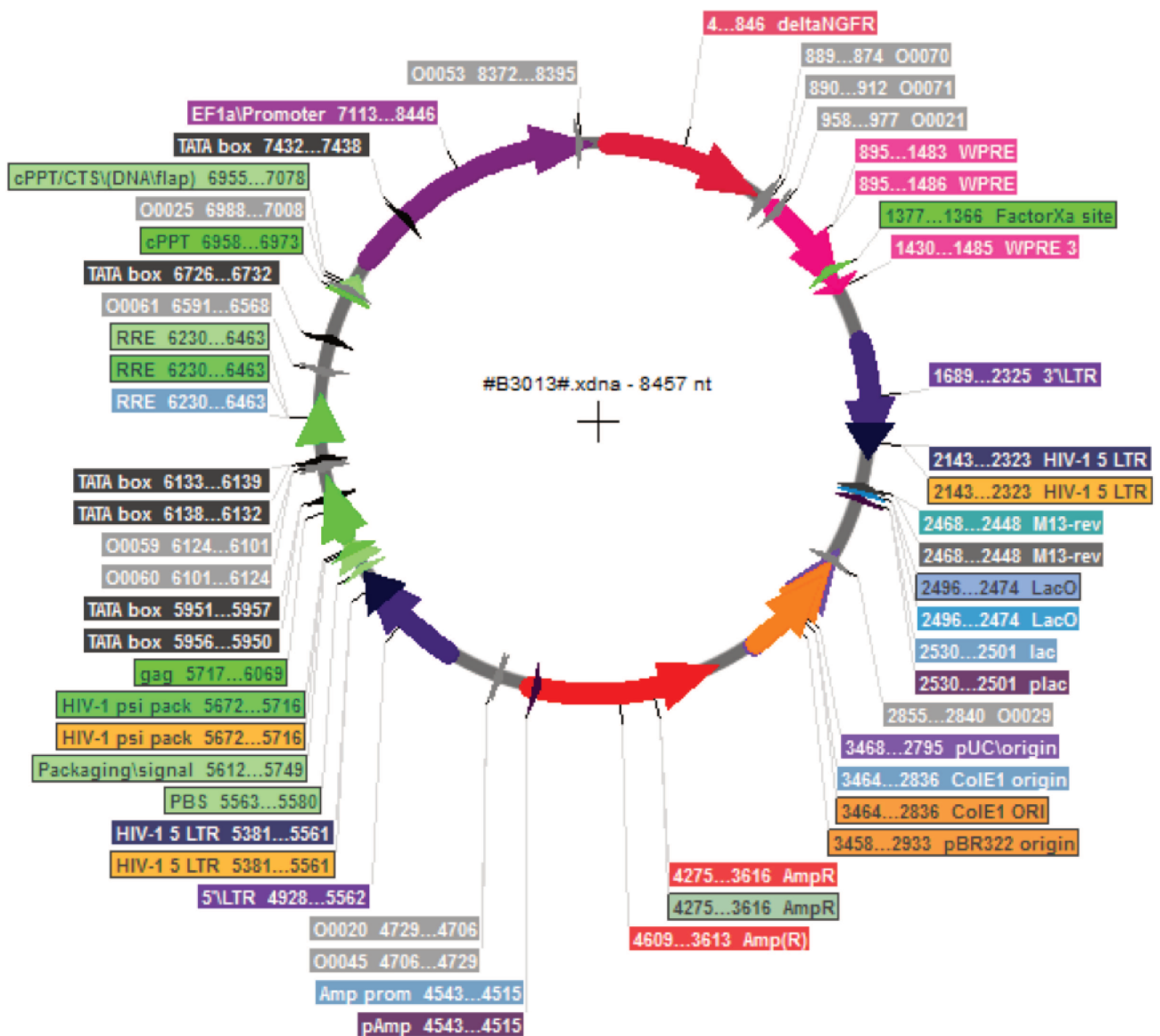
Construction #Bb0014#**Digestion of #B3015#**

Restriction enzymes : AfeI (7566) and EcoRI (7601) Expected fragment sizes : 33 bp and 7606 bp Size of the fragment to keep : 7606 bp

NAME OF THE RESULT :	#BBO014#
----------------------	----------



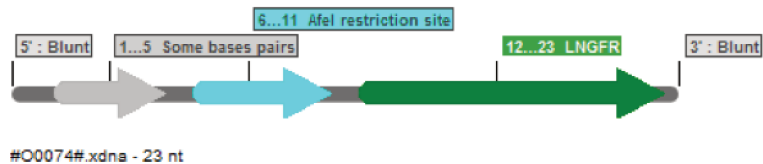
#Bb0014.xdna - 7606 nt

Construction #B3013#**Ligation of #I0020# with #Bb0014**

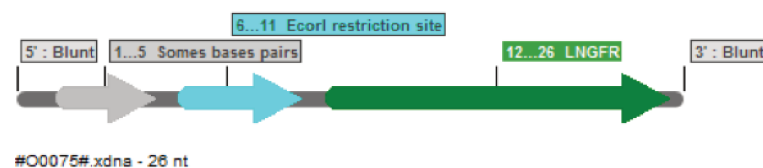
E.2 Design and construction of #B3014#

Construction #I0020#

PCR on #B2009#



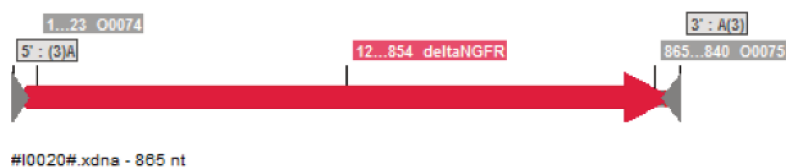
Oligo F : #O0074#



Oligo R : #O0075#

DNA matrix : #B2009# Size of the PCR product : 865 bp

NAME OF THE RESULT :	#I0020-I1#
----------------------	------------



Digestion of #I0020-I1#

Restriction enzymes : AfeI (9) and EcoRI (856) Expected fragment sizes : 8 bp and 11 bp and 847 bp Size of the fragment to keep : 847 bp

NAME OF THE RESULT :	#I0020#
----------------------	---------

E.2.1 Construction #Bb0015#

Digestion of #B3016#

Restriction enzymes : AfeI (557) and EcoRI (576) Expected fragment sizes : 17 bp and 6758 bp Size of the fragment to keep : 6758 bp

NAME OF THE RESULT :	#B0015-I1#
----------------------	------------

Dephosphorylation of #Bb0015-I1#

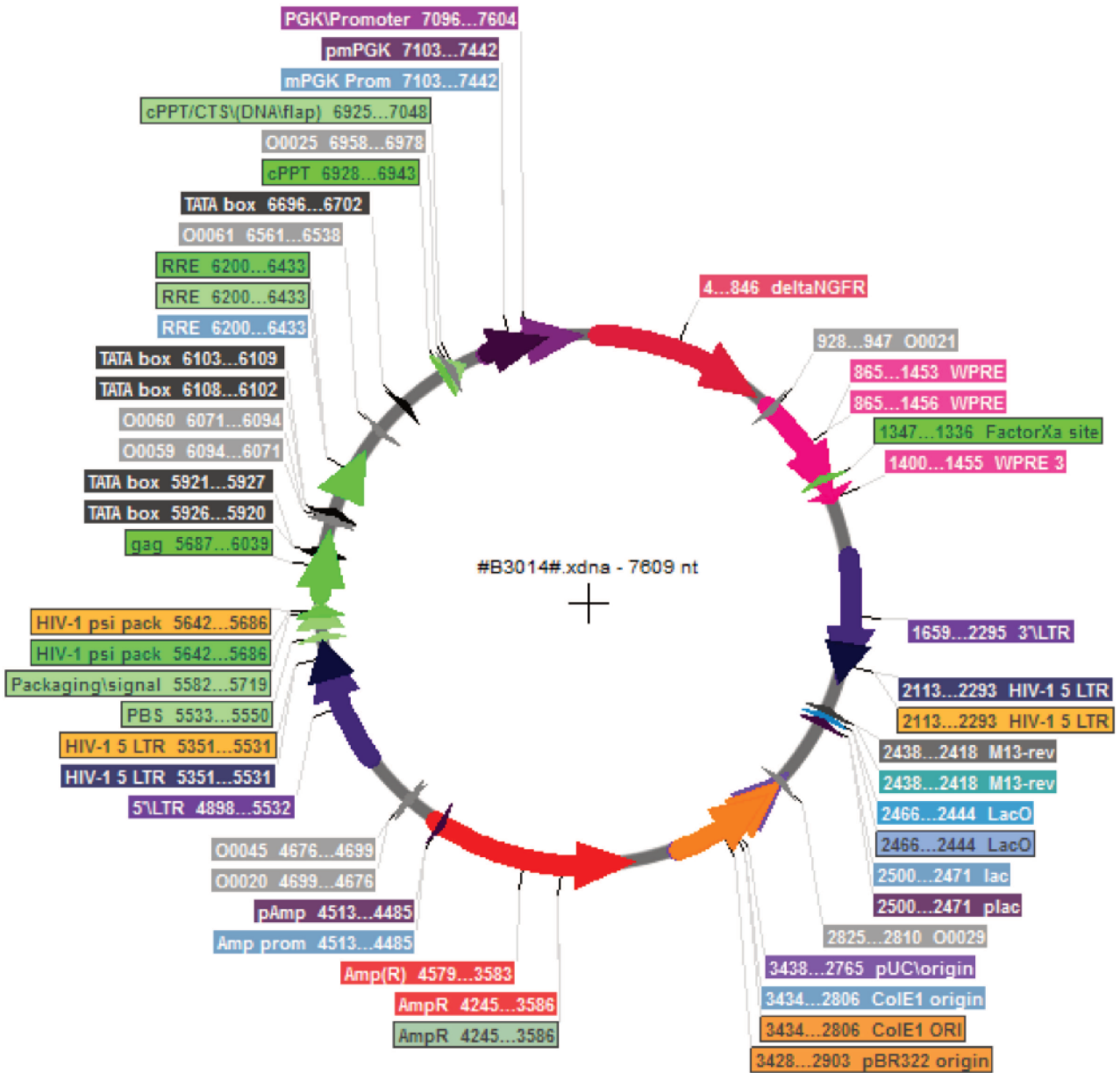
NAME OF THE RESULT :	#B0015#
----------------------	---------



#Bb0015#.xdna - 6758 nt

E.2.2 Construction #B3014#

Ligation of #I0020# with #Bb0015



E.3 Design and construction of #B3039#

Construction #Bb0026#

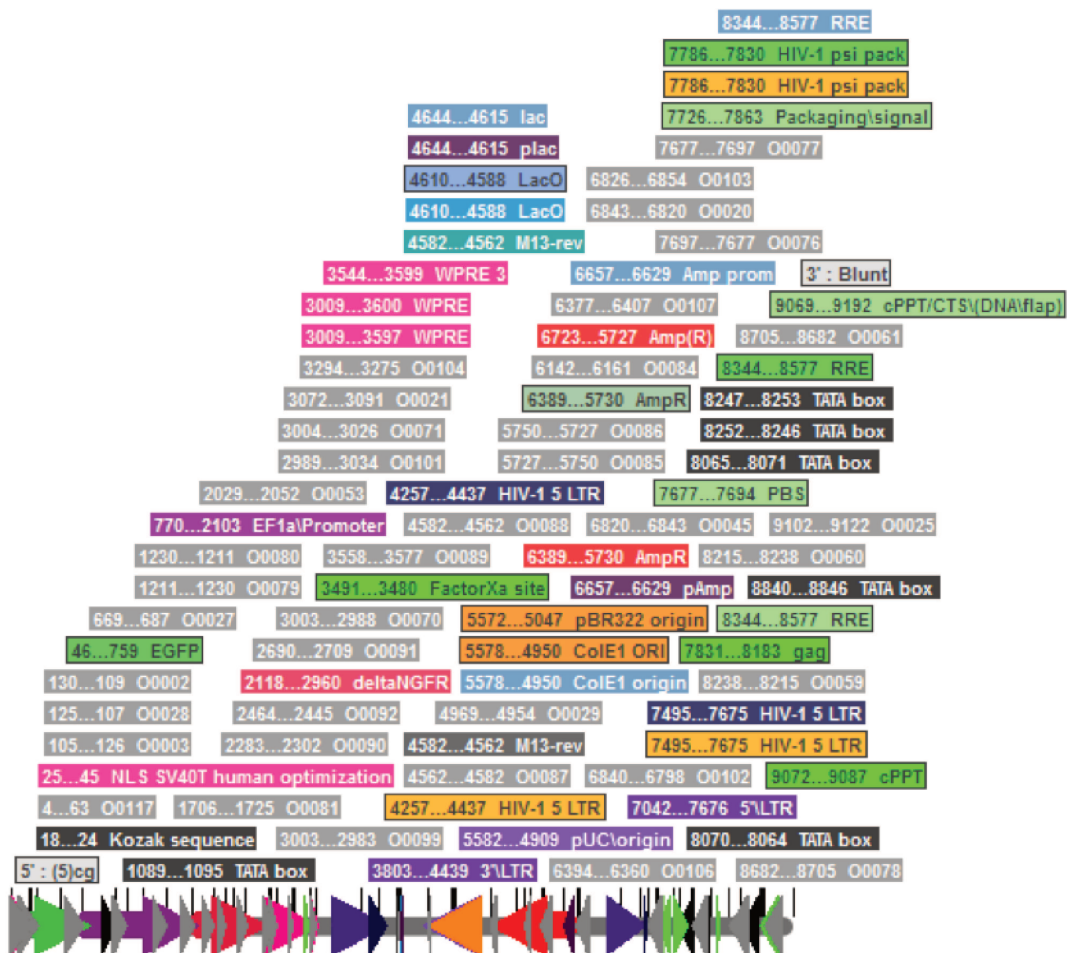
Digestion of #B3019#

Restriction enzymes : BstBI (9253) and HpaI (9247) Expected fragment sizes : 9260bp and 5bp
Size of the fragment to keep : 9260bp

NAME OF THE RESULT :	#BBO026-I1#
----------------------	-------------

Dephosphorylation of #Bb0026-I1#

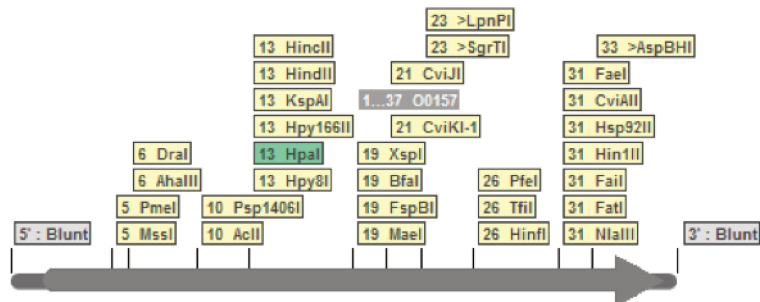
NAME OF THE RESULT :	#BBO026#
----------------------	----------



#Bb0026#.xdna - 9260 nt

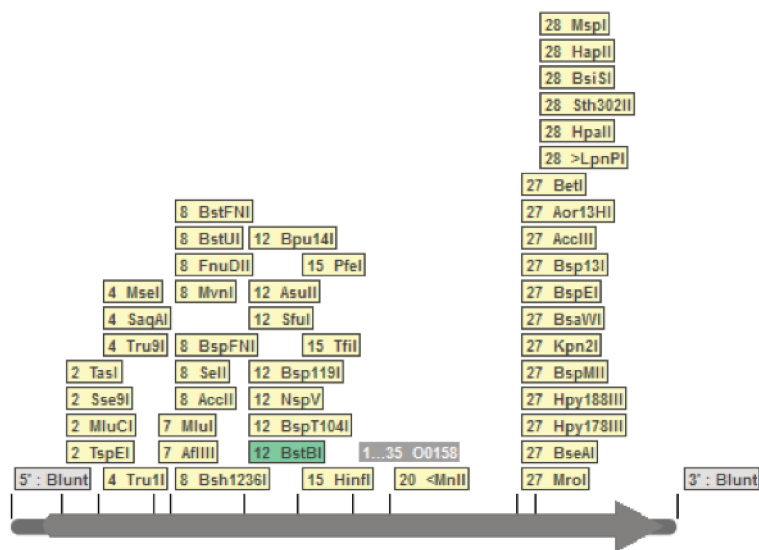
Construction #I0040#

PCR on #A2027#



#O0157#.xdna - 37 nt

Oligo F : #O0157#

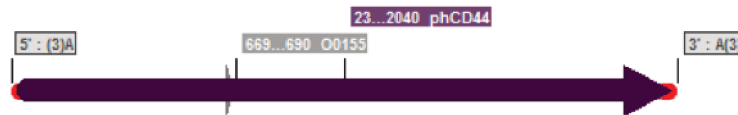


#O0158#.xdna - 35 nt

Oligo R : #O0158#

DNA matrix : #A2027# Size of the PCR product : 2061bp

NAME OF THE RESULT : #I0040#



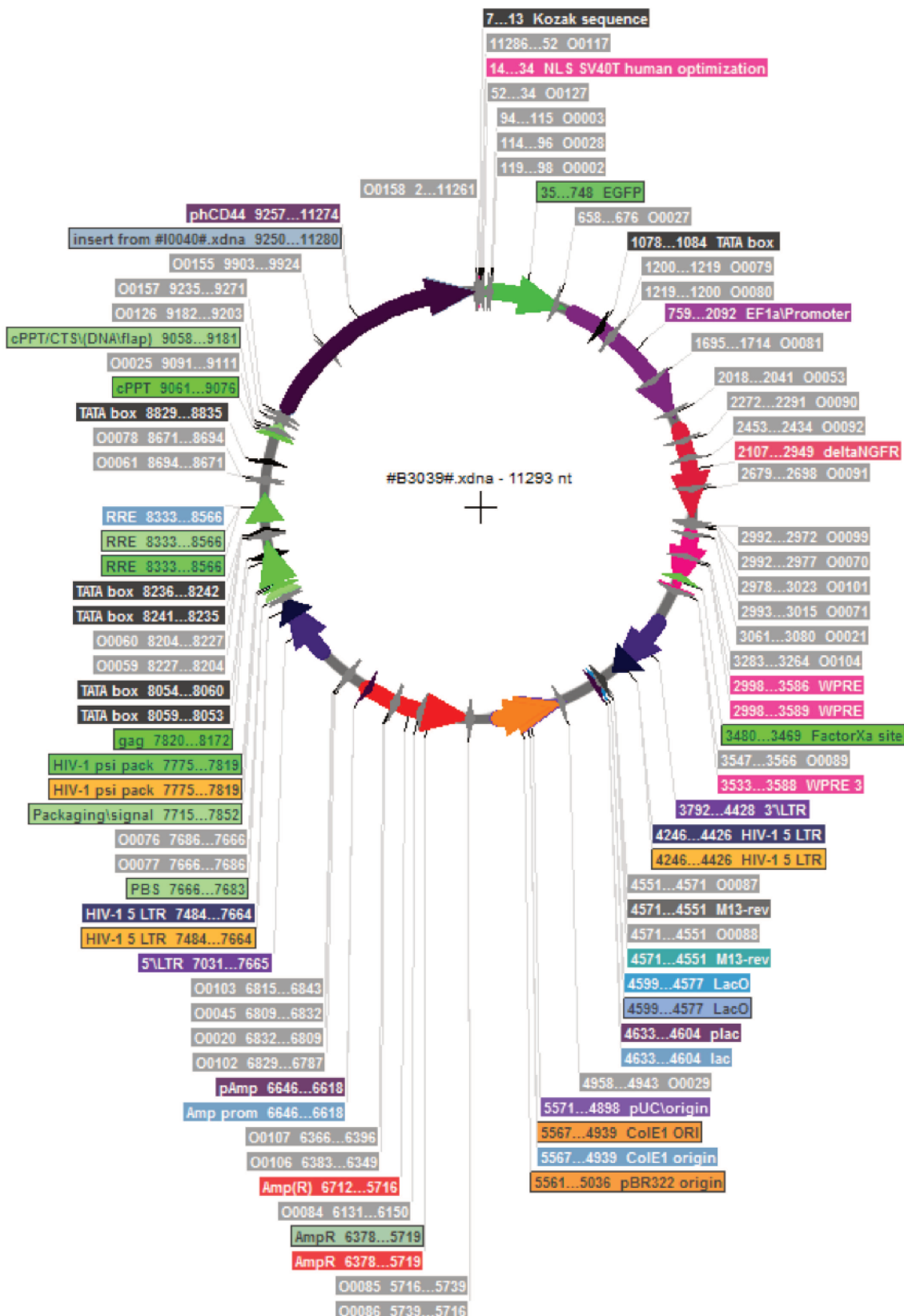
#I0040#.xdna - 2061 nt

E.3.1 Construction #B3039#

Homologous recombination between #I0040# and #Bb0026#

Enzyme for homologous recombination : In-Fusion

NAME OF THE RESULT : #B3039#



E.4 Design and construction of #B3043#

Construction #Bb0026#

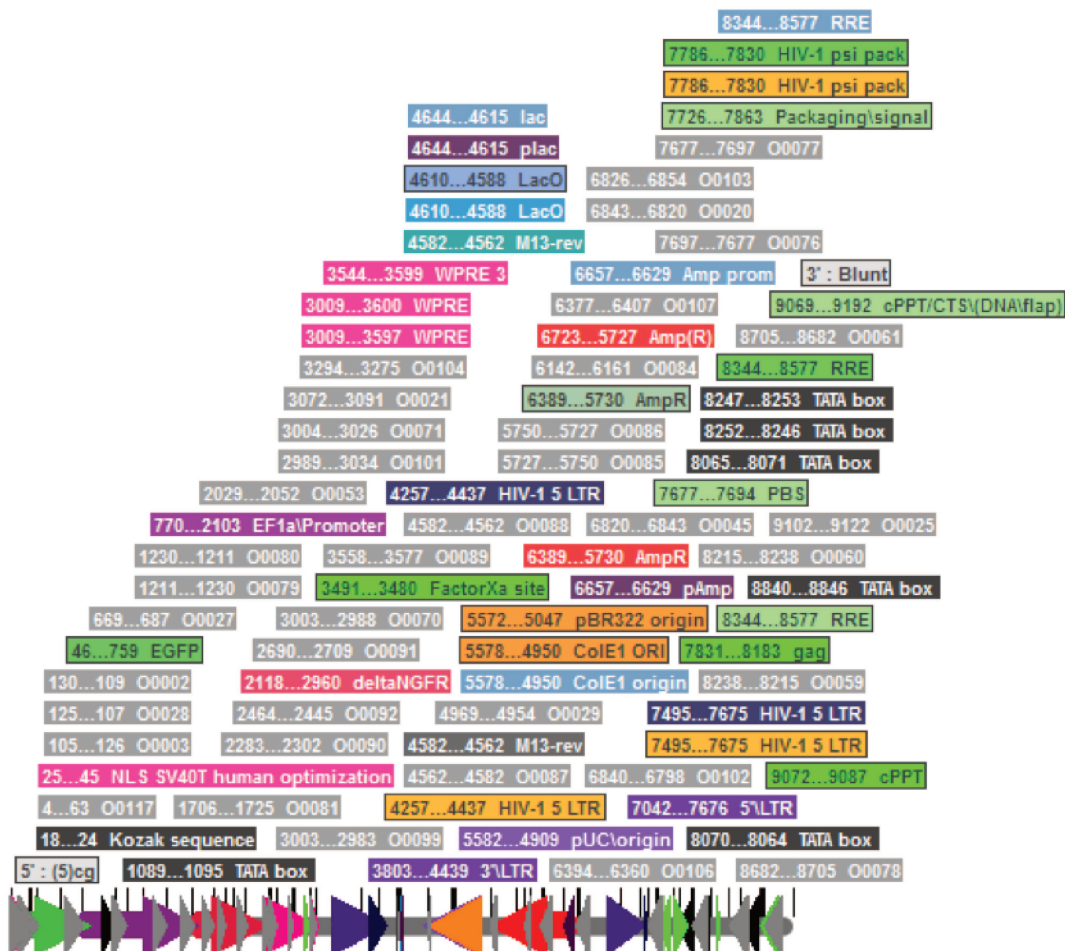
Digestion of #B3019#

Restriction enzymes : BstBI (9253) and HpaI (9247) Expected fragment sizes : 9260bp and 5bp
Size of the fragment to keep : 9260bp

NAME OF THE RESULT :	#BBO026-I1#
----------------------	-------------

Dephosphorylation of #Bb0026-I1#

NAME OF THE RESULT :	#BBO026#
----------------------	----------



#Bb0026#.xdna - 9260 nt

E.4.1 Construction #I0041#

PCR on #A2028#



#O0159#.xdna - 39 nt

Oligo F : #O0159#



#O0160#.xdna - 35 nt

Oligo R : #O0160#

DNA matrix : #A2028# Size of the PCR product : 1411bp

NAME OF THE RESULT :

#I0041#

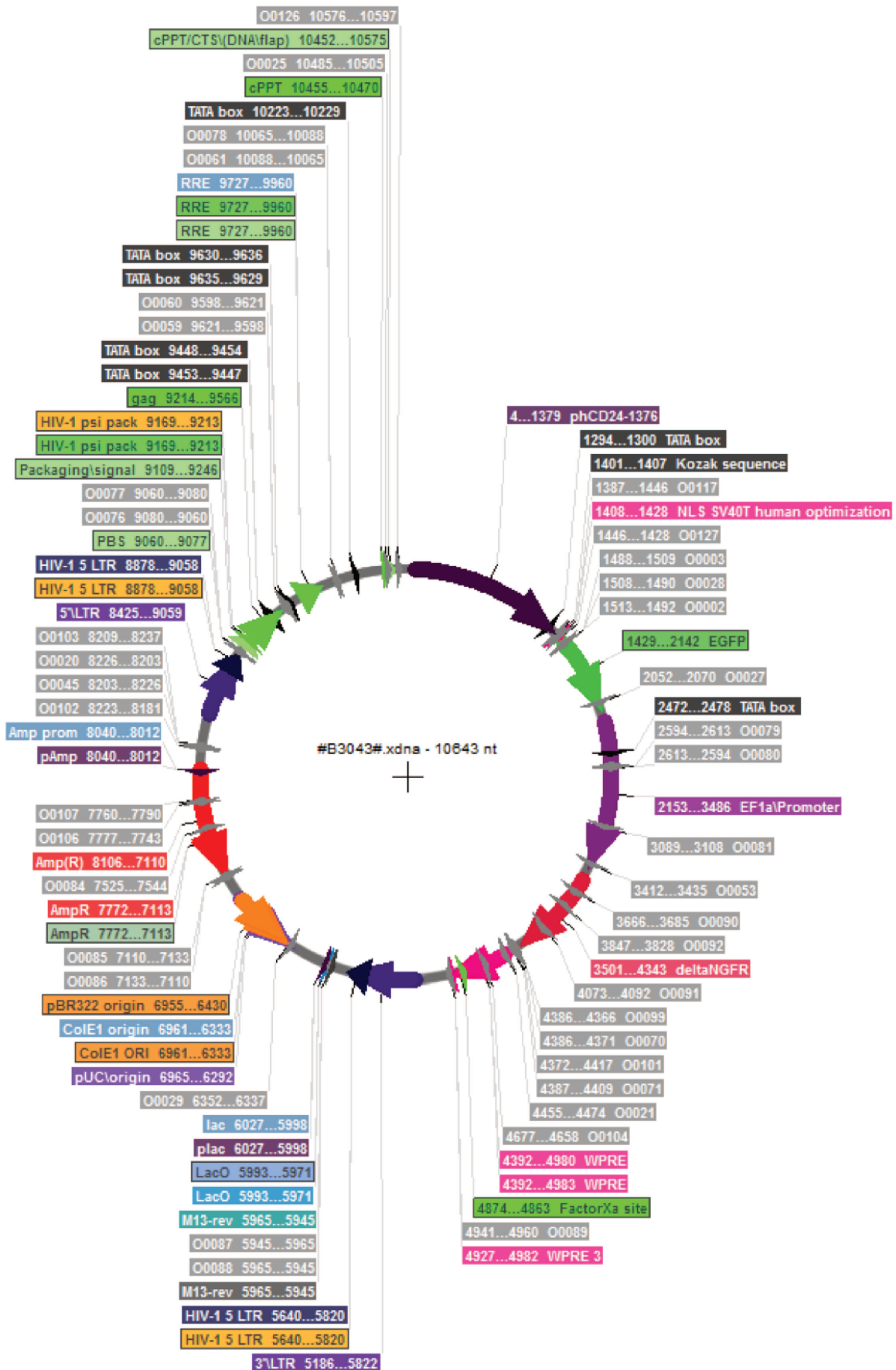
E.4.2 Construction #B3043#

Homologous recombination between #I0041# and #Bb0026#

Enzyme for homologous recombination : In-Fusion

NAME OF THE RESULT :

#B3043#



E.5 Design and construction of #B3047#

Construction #Bb0026#

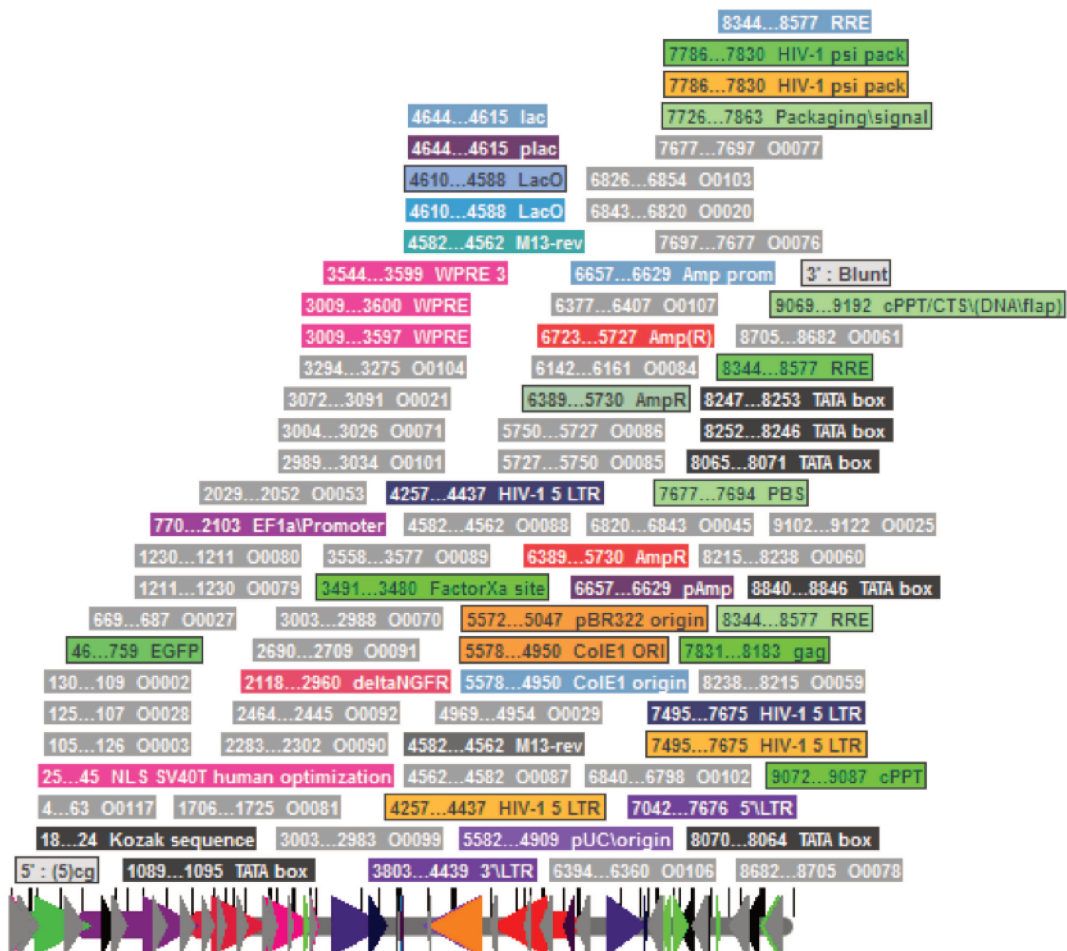
Digestion of #B3019#

Restriction enzymes : BstBI (9253) and HpaI (9247) Expected fragment sizes : 9260bp and 5bp
Size of the fragment to keep : 9260bp

NAME OF THE RESULT :	#BBO026-I1#
----------------------	-------------

Dephosphorylation of #Bb0026-I1#

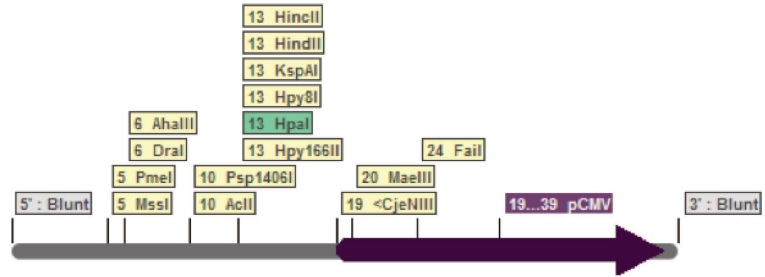
NAME OF THE RESULT :	#BBO026#
----------------------	----------



#Bb0026#.xdna - 9260 nt

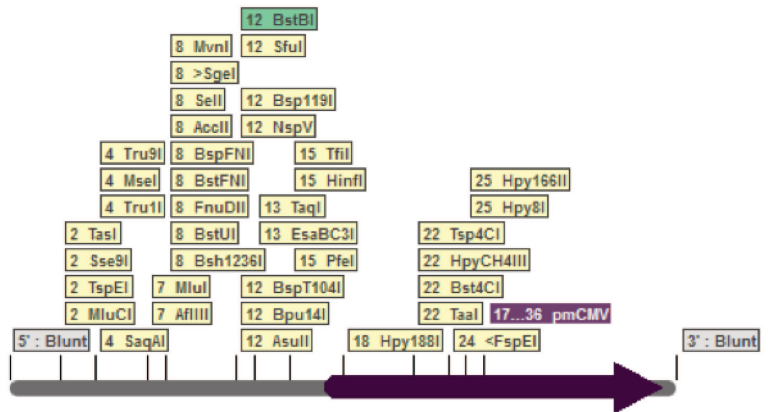
Construction #I0036#

PCR on #A2001#



#O0151#.xdna - 39 nt

Oligo F : #O0151#

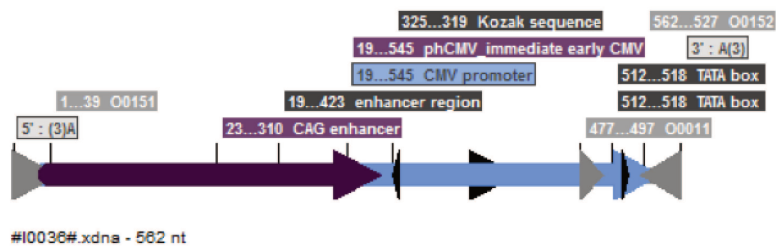


#O0152#.xdna - 36 nt

Oligo R : #O0152#

DNA matrix : #A2001# Size of the PCR product : 562bp

NAME OF THE RESULT :	#Ioo36#
----------------------	---------



#I0036#.xdna - 562 nt

Construction #B3047#

Homologous recombination btw #I0036# and #Bb0026#

Enzyme for homologous recombination : In-Fusion

NAME OF THE RESULT :	#B3047#
----------------------	---------

E.6 Design and construction of #B3051#

Construction #Bb0026#

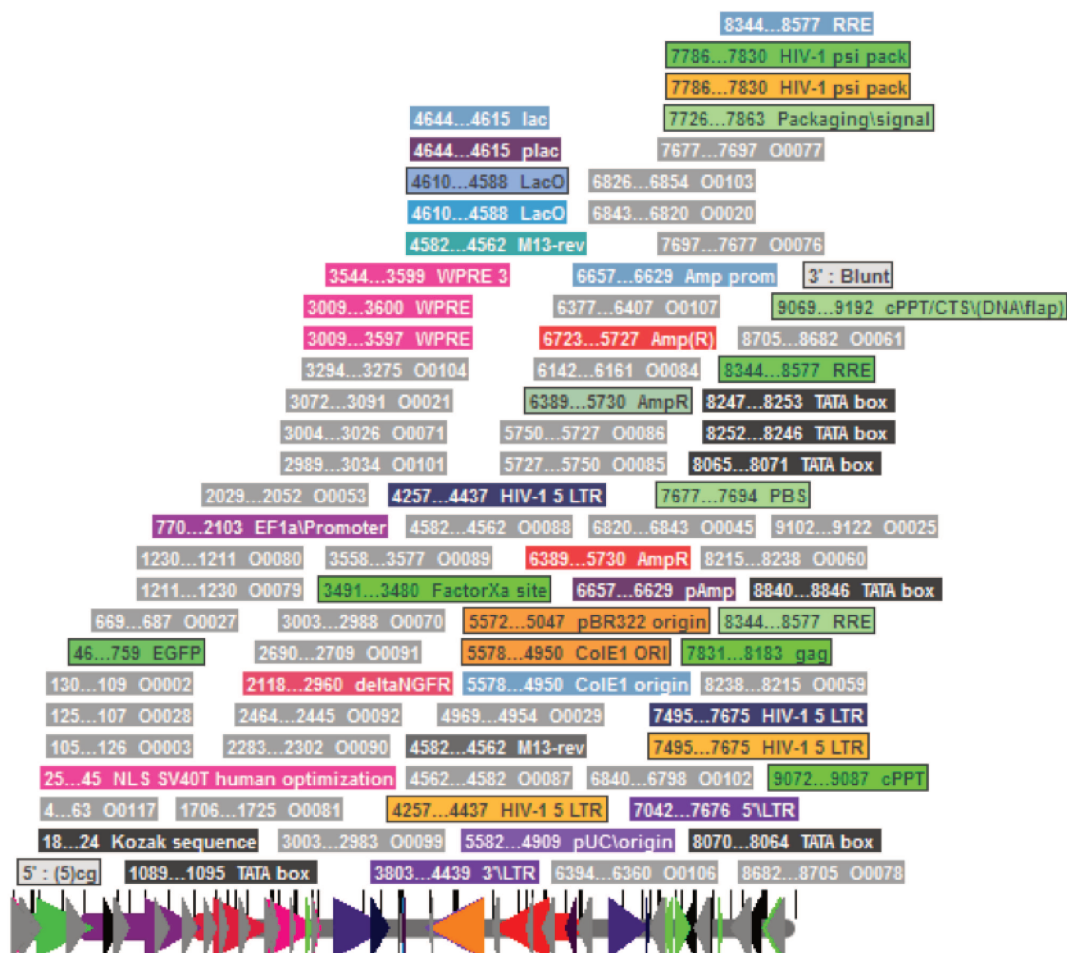
Digestion of #B3019#

Restriction enzymes : BstBI (9253) and HpaI (9247) Expected fragment sizes : 9260bp and 5bp
Size of the fragment to keep : 9260bp

NAME OF THE RESULT :	#BBO026-I1#
----------------------	-------------

Dephosphorylation of #Bb0026-I1#

NAME OF THE RESULT :	#BBO026#
----------------------	----------



#Bb0026#.xdna - 9260 nt

Construction #I0042#**PCR on #A2029#**

#O0159#.xdna - 39 nt

Oligo F : #O0159#

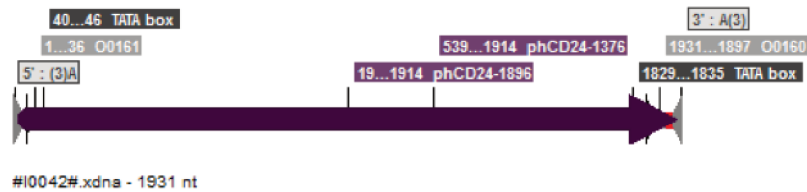


#O0161#.xdna - 36 nt

Oligo R : #O0161#

DNA matrix : #A2029# Size of the PCR product : 1931bp

NAME OF THE RESULT :	#I0042#
----------------------	---------



#I0042#.xdna - 1931 nt

E.6.1 Construction #B3051#**Homologous recombination between #I0042# and #Bb0026#**

Enzyme for homologous recombination : In-Fusion

NAME OF THE RESULT :	#B3051#
----------------------	---------

E.7 Design and construction of #B3056#

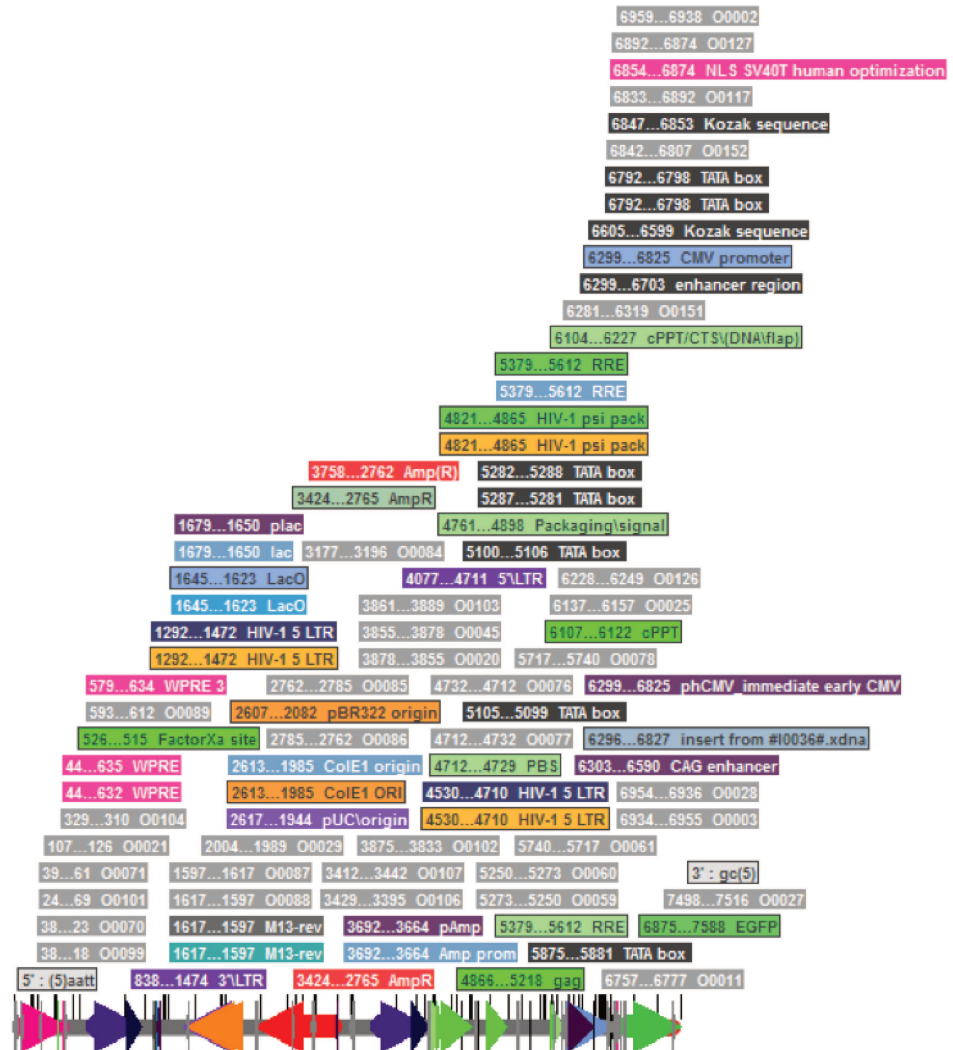
Construction #Bb0039#

Digestion of #B3047#

Restriction enzymes : ClaI (763) and EcoRI (2961) Expected fragment sizes : 7593bp and 2195bp

Size of the fragment to keep : 7593bp #Bb0039-I1#

E.7.1 Blunting procedure of #Bb0039-I1#



#Bb0039#

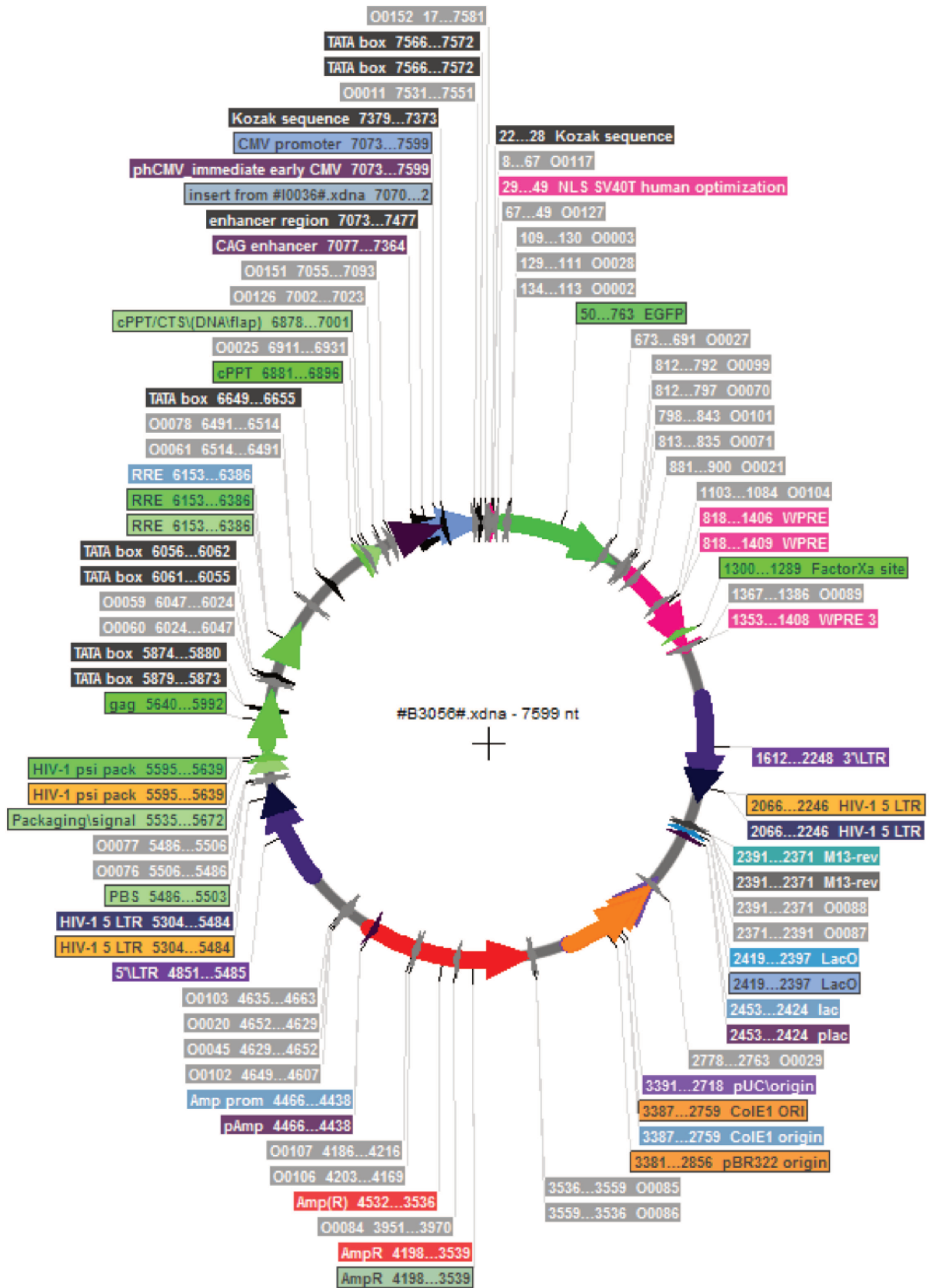
#Bb0039#.xdna - 7593 nt

Construction #B3056#

Ligation of #Bb0039#

NAME OF THE RESULT :

#B3056#



As a note, the following molecular constructs : #B3060#, #B3061# and #B3062# have been construct on the same basis of #B3056# *cf.* appendix E.7 on page 306

Appendix **F**

Various tables

Those three following tables quickly gather the various libraries of :

- Molecular constructs that have been constructed and/or used within the part IV on page 143.
- Selected oligos designed in order to construct or validate by sequencing the above-described molecular constructs.
- Primary and secondary antibodies coupled with fluorochromes which validated SUM149-PT and SUM159-PT cell lines phenotypic heterogeneity, tracked and quantified phenotypic plasticity for both 2D and 3D cell culture conditions, validated intermediate constructs of the PhenotypeTracker molecular platform and initiated microscopy experiments on the spatial organization of cell phenotypes.

F.1 Listing of molecular constructs

F.2 Listing of oligos

name oligos (code)	full name oligos	description (sequence) 5'-->3'	TM (°C)
			calculate on per/primer
#00016	B1001 pEF1alpha Forward	TTTTATCGACTAGT-ATCGAAGCTTTGCAAAGATGGATAAAG	63.04
#00017	B1001 eGFP Reverse	ATTGCGTAGCTCTAGA-TTACTTTGTACAGCTCGTCCATGC	63.30
#00018	Swai SV40latepolyA Forward	TTAAT-ATTTAAAT-CAGACATGATAAGATACATT	50.08
#00019	NotI SV40latepolyA Reverse	TTAAT-GCGGCCGC-TACCACATTTGTAGAGGTTT	54.31
#00020	SV40 polyA Forward	TTGGACAAACCACAACAGAATGC	62.29
#00021	WPRE Forward	TTTACGCTATGTGGATACGC	58.32
#00022	phEF1 alpha Forward	GAGAGGAATCTTTGCAGCTAATGG	62.04
#00023	EcoRI eGFP/mCherry Forward	TTAAT-GAATTC-ATGGTGAAGCAAGGGCAGG	69.30
#00024	BstBI synthetiopolyA_LabOmics Reverse	TTAAT-ITCGAA-ACAGAAAACATCTTACTTGGCATGA	67.83
#00025	cPPT Forward	GCAGGGGAAAAGATAGTAGAC	58.01
#00026	puromycinR Reverse	ACCGTGGGCTTACTCT	58.94
#00027	EGFP Forward	CAAAAGCCCCAAGGAAAG	58.53
#00028	EGFP Reverse	AACCTGTGGCCGTTTACGT	60.62
#00029	ColE1 ORI Forward	AAAAACGCGCAAGCAAGC	59.61
#00030	mCherry Reverse	TACGAAACGGCCGAGG	62.42
#00033	A1004 eGFP Forward	AAAAAGCAGGCTTCGAA-CACAGAAAAGCATCTTA	64.55
#00034	A1004 polyALabOmics Reverse	CATAGTCTTGAATTC-CATGGTGAGCAAGGGCGAG	76.16
#00035	A1005 RFP Forward	AAAAAGCAGGCTTCGAA-CACAGAAAAGCATCTTA	69.42
#00036	A1005 polyALabOmics Reverse	GATTGTCTGGAATTC-CATGGTGAGCAAGGGCGAG	74.54
#00037	B3001-I1 polyALabOmics Forward	AGCGAATTCGAAATTAAT-GAGCAGACATGATAAGATAC	67.51
#00038	B3001-I1 polyALabOmics Reverse	CAGATCTTCGCGCCGCTTTACCACATTTGTAGAG	73.33
#00039	AttL2 Reverse	TGCCAACTTTGTACAAGAAAGC	60.40
#00040	AttL1 Forward	TTTATAATGCCAACTTTGTACAAA	56.02
#00041	AttB1 Forward	ACAAGTTTGTACAAA	44.80
#00042	AttB2 Reverse	TTTTGTACAAAGAAAGCTG	49.45
#00043	LacZ alpha Forward	GGAAAGGGCGATCGGTG	59.87
#00044	Luciferase Reverse	CCGGGCGCTTCTTTATGTTT	58.92
#00045	SV40 polyA Reverse	GCATCTTAGTGTGTGTTGTCCAA	62.29
#00046	Luciferase Forward	CCGGGCGCTTCTTTATGTTT	58.92
#00047	BamHI EF1alpha Forward	TTTAT-GGATCC-AGGC TTTGCAAAGATGGATAAAG	66.55
#00048	BamHI EF1alpha Reverse	TTTAT-GGATCC-GTCGACTGCAGAAATCCTCAC	69.93
#00049	TurboRFP Forward	ATGAGCCGAGCTGATCAAGGA	61.54
#00050	TurboRFP Reverse	GTGCATGTTCTCCTTGATCAGC	62.34
#00051	TurboRFP Reverse	TACACATTTGATCCTAGCAGAAGCA	62.20
#00052	TurboRFP Forward	CGACCTCCCTAGCAAACTGG	62.73
#00053	pEF1alpha Forward	GAGTTTGGATCTTGGTTCATCTCTC	60.12
#00054	ZsGreen1 Reverse	GCAGGAGTTCTTGAAGTAGTCG	61.17
#00055	H2B-eGFP(noATG) Forward	CCACCCTGTCGCACCCTGAGCAAGGGCGAG	80.31
#00056	H2B-eGFP(noATG) Reverse	CTCGCCCTTGTCTCAGGTGGCCAGCCGTGG	80.31
#00057	HIV1 psi packaging Forward	TTGTATTTAAAGTGCTTAGCTCG	58.70
#00058	HIV1 psi packaging Reverse	TGCTTTCAAGTCCCTGTTC	62.60
#00059	GAG Reverse	CTCCAAATGTCCTCATATCTCT	61.79
#00060	GAG Forward	AGGAGATATGAGGACAAATGGAG	61.79
#00061	GAG Reverse	TAATTTCTCTGTCCCACTCCATCC	62.21
#00062	Env Forward	TCTCTGGAACAGATTGGAATCAC	60.57
#00063	eGFP Kozack phK8 Forward	TTGCTCACCATGGTTC CAAGA ACTATGAAAT	68.48
#00064	eGFP Kozack phK8 Reverse	ACTTCATAGTCTTGAACACCATGGTGAGCAA	68.48
#00065	mCherry Kozack phK5 Forward	CCTTGTCAACATGTTTCCAGGACAATCCA	73.26
#00066	mCherry Kozack phK5 Reverse	TGGATTGTCTCTGGAACACCATGGTGAGCAAGG	73.26
#00067	Chloramphenicol Forward	TCCGTATGGCAATGAAAGACG	61.23
#00070	MCS Reverse	ACCGGTGGATCCCGGG	63.57
#00071	WPRE Forward	GGAAACATCAACCTCTGGATTAC	59.56
#00072	MCS/phPGK Reverse	GGTTAACTCGAGCGCT-GTCGAAAGCCCCGA	77.42
#00073	MCS/WPRE Forward	GGGAATTCATG CATCGAT-ACAATCAAACCTCTG GATTAC	70.01
#00074	AjellLNGFR Forward	TTAAT-AGCGCT-ATGGGGGCGAGGT	66.23
#00075	EcoRI/LNGFR Reverse	TTAAT-GAATTC-CTAGAAGATCCCCCT	62.97
#00076	PBS Reverse	CAAGTCCCTGTTCGGGCGCCA	70.13
#00077	PBS Forward	TGGCGCCGAAACAGGGACTTG	70.13
#00078	GAG Forward	GGATGGAGTGGGACAGAGAAATTA	62.21
#00079	pEF1alpha Forward	GCCTTGAATTACTTCCAGC	60.02
#00080	pEF1alpha Reverse	GCCTTGAATTACTTCCAGC	60.02
#00081	pEF1alpha Forward	TTGCGTGGAGCGGAAAGATGG	63.9
#00082	AcGFP Forward	AACTACAAAGCCCAATGTTG	61.91
#00083	puromycinR Forward	GAGTACAAGCCCAAGGT	58.94
#00084	AmpR Forward	AAAGCGGTAGCTCCTTCGG	63.01
#00085	AmpR Forward	TTACCAATGCTTAAATCATGAGAGGC	61.70
#00086	AmpR Reverse	GCCTCACTGATTAAGCATTGGTAA	61.70
#00087	M13 Forward	CATGTCATAGCTGTTCTCTG	59.19
#00088	M13 Reverse	CAGGAAACAGCTATGACCATG	59.19
#00089	WPRE Forward	CCTCAGACGAGTCGGATCTC	61.64
#00090	LNGFR Forward	CCTTGTGGAGC CAACAGAC	63.53
#00091	LNGFR Forward	CCCTGGCCGTTGGATTACAC	63.66
#00092	LNGFR Reverse	TCTGGTAGTAGCCGTAGGC	63.13

Table F.3 : Data list of oligos. - Part 1

name oligos (code)	full name oligos	description (sequence) 5'-->3'	TM (°C)
#00093	#I0021# Forward	CGG GCGCG CCATATGCGG CCGGTATACG TTTAAACGTTAACTTCG AACGCG TTAATTAATTT AAAT	80.19
#00094	#I0021# Reverse	CGATTTAAATTAATTAACGCGTTCGAAGTTAACGTTT AAACGTATACGCGGCCCATATGGCGCGCC	79.65
#00095	#I0022# Forward	CTAGGGCG CGCCATATGCG GCGCG GTATACGTTTAAACGTTAACTT CGAACG CGTTAATTAATTTAAAT	79.58
#00096	#I0022# Reverse	CTAGATTTAAATTAATTAACGCGTTCGAAGTTAACGTTTAAACGTATACGCGGCGCATATGGCGCGCC	79.06
#00097	pmK8HpaI Forward	AAC-GGTG GATCAC TTTGCC CCGCTCGTTT	71.13
#00098	pmK8BstBI Reverse	TTAAT-ITCGAA-GG GACAG CGCCCAGCGAAGGCCCTT	77.22
#00099	MCS Reverse	ACC GGTGG ATCCC GGGCCCGC	76.26
#00100	WPRE Forward	CGTCTGGAACATCAACCTCTGGATTACA	66.91
#00101	WPRE Reverse	CCG GGATC CACCG GT-GGAACAAT CAACCTCTGGATTACAAAATTTG	64.11
#00102	5'LTR Reverse	GACAAACCACAACCTA-GAATGCGAGTAAAAAATCGTTTATTG	60.45
#00103	5'LTR Forward	TAG TTGTG GTTTGT CCAAACTCATCAATG	64.66
#00104	WPRE Reverse	CGAAAGTCGGGAAAGGAGC	63.62
#00105	MCS/WPRE Forward	AGC GCTCG AGTTAAACCGG GAATT CATG CATCGAT-ACAATCAACCTCTGGATTACAAAATTTG TG	63.54
#00106	AmpR Reverse	CAATGATGAGCACTTTAAAGTTCTCTATGTGGC	68.58
#00107	AmpR Forward	TAAAGTGTCTCATCTTTGG AAAACGTTCTTC	65.25
#00108	MCS/phPGK Reverse	GGGTTAACTCGAGCGCT-GTCGAAAGCCCGGAGATGAG	65.35
#00109	phPGK Reverse	TAC CGGTG GATGTG GAATGTG	62.63
#00110	phPGK Forward	ATT CCACATCCACCGGTAGG	61.84
#00111	pmK8 Forward	TGG CATGC AAGACAAACAAATTGAG	63.35
#00112	pmK8 Forward	GGGTCTCACTTTACTGCCT	61.10
#00113	pmK8 Forward	AGG CAAGG AAGGAATATGT	58.88
#00114	pmK8 Forward	CCCATCAGTAATTTAATG CAGG	60.48
#00115	pmK8 Forward	AGAAGAGCTCGCGTTTGAATTCTC	61.93
#00116	pmK8 Forward	TACGTGACTCCCGAGGAGG	63.33
#00117	PacINLSFP Forward	GCG-TTAATTAAC-CGCCACC ATG-CC CAAGA AAGAAGC GCAAG GTG-GTGAGCAAGGGCGAGGAG	62.92
#00118	ClalNLSFP Reverse	GCG-ATCGAT-CTACTGTGACAGCTGTCCATG	60.98
#00119	PacINLSGFPdest Forward	GCG-TTAATTAAC-CGCCACC ATG-CC CAAGA AAGAAGC GCAAG GTG-G CCCAGTCCAAG CACGG C	66.87
#00120	ClalNLSFPdest Reverse	GCG-ATCGAT-CTACACATTTATCTAGCAGAGCA	62.96
#00121	pmK8HpaI Forward	ATA CGTTTAAACGTTTAAAC-CTAAATTTGAAG CGTTAATTT	56.55
#00122	pmK8BstBI Reverse	TAATTAACGCGTTCGAA-GTGGCACTTTTCGGGGAA	66.05
#00123	PacINLSRFPdest Forward	GCG-TTAATTAAC-CGCCACC ATG-CC CAAGA AAGAAGC GCAAG GTG-G AGCGAGCTGATCAAGGA AACTATGC	67.62
#00124	pmK8HpaI Forward	ATACGTTTAAACGTTTAAAC-GTGGATCACTTCCCGCTCCG	69.67
#00125	pmK8BstBI Reverse	TAATTAACGCGTTCGAA-GG GACAG CGCCCCAGCGAAGGC	73.28
#00126	cPPT Forward	ATTACAGG GACAG CAGAGATCC	61.81
#00127	SV40NLS Reverse	CTCCTCGCCTTGTCTCACC	65.40
#00128	NdeINLSFP Forward	GCG-CATATG-CGC CACCATG-CCCAAGAAG AAGCG CAAGGTG-GTGAGCAAGGGCGAGGAG	62.92
#00129	AscINLSFP Reverse	TTA-GGCG CGCC-CTACTGTACAG CTGCTCATG	60.98
#00130	DeltaNP63HpaI Forward	ATACGTTTAAACGTTTAAAC-TATTCACCCCTTAACTCTCCGTG	62.47
#00131	DeltaNP63BstBI Reverse	TAATTAACGCGTTCGAA-TG AGAGCCTTGGCGTGGGA	67.52
#00132	pmK8HpaI Forward	ATACGTTTAAACGTTTAAAC-GTGGATCACTTCCCGCTCCGTTG	71.63
#00133	pmK8BstBI Reverse	TAATTAACGCGTTCGAA-GG GACAG CGCCCCAGCGAAGGCC	76.57
#00134	pmK8 Forward	CGTTTGTAGCGCAGACAATG	63.36
#00135	pmK8 Reverse	CTG GGGGTATCGACCTGAACAG	65.71
#00136	pmK8 Forward	GGAGAACCCTGAACTCGCTCC	65.34
#00137	pmK8 Forward	CTG ACCACTTGGCCCTAGTCCC	66.83
#00138	pmK8 Forward	CCAGGGCTGTACACTCTGTCTG	64.87
#00139	pmK8 Forward	CACACACACACACAGGTAGGG	65.49
#00140	pmK8 Forward	AATTGCAGGTGGAATCTTTGCC	62.51
#00141	pmK8 Forward	TTGTTATTCCTCGAGGCGCTCTG	62.32
#00142	pmK8 Forward	CCTAGTGTGGCTGTTAAAGGC	65.70
#00143	pmK8 Forward	GCTCTCTCTGGATTTTAGGG	60.67
#00144	pmK5 Forward	GGATCCCCGGGTTCTTAAAC	62.99
#00145	pmK5 Reverse	GTTTAGGAACCCGGGATCC	62.99
#00146	pmK5 Forward	CCAGCCCACTTATCATCAG	63.23
#00147	polyA LabOmics Forward	CGCGCC-CTAGAAATAATTCTTACTGTCATGCCAAGTAAGATGCTTTCTG-GG	75.55
#00148	polyA LabOmics Reverse	CGCGCC-CAGAAAAGCATCTTACTGGCATGACAGTAAGAAATATTCTAGA-GG	75.55
#00151	phCMVHpaI Forward	ATACGTTTAAACGTTTAAAC-CCGTTACATAACTTACGGTAAA	56.42
#00152	phCMVBstBI Reverse	TAATTAACGCGTTCGAA-ATCTGACGTTCACTAAACG	58.10
#00153	phK5 Reverse	CTG GTCTG GGTCTTCCAGGGA	65.93
#00154	phK5 Forward	GGTGTCCCATGGCCTTGAGGAC	67.66
#00155	phCD44 Forward	AATATTACA AACTCTTCTCTC	53.47
#00156	phCD44 Reverse	GAGGAGCTGGAGAGAGGGCGA	67.76
#00157	phCD44HpaI Forward	ATACGTTTAAACGTTTAAAC-CTAGCCTGAATCATGCTG	58.14
#00158	phCD44BstBI Reverse	TAATTAACGCGTTCGAA-TCGAGGGTTCGCGGAGCG	66.59
#00159	phCD24-1376 Forward	ATACGTTTAAACGTTTAAAC-ATACTCCACAGTAAAATTAC	53.09
#00160	phCD24 Reverse	TAATTAACGCGTTCGAA-CAGGATGCTGGGTGCTTG	61.21
#00161	phCD24-1896 Forward	ATACGTTTAAACGTTTAAAC-CACGCCCGGCCAAGTAT	63.26
#00162	phCD24 Forward	TGTTCCCAAGCACATCTTAAACG	62.13
#00163	phCD24 Forward	CCATACTAGTTTGTCTCTG	57.66
#00164	phCD24 Forward	GAATGACGGCCACAGAGAC	61.50
#00165	phCD24 Forward	CGCTCACAGAACAAAGCAAGG	62.82
#00166	phCD24 Forward	TGAGCTCGAGTGGCAATG C	63.86
#00167	phCD24 Forward	TCACGGCTATTGTGCTTTCC	63.50
#00168	PacIFP Forward	GCG-TTAATTAAC-CGCCACC ATG-GTGAGCAAGGGCGAGGAG	62.92
#00169	PacIFPdest Forward	GCG-TTAATTAAC-CGCCACC ATG-GCCAGTCCAAGCAGGGC	66.87
#00170	PacIRFPdest Forward	GCG-TTAATTAAC-CGCCACC ATG-AGCGAGCTGATCAAGGAGAACATGC	67.62

Table F.4 : Data list of oligos. - Part II

F.3 Listing of antibodies

This list summarizes each antibodies that were used over the experiments.

The selection of CD24 and CD44 antibodies, used over the above-described experiments, is based upon the antibodies previously used by the team of Piyush B. Gupta. These later have been identified within its publication in Cell in year 2011 [113]. However other antibodies are also available from various suppliers. Thus, I encourage you to follow the below hyperlinked links in order to take into consideration antibodies which have already been reported within publications (manufacturing house, clone, concentration) :

- Data list of CD24 antibodies used over publications.
- Data list of CD44 antibodies reported over publications.

name	I or II	type	clone	isotype	lot	firm	concentration	dilution to use (tested)	tested on
stock									
finished									
AlexaFluor 488_goat anti-rabbit	II	monoclonal Ab	A11008	IgG (H+L)	1073092	Life Technologies	2mg/ml	1/300	FACS/IF
AlexaFluor 555_donkey anti-mouse	II	monoclonal Ab	A31570	IgG (H+L)	737680	Invitrogen	2mg/ml	1/1000	IF
AlexaFluor 555_donkey anti-rabbit	II	monoclonal Ab	A31572	IgG (H+L)	47703A	Invitrogen	2mg/ml	1/1000	IF
AlexaFluor 555_goat antimouse	II	monoclonal Ab	A21137	IgG2a(Y2a)	1541489	Life Technologies	2mg/ml	1/300	IF
AlexaFluor 633_goat anti-rat	II	monoclonal Ab	A31094	IgG	1301835	Life Technologies	2mg/ml	1/300	FACS/IF
CD24-PE_mouse anti-human	I coupled	monoclonal Ab	M15	IgG2aK	2205927	BD	12.5 ug/ml	1/10	FACS/IF
CD24-PE_mouse anti-human	I coupled	monoclonal Ab	M15	IgG2aK	2172554	BD	25 ug/ml	1/10	FACS/IF
CD24-PE_mouse anti-human	I coupled	monoclonal Ab	M15	IgG2aK	3008938	BD	25 ug/ml	1/10	FACS/IF
CD271_mouse anti-human	I need II	monoclonal Ab	C40-1457	IgG1	18111	BD	500ug/ml	1/100	FACS/IF
CD326-FITC_mouse anti-human	I coupled	monoclonal Ab	VU-109	IgG1	1109	AbD Serotec	100ug/ml	1/100 (FACS)	FACS/IF
CD44-APC_mouse anti-human	I coupled	monoclonal Ab	G44-26	IgG2bK	39041	BD	3ug/ml	1/10	FACS/IF
CD44-APC_mouse anti-human	I coupled	monoclonal Ab	G44-26	IgG2bK	17712	BD	3ug/ml	1/10	FACS/IF
CD44-APC_mouse anti-human	I coupled	monoclonal Ab	G44-26	IgG2bK	2203691	BD	3ug/ml	1/10	FACS/IF
KRT5_rabbit anti-mouse/human	I need II	polyclonal Ab	AF 138 (PR8 160 P)	IgG		Covance	1mg/ml	1/1000	FACS/IF
KRT8_mouse anti-human	I need II	monoclonal Ab	E8 (MMS-167P)	IgG2aK		Covance		1/100	IF
KRT8_rat anti-mouse/human	I need II	monoclonal Ab	Troma-1b	IgG2	8/22/13	DHSB	1mg/ml	1/1000	FACS
P63_mouse anti-mouse	I need II	monoclonal Ab	4A4 (559951)	IgG2aK		BD Pharmingen	1mg/ml	1/50	IF
PE-Cy7_goat anti-mouse IgG1 Fab 2	II	monoclonal Ab	sc 3817	Ig	81313	Santa Cruz	400ug/ml	1/300	FACS
PE-Cy7_goat anti-rabbit IgG1 Fab 2	II	monoclonal Ab	sc 3845	Ig	00413	Santa Cruz	400ug/ml	1/300	FACS

Table F.5 : Data list of antibodies

Programs used to analyze experimental data

G.1 Code to read.lsm images - Matlab

- Adapted by Dr Isabelle Bonnet, Institut Curie & Université Pierre et Marie Curie

```
1 function [Data, varargout] = lsmread(fileName, varargin)
2
3 % [Data] = lsmread(filename) will return image data in its original
4 % dimensions
5 % [Data] = lsmread(filename, 'Range', [T0 T1 C0 C1 Z0 Z1]) will return image
6 % data with specified dimension.
7 % [Data LSMinfo] = lsmread(filename) will return image data as well as
8 % LSMinfo
9 % [LSMinfo] = lsmread(filename, 'InfoOnly') will only return LSM infos.
10 %
11 -----
12
13 % Notes for the Zeiss LSM format.
14 % The first 8 bytes are TIFF file header.
15 % Each 12-byte IFD entry has the following format: Bytes 0-1: The Tag that
16 % identifies the field. Byte 2-3: The field type. Bytes 4-7: Count.
17 % Bytes 8-11: Value or offset. Byte count of the indicated types.
18 % Types 1 = BYTE, 2 =ASCII (8-bit), 3= SHORT (16-bit), 4= LONG(32-bit).
19 % If the count is 1, then bytes 8-11 stores the value, otherwise it's offset
20 % pointing to the position in file where value is stored.
21 % Each optical section at every time point (no matter how many channels)
22 % will have two image file directories(IFDs). The first one contains
23 % information for the real image data. The second IFD contains
24 % information for thumbnail images. The LSM file is structured as such :
25 % header -> IFDs -> (real image/thumbnail)s.
26 % The very first IFD has an entry numbered 34412 and points to LSM-specific
27 % data.
28 %
29 -----
30
31 % Version 1.1.0
32 % Copyright(C) 2014 Chaoyuan Yeh.
```

```

25 % The script is inspired by LSM File Toolbox by Peter Li and tiffird by
    Francois Nedelec.
26
27 fID=fopen(fileName);
28 byteOrder=fread(fID,2,'*char');
29 if (strcmp(byteOrder,'II'))
30     byteOrder = 'ieee-le';
31 elseif (strcmp(byteOrder,'MM'))
32     byteOrder = 'ieee-be';
33 else
34     error('This is not a correct TIFF file');
35 end
36
37 tiffID = fread(fID, 1, 'uint16', byteOrder);
38 if (tiffID ~= 42)
39     error('This is not a correct TIFF file');
40 end
41
42 fseek(fID, 4, 'bof');
43 ifdPos = fread(fID, 1, 'uint32', byteOrder);
44 fseek(fID, ifdPos, 'bof');
45
46 IFDIdx=0;
47 while ifdPos ~= 0
48     IFDIdx=IFDIdx+1;
49     fseek(fID, ifdPos, 'bof');
50     numEntries = fread(fID, 1, 'uint16', byteOrder);
51     entryPos = ifdPos+2; % The first two bytes of each IFD specify number
        of IFD entries.
52     fseek(fID, ifdPos+12*numEntries+2, 'bof'); % Each IFD entry is 12-byte
        long.
53     ifdPos = fread(fID, 1, 'uint32', byteOrder); % The last four bytes of
        an IFD specifies offset to next IFD. If this is zero, it means
        there's no other IFDs.
54     % IFD is structured like this: bytes 1-2 : tag, bytes 3-4: type,
55     % bytes 5-8: count, bytes 9-12: value/offset
56     for ii = 1:numEntries
57         fseek(fID, entryPos+12*(ii-1), 'bof');
58         IFD{IFDIdx}(1, ii) = fread(fID, 1, 'uint16', byteOrder);
59         IFD{IFDIdx}(2, ii) = fread(fID, 1, 'uint16', byteOrder);
60         IFD{IFDIdx}(3, ii) = fread(fID, 1, 'uint32', byteOrder);
61         IFD{IFDIdx}(4, ii) = fread(fID, 1, 'uint32', byteOrder);
62     end
63     if strcmpi(varargin, 'InfoOnly'), break; end
64 end
65 %Reading LSMinfo
66 if IFD{1}(3, IFD{1}(1, :))==258==1
67     LSMinfo.bitDepth = IFD{1}(4, IFD{1}(1, :))==258);
68 else
69     fseek(fID, IFD{1}(4, IFD{1}(1, :))==258), 'bof');
70     LSMinfo.bitDepth = fread(fID, 1, 'uint16', byteOrder);
71 end
72 offsetLSMinfo = IFD{1}(4, IFD{1}(1, :))==34412)+8;
73 fseek(fID, offsetLSMinfo, 'bof');
74 LSMinfo.dimX = fread(fID, 1, 'uint32', byteOrder);
75 LSMinfo.dimY = fread(fID, 1, 'uint32', byteOrder);
76 LSMinfo.dimZ = fread(fID, 1, 'uint32', byteOrder);
77 LSMinfo.dimC = fread(fID, 1, 'uint32', byteOrder);
78 LSMinfo.dimT = fread(fID, 1, 'uint32', byteOrder);
79 fseek(fID, 12, 'cof');

```

```

80 LSMInfo.voxSizeX = fread(fid, 1, 'float64', byteOrder);
81 LSMInfo.voxSizeY = fread(fid, 1, 'float64', byteOrder);
82 LSMInfo.voxSizeZ = fread(fid, 1, 'float64', byteOrder);
83 fseek(fid, 26, 'cof');
84 LSMInfo.specScan = fread(fid, 1, 'uint16', byteOrder);
85
86 if strcmpi(varargin, 'InfoOnly')
87     Data = LSMInfo;
88     fclose(fid);
89 else
90     % Creating offset list for every IFD
91     offset = zeros(1,IFDIdx/2);
92     for ii=1:IFDIdx/2
93         if IFD{2*ii-1}(3,7) == 1 && IFD{2*ii-1}(4,7) < 4294967296
94             offset(ii) = IFD{2*ii-1}(4,7);
95         else
96             fseek(fid, IFD{2*ii-1}(4,7), 'bof');
97             offset(ii) = fread(fid, 1, 'uint32', byteOrder);
98         end
99     end
100
101     if any(strcmpi(varargin, 'Range'))
102         Range = varargin{find(strcmpi(varargin, 'Range'))+1};
103         if any([Range(1)<1, Range(2)>LSMInfo.dimT, Range(3)<1, Range(4)>
104             LSMInfo.dimC,...
105             Range(5)<1, Range(6)>LSMInfo.dimZ])
106             error('Input range exceeds data range');
107         else
108             dimT0 = Range(1); dimT1 = Range(2);
109             dimC0 = Range(3); dimC1 = Range(4);
110             dimZ0 = Range(5); dimZ1 = Range(6);
111         end
112     else
113         dimT0 = 1; dimT1 = LSMInfo.dimT;
114         dimZ0 = 1; dimZ1 = LSMInfo.dimZ;
115         dimC0 = 1; dimC1 = LSMInfo.dimC;
116     end
117     bitDepth = strcat('uint', num2str(LSMInfo.bitDepth));
118     Data = zeros(dimT1-dimT0+1, dimC1-dimC0+1, dimZ1-dimZ0+1, LSMInfo.dimX,
119         LSMInfo.dimY, bitDepth);
120     for indT=dimT0:dimT1
121         for indZ = dimZ0:dimZ1
122             fseek(fid, offset((indT-1)*LSMInfo.dimZ + indZ), 'bof');
123             switch bitDepth
124                 case 'uint16'
125                     fseek(fid, (dimC0-1)*LSMInfo.dimX*LSMInfo.dimY*2, 'cof');
126                     for indC = dimC0:dimC1
127                         Data(indT-dimT0+1, indC-dimC0+1, indZ-dimZ0+1, :, :) = reshape
128                             (uint16(fread(fid, LSMInfo.dimX*LSMInfo.dimY, bitDepth,
129                                 , byteOrder)), LSMInfo.dimX, LSMInfo.dimY)';
130                     end
131                 case 'uint8'
132                     fseek(fid, (dimC0-1)*LSMInfo.dimX*LSMInfo.dimY, 'cof');
133                     for indC = dimC0:dimC1
134                         Data(indT-dimT0+1, indC-dimC0+1, indZ-dimZ0+1, :, :) = reshape
135                             (uint8(fread(fid, LSMInfo.dimX*LSMInfo.dimY, bitDepth,
136                                 , byteOrder)), LSMInfo.dimX, LSMInfo.dimY)';
137                     end
138             end
139         end
140     end

```



```

134     end
135     fclose (fID);
136     if nargout==2
137         varargout{1} = LSMInfo;
138     end
139 end
140 end

```

G.2 Code to compute fluorescence intensity of single cells - Matlab

- Written by Dr Isabelle Bonnet, Institut Curie & Université Pierre et Marie Curie

```

1 % Isabelle Bonnet
2 % November, 2014
3
4 % Purpose: Marie André - compute fluorescence intensity of single cells
5 % Extract the red channel (channel n°2) in a lsm slice-image
6 % Z-projection along slices (Sum)
7 % User indicates the number of cell she wants to analyse within the image
8 % For each cell, she draws a polygon around the cell = ROI
9 % Within each ROI, intensity is computed as the average of the pixels above
10 % the 10th
11 % percentile
12
13 %----- 01/12/2014 : v1 (on Marie Andre's demand)
14 %-----
15 %-- Selection of the backup folder
16 %-- ROI drawing :
17 %     backspace button removes the latest point when ROI is drawn
18 %     user has to validate the roi, he can also cancel it to draw a new
19 %     one
20 %-- Reference ROI - for each ROI
21 %     user has to indicate the background area that has the same shape
22 %     as the cell's ROI - this reference ROI will be used to compute the
23 %     background signal
24 %-- Saving format has been changed so that results are appended in the same
25 %     file for a given backup folderchanged so as to append data
26 %----- End of v1 -----
27
28 %----- 11/02/2015 : v2 (on Marie Andre's demand)
29 %-----
30 %--Calculate mean intensity of drawn ROI named as background area: "meanbg"
31 %--For each cell ROI take pixel which value is up to "meanbg". Function
32 %     pixroi with (pixroi>meanbg)
33 %--And on these specific pixels, calculate sum intensity (int_sum) and mean
34 %     intensity (int_mean) of the cell ROI
35 %----- End of v2 -----
36
37 %-- Sanity checking
38 clear all; close all; clc

```

```

37 %--- Must be filled by the user -----
38 default_folder='TO FILL';
39 perc=10; % If perc=10, intensity would be compute on pixels whose intensity
    is greater than the 10 percentile (90prtil)
40 %-----
41
42
43 %-- Load Data from the selected .lsm file
44 [fn ,pn]=uigetfile('*.lsm','Select the .lsm Stack file to analyse',
    default_folder);
45 if fn==0
46     display('You did not choose a .LSM file');
47     return;
48 else
49
50     %--- Create an output folder
51     Dest_dir= uigetdir(pn,'Select the directory for saving');
52     if Dest_dir==0
53         display('You did not select an output FOLDER');
54         return;
55     else
56
57         fn_string=fn(1:strfind(fn, '.lsm')-1); % remove .lsm from the filename
58
59         [Data LSMinfo]= lsmread([pn, fn]);
60         %%!!! LSM dim are as followx: T,C,Z,X,Y
61         im_chan2=squeeze(Data(1,2, :, :, :));
62
63         % Project along the different Z-slices
64         im_proj=squeeze(sum(im_chan2,1));
65         [HEIGHT,WIDTH]= size(im_proj);
66
67         % Via a dialog boxe, user indicates how many cells he/she wants to
            analyse
68         h0=figure();
69         imshow(im_proj,[])
70         prompt1={'How many cells do you want to analyse?'};
71         dlg1_title='Cells number';
72         num_lines=1;
73         def={'1'};
74         answer1=inputdlg(prompt1, dlg1_title, num_lines, def);
75         nb_cells=str2double(cell2mat(answer1(1)));
76         close(h0)
77
78         %-- initialization for the loop over the number of Cells
79         [X,Y] = meshgrid(1:HEIGHT,1:WIDTH); % needed for the inpolygon
            function called later
80         color = colormap(lines); % load a colormap to have a panel of color
81         colormap(gray) % load an appropriate colormap for displaying images
82
83         % Analyse trough the cells
84         h1=figure;
85         imshow(im_proj,[]);
86         for i=1:nb_cells
87             valid=0;
88
89             while valid==0
90                 clear xv yv
91                 figure(h1)

```

```

92     title(['Select the cell n°' num2str(i)], 'FontSize', 16, '
93         Color', color(i,:));
94     command=['Double-click adds a point AND ends the selection', '
95             ', 'Backspace removes the latest point'];
96     text(HEIGHT/100,WIDTH/100,command, 'FontSize', 9, 'Color','w'
97         );
98
99     % User draws it roi
100    [xv,yv]=getline(h1,'closed');
101    hline=line(xv,yv,'Color',color(i,:));
102
103    % User validates the roi
104    prompt2={'Press 1 to validate selection ' num2str(i) ', 0
105            otherwise'};
106    dlg2_title='Validation';
107    def2={'1'};
108    answer2=inputdlg(prompt2,dlg2_title,num_lines,def2);
109    valid=str2double(cell2mat(answer2(1)));
110    delete(hline)
111 end
112
113 %--Display the validated ROI
114 hold on
115 line(xv,yv,'Color',color(i,:))
116
117 %--- RegionProperties of the ROI
118 cell_bw=inpolygon(X,Y,xv,yv);
119 stats_roi = regionprops(cell_bw,im_proj, 'Centroid', 'Area', '
120     PixelValues');
121
122 %-- Display the centroid and the corresponding cell number
123 scatter(stats_roi.Centroid(1),stats_roi.Centroid(2),30,color(i,:),
124         'filled')
125 text(stats_roi.Centroid(1)+0.01*HEIGHT,stats_roi.Centroid(2)
126     +0.01*WIDTH,num2str(i),'Color',color(i,:), 'FontSize',15);
127 %--- Draggable roi that user positions to background
128 poly_bg = impoly(gca, [xv,yv]);
129 setColor(poly_bg,'yellow');
130 h_text = text(HEIGHT/20,WIDTH/20, 'Drag me and press any key to
131     validate');
132 set(h_text,'BackgroundColor',[1 1 0]);
133 pause() % Stops when a key is pressed
134 xy_bg=getPosition(poly_bg); % store lcoation of the background
135 delete(poly_bg)
136 delete(h_text)
137
138 %--- Statistics within the BackGround
139 %hold on
140 %line(xy_bg(:,2),xy_bg(:,1),'Color','y')
141 bg_bw=inpolygon(X,Y,xy_bg(:,2),xy_bg(:,1));
142 stats_bg = regionprops(bg_bw,im_proj, 'Centroid', 'PixelValues');
143 pixbg=stats_bg.PixelValues; % Take all the pixels within the
144     drawn polygon
145 meanbg = mean(pixbg); % Compute the average intensity over the
146     background pixels
147
148 %--- Statistics within the ROI
149 pixroi=stats_roi.PixelValues; % Take all the pixels within the
150     drawn polygon

```

```

141     %treshroi = prctile(pixroi,perc); % Tresholding: compute the
142         percentile corresponding to perc
143     int_mean=mean(pixroi(pixroi>meanbg)); % Compute the average
144         intensity over the pixels whose intensity is greater than the
145         preceeding threshold
146     int_sum=sum(pixroi(pixroi>meanbg));% Compute the sum intensity
147         over the pixels whose intensity is greater than the
148         preceeding threshold
149
150
151     %--- Store the results
152     %Results(i,:) = [i perc treshroi int_mean int_sum min(pixroi) max
153         (pixroi) mean(pixroi) ];
154     Results(i,:) = [i perc meanbg int_mean int_sum min(pixroi) max(
155         pixroi) mean(pixroi) ];
156     clear xv yv cell_bw stats_roi treshroi int_mean int_sum
157
158 end
159 title('Cells number','FontSize',16);
160 saveas(hl,[Dest_dir '\Cell_' fn_string '.png'])
161 close(hl)
162
163 %--- Saving
164 %dmlmwrite([Dest_dir '\Results_' fn_string '.txt'], Results, '
165     delimiter','\t','newline','pc');
166 %--- set format for saving with fprintf
167 formatLegend='%6s %6s %6s %6s %6s %6s %6s %6s n\n';
168 formatDataSpec = '\n %d %d %0.2f %0.2f %0.2f %0.2f %0.2f \n';
169 formatFile='\n %s';
170 %--- open the file with authorization to append
171 fileID = fopen([Dest_dir '\Results.dat'],'a');
172 if ftell(fileID)==0 % file is empty
173     fprintf(fileID,formatLegend,'Cell number','Mean BG','Mean
174         intensity','Sum intensity','Min on treshold','Max on
175         treshold','Mean on treshold');
176 end
177 fprintf(fileID,formatFile,fn_string);
178 for i=1:nb_cells
179     fprintf(fileID,formatDataSpec,Results(i,:));
180 end
181 fclose(fileID);
182
183 close all
184 end %--- Create an output directory
185 end %--- Select a file to analyse

```

G.3 Code to perform dataset from flow cytometry analysis data - Matlab

- Written by Dr Timo Betz and Marie Pochitaloff-Huvale, Institut Curie

```

1 % Marie André, Marie Pochitaloff and Timo Betz
2 % April, 2015
3
4 % Purpose: Marie André -

```

```

5 % Dataset data from flow cytometry analysis
6
7 %-- Sanity checking
8 clear all; close all; clc;
9
10
11 %--- Must be filled by the user -----
12 default_folder='TO FILL';
13 %-----
14
15 %-- Select correct folder to analyse
16 folder_name = uigetdir(default_folder, 'Select the correct folder to analyse')
17 ;
18 %-- Load all data from .tab file and create a folder_dataset within the
19 %-- folder to analyse
20 tab_files=dir([folder_name, filesep, '*.tab']) % Take all .tab files with the
21 folder_name to do dataset
22
23 int_folder=findstr(folder_name, filesep) % Go back the [folder_name-1]
24 mkdir([folder_name(1:int_folder(end)), 'DATA_flow_cytometry_analysis_viable
25 cells_dataset']); % Create a new folder within the [folder_name-1] named
26 DATA_flowcytometry_dataset
27 save_dir=[folder_name(1:int_folder(end)), 'DATA_flow_cytometry_analysis_viable
28 cells_dataset']; % Save direction for the dataset files
29
30 %-- Loop to create dataset and 2 save procedures
31 for i=1:length(tab_files) % Loop done for all .tab files within the
32 folder_name
33
34 ds=dataset('File', [folder_name, filesep, tab_files(i).name]); % Create
35 dataset for all .tab files within the folder_name selected
36
37 % Save procedure of each dataset within a subfolder_dataset into the
38 % [folder_name-1]
39 var_int=tab_files(i).name(1:end-8); % Name each dataset using .tab file
40 name without 8 characters (.fcs.tab)
41 save([save_dir, filesep, var_int, '.mat'], 'ds'); % Save each dataset (ds)
42 named var_int within the save_dir
43
44 % Save procedure of each dataset within TO FILL
45 name_int=strrep([folder_name, filesep, tab_files(i).name], filesep, '_'); %
46 Define the name_int of each dataset to save as the entire pathway and
47 replace within the name all \ per _
48 name_int=strrep(name_int, '#', ''); % Within the pathway replace all # with
49 _
50 name_int=strrep(name_int, ':', ''); % Within the pathway replace all : with
51 _
52 save(['TO FILL\ ', name_int(1:end-8), '.mat'], 'ds'); % Save each dataset
53 named name_int minus 8 characters (.fcs.tab) within TO FILL with a .
54 mat format
55
56 end

```

G.4 Code to calculate MFI, CV and SD from flow cytometry analysis based on the CD44 marker threshold - Matlab

**- Written by Mathieu Richard (PhD student, Institut Curie),
Dr Delphine Ladarre and Mamy Ratsimbazafy**

```

1 % Mathieu Richard and Delphine Ladarre -- THANK YOU SO MUCH, YOU SAVED MY
2 % PhD ANALYSIS!
3 % April , 2015
4
5 % Mamy Ratsimbazafy , v2 update July , 2015
6
7 % Purpose: Marie André -
8 % MFI, CV and SD from flow cytometry analysis. Threshold on CD44-APC marker.
9 % Ready for normalization on beads fluorescence and UNST cells
10
11 function [] = Fluorescence_v2()
12
13     %--- Must be filled by the user -----
14     %default_folder='TO FILL';
15     default_folder= pwd;
16     cd( default_folder);
17     %-----
18
19     %% Choose folder
20     [filename , pathname] = uigetfile( '*.mat' );
21     cd(pathname);
22
23     disp( strcat( 'Start processing: ',pathname))
24     disp( 'Starting SUM149 processing ' )
25
26     try
27         fluoproc( 'SUM149' );
28     catch ME
29         disp( 'Error while processing SUM149' )
30         disp( ME)
31     end
32
33     disp( 'Starting SUM159 processing ' )
34     try
35         fluoproc( 'SUM159' );
36     catch ME
37         disp( 'Error while processing SUM159' )
38         disp( ME)
39     end
40
41     disp( strcat( 'End processing: ',pathname))
42
43     %%Back to default folder
44     cd( default_folder);
45
46 end
47
48 function [] = fluoproc( SubName)
49     Listing=dir( strcat( '*', SubName, '*.mat' ));
50
51     %% Import threshold
52     TextFile=dir( strcat( '*', SubName, '*.txt' ));
53     Threshold=importdata( TextFile(1).name);
54     Threshold_APC_A=Threshold(2);
55
56     %% Import Data and analysis

```

```

57     Result=transformAsinh ( Listing , Threshold_APC_A );
58
59     %% Save result in xls file
60     %xlwrite ( strcat ( ' Result_ ', SubName, '.xlsx ' ), Result , 1, 'A1' ); %use
        xlwrite on mac instead of xlswrite
61     export ( cell2dataset ( Result ), ' file ', strcat ( ' Result_ ', SubName, '.csv ' ), '
        delimiter ', ';' );
62 end
63 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
64
65 function out=transformAsinh ( Listing , Threshold_APC_A)
66     Result=cell ( size ( Listing , 1) +1, 11);
67     Result ( 1 , :) = [ { ' Name ' }, { ' perc_data_neg_APC ' }, { ' perc_data_pos_APC ' }, ...
68         { ' mynAPC_data_neg_APC ' }, ...
69         { ' SDasinhAPC_data_neg_APC ' }, { ' mynPE_data_neg_APC ' }, ...
70         { ' SDasinhPE_data_neg_APC ' }, ...
71         { ' mynAPC_data_pos_APC ' }, ...
72         { ' SDasinhAPC_data_pos_APC ' }, { ' mynPE_data_pos_APC ' }, ...
73         { ' SDasinhPE_data_pos_APC ' } ];
74     for i = 1 : size ( Listing , 1)
75         [ ~ , name , ~ ] = fileparts ( Listing ( i ) . name );
76         Data = importdata ( Listing ( i ) . name );
77
78         % Threshold Calcul
79         IndiceDataNegAPC = Data . APC_A < Threshold_APC_A ;
80         IndiceDataPosAPC = Data . APC_A > Threshold_APC_A ;
81
82         % Percentage
83         perc_data_neg_APC = sum ( IndiceDataNegAPC ) / size ( Data . APC_A , 1 ) * 100 ;
84         perc_data_pos_APC = 100 - perc_data_neg_APC ;
85
86         % Argsh
87         dataargshAPC_neg_APC = asinh ( Data . APC_A ( IndiceDataNegAPC ) );
88         dataargshAPC_pos_APC = asinh ( Data . APC_A ( IndiceDataPosAPC ) );
89
90         dataargshPE_neg_APC = asinh ( Data . PE_A ( IndiceDataNegAPC ) );
91         dataargshPE_pos_APC = asinh ( Data . PE_A ( IndiceDataPosAPC ) );
92
93         % Mean Arsh
94         mynargshAPC_data_neg_APC = mean ( dataargshAPC_neg_APC );
95         mynargshAPC_data_pos_APC = mean ( dataargshAPC_pos_APC );
96
97         mynargshPE_data_neg_APC = mean ( dataargshPE_neg_APC );
98         mynargshPE_data_pos_APC = mean ( dataargshPE_pos_APC );
99
100        % Sinh
101        mynAPC_data_neg_APC = sinh ( mynargshAPC_data_neg_APC );
102        mynAPC_data_pos_APC = sinh ( mynargshAPC_data_pos_APC );
103
104        mynPE_data_neg_APC = sinh ( mynargshPE_data_neg_APC );
105        mynPE_data_pos_APC = sinh ( mynargshPE_data_pos_APC );
106
107        %SD asinh
108        SDasinhAPC_data_neg_APC = std ( dataargshAPC_neg_APC );
109        SDasinhAPC_data_pos_APC = std ( dataargshAPC_pos_APC );
110
111        SDasinhPE_data_neg_APC = std ( dataargshPE_neg_APC );
112        SDasinhPE_data_pos_APC = std ( dataargshPE_pos_APC );
113
114        Result ( i + 1 , :) = [ { name } , ...

```

```

115         {perc_data_neg_APC }, {perc_data_pos_APC }, ...
116         {mynAPC_data_neg_APC }, {SDasinhAPC_data_neg_APC }, ...
117         {mynPE_data_neg_APC }, {SDasinhPE_data_neg_APC }, ...
118         {mynAPC_data_pos_APC }, {SDasinhAPC_data_pos_APC }, ...
119         {mynPE_data_pos_APC }, {SDasinhPE_data_pos_APC }];
120     end
121 out = Result;
122 end

```

G.5 Code to consolidate daily data into a single Excel file, formatted for GraphPad's graphs - Python 3

- Written by Mamy Ratsimbazafy

```

1 import pandas as pd
2 import numpy as np
3 from pathlib import Path
4 import re
5 from functional import compose
6 from functools import reduce, partial
7 import itertools as it
8
9
10 # Speed check
11 # Check that numpy is compiled with OpenBLAS, Apple Accelerate or Intel MKL
    support
12 #np.__config__.show()
13
14 #####
15 #Helper High Order Functions
16 #map that return a list
17 def lmap(func, *iterable): return list(map(func, *iterable))
18
19 #compose list of functions
20 def mcompose(*func): return reduce(compose, func)
21
22 #####
23 #Part I: Functions for retrieving and dumping the data
24 #Composed functions / Functions dependant of data:
25
26 #Filter unused columns from Data Frame
27 def dfFilter(df, regexpKeepCol):
28     return df.select(
29         lambda colname:
30             regexpKeepCol.search(colname), axis=1)
31
32 #1. Parse file path, extract (Experiment,Day,Cell line) from folder names
33 #3. Convert all values from string to integer
34 def xpparam(stringpath, regexpFolder):
35     return compose(
36         partial(lmap, int),
37         regexpFolder.findall
38     )(stringpath)
39
40 # 1. Read Excel/Csv
41 # 2. Filter out unused columns

```



```

42 # 3. Transform DataFrame to narrow format.
43 # Format is more suitable for automatic processing/Pivot tables and widely
    used for databases.
44 # Also format allows not to care about order of the data.
45 #   A B           Data  Value
46 # 1  p  q   ==>  1   A    p
47 # 2  r  s           2   A    r
48 #                   1   B    q
49 #                   2   B    s
50 def xptonarrow(stringpath , regexpKeepCol):
51     return mcompose(
52         partial(pd.melt, id_vars="Name", var_name="Type"),
53         partial(dfFilter, regexpKeepCol=reKeepCol),
54         partial(pd.read_csv, delimiter=";")
55     )(stringpath)
56
57 #Create tuple of ((Experiment, Day, Cell line), Data Frame(Std Deviation)
58 #in record/narrow format from list of Excel files
59 #Arg: list of paths, Regexp to extract metadata from path, Regexp to keep
    specific column names
60 def edcl(paths, reFpath, reFcol):
61     return [(xpparam(str(path), reFpath), xptonarrow(str(path), reFcol)) for
        path in paths]
62
63 #Heavy lifting of the +/-/low/high/neg values
64 #Additional lifting: convert Replicat to Integer
65 def mapDico(*lconfig):
66     if len(*lconfig) != 4:
67         return None, None, None, None
68     else:
69         lconfig = tuple(*lconfig) #Can't use list directly as it's an
            unhashable type
70         #CD24, CD44, Replicat, Comments/error
71         return dicoRef[lconfig[0].lower()], dicoRef[lconfig[1].lower()], int(
            lconfig[2]), lconfig[3]
72
73 #Transform ((a,b,c), DataFrame) into a DataFrame with new columns A, B, C
    filled with values a,b,c
74 def mapMetaData(data):
75     experiment = data[0][0]
76     day = data[0][1]
77     cell = data[0][2]
78     DF = data[1]
79
80     DF["Experiment"], DF["Day"], DF["Cell_Line"] = experiment, day, cell
81
82     return DF
83
84 #Column "Name" holds CD24/CD44 +/-, Replicat ID
85 def extractMetaFromName(df, reParseName):
86     DF = df #use a temp variable to avoid side effects
87
88     DF["CD24"], DF["CD44"], DF["Replica"], DF["Comment_Error"] = \
89     zip(*DF["Name"].map(compose(mapDico, reParseName.findall)))
90
91     return DF
92
93 #####
94 # Functions for part II
95

```

```

96 #Add empty column
97 def add_none_col(df, i):
98     df[i] = np.nan
99     return df
100
101
102 #####
103 # Setup
104
105 #Regexp to parse info from folder path (probably working only on Linux, Mac
    and not on windows):
106 reFolder = re.compile(r'(?<=#0)\d+(?=/)|(?<=_D)\d+(?=/)|(?<=Result_SUM)\d
    +(?=\.csv)')
107 #Regexp to filter kept columns
108 reKeepCol = re.compile(r'^SD*|^Name')
109 #Regexp to parse info from column "Name"
110 reParseName = re.compile(r'((?<=CD24)[^C]+(?=CD44)|(?<=CD44)[^_]+(?=_)|(?<=_
    )\d|(?<=_\d).*(?=$))')
111
112 #Dictionary for consistent referencing (transform input to lower case for
    insensitive use)
113 dicoRef = {"+": "+", "high": "+", "med": "med", "low": "-", "-": "-", "neg": "-"}
114
115 #Order of the resulting columns
116 dicoSort = {'SDasinhAPC_data_neg_APC':1, 'SDasinhPE_data_neg_APC':2, '
    SDasinhAPC_data_pos_APC':3, 'SDasinhPE_data_pos_APC':4}
117
118 #CSVseparator
119 csvsep = ";"
120
121 #####
122 # Part I - Proc retrieve all the data and dump them
123
124 files_path = Path('dataset_only_transition').glob('**/*.csv')
125 data = edcl(files_path, reFolder, reKeepCol)
126 data_transf = pd.concat([mapMetaData(dataset) for dataset in data],
    ignore_index=True)
127 dtset = extractMetaFromName(data_transf, reParseName)
128
129 #Errors correction, potential side-effects due to in-place modification
130 #Error 1. med/high combo is actually high/high
131 dtset.loc[(dtset.CD24 == "med") & (dtset.CD44 == "+"), ["CD24", "
    Comment_Error"]] = \
132     "+", "CD24 changed from med to +"
133 #Error 2. tagged data named SUM149 with tag "_error SUM159" are actually for
    159 cell line
134 dtset.loc[(dtset.Cell_Line == 149) & (dtset.Comment_Error == "_error SUM159")
    , ["Cell_Line", "Comment_Error"]] = \
135     159, "_error SUM159. Corrected changed cell line from 149 to 159"
136
137 #Export the raw records
138 writer = pd.ExcelWriter("GraphPad_formatted_data.xlsx", engine='xlsxwriter')
139 dtset.to_excel(writer, "Raw data Std Dev in Asinh space")
140
141 #Retransform to linear space and export
142 dtlin = dtset
143 dtlin['value'] = dtlin['value'].map(np.sinh)
144 dtlin.to_excel(writer, "Geom-like sinh SD asinh")
145
146

```

```

147 #####
148 # Part II - Pivot the data for copy/paste into GraphPad
149 #Clean data from control data rows, i.e. if CD24 or CD44 or Replicat are
    empty, it's not a regular experiment
150 clean_dt = dtset[dtset.CD24 != ""]
151
152 #Pivot the data and generate control data before reformatting
153 pivot = clean_dt.pivot_table(values = ["value"], \
154     index = ["Cell_Line", "CD24", "CD44", "Day"], \
155     columns = ["Type", "Experiment", "Replica"])
156 pivot.to_excel(writer, "Std Dev Asinh - Raw Pivot")
157 pivot.applymap(np.sinh).to_excel(writer, "Sinh SD Asinh - Raw Pivot")
158
159 #Data needs all 139 days even if no experiment -> rebuild the row indexes
160 fulldays_combo = it.product((149,159), ("+", "-", "med"), ("+", "-", "med"),
    range(1, 1+max(clean_dt['Day'])))
161
162 #Filter out (CD24+,CD44-) combination which doesn't exist
163 final_combo = [(a,b,c,d) for (a,b,c,d) in fulldays_combo if (b,c) not in [( "+
    ", "-"), ("-", "med"), ("+", "med"), ("med", "-"), ("med", "+")]]
164
165 ### Rebuild Indexes and export to csv
166 #Reindex rows
167 pivot_tosort = pivot.reindex(final_combo)
168
169 #Add empty columns 'Replica 7' after each Experiment 59 (GraphPad quirk)
170 cl = lmap(lambda x:x.tolist(), pivot_tosort.columns.levels)[-2] + [[59],[7]]
    #Build list of column level names + Experiment 59 + Replica 7
171 pivot_tosort = reduce(add_none_col, it.product(*cl), pivot_tosort) #Create
    all empty Experiment 59/Replica 7 columns
172
173 #Sort columns
174 pivot_final = pivot_tosort.reindex_axis( \
175     sorted(pivot_tosort.columns, \
176         key=lambda x: (dicoSort[x[1]], x[2], x[3])), axis=1)
177
178 #Export
179 pivot_final.to_excel(writer, "SD Asinh Formatted Final Pivot")
180 pivot_final.applymap(np.sinh).to_excel(writer, "Sinh SD Asinh - Formatted
    Final")
181
182 #Highlight the important table
183 writer.sheets["Sinh SD Asinh - Formatted Final"].set_tab_color('red')
184
185 writer.save()
186
187 # END
188 #####

```

Bibliography

- [1] Jerry M Adams and Suzanne Cory. “Bcl-2-regulated apoptosis : mechanism and therapeutic potential.” In : *Curr. Opin. Immunol.* 19.5 (Oct. 2007), pp. 488–96. ISSN : 0952-7915. DOI : 10.1016/j.coi.2007.05.004. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0952791507000994><http://www.ncbi.nlm.nih.gov/pubmed/17629468><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2754308>.
- [2] Justin W Ady et al. “Intercellular communication in malignant pleural mesothelioma : properties of tunneling nanotubes.” In : *Front. Physiol.* 5.October (Jan. 2014), p. 400. ISSN : 1664-042X. DOI : 10.3389/fphys.2014.00400. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4215694>[%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4215694%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract).
- [3] Julie Agaëse. “Les cellules souches cancéreuses : Une cible thérapeutique ?” Synthèse Bibliographique en Biologie et Biotechnologie Master 2 Biologie Gestion. Université de Rennes 1, 2013, pp. 1–38. URL : <https://etudes.univ-rennes1.fr/digitalAssets/39/39581%7B%5C%7DJ%7B%5C%7DAgaesse%7B%5C%7DVLcorrige.pdf>.
- [4] Nima Aghaeepour et al. “Critical assessment of automated flow cytometry data analysis techniques.” In : *Nat. Methods* 10.3 (2013), pp. 228–38. ISSN : 1548-7105. DOI : 10.1038/nmeth.2365. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3906045>[%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3906045%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract).
- [5] François Amblard and Sylvie Coscoy. “Time shapes the cell.” In : *Med. Sci. (Paris)*. 27.4 (Apr. 2011), pp. 425–32. ISSN : 0767-0974. DOI : 10.1051/medsci/2011274020. URL : <http://www.medicinesciences.org/10.1051/medsci/2011274425><http://www.ncbi.nlm.nih.gov/pubmed/21524409>.
- [6] American Cancer Society. *Cancer staging*. 2015. URL : <http://www.cancer.org/treatment/understandingyourdiagnosis/staging>.
- [7] American Cancer Society. “The History of Cancer”. In : *Am. Cancer Soc.* (2014). URL : <http://www.cancer.org/acs/groups/cid/documents/webcontent/002048-pdf.pdf>.
- [8] American Cancer Society. *Types of targeted therapies used to treat cancer*. 2014. URL : <http://www.cancer.org/treatment/treatmentsandsideeffects/treatmenttypes/targetedtherapy/targeted-therapy-types>.
- [9] Lilia Antonova and Christopher R Mueller. “Hydrocortisone down-regulates the tumor suppressor gene BRCA1 in mammary cells : a possible molecular link between stress and breast cancer.” In : *Genes. Chromosomes Cancer* 47.4 (Apr. 2008), pp. 341–52. ISSN : 1098-2264. DOI : 10.1002/gcc.20538. URL : <http://www.ncbi.nlm.nih.gov/pubmed/18196591>.

- [10] H Arakawa. “p53 apoptosis and axon-guidance molecules.” In : *Cell Death Differ.* 12.8 (Aug. 2005), pp. 1057–65. ISSN : 1350-9047. DOI : 10.1038/sj.cdd.4401601. URL : <http://www.nature.com/doifinder/10.1038/sj.cdd.4401601%20http://www.ncbi.nlm.nih.gov/pubmed/15818407>.
- [11] W R ASHBY. “Principles of the self-organizing dynamic system.” In : *J. Gen. Psychol.* 37.2 (Oct. 1947), pp. 125–8. ISSN : 0022-1309. DOI : 10.1080/00221309.1947.9918144. URL : <http://www.ncbi.nlm.nih.gov/pubmed/20270223>.
- [12] Avi Ashkenazi and Vishva M. Dixit. “Apoptosis control by death and decoy receptors”. In : *Curr. Opin. Cell Biol.* 11.2 (Apr. 1999), pp. 255–260. ISSN : 09550674. DOI : 10.1016/S0955-0674(99)80034-9. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0955067499800349>.
- [13] Atelier Parisien d’Urbanisme. *La population étrangère à Paris - Éléments de diagnostic sociodémographique à partir des données du recensement*. Paris, 2002. URL : <https://www.google.com/url?sa=t%7B%5C%26%7Drct=j%7B%5C%26%7Dq=%7B%5C%26%7Dsrc=s%7B%5C%26%7Dsource=web%7B%5C%26%7Dcd=5%7B%5C%26%7Dcad=rja%7B%5C%26%7Duact=8%7B%5C%26%7Dved=0CEAQFjAEahUKEwiwh73ZwsfIAhXBiRoKHZajDEA%7B%5C%26%7Dur1=http://www.apur.org/sites/default/files/documents/133.pdf%7B%5C%26%7Dusg=AFQjCNEujQz1z0PZRPIFRHP22BUiqcAuDw%7B%5C%26%7Dsig2=BjFjEbakp>.
- [14] Robert Axelrod, David E. Axelrod, and Kenneth J. Pienta. “Evolution of cooperation among tumor cells.” In : *Proc. Natl. Acad. Sci. U. S. A.* 103.36 (Sept. 2006), pp. 13474–9. ISSN : 0027-8424. DOI : 10.1073/pnas.0606053103. URL : <http://www.pnas.org/content/103/36/13474.abstract%20http://www.pnas.org/content/103/36/13474.full%20http://www.ncbi.nlm.nih.gov/pubmed/16938860%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1557388>.
- [15] M. Madan Babu et al. “Intrinsically disordered proteins : regulation and disease”. In : *Curr. Opin. Struct. Biol.* 21.3 (June 2011), pp. 432–440. ISSN : 0959440X. DOI : 10.1016/j.sbi.2011.03.011. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0959440X11000637>.
- [16] M Barak et al. “Evaluation of prostate-specific antigen as a marker for adenocarcinoma of the prostate.” In : *J. Lab. Clin. Med.* 113.5 (May 1989), pp. 598–603. ISSN : 0022-2143. URL : <http://www.ncbi.nlm.nih.gov/pubmed/2469756>.
- [17] Rita Barallon et al. “Recommendation of short tandem repeat profiling for authenticating human cell lines, stem cells, and tissues”. In : *Vitr. Cell. Dev. Biol. - Anim.* 46.9 (Oct. 2010), pp. 727–732. ISSN : 1071-2690. DOI : 10.1007/s11626-010-9333-z. URL : <http://link.springer.com/10.1007/s11626-010-9333-z>.
- [18] Isabelle Barde, Patrick Salmon, and Didier Trono. “Production and titration of lentiviral vectors.” In : *Curr. Protoc. Neurosci.* Chapter 4 (Oct. 2010), Unit 4.21. ISSN : 1934-8576. DOI : 10.1002/0471142301.ns0421s53. URL : <http://www.ncbi.nlm.nih.gov/pubmed/20938923%20http://www.ncbi.nlm.nih.gov/pubmed/23681623>.
- [19] Xavier BARLOVATZ-MEIMON, Georgia; RONOT. *Culture de cellules animales*. 3^o édition. Lavoisier, 2014. ISBN : 9782743069896. URL : <https://books.google.fr/books?id=YQsIBQAAQBAJ>.

- [20] Nandita Barnabas and Dalia Cohen. “Phenotypic and Molecular Characterization of MCF10DCIS and SUM Breast Cancer Cell Lines.” In : *Int. J. Breast Cancer* 2013 (Jan. 2013), p. 872743. ISSN : 2090-3170. DOI : 10.1155/2013/872743. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3562669%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://www.ncbi.nlm.nih.gov/pubmed/23401782%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3562669>.
- [21] Yaneer Bar-Yam. *Dynamics of Complex Systems*. Addison-Wesley, 1997. ISBN : 9780201557480. URL : <http://necsi.edu/publications/dcs/%7B%5C%7Dfulltext>.
- [22] Eshel Ben-Jacob, Donald S Coffey, and Herbert Levine. “Bacterial survival strategies suggest rethinking cancer cooperativity.” In : *Trends Microbiol.* 20.9 (Sept. 2012), pp. 403–10. ISSN : 1878-4380. DOI : 10.1016/j.tim.2012.06.001. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0966842X12001011%20http://www.ncbi.nlm.nih.gov/pubmed/22750098>.
- [23] Hanna Berglind et al. “Analysis of p53 mutation status in human cancer cell lines : a paradigm for cell line cross-contamination.” In : *Cancer Biol. Ther.* 7.5 (2008), pp. 699–708. ISSN : 1538-4047. DOI : 10.4161/cbt.7.5.5712.
- [24] Neil A Bhowmick and Harold L Moses. “Tumor-stroma interactions.” In : *Curr. Opin. Genet. Dev.* 15.1 (Feb. 2005), pp. 97–101. ISSN : 0959-437X. DOI : 10.1016/j.gde.2004.12.003. URL : <http://www.ncbi.nlm.nih.gov/pubmed/15661539%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2819733>.
- [25] Neil A. Bhowmick, Eric G. Neilson, and Harold L. Moses. “Stromal fibroblasts in cancer initiation and progression.” In : *Nature* 432.7015 (Nov. 2004), pp. 332–7. ISSN : 1476-4687. DOI : 10.1038/nature03096. URL : <http://www.nature.com/doifinder/10.1038/nature03096%20http://www.ncbi.nlm.nih.gov/pubmed/15549095%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3050735>.
- [26] Eldad S Bialecki and Adrian M Di Bisceglie. “Diagnosis of hepatocellular carcinoma.” In : *HPB (Oxford)*. 7.1 (2005), pp. 26–34. ISSN : 1365-182X. DOI : 10.1080/13651820410024049. URL : <http://www.ncbi.nlm.nih.gov/pubmed/18333158%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2023919>.
- [27] Craig Biddle. “Individualism vs. Collectivism : Our Future, Our Choice”. In : *Object. Stand. Vol. 7, No. 1.* (2012). URL : <https://www.theobjectivestandard.com/issues/2012-spring/>.
- [28] Krastan B Blagoev. “Organ aging and susceptibility to cancer may be related to the geometry of the stem cell niche.” In : *Proc. Natl. Acad. Sci. U. S. A.* 108.48 (Nov. 2011), pp. 19216–21. ISSN : 1091-6490. DOI : 10.1073/pnas.1106105108. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3228473%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [29] Cédric Blanpain. “Tracing the cellular origin of cancer.” In : *Nat. Cell Biol.* 15.2 (Feb. 2013), pp. 126–34. ISSN : 1476-4679. DOI : 10.1038/ncb2657. URL : <http://www.nature.com/doifinder/10.1038/ncb2657%20http://www.ncbi.nlm.nih.gov/pubmed/23334500>.
- [30] Blausen. “Blausen gallery 2014”. In : *Wikiversity J. Med.* (2014). ISSN : 20018762. DOI : 10.15347/wjm/2014.010. URL : <https://en.wikiversity.org/wiki/Blausen%7B%5C%7Dgallery%7B%5C%7D2014>.

- [31] D Bonnet and J E Dick. "Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell." In : *Nat. Med.* 3.7 (July 1997), pp. 730–7. ISSN : 1078-8956. URL : <http://www.ncbi.nlm.nih.gov/pubmed/9212098>.
- [32] L. Borradori and A. Sonnenberg. "Structure and function of hemidesmosomes : more than simple adhesion complexes." In : *J. Invest. Dermatol.* 112.4 (Apr. 1999), pp. 411–8. ISSN : 0022-202X. DOI : 10.1046/j.1523-1747.1999.00546.x. URL : <http://www.ncbi.nlm.nih.gov/pubmed/10201522>.
- [33] Eva Brauchle and Katja Schenke-Layland. "Raman spectroscopy in biomedicine - non-invasive in vitro analysis of cells and extracellular matrix components in tissues." In : *Biotechnol. J.* 8.3 (2013), pp. 288–97. ISSN : 1860-7314. DOI : 10.1002/biot.201200163. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3644878&%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://www.ncbi.nlm.nih.gov/pubmed/23161832%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3644878>.
- [34] Dale E Bredesen, Patrick Mehlen, and Shahrooz Rabizadeh. "Apoptosis and dependence receptors : a molecular basis for cellular addiction." In : *Physiol. Rev.* 84.2 (2004), pp. 411–430. ISSN : 0031-9333. DOI : 10.1152/physrev.00027.2003.
- [35] D A Bronzert et al. "Synthesis and secretion of platelet-derived growth factor by human breast cancer cell lines." In : *Proc. Natl. Acad. Sci. U. S. A.* 84.16 (Aug. 1987), pp. 5763–7. ISSN : 0027-8424. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=298943&%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://www.ncbi.nlm.nih.gov/pubmed/3039506%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC298943>.
- [36] E M Brown et al. "Cloning and characterization of an extracellular Ca(2+)-sensing receptor from bovine parathyroid." In : *Nature* 366.6455 (Dec. 1993), pp. 575–80. ISSN : 0028-0836. DOI : 10.1038/366575a0. URL : <http://www.ncbi.nlm.nih.gov/pubmed/8255296>.
- [37] T M Bryan et al. "Telomere elongation in immortal human cells without detectable telomerase activity." In : *EMBO J.* 14.17 (Sept. 1995), pp. 4240–8. ISSN : 0261-4189. DOI : papers3://publication/uuid/E553EAC0-7D77-471F-A996-D41412778FFB. URL : <http://www.ncbi.nlm.nih.gov/pubmed/7556065%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC394507>.
- [38] Tim Burton and Neil B Metcalfe. "Can environmental conditions experienced in early life influence future generations?" In : *Proc. R. Soc. B Biol. Sci.* 281.1785 (May 2014), pp. 20140311–20140311. ISSN : 0962-8452. DOI : 10.1098/rspb.2014.0311. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4024293&%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%5Cbackslash%20http://www.ncbi.nlm.nih.gov/pubmed/24807254%20http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2014.0311>.
- [39] John M. Butler. *Genetics and genomics of core short tandem repeat loci used in human identity testing*. 2006. DOI : 10.1111/j.1556-4029.2006.00046.x.
- [40] J. Cairns. "Mutation and cancer : the antecedents to our studies of adaptive mutation." In : *Genetics* 148.4 (Apr. 1998), pp. 1433–40. ISSN : 0016-6731. URL : <http://www.ncbi.nlm.nih.gov/pubmed/9560363%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1460072>.

- [41] Christophe Y. Calvet, Franck M. André, and Lluís M. Mir. “The Culture of Cancer Cell Lines as Tumorspheres Does Not Systematically Result in Cancer Stem Cell Enrichment”. In : *PLoS One* 9.2 (Feb. 2014). Ed. by Anita B. Hjelmeland, e89644. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0089644. URL : <http://dx.plos.org/10.1371/journal.pone.0089644>.
- [42] Fernando Calvo and Erik Sahai. “Cell communication networks in cancer invasion.” In : *Curr. Opin. Cell Biol.* 23.5 (Oct. 2011), pp. 621–9. ISSN : 1879-0410. DOI : 10.1016/j.ceb.2011.04.010. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21570276>.
- [43] Walter B. Cannon. *The Wisdom Of The Body*. 2nd editio. W.W. Norton & Company, inc, 1967. ISBN : 978-0393002058. URL : <https://books.google.fr/books?id=PTYjQQAACAAJ>.
- [44] Lu Cao et al. “Sphere-forming cell subpopulations with cancer stem cell properties in human hepatoma cell lines”. In : *BMC Gastroenterol.* 11.1 (2011), p. 71. ISSN : 1471-230X. DOI : 10.1186/1471-230X-11-71. URL : <http://www.biomedcentral.com/1471-230X/11/71>.
- [45] Patrícia Carneiro et al. “E-cadherin dysfunction in gastric cancer—cellular consequences, clinical applications and open questions.” In : *FEBS Lett.* 586.18 (Aug. 2012), pp. 2981–9. ISSN : 1873-3468. DOI : 10.1016/j.febslet.2012.07.045. URL : <http://doi.wiley.com/10.1016/j.febslet.2012.07.045>
<http://www.ncbi.nlm.nih.gov/pubmed/22841718>.
- [46] Sandra N Catlin et al. “The replication rate of human hematopoietic stem cells in vivo.” In : *Blood* 117.17 (Apr. 2011), pp. 4460–6. ISSN : 1528-0020. DOI : 10.1182/blood-2010-08-303537. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3099568&tool=pmcentrez&rendertype=abstract>
<http://www.ncbi.nlm.nih.gov/pubmed/21343613>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3099568>.
- [47] William A Catterall et al. “Regulation of sodium and calcium channels by signaling complexes.” In : *J. Recept. Signal Transduct. Res.* 26.5-6 (2006), pp. 577–98. ISSN : 1079-9893. DOI : 10.1080/10799890600915100. URL : <http://www.ncbi.nlm.nih.gov/pubmed/17118799>
<http://informahealthcare.com/doi/pdfplus/10.1080/10799890600915100>
<http://www.ncbi.nlm.nih.gov/pubmed/17118799>.
- [48] Christine L Chaffer et al. “Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state.” In : *Proc. Natl. Acad. Sci. U. S. A.* 108.19 (May 2011), pp. 7950–5. ISSN : 1091-6490. DOI : 10.1073/pnas.1102454108. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3093533&tool=pmcentrez&rendertype=abstract>.
- [49] Emmanuelle Charafe-Jauffret, Christophe Ginestier, and Daniel Birnbaum. “Breast cancer stem cells : tools and models to rely on.” In : *BMC Cancer* 9 (Jan. 2009), p. 202. ISSN : 1471-2407. DOI : 10.1186/1471-2407-9-202. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2708191&tool=pmcentrez&rendertype=abstract>.

- [50] Emmanuelle Charafe-Jauffret et al. "Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature." In : *Cancer Res.* 69.4 (Feb. 2009), pp. 1302–1313. ISSN : 0008-5472. DOI : 10.1158/0008-5472.CAN-08-2741. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2819227%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://cancerres.aacrjournals.org/cgi/doi/10.1158/0008-5472.CAN-08-2741>.
- [51] Chang-Zheng Chen et al. "MicroRNAs modulate hematopoietic lineage differentiation." In : *Science* 303.5654 (Jan. 2004), pp. 83–6. ISSN : 1095-9203. DOI : 10.1126/science.1091903. URL : <http://www.ncbi.nlm.nih.gov/pubmed/14657504>.
- [52] Chiann-mun Chen et al. "A Comparison of Exogenous Promoter Activity at the ROSA26 Locus Using a PhiC31 Integrase Mediated Cassette Exchange Approach in Mouse ES Cells". In : *PLoS One* 6.8 (Aug. 2011). Ed. by Andre Van Wijnen, e23376. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0023376. URL : <http://dx.plos.org/10.1371/journal.pone.0023376>.
- [53] Zhengshan Chen and Jiang Gu. "Immunoglobulin G expression in carcinomas and cancer cell lines." In : *FASEB J.* 21.11 (Sept. 2007), pp. 2931–8. ISSN : 1530-6860. DOI : 10.1096/fj.07-8073com. URL : <http://www.ncbi.nlm.nih.gov/pubmed/17475920>.
- [54] Tom H. Cheung and Thomas A. Rando. "Molecular regulation of stem cell quiescence." In : *Nat. Rev. Mol. Cell Biol.* 14.6 (May 2013), pp. 329–40. ISSN : 1471-0080. DOI : 10.1038/nrm3591. arXiv : NIHMS150003. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0734975011001534%20http://www.nature.com/doifinder/10.1038/nrm3591%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3808888%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [55] Mi-Joo Chung et al. "Prognostic significance of serum carcinoembryonic antigen normalization on survival in rectal cancer treated with preoperative chemoradiation." In : *Cancer Res. Treat.* 45.3 (Sept. 2013), pp. 186–92. ISSN : 1598-2998. DOI : 10.4143/crt.2013.45.3.186. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3804730%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://www.ncbi.nlm.nih.gov/pubmed/24155677%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3804730>.
- [56] Michael F. Clarke et al. "Cancer stem cells - Perspectives on current status and future directions : AACR workshop on cancer stem cells". In : *Cancer Res.* 66.19 (2006), pp. 9339–9344. ISSN : 00085472. DOI : 10.1158/0008-5472.CAN-06-3126.
- [57] Núria Coll-Bonfill et al. "Transdifferentiation of endothelial cells to smooth muscle cells play an important role in vascular remodelling." In : *Am. J. Stem Cells* 4.1 (2015), pp. 13–21. ISSN : 2160-4150. URL : <http://www.ncbi.nlm.nih.gov/pubmed/25973327%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4396157>.
- [58] Francesco Colotta et al. "Cancer-related inflammation, the seventh hallmark of cancer : links to genetic instability." In : *Carcinogenesis* 30.7 (July 2009), pp. 1073–81. ISSN : 1460-2180. DOI : 10.1093/carcin/bgp127. URL : <http://www.ncbi.nlm.nih.gov/pubmed/19468060>.
- [59] J S Connexin, P D Connexin, and A L Connexin. "Cell line misidentification : the beginning of the end." In : *Nat. Rev. Cancer* 10.6 (June 2010), pp. 441–448. ISSN : 1474-175X. DOI : 10.1038/nrc2852.

- [60] Adam P Cribbs et al. "Simplified production and concentration of lentiviral vectors to achieve high transduction in primary human T cells." In : *BMC Biotechnol.* 13.1 (Jan. 2013), p. 98. ISSN : 1472-6750. DOI : 10.1186/1472-6750-13-98. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3830501%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [61] F Crick. "Central dogma of molecular biology." In : *Nature* 227.5258 (Aug. 1970), pp. 561–3. ISSN : 0028-0836. URL : <http://www.ncbi.nlm.nih.gov/pubmed/4913914>.
- [62] Cécile Crosnier, Despina Stamatakis, and Julian Lewis. "Organizing cell renewal in the intestine : stem cells, signals and combinatorial control." In : *Nat. Rev. Genet.* 7.5 (2006), pp. 349–359. ISSN : 1471-0056. DOI : 10.1038/nrg1840.
- [63] Deanna Cross and James K Burmester. "Gene therapy for cancer treatment : past, present and future." In : *Clin. Med. Res.* 4.3 (Sept. 2006), pp. 218–27. ISSN : 1539-4182. DOI : 10.3121/cmr.4.3.218. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1570487%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://www.ncbi.nlm.nih.gov/pubmed/16988102%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1570487>.
- [64] Chi V Dang. "Links between metabolism and cancer." In : *Genes Dev.* 26.9 (2012), pp. 877–90. ISSN : 1549-5477. DOI : 10.1101/gad.189365.112. URL : <http://genesdev.cshlp.org/content/26/9/877.full>.
- [65] M A Davies and Y Samuels. "Analysis of the genome to personalize therapy for melanoma." In : *Oncogene* 29.41 (Oct. 2010), pp. 5545–55. ISSN : 1476-5594. DOI : 10.1038/onc.2010.323. URL : <http://www.ncbi.nlm.nih.gov/pubmed/20697348%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3169242>.
- [66] Penny K. Davis and Rainer K. Brachmann. *Chromatic remodeling and cancer*. 2003.
- [67] Lucia Daxinger and Emma Whitelaw. "Transgenerational epigenetic inheritance : more questions than answers." In : *Genome Res.* 20.12 (2010), pp. 1623–8. ISSN : 1549-5469. DOI : 10.1101/gr.106138.110. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2989988%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [68] Farshid Dayyani et al. "Meta-analysis of the impact of human papillomavirus (HPV) on cancer risk and overall survival in head and neck squamous cell carcinomas (HNSCC)". In : *Head Neck Oncol.* 2.1 (2010), p. 15. ISSN : 1758-3284. DOI : 10.1186/1758-3284-2-15. URL : <http://www.ncbi.nlm.nih.gov/pubmed/20587061%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2908081%20http://www.headandneckoncology.org/content/2/1/15>.
- [69] Bisrat G Debeb et al. "Pre-clinical studies of Notch signaling inhibitor RO4929097 in inflammatory breast cancer cells." In : *Breast Cancer Res. Treat.* 134.2 (July 2012), pp. 495–510. ISSN : 1573-7217. DOI : 10.1007/s10549-012-2075-8. URL : <http://www.ncbi.nlm.nih.gov/pubmed/22547109>.
- [70] Leif Dehmelt and Philippe I H Bastiaens. "Spatial organization of intracellular communication : insights from imaging." In : *Nat. Rev. Mol. Cell Biol.* 11.6 (June 2010), pp. 440–52. ISSN : 1471-0080. DOI : 10.1038/nrm2903. URL : <http://dx.doi.org/10.1038/nrm2903%20http://www.ncbi.nlm.nih.gov/pubmed/20485292>.

- [71] Devaveena Dey et al. "Phenotypic and Functional Characterization of Human Mammary Stem/Progenitor Cells in Long Term Culture". In : *PLoS One* 4.4 (Apr. 2009). Ed. by Gianni Parise, e5329. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0005329. URL : <http://dx.plos.org/10.1371/journal.pone.0005329>.
- [72] Riccardo Di Fiore et al. "RB1 in cancer : different mechanisms of RB1 inactivation and alterations of pRb pathway in tumorigenesis." In : *J. Cell. Physiol.* 228.8 (Aug. 2013), pp. 1676–87. ISSN : 1097-4652. DOI : 10.1002/jcp.24329. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23359405>.
- [73] Daniel Djakiew et al. "Role of the low-affinity nerve growth factor receptor and the high-affinity trk nerve growth factor receptor in human prostate carcinogenesis". In : *Radiat. Oncol. Investig.* 3.6 (1995), pp. 333–339. ISSN : 10657541. DOI : 10.1002/roi.2970030617. URL : <http://doi.wiley.com/10.1002/roi.2970030617>.
- [74] Xianjun Dong et al. "Modeling gene expression using chromatin features in various cellular contexts." In : *Genome Biol.* 13.9 (2012), R53. ISSN : 1474-760X. DOI : 10.1186/gb-2012-13-9-r53. URL : <http://genomebiology.com/2012/13/9/R53>
<http://www.ncbi.nlm.nih.gov/pubmed/22950368>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3491397>.
- [75] Gabriela Dontu et al. "Stem cells in normal breast development and breast cancer." In : *Cell Prolif.* 36 Suppl 1 (2003), pp. 59–72. ISSN : 0960-7722. DOI : 274[pii].
- [76] Kimberly J Dougherty, Michael A Sawchuk, and Shawn Hochman. "Phenotypic diversity and expression of GABAergic inhibitory interneurons during postnatal development in lumbar spinal cord of glutamic acid decarboxylase 67-green fluorescent protein mice." In : *Neuroscience* 163.3 (Oct. 2009), pp. 909–19. ISSN : 1873-7544. DOI : 10.1016/j.neuroscience.2009.06.055. URL : <http://www.sciencedirect.com/science/article/B6T0F-4WM756N-1/2/05ecaf74f204be9ddcfae596141b9fe1>
<http://www.ncbi.nlm.nih.gov/pubmed/19560523>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2746868>.
- [77] Asuka Eguchi et al. "Controlling gene networks and cell fate with precision-targeted DNA-binding proteins and small-molecule-based genome readers". In : *Biochem. J.* 462.3 (Sept. 2014), pp. 397–413. ISSN : 0264-6021. DOI : 10.1042/BJ20140400. URL : <http://www.ncbi.nlm.nih.gov/pubmed/25145439>
<http://www.biochemj.org/bj/462/bj4620397.htm>.
- [78] J Ehnfors et al. "Horizontal transfer of tumor DNA to endothelial cells in vivo." In : *Cell Death Differ.* 16.5 (May 2009), pp. 749–57. ISSN : 1476-5403. DOI : 10.1038/cdd.2009.7. URL : <http://dx.doi.org/10.1038/cdd.2009.7>
<http://www.ncbi.nlm.nih.gov/pubmed/19219067>.
- [79] Eli Eisenberg and Erez Y. Levanon. "Human housekeeping genes, revisited". In : *Trends Genet.* 29.10 (Oct. 2013), pp. 569–574. ISSN : 01689525. DOI : 10.1016/j.tig.2013.05.010. URL : <http://dx.doi.org/10.1016/j.tig.2013.05.010>
<http://linkinghub.elsevier.com/retrieve/pii/S0168952513000899>.
- [80] Fons Elstrodt et al. "BRCA1 mutation analysis of 41 human breast cancer cell lines reveals three new deleterious mutants." In : *Cancer Res.* 66.1 (Jan. 2006), pp. 41–5. ISSN : 0008-5472. DOI : 10.1158/0008-5472.CAN-05-2853. URL : <http://www.ncbi.nlm.nih.gov/pubmed/16397213>.

- [81] S P Ethier et al. "The influence of growth factors on the proliferative potential of normal and primary breast cancer-derived human breast epithelial cells." In : *Breast Cancer Res. Treat.* 17.3 (1991), pp. 221–230.
- [82] Stephen P Ethier et al. "Differential isolation of normal luminal mammary epithelial cells and breast cancer cells from primary and metastatic sites using selective media". In : *Cancer Res.* 53.3 (1993), pp. 627–635. ISSN : 00085472.
- [83] Sandrine Etienne-Manneville. "Neighborly relations during collective migration." In : *Curr. Opin. Cell Biol.* 30 (Oct. 2014), pp. 51–9. ISSN : 1879-0410. DOI : 10.1016/j.ceb.2014.06.004. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24997300>.
- [84] Annarita Favia et al. "VEGF-induced neoangiogenesis is mediated by NAADP and two-pore channel-2-dependent Ca²⁺ signaling." In : *Proc. Natl. Acad. Sci. U. S. A.* 111.44 (Nov. 2014), E4706–15. ISSN : 1091-6490. DOI : 10.1073/pnas.1406029111. URL : <http://www.ncbi.nlm.nih.gov/pubmed/25331892><http://www.pnas.org/cgi/doi/10.1073/pnas.1406029111>.
- [85] Boris Fehse et al. "Selective Immunoaffinity-Based Enrichment of CD34 + Cells Transduced with Retroviral Vectors Containing an Intracytoplasmatically Truncated Version of the Human Low-Affinity Nerve Growth Factor Receptor (Δ LNGFR) Gene". In : *Hum. Gene Ther.* 8.15 (Oct. 1997), pp. 1815–1824. ISSN : 1043-0342. DOI : 10.1089/hum.1997.8.15-1815. URL : <http://www.ncbi.nlm.nih.gov/pubmed/9358031><http://www.liebertonline.com/doi/abs/10.1089/hum.1997.8.15-1815>.
- [86] D. Fenistein et al. "A fast, fully automated cell segmentation algorithm for high-throughput and high-content screening". In : *Cytom. Part A* 73.7 (2008), pp. 958–964. ISSN : 15524922. DOI : 10.1002/cyto.a.20627.
- [87] I. S. Fentiman. "Cell communication in breast cancer." In : *Ann. R. Coll. Surg. Engl.* 62.4 (July 1980), pp. 280–6. ISSN : 0035-8843. URL : <http://www.ncbi.nlm.nih.gov/pubmed/7396358><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2492285>.
- [88] I J Fidler and M L Kripke. "Metastasis results from preexisting variant cells within a malignant tumor." In : *Science* 197.4306 (Aug. 1977), pp. 893–5. ISSN : 0036-8075. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3898730><http://www.ncbi.nlm.nih.gov/pubmed/887927><http://www.ncbi.nlm.nih.gov/pubmed/887927>.
- [89] Christine M Fillmore and Charlotte Kuperwasser. "Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy." In : *Breast Cancer Res.* 10.2 (Jan. 2008), R25. ISSN : 1465-5411. DOI : 10.1186/bcr1982. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2397524><http://breast-cancer-research.com/content/10/2/R25>.
- [90] Christine Fillmore and Charlotte Kuperwasser. "Human breast cancer stem cell markers CD44 and CD24 : enriching for cells with functional properties in mice or in man?" In : *Breast Cancer Res.* 9.3 (2007), p. 303. ISSN : 14655411. DOI : 10.1186/bcr1673. URL : <http://breast-cancer-research.com/content/9/3/303>.
- [91] Henri J Folse and Joan Roughgarden. "What is an individual organism? A multilevel selection perspective." In : *Q. Rev. Biol.* 85.4 (Dec. 2010), pp. 447–72. ISSN : 0033-5770. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21243964>.

- [92] F Forozan et al. “Molecular cytogenetic analysis of 11 new breast cancer cell lines.” In : *Br. J. Cancer* 81.8 (Dec. 1999), pp. 1328–34. ISSN : 0007-0920. DOI : 10.1038/sj.bjc.6695007. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2362964%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [93] K Francis and B O Palsson. “Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion.” In : *Proc. Natl. Acad. Sci. U. S. A.* 94.23 (Nov. 1997), pp. 12258–62. ISSN : 0027-8424. URL : <http://www.ncbi.nlm.nih.gov/pubmed/9356436%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC24899>.
- [94] Douglas E. Friesen, Vickie E. Baracos, and Jack a. Tuszynski. “Modeling the energetic cost of cancer as a result of altered energy metabolism : implications for cachexia.” In : *Theor. Biol. Med. Model.* 12.1 (2015), p. 17. ISSN : 1742-4682. DOI : 10.1186/s12976-015-0015-0. URL : <http://www.tbiomed.com/content/12/1/17%20http://www.ncbi.nlm.nih.gov/pubmed/26370269%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4570294>.
- [95] Hirota Fujiki. “Gist of Dr. Katsusaburo Yamagiwa’s papers entitled ‘Experimental study on the pathogenesis of epithelial tumors’ (I to VI reports).” In : *Cancer Sci.* 105.2 (2014), pp. 143–9. ISSN : 1349-7006. DOI : 10.1111/cas.12333. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24313817>.
- [96] Thomas F Gajewski, Hans Schreiber, and Yang-Xin Fu. “Innate and adaptive immune cells in the tumor microenvironment”. In : *Nat. Immunol.* 14.10 (Sept. 2013), pp. 1014–1022. ISSN : 1529-2908. DOI : 10.1038/ni.2703. arXiv : NIHMS150003. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24048123%20http://www.nature.com/doi/10.1038/ni.2703>.
- [97] Lorenzo Galluzzi et al. “Metabolic targets for cancer therapy”. In : *Nat. Publ. Gr.* 12.11 (Oct. 2013), pp. 829–846. ISSN : 1474-1776. DOI : 10.1038/nrd4145. URL : [http://www.nature.com/doi/10.1038/nrd4145%20http://dx.doi.org/10.1038/nrd4145%5Cbackslash\\$npapers2://publication/doi/10.1038/nrd4145](http://www.nature.com/doi/10.1038/nrd4145%20http://dx.doi.org/10.1038/nrd4145%5Cbackslash$npapers2://publication/doi/10.1038/nrd4145).
- [98] Milena Gasco, Isik G. Yulug, and Tim Crook. “TP53 mutations in familial breast cancer : functional aspects.” In : *Hum. Mutat.* 21.3 (Mar. 2003), pp. 301–6. ISSN : 1098-1004. DOI : 10.1002/humu.10173. URL : <http://www.ncbi.nlm.nih.gov/pubmed/12619116>.
- [99] Bernhard Gentner et al. “Identification of hematopoietic stem cell-specific miRNAs enables gene therapy of globoid cell leukodystrophy.” In : *Sci. Transl. Med.* 2.58 (Nov. 2010), 58ra84. ISSN : 1946-6242. DOI : 10.1126/scitranslmed.3001522. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21084719>.
- [100] Efstathia Giannopoulou et al. “Epidermal growth factor receptor status and Notch inhibition in non-small cell lung cancer cells.” In : *J. Biomed. Sci.* 22.1 (2015), p. 98. ISSN : 1423-0127. DOI : 10.1186/s12929-015-0196-1. URL : <http://www.jbiomedsci.com/content/22/1/98%20http://www.ncbi.nlm.nih.gov/pubmed/26497899%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4619334>.
- [101] Richard J Gilbertson and Jeremy N Rich. “Making a tumour’s bed : glioblastoma stem cells and the vascular niche.” In : *Nat. Rev. Cancer* 7.10 (Oct. 2007), pp. 733–6. ISSN : 1474-1768. DOI : 10.1038/nrc2246. URL : <http://www.ncbi.nlm.nih.gov/pubmed/17882276>.
- [102] Christophe Ginestier et al. “La cellule souche cancéreuse Un pilote aux commandes”. In : *Med. Sci.* 23.1133 (2007), p. 9. ISSN : 0767-0974.

- [103] Mario Giordano. “Homeostasis : an underestimated focal point of ecology and evolution.” In : *Plant Sci.* 211 (Oct. 2013), pp. 92–101. ISSN : 1873-2259. DOI : 10.1016/j.plantsci.2013.07.008. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23987815>.
- [104] Samuel Godar et al. “Growth-inhibitory and tumor- suppressive functions of p53 depend on its repression of CD44 expression.” In : *Cell* 134.1 (July 2008), pp. 62–73. ISSN : 1097-4172. DOI : 10.1016/j.cell.2008.06.006. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3222460%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [105] Yann Godet et al. “Analysis of spontaneous tumor-specific CD4 T-cell immunity in lung cancer using promiscuous HLA-DR telomerase-derived epitopes : potential synergistic effect with chemotherapy response.” In : *Clin. Cancer Res.* 18.10 (May 2012), pp. 2943–53. ISSN : 1078-0432. DOI : 10.1158/1078-0432.CCR-11-3185. URL : <http://www.ncbi.nlm.nih.gov/pubmed/22407833>.
- [106] Céline Gomez et al. “Control of segment number in vertebrate embryos.” In : *Nature* 454.7202 (July 2008), pp. 335–9. ISSN : 1476-4687. DOI : 10.1038/nature07020. URL : <http://www.nature.com/doi/10.1038/nature07020%20http://www.ncbi.nlm.nih.gov/pubmed/18563087>.
- [107] Michael M Gottesman, Tito Fojo, and Susan E Bates. “Multidrug resistance in cancer : role of ATP-dependent transporters.” In : *Nat. Rev. Cancer* 2.1 (Jan. 2002), pp. 48–58. ISSN : 1474-175X. DOI : 10.1038/nrc706. URL : <papers3://publication/doi/10.1038/nrc706%20http://www.nature.com/doi/10.1038/nrc706%20http://www.ncbi.nlm.nih.gov/pubmed/11902585>.
- [108] Caroline Grabbe, Koraljka Husnjak, and Ivan Dikic. “The spatial and temporal organization of ubiquitin networks.” In : *Nat. Rev. Mol. Cell Biol.* 12.5 (May 2011), pp. 295–307. ISSN : 1471-0080. DOI : 10.1038/nrm3099. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21448225%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3654194>.
- [109] Anita Grigoriadis et al. “Molecular characterisation of cell line models for triple-negative breast cancers.” In : *BMC Genomics* 13.1 (2012), p. 619. ISSN : 1471-2164. DOI : 10.1186/1471-2164-13-619. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3546428%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://www.ncbi.nlm.nih.gov/pubmed/23151021%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3546428>.
- [110] Charles G. Gross. “Genealogy of the Grandmother Cell”. In : *Neurosci.* 8.5 (Oct. 2002), pp. 512–518. ISSN : 1073-8584. DOI : 10.1177/107385802237175. URL : <http://nro.sagepub.com/cgi/doi/10.1177/107385802237175%20http://www.ncbi.nlm.nih.gov/pubmed/12374433>.
- [111] M. C. Guadamillas, A. Cerezo, and M. a. del Pozo. “Overcoming anoikis - pathways to anchorage-independent growth in cancer”. In : *J. Cell Sci.* 124.19 (Oct. 2011), pp. 3189–3197. ISSN : 0021-9533. DOI : 10.1242/jcs.072165. URL : <http://jcs.biologists.org/cgi/doi/10.1242/jcs.072165>.
- [112] Pranesh Gunjal et al. “An emerging question about putative cancer stem cells in established cell lines— are they true stem cells or a fluctuating cell phenotype ?” In : *J. Cancer Stem Cell Res.* 3.4 (2015), p. 1. ISSN : 2329-5872. DOI : 10.14343/JCSCR.2015.3e1004. URL : <http://cancerstemcellsresearch.com>.

- [113] Piyush B. Gupta et al. “Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells”. In : *Cell* 146.4 (Aug. 2011), pp. 633–644. ISSN : 00928674. DOI : 10.1016/j.cell.2011.07.026. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21854987> <http://linkinghub.elsevier.com/retrieve/pii/S0092867411008245>.
- [114] I K Guttilla et al. “Prolonged mammosphere culture of MCF-7 cells induces an EMT and repression of the estrogen receptor by microRNAs.” In : *Breast Cancer Res. Treat.* 132.1 (Feb. 2012), pp. 75–85. ISSN : 1573-7217. DOI : 10.1007/s10549-011-1534-y. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21553120>.
- [115] Patrice G Guyenet and Douglas A Bayliss. “Neural Control of Breathing and CO₂ Homeostasis.” In : *Neuron* 87.5 (Sept. 2015), pp. 946–61. ISSN : 1097-4199. DOI : 10.1016/j.neuron.2015.08.001. URL : <http://www.ncbi.nlm.nih.gov/pubmed/26335642> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4559867>.
- [116] Florian Hahne et al. “Per-channel basis normalization methods for flow cytometry data”. In : *Cytom. Part A* 77.2 (2010), pp. 121–131. ISSN : 15524922. DOI : 10.1002/cyto.a.20823.
- [117] Muhammad Al-Hajj et al. “Prospective identification of tumorigenic breast cancer cells.” In : *Proc. Natl. Acad. Sci. U. S. A.* 100.7 (Apr. 2003), pp. 3983–8. ISSN : 0027-8424. DOI : 10.1073/pnas.0530291100. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=153034%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [118] Fujun Han and Biliang Zhang. “Characterizing cell-cell interactions induced spatial organization of cell phenotypes : application to density-dependent protein nucleocytoplasmic distribution.” In : *Cell Biochem. Biophys.* 65.2 (Mar. 2013), pp. 163–72. ISSN : 1559-0283. DOI : 10.1007/s12013-012-9412-8. URL : <http://www.ncbi.nlm.nih.gov/pubmed/22915253>.
- [119] D Hanahan and R A Weinberg. “The hallmarks of cancer.” In : *Cell* 100.1 (Jan. 2000), pp. 57–70. ISSN : 0092-8674. URL : <http://www.ncbi.nlm.nih.gov/pubmed/10647931>.
- [120] Douglas Hanahan and Judah Folkman. “Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis.” In : *Cell* 86.3 (Aug. 1996), pp. 353–64. ISSN : 0092-8674. DOI : 10.1016/S0092-8674(00)80108-7. URL : <http://www.ncbi.nlm.nih.gov/pubmed/8756718>.
- [121] Douglas Hanahan and Robert a Weinberg. “Hallmarks of cancer : the next generation.” In : *Cell* 144.5 (Mar. 2011), pp. 646–74. ISSN : 1097-4172. DOI : 10.1016/j.cell.2011.02.013. arXiv : 0208024 [gr-qc]. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21376230>.
- [122] Pietradewi Hartrianti et al. “Modulating Mesenchymal Stem Cell Behavior Using Human Hair Keratin-Coated Surfaces.” In : *Stem Cells Int.* 2015 (2015), p. 752424. ISSN : 1687-966X. DOI : 10.1155/2015/752424. URL : <http://www.ncbi.nlm.nih.gov/pubmed/26124842> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4466490>.

- [123] Rippei Hayashi, Nono Takeuchi, and Takuya Ueda. “Nuclear Respiratory Factor 2β (NRF- 2β) recruits NRF- 2α to the nucleus by binding to importin- α : β via an unusual monopartite-type nuclear localization signal.” In : *J. Mol. Biol.* 425.18 (Sept. 2013), pp. 3536–48. ISSN : 1089-8638. DOI : 10.1016/j.jmb.2013.07.007. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23856623>.
- [124] L Hayflick. “Intracellular determinants of cell aging.” In : *Mech. Ageing Dev.* 28.2-3 (Dec. 1984), pp. 177–85. ISSN : 0047-6374. URL : <http://www.ncbi.nlm.nih.gov/pubmed/6521502>.
- [125] L HAYFLICK and P S MOORHEAD. “The serial cultivation of human diploid cell strains.” In : *Exp. Cell Res.* 25 (Dec. 1961), pp. 585–621. ISSN : 0014-4827. URL : <http://www.ncbi.nlm.nih.gov/pubmed/13905658>.
- [126] National Institutes of Health. *Stem cells basics*. URL : <http://stemcells.nih.gov/info/basics/pages/basics2.aspx>.
- [127] Leonore a Herzenberg et al. “Interpreting flow cytometry data : a guide for the perplexed.” In : *Nat. Immunol.* 7.7 (2006), pp. 681–685. ISSN : 1529-2908. DOI : 10.1038/ni0706-681.
- [128] D Heuss. “Light microscopy study of low-affinity nerve growth factor receptor and phosphoprotein B-50/neuromodulin in inflammatory myopathies.” In : *Acta Neuropathol.* 91.4 (1996), pp. 409–415. ISSN : 0001-6322.
- [129] R K Hirata et al. “Transduction of hematopoietic cells by foamy virus vectors.” In : *Blood* 88.9 (Nov. 1996), pp. 3654–61. ISSN : 0006-4971. URL : <http://www.ncbi.nlm.nih.gov/pubmed/8896436>.
- [130] Hsiu J Ho et al. “Parametric modeling of cellular state transitions as measured with flow cytometry.” In : *BMC Bioinformatics* 13 Suppl 5.Suppl 5 (Jan. 2012), S5. ISSN : 1471-2105. DOI : 10.1186/1471-2105-13-S5-S5. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3358665%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [131] Amelia Hochreiter-Hufford and Kodi S. Ravichandran. “Clearing the dead : apoptotic cell sensing, recognition, engulfment, and digestion.” In : *Cold Spring Harb. Perspect. Biol.* 5.1 (Jan. 2013), a008748. ISSN : 1943-0264. DOI : 10.1101/cshperspect.a008748. URL : <http://cshperspectives.cshlp.org/lookup/doi/10.1101/cshperspect.a008748%20http://www.ncbi.nlm.nih.gov/pubmed/23284042%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3579390>.
- [132] M Höckel and P Vaupel. “Tumor hypoxia : definitions and current clinical, biologic, and molecular aspects.” In : *J. Natl. Cancer Inst.* 93.4 (Feb. 2001), pp. 266–76. ISSN : 0027-8874. DOI : 10.1093/jnci/93.4.266. URL : <http://www.ncbi.nlm.nih.gov/pubmed/11181773>.
- [133] Dirk Hockemeyer et al. “Genetic engineering of human pluripotent cells using TALE nucleases.” In : *Nat. Biotechnol.* 29.8 (Aug. 2011), pp. 731–4. ISSN : 1546-1696. DOI : 10.1038/nbt.1927. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3152587%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [134] B Hoffman and D a Liebermann. “Apoptotic signaling by c-MYC.” In : *Oncogene* 27.50 (Oct. 2008), pp. 6462–72. ISSN : 1476-5594. DOI : 10.1038/onc.2008.312. URL : <http://www.ncbi.nlm.nih.gov/pubmed/18955973>.

- [135] Antoinette Hollestelle et al. “Distinct gene mutation profiles among luminal-type and basal-type breast cancer cell lines”. In : *Breast Cancer Res. Treat.* 121 (2010), pp. 53–64. ISSN : 01676806. DOI : 10.1007/s10549-009-0460-8.
- [136] Sunghoi Hong et al. “Functional Analysis of Various Promoters in Lentiviral Vectors at Different Stages of In Vitro Differentiation of Mouse Embryonic Stem Cells”. In : *Mol. Ther.* 15.9 (Sept. 2007), pp. 1630–1639. ISSN : 1525-0016. DOI : 10.1038/sj.mt.6300251. URL : <http://www.nature.com/mt/journal/vaop/ncurrent/full/6300251a.html> <http://www.nature.com/doi/10.1038/sj.mt.6300251>.
- [137] Robert Hooke. *Micrographia : Or Some Physiological Descriptions Of Minute Bodies Made By Magnifying Glasses. With Observations and Inquiries Thereupon.* John Martyn, 1667. URL : <https://books.google.fr/books?id=ZNBWAAAACAAJ>.
- [138] Frederik Memorial Hospital. *Tumor markers.* 2016. URL : <http://www.fmh.org/body.cfm?id=598>.
- [139] Sui Huang. “Tumor progression : Chance and necessity in Darwinian and Lamarckian somatic (mutationless) evolution”. In : *Prog. Biophys. Mol. Biol.* 110.1 (Sept. 2012), pp. 69–86. ISSN : 00796107. DOI : 10.1016/j.pbiomolbio.2012.05.001. URL : <http://www.ncbi.nlm.nih.gov/pubmed/22579660> <http://linkinghub.elsevier.com/retrieve/pii/S0079610712000247>.
- [140] Taro Ichimura et al. “Visualizing cell state transition using Raman spectroscopy.” In : *PLoS One* 9.1 (2014), e84478. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0084478. URL : <http://dx.plos.org/10.1371/journal.pone.0084478> <http://www.ncbi.nlm.nih.gov/pubmed/24409302> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3883674>.
- [141] Satoko Inoda et al. “Cytotoxic T lymphocytes efficiently recognize human colon cancer stem-like cells.” In : *Am. J. Pathol.* 178.4 (Apr. 2011), pp. 1805–13. ISSN : 1525-2191. DOI : 10.1016/j.ajpath.2011.01.004. URL : <http://dx.doi.org/10.1016/j.ajpath.2011.01.004> <http://www.ncbi.nlm.nih.gov/pubmed/21435460> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3078439>.
- [142] Haruhisa Inoue et al. “iPS cells : a game changer for future medicine”. In : *EMBO J.* 33.5 (Mar. 2014), pp. 409–417. ISSN : 0261-4189. DOI : 10.1002/embj.201387098. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24500035> <http://emboj.embopress.org/cgi/doi/10.1002/embj.201387098>.
- [143] Institut Jules Bordet. *Origine du cancer.* URL : <http://www.bordet.be/fr/presentation/historic/cancer/cancer2.htm>.
- [144] S Isoe et al. “Resistance to growth inhibition by transforming growth factor-beta in malignant glioma cells with functional receptors”. In : *J Neurosurg* 88.3 (1998), pp. 529–534. ISSN : 0022-3085. DOI : 10.3171/jns.1998.88.3.0529.
- [145] Marie Jacolot. “Les cellules souches cancéreuses origines du cancer ? hypothèses, caractéristiques et implications thérapeutiques en médecine humaine et vétérinaire”. Doctorat Vétérinaire. Faculté de Médecine de Créteil, 2013. URL : <https://www.google.fr/url?sa=t&drct=j&dq=%7B%5C%7Desrc=s%7B%5C%7Dsource=web%7B%5C%7Dcd=8%7B%5C%7Dved=0CF0QFjAHahUKewi%7B%5C%7DpbH3oq3HAhVDqxoKHyzABsA%7B%5C%7Durl=http://theses.vet-alfort.fr/telecharger.php?id=1471%7B%5C%7Dei=plXQVb-6IMPWaoYBm4AM%7B%5C%7Dusg=AFQjCNG9dyiWS38AX0x5WmpcD65QGQfKBQ%7B%5C%7Dsig2=eKk4CkUCgk50r1HM4OKbCw%7B%5C%7D>.

- [146] N Jahroudi and D C Lynch. "Endothelial-cell-specific regulation of von Willebrand factor gene expression." In : *Mol. Cell. Biol.* 14.2 (Dec. 1994), pp. 999–1008. ISSN : 0270-7306. DOI : 10.1016/0378-1119(91)90434-D. URL : <http://linkinghub.elsevier.com/retrieve/pii/037811199190434D>.
- [147] Paul A. Janmey and Christopher A. McCulloch. "Cell mechanics : integrating cell responses to mechanical stimuli." In : *Annu. Rev. Biomed. Eng.* 9.1 (2007), pp. 1–34. ISSN : 1523-9829. DOI : 10.1146/annurev.bioeng.9.060906.151927. URL : <http://www.annualreviews.org/doi/abs/10.1146/annurev.bioeng.9.060906.151927> %20<http://www.ncbi.nlm.nih.gov/pubmed/17461730>.
- [148] Sophie Jarriault, Yannick Schwab, and Iva Greenwald. "A *Caenorhabditis elegans* model for epithelial-neuronal transdifferentiation". In : *Proc. Natl. Acad. Sci.* 105.10 (Mar. 2008), pp. 3790–3795. ISSN : 0027-8424. DOI : 10.1073/pnas.0712159105. URL : <http://www.pnas.org/cgi/doi/10.1073/pnas.0712159105>.
- [149] W. Johannsen. "The genotype conception of heredity. 1911." In : *Int. J. Epidemiol.* 43.4 (Aug. 2014), pp. 989–1000. ISSN : 1464-3685. DOI : 10.1093/ije/dyu063. URL : <http://www.ije.oxfordjournals.org/cgi/doi/10.1093/ije/dyu063> %20<http://www.ncbi.nlm.nih.gov/pubmed/24691957> %20<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4258772>.
- [150] Hexin ; Johnson, Sara ; Chen and Pang-Kuo Lo. *In vitro Tumorsphere Formation Assays*. 2013. URL : <http://www.bio-protocol.org/e325>.
- [151] Chris Jopling, Stephanie Boue, and Juan Carlos Izpisua Belmonte. "Dedifferentiation, transdifferentiation and reprogramming : three routes to regeneration." In : *Nat. Rev. Mol. Cell Biol.* 12.2 (2011), pp. 79–89. ISSN : 1471-0072. DOI : 10.1038/nrm3043. URL : <http://dx.doi.org/10.1038/nrm3043>.
- [152] Miyazaki Jun-ichi et al. "Expression vector system based on the chicken β -actin promoter directs efficient production of interleukin-5". In : *Gene* 79.2 (July 1989), pp. 269–277. ISSN : 03781119. DOI : 10.1016/0378-1119(89)90209-6. URL : <http://linkinghub.elsevier.com/retrieve/pii/0378111989902096>.
- [153] D Kalderon et al. "A short amino acid sequence able to specify nuclear location." In : *Cell* 39.3 Pt 2 (Dec. 1984), pp. 499–509. ISSN : 0092-8674. URL : <http://www.ncbi.nlm.nih.gov/pubmed/6096007>.
- [154] O P Kallioniemi et al. "Serum CA 15-3 assay in the diagnosis and follow-up of breast cancer." In : *Br. J. Cancer* 58.2 (Aug. 1988), pp. 213–5. ISSN : 0007-0920. URL : <http://www.ncbi.nlm.nih.gov/pubmed/3166913> %20<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2246767>.
- [155] S Kamel-Reid and J E Dick. "Engraftment of immune-deficient mice with human hematopoietic stem cells." In : *Science* 242.4886 (1988), pp. 1706–1709. ISSN : 0036-8075.
- [156] Yukio Kamohara et al. "The search for cancer stem cells in hepatocellular carcinoma". In : *Surgery* 144.2 (Aug. 2008), pp. 119–124. ISSN : 00396060. DOI : 10.1016/j.surg.2008.04.008. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0039606008002572>.
- [157] Atsushi Kaneda et al. "Epstein-Barr Virus Infection as an Epigenetic Driver of Tumorigenesis". In : *Cancer Res.* 72.14 (July 2012), pp. 3445–3450. ISSN : 0008-5472. DOI : 10.1158/0008-5472.CAN-11-3919. URL : <http://cancerres.aacrjournals.org/cgi/doi/10.1158/0008-5472.CAN-11-3919>.

- [158] S A Kauffman. "Metabolic stability and epigenesis in randomly constructed genetic nets." In : *J. Theor. Biol.* 22.3 (Mar. 1969), pp. 437–67. ISSN : 0022-5193. URL : <http://www.ncbi.nlm.nih.gov/pubmed/5803332>.
- [159] Benjamin B Kaufmann et al. "Heritable stochastic switching revealed by single-cell genealogy." In : *PLoS Biol.* 5.9 (Sept. 2007), e239. ISSN : 1545-7885. DOI : 10.1371/journal.pbio.0050239. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1964776%7B%5C%26%7Dtool=pmcentrez%7B%5C%26%7Drendertype=abstract>.
- [160] J F Kerr, A H Wyllie, and A R Currie. "Apoptosis : a basic biological phenomenon with wide-ranging implications in tissue kinetics." In : *Br. J. Cancer* 26.4 (Aug. 1972), pp. 239–57. ISSN : 0007-0920. URL : <http://www.ncbi.nlm.nih.gov/pubmed/4561027%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2008650>.
- [161] Boris N Kholodenko. "Cell-signalling dynamics in time and space." In : *Nat. Rev. Mol. Cell Biol.* 7.3 (Mar. 2006), pp. 165–76. ISSN : 1471-0072. DOI : 10.1038/nrm1838. URL : <http://www.ncbi.nlm.nih.gov/pubmed/16482094%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1679905>.
- [162] Noriko Kimura et al. "Tissue Localization of Nerve Growth Factor Receptors : trk A and Low-Affinity Nerve Growth Factor Receptor in Neuroblastoma, Pheochromocytoma, and Retinoblastoma." In : *Endocr. Pathol.* 7.4 (1996), pp. 281–289. ISSN : 1559-0097. DOI : EP0704281[pii]. URL : <http://www.ncbi.nlm.nih.gov/pubmed/12114799>.
- [163] Juergen a. Knoblich. "Mechanisms of Asymmetric Stem Cell Division". In : *Cell* 132.4 (2008), pp. 583–597. ISSN : 00928674. DOI : 10.1016/j.cell.2008.02.007. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0092867408002080>.
- [164] Natália Cristina Ciufa Kobayashi and Samuel Marcos Ribeiro De Noronha. "Cancer stem cells : a new approach to tumor development". In : *Rev. Assoc. Med. Bras.* 61.1 (Feb. 2015), pp. 86–93. ISSN : 0104-4230. DOI : 10.1590/1806-9282.61.01.086.
- [165] Juliana Korah, Lucie Canaff, and Jean-Jacques Lebrun. "The Retinoblastoma Tumor Suppressor Protein (pRb)/E2 Promoter Binding Factor 1 (E2F1) Pathway as a Novel Mediator of TGF β -induced Autophagy." In : *J. Biol. Chem.* 291.5 (Jan. 2016), pp. 2043–54. ISSN : 1083-351X. DOI : 10.1074/jbc.M115.678557. URL : <http://www.jbc.org/lookup/doi/10.1074/jbc.M115.678557%20http://www.ncbi.nlm.nih.gov/pubmed/26598524%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4732193>.
- [166] Antonija Kreso and John E. Dick. "Evolution of the Cancer Stem Cell Model". In : *Cell Stem Cell* 14.3 (Mar. 2014), pp. 275–291. ISSN : 19345909. DOI : 10.1016/j.stem.2014.02.006. URL : <http://dx.doi.org/10.1016/j.stem.2014.02.006%20http://linkinghub.elsevier.com/retrieve/pii/S1934590914000575%20http://www.ncbi.nlm.nih.gov/pubmed/24607403>.
- [167] Anna Krzeslak et al. "Expression of GLUT1 and GLUT3 glucose transporters in endometrial and breast cancers." In : *Pathol. Oncol. Res.* 18.3 (July 2012), pp. 721–8. ISSN : 1532-2807. DOI : 10.1007/s12253-012-9500-5. URL : <http://www.ncbi.nlm.nih.gov/pubmed/22270867%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3342495>.
- [168] La Société des Chercheurs sur le Cancer. *Historique du Cancer*. URL : <http://www.societederecherchesurlecancer.ca/page.aspx?pid=1494>.

- [169] Marc Lacroix. “Persistent use of ‘false’ cell lines.” In : *Int. J. Cancer* 122.1 (Jan. 2008), pp. 1–4. ISSN : 1097-0215. DOI : 10.1002/ijc.23233. URL : <http://www.ncbi.nlm.nih.gov/pubmed/17960586>.
- [170] Chrystelle Lamagna and Gabriele Bergers. “The bone marrow constitutes a reservoir of pericyte progenitors.” In : *J. Leukoc. Biol.* 80.4 (Oct. 2006), pp. 677–81. ISSN : 0741-5400. DOI : 10.1189/jlb.0506309. URL : <http://www.ncbi.nlm.nih.gov/pubmed/16888086>.
- [171] Arthur D Lander. “The ‘stem cell’ concept : is it holding us back?” In : *J. Biol.* 8.8 (Jan. 2009), p. 70. ISSN : 1475-4924. DOI : 10.1186/jbiol177. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2776917&B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [172] Allison Lange et al. “Classical nuclear localization signals : Definition, function, and interaction with importin α ”. In : *J. Biol. Chem.* 282.8 (Feb. 2007), pp. 5101–5105. ISSN : 00219258. DOI : 10.1074/jbc.R600026200. URL : <http://www.jbc.org/cgi/doi/10.1074/jbc.R600026200>.
- [173] Eszter Lázár-Molnár et al. “Autocrine and paracrine regulation by cytokines and growth factors in melanoma.” In : *Cytokine* 12.6 (June 2000), pp. 547–54. ISSN : 1043-4666. DOI : 10.1006/cyto.1999.0614. URL : <http://linkinghub.elsevier.com/retrieve/pii/S104346669906142%20http://www.ncbi.nlm.nih.gov/pubmed/10843728>.
- [174] Yuri Lazebnik. “What are the hallmarks of cancer?” In : *Nat. Rev. Cancer* 10.4 (Mar. 2010), pp. 232–233. ISSN : 1474-175X. DOI : 10.1038/nrc2827. URL : <http://www.nature.com/doifinder/10.1038/nrc2827>.
- [175] Eric R Lechman et al. “Attenuation of miR-126 activity expands HSC in vivo without exhaustion.” In : *Cell Stem Cell* 11.6 (Dec. 2012), pp. 799–811. ISSN : 1875-9777. DOI : 10.1016/j.stem.2012.09.001. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3517970&B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [176] J LEDERBERG and E L TATUM. “Gene recombination in *Escherichia coli*.” In : *Nature* 158.4016 (Oct. 1946), p. 558. ISSN : 0028-0836. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21001945>.
- [177] Gyemin Lee, L Stoolman, and C Scott. “Transfer Learning for Auto-gating of Flow Cytometry Data.” In : *JMLR(workshop)* (2012), pp. 155–166. URL : <http://www.clopinet.com/isabelle/Projects/ICML2011/slides/lee11.pdf>.
- [178] Alexander Levitzki and Shoshana Klein. “Signal transduction therapy of cancer.” In : *Mol. Aspects Med.* 31.4 (Aug. 2010), pp. 287–329. ISSN : 1872-9452. DOI : 10.1016/j.mam.2010.04.001. URL : <http://dx.doi.org/10.1016/j.mam.2010.04.001%20http://www.ncbi.nlm.nih.gov/pubmed/20451549>.
- [179] Jing Li et al. “Quantitative analysis of histone H3 and H4 post-translational modifications in doxorubicin resistant leukemia cells.” In : *Biomed. Chromatogr.* (2015). ISSN : 1099-0801. DOI : 10.1002/bmc.3608. URL : <http://www.ncbi.nlm.nih.gov/pubmed/26317407>.
- [180] Xianqiang Li. “Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter”. In : *J. Biol. Chem.* 273.52 (Dec. 1998), pp. 34970–34975. ISSN : 00219258. DOI : 10.1074/jbc.273.52.34970. URL : <http://www.ncbi.nlm.nih.gov/pubmed/9857028%20http://www.jbc.org/cgi/doi/10.1074/jbc.273.52.34970>.

- [181] Michelle S. Liberio et al. “Differential effects of tissue culture coating substrates on prostate cancer cell adherence, morphology and behavior.” In : *PLoS One* 9.11 (2014), e112122. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0112122. URL : <http://www.ncbi.nlm.nih.gov/pubmed/25375165><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4223027>.
- [182] Peter Lichter et al. “Obligation for cell line authentication : Appeal for concerted action”. In : *Int. J. Cancer* 126.1 (Jan. 2010), pp. 1–1. ISSN : 00207136. DOI : 10.1002/ijc.24985. URL : <http://doi.wiley.com/10.1002/ijc.24985>.
- [183] Ligue contre le Cancer. *Brève histoire du cancer du sein*. URL : http://www.ligue-cancer.net/vivre/article/26490%7B%5C_%7Dbrève-histoire-du-cancer-du-sein.
- [184] Ligue contre le Cancer. *Le Cancer à travers les siècles*. URL : http://www.ligue-cancer.net/article/105%7B%5C_%7Dle-cancer-a-travers-les-siecles.
- [185] D. Linon-Chipon, S. ;Vaj. *Relations savantes : voyages et discours scientifiques*. Imago mund. Presses de l’Université Paris-Sorbonne, 2006, pp. 217–218. ISBN : 9782840504696. URL : <https://books.google.fr/books?id=QZTCPeWnj70C>.
- [186] Allan Yi Liu et al. “Twist2 promotes self-renewal of liver cancer stem-like cells by regulating CD24.” In : *Carcinogenesis* 35.3 (Mar. 2014), pp. 537–45. ISSN : 1460-2180. DOI : 10.1093/carcin/bgt364. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24193512>.
- [187] C Lkhagvasuren et al. “A comparative analysis of constitutive promoters located in adeno-associated viral vectors.” In : *PLoS One* 9.8 (Aug. 2014). Ed. by Jitesh Pratap, e106472. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0106472. URL : <http://dx.plos.org/10.1371/journal.pone.0106472><http://www.ncbi.nlm.nih.gov/pubmed/25170953><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4149579>.
- [188] K R Loeb and L a Loeb. “Significance of multiple mutations in cancer.” In : *Carcinogenesis* 21.3 (2000), pp. 379–385. ISSN : 0143-3334. DOI : 10.1093/carcin/21.3.379.
- [189] Aaron C Logan, C Lutzko, and D B Kohn. “Advances in lentiviral vector design for gene- modification of hematopoietic stem cells”. In : *Curr. Opin. Biotechnol.* 13.5 (2002), pp. 429–436.
- [190] Angelo Lombardo et al. “Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery”. In : *Nat. Biotechnol.* 25.11 (Nov. 2007), pp. 1298–1306. ISSN : 1087-0156. DOI : 10.1038/nbt1353. URL : <http://www.ncbi.nlm.nih.gov/pubmed/17965707><http://www.ncbi.nlm.nih.gov/gate1.inist.fr/pubmed/17965707><http://www.nature.com/doifinder/10.1038/nbt1353>.
- [191] Y. A. Luqmani. “Mechanisms of drug resistance in cancer chemotherapy.” In : *Med. Princ. Pract.* 14 Suppl 1.SUPPL. 1 (2005), pp. 35–48. ISSN : 1011-7571. DOI : 10.1159/000086183. URL : <http://www.ncbi.nlm.nih.gov/pubmed/16103712>.
- [192] M Mahner and M Kary. “What exactly are genomes, genotypes and phenotypes? And what about phenomes?” In : *J. Theor. Biol.* 186.1 (May 1997), pp. 55–63. ISSN : 0022-5193. DOI : 10.1006/jtbi.1996.0335. URL : <http://www.sciencedirect.com/science/article/pii/S0022519396903355><http://www.ncbi.nlm.nih.gov/pubmed/9176637>.

- [193] Christophe Malaterre. “Organicism and Reductionism in Cancer Research : Towards a Systemic Approach”. In : *Int. Stud. Philos. Sci.* 21.1 (Mar. 2007), pp. 57–73. ISSN : 0269-8595. DOI : 10.1080/02698590701305792. URL : <http://www.tandfonline.com/doi/abs/10.1080/02698590701305792>.
- [194] Marta Maleszewska and Bozena Kaminska. “Deregulation of histone-modifying enzymes and chromatin structure modifiers contributes to glioma development.” In : *Future Oncol.* (2015), pp. 1–15. ISSN : 1744-8301. DOI : 10.2217/fon.15.171. URL : <http://www.ncbi.nlm.nih.gov/pubmed/26289459>.
- [195] Jessica C Mar et al. “attract : A method for identifying core pathways that define cellular phenotypes.” In : *PLoS One* 6.10 (2011), e25445. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0025445. URL : <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0025445> <http://www.ncbi.nlm.nih.gov/pubmed/22022396> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3194807>.
- [196] O Markovic and Nenad Markovic. “Cell cross-contamination in cell cultures : the silent and neglected danger.” In : *In Vitro Cell. Dev. Biol. Anim.* 34.1 (1998), pp. 1–8. ISSN : 1071-2690. DOI : 10.1007/s11626-998-0040-y.
- [197] L. Martinez-Lostao, Alberto Anel, and Julián Pardo. “How Do Cytotoxic Lymphocytes Kill Cancer Cells?” In : *Clin. Cancer Res.* 21.22 (Nov. 2015), pp. 5047–5056. ISSN : 1078-0432. DOI : 10.1158/1078-0432.CCR-15-0685. URL : <http://clincancerres.aacrjournals.org/cgi/doi/10.1158/1078-0432.CCR-15-0685> <http://www.ncbi.nlm.nih.gov/pubmed/26567364>.
- [198] Inigo Martincorena and Peter J Campbell. “Somatic mutation in cancer and normal cells”. In : *Science (80-.)*. 349.6255 (2015), pp. 1483–1489. ISSN : 0036-8075. DOI : 10.1126/science.aab4082.
- [199] Andriy Marusyk and Kornelia Polyak. “Tumor heterogeneity : causes and consequences.” In : *Biochim. Biophys. Acta* 1805.1 (Jan. 2010), pp. 105–17. ISSN : 0006-3002. DOI : 10.1016/j.bbcan.2009.11.002. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2814927%7B%5C%26%7Dtool=pmcentrez%7B%5C%26%7Drendertype=abstract> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3248630%7B%5C%26%7Dtool=pmcentrez%7B%5C%26%7Drendertype=abstract>.
- [200] Andriy Marusyk et al. “Non-cell-autonomous driving of tumour growth supports subclonal heterogeneity”. In : *Nature* 514.7520 (July 2014), pp. 54–8. ISSN : 0028-0836. DOI : 10.1038/nature13556. URL : <http://www.nature.com/doi/10.1038/nature13556> <http://www.ncbi.nlm.nih.gov/pubmed/25079331>.
- [201] Guilhem Mascré et al. “Distinct contribution of stem and progenitor cells to epidermal maintenance.” In : *Nature* 489.7415 (Sept. 2012), pp. 257–62. ISSN : 1476-4687. DOI : 10.1038/nature11393. URL : <http://www.ncbi.nlm.nih.gov/pubmed/22940863>.
- [202] J R Masters et al. “Short tandem repeat profiling provides an international reference standard for human cell lines.” In : *Proc. Natl. Acad. Sci. U. S. A.* 98.14 (July 2001), pp. 8012–7. ISSN : 0027-8424. DOI : 10.1073/pnas.121616198. URL : <http://www.ncbi.nlm.nih.gov/pubmed/11416159> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC35459>.

- [203] Jennie Powell Mather. "In vitro models." In : *Stem Cells* 30.2 (Feb. 2012), pp. 95–9. ISSN : 1549-4918. DOI : 10.1002/stem.774. URL : <http://www.ncbi.nlm.nih.gov/pubmed/22076915>.
- [204] Takaaki Matsui, Hiroshi Ishikawa, and Yasumasa Bessho. "Cell collectivity regulation within migrating cell cluster during Kupffer's vesicle formation in zebrafish." In : *Front. cell Dev. Biol.* 3.May (2015), p. 27. ISSN : 2296-634X. DOI : 10.3389/fcell.2015.00027. URL : http://www.frontiersin.org/Cell%7B%5C_%7DAdhesion%7B%5C_%7Dand%7B%5C_%7DMigration/10.3389/fcell.2015.00027/abstract%20http://www.ncbi.nlm.nih.gov/pubmed/26000276%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4423447.
- [205] W J McAleer et al. "Human hepatitis B vaccine from recombinant yeast. 1984." In : *Biotechnology* 24 (1992), pp. 500–2. ISSN : 0740-7378. URL : <http://www.ncbi.nlm.nih.gov/pubmed/1422061>.
- [206] M W McBurney et al. "Murine PGK-1 promoter drives widespread but not uniform expression in transgenic mice." In : *Dev. Dyn.* 200.4 (Feb. 1994), pp. 278–293. ISSN : 1058-8388. DOI : 10.1002/aja.1002000403.
- [207] Joby L McKenzie et al. "Individual stem cells with highly variable proliferation and self-renewal properties comprise the human hematopoietic stem cell compartment". In : *Nat. Immunol.* 7.11 (Nov. 2006), pp. 1225–1233. ISSN : 1529-2908. DOI : 10.1038/ni1393. URL : <http://www.nature.com/doifinder/10.1038/ni1393>.
- [208] Corbin E Meacham and Sean J Morrison. "Tumour heterogeneity and cancer cell plasticity." In : *Nature* 501.7467 (Sept. 2013), pp. 328–37. ISSN : 1476-4687. DOI : 10.1038/nature12624. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24048065>.
- [209] R H Medema and J L Bos. "The role of p21ras in receptor tyrosine kinase signaling." In : *Crit. Rev. Oncog.* 4.6 (1993), pp. 615–61. ISSN : 0893-9675. URL : <http://www.ncbi.nlm.nih.gov/pubmed/8286433>.
- [210] Patrick Mehlen and Alain Puisieux. "Metastasis : a question of life or death." In : *Nat. Rev. Cancer* 6.6 (June 2006), pp. 449–58. ISSN : 1474-175X. DOI : 10.1038/nrc1886. URL : <http://www.ncbi.nlm.nih.gov/pubmed/16723991>.
- [211] Miriam Merad and Florent Ginhoux. "Dendritic cell genealogy : a new stem or just another branch?" In : *Nat. Immunol.* 8.11 (2007), pp. 1199–201. ISSN : 1529-2908. DOI : 10.1038/ni1107-1199. URL : <http://www.ncbi.nlm.nih.gov/pubmed/17952047> <http://www.nature.com/ni/journal/v8/n11/pdf/ni1107-1199.pdf%20http://www.ncbi.nlm.nih.gov/pubmed/17952047>.
- [212] R. E. Michod and D. Roze. "Cooperation and conflict in the evolution of multicellularity." In : *Heredity (Edinb)*. 86.Pt 1 (Jan. 2001), pp. 1–7. ISSN : 0018-067X. DOI : 10.1046/j.1365-2540.2001.00808.x. URL : <http://www.ncbi.nlm.nih.gov/pubmed/11298810>.
- [213] Anna von Mikecz. "The nuclear ubiquitin-proteasome system". In : *J. Cell Sci.* 119.10 (May 2006), pp. 1977–1984. ISSN : 0021-9533. DOI : 10.1242/jcs.03008. URL : <http://jcs.biologists.org/cgi/doi/10.1242/jcs.03008>.

- [214] Maria Mittelbrunn and Francisco Sánchez-Madrid. “Intercellular communication : diverse structures for exchange of genetic information.” In : *Nat. Rev. Mol. Cell Biol.* 13.5 (May 2012), pp. 328–35. ISSN : 1471-0080. DOI : 10.1038/nrm3335. URL : <http://dx.doi.org/10.1038/nrm3335><http://www.ncbi.nlm.nih.gov/pubmed/22510790><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3738855>.
- [215] Cruz Morenilla-Palao et al. “Ion channel profile of TRPM8 cold receptors reveals a role of TASK-3 potassium channels in thermosensation.” In : *Cell Rep.* 8.5 (Sept. 2014), pp. 1571–82. ISSN : 2211-1247. DOI : 10.1016/j.celrep.2014.08.003. URL : <http://www.ncbi.nlm.nih.gov/pubmed/25199828>.
- [216] Sean J Morrison and Judith Kimble. “Asymmetric and symmetric stem-cell divisions in development and cancer.” In : *Nature* 441.7097 (June 2006), pp. 1068–1074. ISSN : 0028-0836. DOI : 10.1038/nature04956. URL : <http://www.nature.com/doifinder/10.1038/nature04956>.
- [217] Joseph H. Nadeau. “Transgenerational genetic effects on phenotypic variation and disease risk”. In : *Hum. Mol. Genet.* 18.R2 (2009), pp. 202–210. ISSN : 09646906. DOI : 10.1093/hmg/ddp366. URL : <http://www.hmg.oxfordjournals.org/cgi/doi/10.1093/hmg/ddp366>.
- [218] Sunil Nagpal et al. “Promoter context- and response element-dependent specificity of the transcriptional activation and modulating functions of retinoic acid receptors”. In : *Cell* 70.6 (Sept. 1992), pp. 1007–1019. ISSN : 00928674. DOI : 10.1016/0092-8674(92)90250-G. URL : <http://linkinghub.elsevier.com/retrieve/pii/S009286749290250G>.
- [219] L Naldini et al. “Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector.” In : *Proc. Natl. Acad. Sci. U. S. A.* 93.21 (Oct. 1996), pp. 11382–8. ISSN : 0027-8424. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=38066%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [220] L Naldini et al. “In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector.” In : *Science* 272.5259 (Apr. 1996), pp. 263–7. ISSN : 0036-8075. URL : <http://www.ncbi.nlm.nih.gov/pubmed/8602510>.
- [221] Simona Negrini, Vassilis G Gorgoulis, and Thanos D Halazonetis. “Genomic instability—an evolving hallmark of cancer.” In : *Nat. Rev. Mol. Cell Biol.* 11.3 (Mar. 2010), pp. 220–8. ISSN : 1471-0080. DOI : 10.1038/nrm2858. URL : <http://www.ncbi.nlm.nih.gov/pubmed/20177397><http://www.nature.com/doifinder/10.1038/nrm2858>.
- [222] D M Nettelbeck, V Jérôme, and R Müller. “A strategy for enhancing the transcriptional activity of weak cell type-specific promoters.” In : *Gene Ther.* 5.12 (Dec. 1998), pp. 1656–64. ISSN : 0969-7128. DOI : 10.1038/sj.gt.3300778. URL : <http://www.ncbi.nlm.nih.gov/pubmed/10023445>.
- [223] Ralph a Neumüller and Juergen a Knoblich. “Dividing cellular asymmetry : asymmetric cell division and its implications for stem cells and cancer.” In : *Genes Dev.* 23.23 (2009), pp. 2675–99. ISSN : 1549-5477. DOI : 10.1101/gad.1850809. URL : <http://genesdev.cshlp.org/content/23/23/2675.full><http://www.ncbi.nlm.nih.gov/pubmed/19952104><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2788323>.

- [235] Catherine A. O'Brien et al. "A human colon cancer cell capable of initiating tumour growth in immunodeficient mice". In : *Nature* 445.7123 (Jan. 2007), pp. 106–110. ISSN : 0028-0836. DOI : 10.1038/nature05372. URL : <http://www.nature.com/doi/10.1038/nature05372>.
- [236] Hideyuki Okano et al. "Steps Toward Safe Cell Therapy Using Induced Pluripotent Stem Cells". In : *Circ. Res.* 112.3 (Feb. 2013), pp. 523–533. ISSN : 0009-7330. DOI : 10.1161/CIRCRESAHA.111.256149. URL : <http://circres.ahajournals.org/cgi/doi/10.1161/CIRCRESAHA.111.256149>.
- [237] Breast Cancer Organisation. *Treatments breast cancer*. URL : <http://www.breastcancer.org/treatment>.
- [238] N Panza et al. "Cancer antigen 125, tissue polypeptide antigen, carcinoembryonic antigen, and beta-chain human chorionic gonadotropin as serum markers of epithelial ovarian carcinoma." In : *Cancer* 61.1 (Jan. 1988), pp. 76–83. ISSN : 0008-543X. URL : <http://ovidsp.ovid.com/ovidweb.cgi?T=JS%7B%5C%7DPAGE=reference%7B%5C%7DD=emed88%7B%5C%7DNEWS=N%7B%5C%7DAN=1988023113%5Cbackslash%7Dhttp://ovidsp.ovid.com/ovidweb.cgi?T=JS%7B%5C%7DPAGE=reference%7B%5C%7DD=med3%7B%5C%7DNEWS=N%7B%5C%7DAN=2446734%20http://www.ncbi.nlm.nih.gov/pubmed/2446734>.
- [239] Sung-Suk Park et al. "Esculetin inhibits cell proliferation through the Ras/ERK1/2 pathway in human colon cancer cells." In : *Oncol. Rep.* 25 (2011), pp. 223–230. ISSN : 1791-2431. DOI : 10.3892/or.
- [240] Flavius C. Pascut et al. "Non-invasive label-free monitoring the cardiac differentiation of human embryonic stem cells in-vitro by Raman spectroscopy." In : *Biochim. Biophys. Acta* 1830.6 (2013), pp. 3517–24. ISSN : 0006-3002. DOI : 10.1016/j.bbagen.2013.01.030. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0304416513000433%20http://www.ncbi.nlm.nih.gov/pubmed/23403134>.
- [241] Erika Pastrana, Violeta Silva-Vargas, and Fiona Doetsch. "Eyes wide open : a critical review of sphere-formation as an assay for stem cells." In : *Cell Stem Cell* 8.5 (2011), pp. 486–98. ISSN : 1875-9777. DOI : 10.1016/j.stem.2011.04.007. URL : <http://linkinghub.elsevier.com/retrieve/pii/S193459091100172X%20http://www.ncbi.nlm.nih.gov/pubmed/21549325%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3633588>.
- [242] Samuel B Pattle and Paul J Farrell. "The role of Epstein-Barr virus in cancer." In : *Expert Opin. Biol. Ther.* 6.11 (Nov. 2006), pp. 1193–205. ISSN : 1744-7682. DOI : 10.1517/14712598.6.11.1193. URL : <http://www.ncbi.nlm.nih.gov/pubmed/17049016>.
- [243] David J Pearton, Corinne Ferraris, and Danielle Dhouailly. "Transdifferentiation of corneal epithelium : evidence for a linkage between the segregation of epidermal stem cells and the induction of hair follicles during embryogenesis." In : *Int. J. Dev. Biol.* 48.2-3 (2004), pp. 197–201. ISSN : 0214-6282. DOI : 10.1387/ijdb.15272385. URL : <http://www.intjdevbiol.com/paper.php?doi=15272385>.
- [244] C M Perou et al. "Molecular portraits of human breast tumours." In : *Nature* 406.6797 (2000), pp. 747–752. ISSN : 0028-0836. DOI : 10.1038/35021093.
- [245] Kristian Pietras and Arne Ostman. "Hallmarks of cancer : interactions with the tumor stroma." In : *Exp. Cell Res.* 316.8 (May 2010), pp. 1324–31. ISSN : 1090-2422. DOI : 10.1016/j.yexcr.2010.02.045. URL : <http://www.ncbi.nlm.nih.gov/pubmed/20211171>.

- [246] Dario Ponti et al. "Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties". In : *Cancer Res.* 65.13 (2005), pp. 5506–5511. ISSN : 00085472. DOI : 10.1158/0008-5472.CAN-05-0626.
- [247] William B Porterfield and Jennifer a Prescher. "Tools for visualizing cell-cell 'interactomes'." In : *Curr. Opin. Chem. Biol.* 24 (Feb. 2015), pp. 121–30. ISSN : 1879-0402. DOI : 10.1016/j.cbpa.2014.11.006. URL : <http://www.ncbi.nlm.nih.gov/pubmed/25461730><http://linkinghub.elsevier.com/retrieve/pii/S1367593114001707>.
- [248] Saumyadipta Pyne et al. "Automated high-dimensional flow cytometric data analysis." In : *Proc. Natl. Acad. Sci. U. S. A.* 106.21 (2009), pp. 8519–8524. ISSN : 0027-8424. DOI : 10.1073/pnas.0903028106.
- [249] Jane Yuxia Qin et al. "Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter". In : *PLoS One* 5.5 (May 2010). Ed. by Immo A. Hansen, e10611. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0010611. URL : <http://dx.plos.org/10.1371/journal.pone.0010611>.
- [250] Stefanie Raab et al. "A Comparative View on Human Somatic Cell Sources for iPSC Generation." In : *Stem Cells Int.* 2014 (2014), p. 768391. ISSN : 1687-966X. DOI : 10.1155/2014/768391. URL : <http://www.hindawi.com/journals/sci/2014/768391/><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4241335%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [251] S Rabizadeh et al. "Induction of apoptosis by the low-affinity NGF receptor." In : *Science* 261.5119 (1993), pp. 345–348. ISSN : 0036-8075. DOI : 10.1126/science.8332899.
- [252] Fa Ran et al. "Genome engineering using the CRISPR-Cas9 system". In : *Nat. Protoc.* 8.11 (2013), pp. 2281–308. ISSN : 1750-2799. DOI : 10.1038/nprot.2013.143. arXiv : NIHMS150003. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24157548><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3969860%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract><http://www.ncbi.nlm.nih.gov/pubmed/24157548><http://www.nature.com/nprot/journal/v8/n11/abs/nprot.2013.143.h>.
- [253] Yael Raz and Neta Erez. "An inflammatory vicious cycle : Fibroblasts and immune cell recruitment in cancer." In : *Exp. Cell Res.* 319.11 (July 2013), pp. 1596–603. ISSN : 1090-2422. DOI : 10.1016/j.yexcr.2013.03.022. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0014482713001304><http://www.ncbi.nlm.nih.gov/pubmed/23567181>.
- [254] Sandra Rebouissou et al. "EGFR as a potential therapeutic target for a subset of muscle-invasive bladder cancers presenting a basal-like phenotype." In : *Sci. Transl. Med.* 6.244 (July 2014), 244ra91. ISSN : 1946-6242. DOI : 10.1126/scitranslmed.3008970. URL : <http://www.ncbi.nlm.nih.gov/pubmed/25009231>.
- [255] C. R a Regenbrecht, H. Lehrach, and J. Adjaye. "Stemming Cancer : Functional Genomics of Cancer Stem Cells in Solid Tumors". In : *Stem Cell Rev.* 4.4 (Dec. 2008), pp. 319–328. ISSN : 1550-8943. DOI : 10.1007/s12015-008-9034-0. URL : <http://link.springer.com/10.1007/s12015-008-9034-0>.

- [256] Yvonne Reid et al. *Authentication of Human Cell Lines by STR DNA Profiling Analysis*. 2004. URL : [\\$%5Cbackslash\\$http://www.ncbi.nlm.nih.gov/pubmed/23805434](http://www.ncbi.nlm.nih.gov/books/NBK144066/$%5Cbackslash$http://www.ncbi.nlm.nih.gov/books/NBK144066/?report=printable) %20http://www.ncbi.nlm.nih.gov/pubmed/23805434.
- [257] T Reya et al. "Stem cells, cancer, and cancer stem cells." In : *Nature* 414.6859 (2001), pp. 105–11. ISSN : 0028-0836. DOI : 10.1038/35102167. URL : http://dx.doi.org/10.1007/978-1-60327-933-8%7B%5C_%7D15%20http://www.ncbi.nlm.nih.gov/pubmed/11689955.
- [258] Muhammad Riaz et al. "miRNA expression profiling of 51 human breast cancer cell lines reveals subtype and driver mutation-specific miRNAs." In : *Breast Cancer Res.* 15 (2013), R33. ISSN : 1465-542X. DOI : 10.1186/bcr3415. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3672661%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [259] Daniela Riccardi and Edward M Brown. "Physiology and pathophysiology of the calcium-sensing receptor in the kidney." In : *Am. J. Physiol. Renal Physiol.* 298.3 (Mar. 2010), F485–99. ISSN : 1522-1466. DOI : 10.1152/ajprenal.00608.2009. URL : <http://www.ncbi.nlm.nih.gov/pubmed/19923405%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2838589%20http://ajprenal.physiology.org/cgi/doi/10.1152/ajprenal.00608.2009>.
- [260] Jai Prakash Richard et al. "Direct in vivo cellular reprogramming involves transition through discrete, non-pluripotent steps". In : *J. Cell Sci.* 124.8 (Apr. 2011), e1–e1. ISSN : 0021-9533. DOI : 10.1242/jcs.090043. URL : <http://jcs.biologists.org/cgi/doi/10.1242/jcs.090043>.
- [261] Laila Ritsma et al. "Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging." In : *Nature* 507.7492 (Mar. 2014), pp. 362–5. ISSN : 1476-4687. DOI : 10.1038/nature12972. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24531760%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3964820>.
- [262] Nelly Robillard, Catherine Pellat-Deceunynck, and Régis Bataille. "Phenotypic characterization of the human myeloma cell growth fraction." In : *Blood* 105.12 (June 2005), pp. 4845–8. ISSN : 0006-4971. DOI : 10.1182/blood-2004-12-4700. URL : <http://www.bloodjournal.org/cgi/doi/10.1182/blood-2004-12-4700%20http://www.ncbi.nlm.nih.gov/pubmed/15741217>.
- [263] Alex Rosenberg. "Defending information-free genocentrism." In : *Hist. Philos. Life Sci.* 27.3-4 (2005), 345–59, discussion 361–4. ISSN : 0391-9714. URL : <http://www.ncbi.nlm.nih.gov/pubmed/16898205>.
- [264] Harry Rubin. "The suppression of morphological alterations in cells infected with Rous sarcoma virus." In : *Virology* 12 (1960), pp. 14–31. ISSN : 0042-6822. URL : <http://www.ncbi.nlm.nih.gov/pubmed/13744356>.
- [265] Harry Rubin and C Colby. "Early release of growth inhibition in cells infected with Rous sarcoma virus." In : *Proc. Natl. Acad. Sci. U. S. A.* 60.2 (June 1968), pp. 482–8. ISSN : 0027-8424. URL : <http://www.ncbi.nlm.nih.gov/pubmed/4302638%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC225073>.

- [266] R.J Ruch et al. “Defective gap junctional intercellular communication in lung cancer : loss of an important mediator of tissue homeostasis and phenotypic regulation.” In : *Exp. Lung Res.* 27.3 (), pp. 231–43. ISSN : 0190-2148. URL : <http://www.ncbi.nlm.nih.gov/pubmed/11293326>.
- [267] Adrian P. Rybak et al. “Characterization of sphere-propagating cells with stem-like properties from DU145 prostate cancer cells”. In : *Biochim. Biophys. Acta - Mol. Cell Res.* 1813.5 (May 2011), pp. 683–694. ISSN : 01674889. DOI : 10.1016/j.bbamcr.2011.01.018. URL : <http://dx.doi.org/10.1016/j.bbamcr.2011.01.018> <http://linkinghub.elsevier.com/retrieve/pii/S0167488911000267>.
- [268] Asako Sakaue-Sawano et al. “Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression”. In : *Cell* 132.3 (Feb. 2008), pp. 487–498. ISSN : 00928674. DOI : 10.1016/j.cell.2007.12.033. URL : <http://www.ncbi.nlm.nih.gov/pubmed/18267078>.
- [269] Catherine Sanchez et al. “Grasping at molecular interactions and genetic networks in Drosophila melanogaster using FlyNets, an Internet database.” In : *Nucleic Acids Res.* 27.1 (Jan. 1999), pp. 89–94. ISSN : 0305-1048. DOI : 10.1093/nar/27.1.89. URL : <http://www.ncbi.nlm.nih.gov/pubmed/9847149> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC148104>.
- [270] Gilda da Cunha Santos et al. “Multiplex sequencing for EZH2, CD79B, and MYD88 mutations using archival cytospin preparations from B-cell non-Hodgkin lymphoma aspirates previously tested for MYC rearrangement and IGH/BCL2 translocation.” In : *Cancer Cytopathol.* 123.7 (July 2015), pp. 413–20. ISSN : 1934-6638. DOI : 10.1002/cncy.21535. URL : <http://doi.wiley.com/10.1002/cncy.21535> <http://www.ncbi.nlm.nih.gov/pubmed/25807917>.
- [271] Tobias Schatton, Natasha Y Frank, and Markus H Frank. “Identification and targeting of cancer stem cells.” In : *Bioessays* 31.10 (Oct. 2009), pp. 1038–49. ISSN : 1521-1878. DOI : 10.1002/bies.200900058. URL : <http://www.ncbi.nlm.nih.gov/pubmed/19708024> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2887758>.
- [272] Michael R Schlabach et al. “Synthetic design of strong promoters.” In : *Proc. Natl. Acad. Sci. U. S. A.* 107.6 (2010), pp. 2538–2543. ISSN : 0027-8424. DOI : 10.1073/pnas.0914803107.
- [273] Felix Scholkmann, Daniel Fels, and Michal Cifra. “Non-chemical and non-contact cell-to-cell communication : a short review”. In : *Am J Transl Res* 5.6 (2013), pp. 586–593. ISSN : 1943-8141. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3786266%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20www.ajtr.org>.
- [274] Jun Seita and Irving L. Weissman. “Hematopoietic stem cell : self-renewal versus differentiation”. In : *Wiley Interdiscip. Rev. Syst. Biol. Med.* 2.6 (Nov. 2010), pp. 640–653. ISSN : 19395094. DOI : 10.1002/wsbm.86. URL : <http://doi.wiley.com/10.1002/wsbm.86>.
- [275] Mark a Sellmyer et al. “Visualizing cellular interactions with a generalized proximity reporter.” In : *Proc. Natl. Acad. Sci. U. S. A.* 110.21 (May 2013), pp. 8567–72. ISSN : 1091-6490. DOI : 10.1073/pnas.1218336110. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3666728%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://www.ncbi.nlm.nih.gov/pubmed/>

- 23650381%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3666728.
- [276] Tracy Seymour, Anna Nowak, and Foteini Kakulas. "Targeting Aggressive Cancer Stem Cells in Glioblastoma." In : *Front. Oncol.* 5, July (July 2015), p. 159. ISSN : 2234-943X. DOI : 10.3389/fonc.2015.00159. URL : <http://journal.frontiersin.org/Article/10.3389/fonc.2015.00159/abstract%20http://www.ncbi.nlm.nih.gov/pubmed/26258069%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4507454>.
- [277] Mark Shackleton et al. "Generation of a functional mammary gland from a single stem cell." In : *Nature* 439.7072 (2006), pp. 84–88. ISSN : 0028-0836. DOI : 10.1038/nature04372.
- [278] Masood A Shammas. "Telomeres, lifestyle, cancer, and aging." In : *Curr. Opin. Clin. Nutr. Metab. Care* 14.1 (Jan. 2011), pp. 28–34. ISSN : 1473-6519. DOI : 10.1097/MCO.0b013e32834121b1. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3370421%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://www.ncbi.nlm.nih.gov/pubmed/21102320%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3370421>.
- [279] Benjamin Sharpe et al. "Searching for prostate cancer stem cells : markers and methods." In : *Stem Cell Rev.* 9.5 (Oct. 2013), pp. 721–30. ISSN : 1558-6804. DOI : 10.1007/s12015-013-9453-4. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23775699>.
- [280] J W Shay and S Bacchetti. "A survey of telomerase activity in human cancer." In : *Eur. J. Cancer* 33.5 (Apr. 1997), pp. 787–91. ISSN : 0959-8049. DOI : 10.1016/S0959-8049(97)00062-2. URL : <http://www.ncbi.nlm.nih.gov/pubmed/9282118>.
- [281] Nathan M Sherer and Walther Mothes. "Cytosomes and tunneling nanotubules in cell-cell communication and viral pathogenesis." In : *Trends Cell Biol.* 18.9 (Sept. 2008), pp. 414–20. ISSN : 1879-3088. DOI : 10.1016/j.tcb.2008.07.003. URL : <http://www.ncbi.nlm.nih.gov/pubmed/18703335%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2628975>.
- [282] Carol Sheridan et al. "CD44+/CD24- breast cancer cells exhibit enhanced invasive properties : an early step necessary for metastasis." In : *Breast Cancer Res.* 8.5 (2006), R59. ISSN : 1465-542X. DOI : 10.1186/bcr1610. URL : <http://breast-cancer-research.com/content/8/5/R59>.
- [283] K. Shi et al. "Protease-activated receptor-2 induces migration of pancreatic cancer cells in an extracellular ATP-dependent manner." In : *J. Thromb. Haemost.* 11.10 (Oct. 2013), pp. 1892–902. ISSN : 1538-7836. DOI : 10.1111/jth.12361. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23899344>.
- [284] Quanming Shi et al. "Rapid disorganization of mechanically interacting systems of mammary acini". In : *Proc. Natl. Acad. Sci.* 111.2 (Jan. 2014), pp. 658–663. ISSN : 0027-8424. DOI : 10.1073/pnas.1311312110. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3896145%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract%20http://www.pnas.org/cgi/doi/10.1073/pnas.1311312110>.
- [285] Michail Shipitsin et al. "Molecular Definition of Breast Tumor Heterogeneity". In : *Cancer Cell* 11.3 (Mar. 2007), pp. 259–273. ISSN : 15356108. DOI : 10.1016/j.ccr.2007.01.013. URL : <http://linkinghub.elsevier.com/retrieve/pii/S1535610807000293>.

- [286] Sheila K Singh et al. "Identification of human brain tumour initiating cells". In : *Nature* 432.7015 (Nov. 2004), pp. 396–401. ISSN : 0028-0836. DOI : 10.1038/nature03128. URL : <http://www.nature.com/doi/10.1038/nature03128>.
- [287] Sunita Sinha and Rosemary J. Redfield. "Natural DNA uptake by *Escherichia coli*." In : *PLoS One* 7.4 (2012), e35620. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0035620. URL : <http://dx.plos.org/10.1371/journal.pone.0035620><http://www.ncbi.nlm.nih.gov/pubmed/22532864><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3330819>.
- [288] Incilay Sinici et al. "Comparison of HCMV IE and EF-1 Promoters for the Stable Expression of β -Subunit of Hexosaminidase in CHO Cell Lines". In : *Biochem. Genet.* 44.3-4 (Sept. 2006), pp. 168–175. ISSN : 0006-2928. DOI : 10.1007/s10528-006-9016-3. URL : <http://link.springer.com/10.1007/s10528-006-9016-3>.
- [289] Dubravka Škalamera et al. "Generation of a Genome Scale Lentiviral Vector Library for EF1 α Promoter-Driven Expression of Human ORFs and Identification of Human Genes Affecting Viral Titer". In : *PLoS One* 7.12 (Dec. 2012). Ed. by Ranjit Ray, e51733. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0051733. URL : <http://dx.plos.org/10.1371/journal.pone.0051733>.
- [290] M Skobe and N E Fusenig. "Tumorigenic conversion of immortal human keratinocytes through stromal cell activation." In : *Proc. Natl. Acad. Sci. U. S. A.* 95.3 (Feb. 1998), pp. 1050–5. ISSN : 0027-8424. DOI : 10.1073/pnas.95.3.1050. URL : <http://www.sciencedirect.com/science/article/pii/S0898656815001114><http://linkinghub.elsevier.com/retrieve/pii/S0898656815001114><http://www.ncbi.nlm.nih.gov/pubmed/9448283><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC18668>.
- [291] Chanel E. Smart et al. "In Vitro Analysis of Breast Cancer Cell Line Tumourspheres and Primary Human Breast Epithelia Mammospheres Demonstrates Inter- and Intrasphere Heterogeneity". In : *PLoS One* 8.6 (2013). ISSN : 19326203. DOI : 10.1371/journal.pone.0064388.
- [292] Shannon M. Smith and Li Cai. "Cell Specific CD44 Expression in Breast Cancer Requires the Interaction of AP-1 and NF κ B with a Novel cis-Element". In : *PLoS One* 7.11 (2012), pp. 1–12. ISSN : 19326203. DOI : 10.1371/journal.pone.0050867.
- [293] Berend Snijder and Lucas Pelkmans. "Origins of regulated cell-to-cell variability." In : *Nat. Rev. Mol. Cell Biol.* 12.2 (Feb. 2011), pp. 119–25. ISSN : 1471-0080. DOI : 10.1038/nrm3044. URL : <http://www.nature.com/doi/10.1038/nrm3044><http://www.ncbi.nlm.nih.gov/pubmed/21224886>.
- [294] Eric L Snyder et al. "Identification of CD44v6(+)/CD24- breast carcinoma cells in primary human tumors by quantum dot-conjugated antibodies." In : *Lab. Invest.* 89.8 (2009), pp. 857–66. ISSN : 1530-0307. DOI : 10.1038/labinvest.2009.54. URL : <http://www.nature.com/doi/10.1038/labinvest.2009.54><http://www.ncbi.nlm.nih.gov/pubmed/19488035>.
- [295] Carlos Sonnenschein, Barbara Davis, and Ana M Soto. "A novel pathogenic classification of cancers". In : *Cancer Cell Int.* 14.1 (2014), p. 113. ISSN : 1475-2867. DOI : 10.1186/s12935-014-0113-9. URL : <http://www.cancerci.com/content/14/1/113>.

- [296] Carlos Sonnenschein and Ana M Soto. “Le cancer et ses gènes insaisissables”. In : *médecine/sciences* 30.6-7 (June 2014), pp. 688–692. ISSN : 0767-0974. DOI : 10.1051/medsci/20143006022. URL : www.ncbi.nlm.nih.gov/pubmed/25014463<http://www.medecinesciences.org/articles/medsci/abs/2014/07/medsci2014306-7p688/medsci2014306-7p688.html><http://www.medecinesciences.org/10.1051/medsci/20143006022>.
- [297] T Sørli et al. “Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.” In : *Proc. Natl. Acad. Sci. U. S. A.* 98.19 (Sept. 2001), pp. 10869–74. ISSN : 0027-8424. DOI : 10.1073/pnas.191367098. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=58566%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [298] Jeffrey L Spees et al. “Mitochondrial transfer between cells can rescue aerobic respiration.” In : *Proc. Natl. Acad. Sci. U. S. A.* 103.5 (Jan. 2006), pp. 1283–8. ISSN : 0027-8424. DOI : 10.1073/pnas.0510511103. URL : <http://www.pnas.org/content/103/5/1283.full.pdf><http://www.ncbi.nlm.nih.gov/pubmed/16432190><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1345715>.
- [299] H Steller. “Mechanisms and genes of cellular suicide.” In : *Science* 267.5203 (Mar. 1995), pp. 1445–9. ISSN : 0036-8075. DOI : 10.1126/science.7878463. URL : <http://www.ncbi.nlm.nih.gov/pubmed/7878463><http://www.sciencemag.org/cgi/doi/10.1126/science.7878463>.
- [300] John E. Stewart. “The direction of evolution : The rise of cooperative organization”. In : *BioSystems* 123 (2014), pp. 27–36. ISSN : 18728324. DOI : 10.1016/j.biosystems.2014.05.006. URL : <http://dx.doi.org/10.1016/j.biosystems.2014.05.006>.
- [301] John Stingl et al. “Purification and unique properties of mammary epithelial stem cells.” In : *Nature* 439.7079 (2006), pp. 993–997. ISSN : 0028-0836. DOI : 10.1038/nature04496.
- [302] D L Stoler et al. “The onset and extent of genomic instability in sporadic colorectal tumor progression.” In : *Proc. Natl. Acad. Sci. U. S. A.* 96.26 (Dec. 1999), pp. 15121–6. ISSN : 0027-8424. DOI : 10.1073/pnas.96.26.15121. URL : [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve%7B%5C%7Ddb=PubMed%7B%5C%7Ddopt=Citation%7B%5C%7Dlist%7B%5C%7Duids=10611348%5Cbackslash\\$n](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve%7B%5C%7Ddb=PubMed%7B%5C%7Ddopt=Citation%7B%5C%7Dlist%7B%5C%7Duids=10611348%5Cbackslash$n)<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC24783/pdf/pq015121.pdf><http://www.ncbi.nlm.nih.gov/pubmed/10611348><http://www.pubmedcentral.nih.gov/articler>.
- [303] Karola Stotz. “Molecular epigenesis : distributed specificity as a break in the central dogma.” In : *Hist. Philos. Life Sci.* 28.4 (2006), pp. 533–48. ISSN : 0391-9714. URL : <http://www.ncbi.nlm.nih.gov/pubmed/18351051>.
- [304] K. Streckfuss-Bomeke et al. “Comparative study of human-induced pluripotent stem cells derived from bone marrow cells, hair keratinocytes, and skin fibroblasts”. In : *Eur. Heart J.* 34.33 (Sept. 2013), pp. 2618–2629. ISSN : 0195-668X. DOI : 10.1093/eurheartj/ehs203. URL : <http://eurheartj.oxfordjournals.org/cgi/doi/10.1093/eurheartj/ehs203>.
- [305] Gürol M Süel et al. “Tunability and noise dependence in differentiation dynamics.” In : *Science* 315.5819 (Mar. 2007), pp. 1716–9. ISSN : 1095-9203. DOI : 10.1126/science.1137455. URL : <http://www.ncbi.nlm.nih.gov/pubmed/17379809>.

- [306] Kaoru Sugimura and Shuji Ishihara. "The mechanical anisotropy in a tissue promotes ordering in hexagonal cell packing." In : *Development* 140.19 (Oct. 2013), pp. 4091–101. ISSN : 1477-9129. DOI : 10.1242/dev.094060. URL : <http://dev.biologists.org/cgi/doi/10.1242/dev.094060><http://www.ncbi.nlm.nih.gov/pubmed/24046322>.
- [307] Jai-Yoon Sul et al. "Transcriptome transfer produces a predictable cellular phenotype." In : *Proc. Natl. Acad. Sci. U. S. A.* 106.18 (2009), pp. 7624–9. ISSN : 1091-6490. DOI : 10.1073/pnas.0902161106. URL : <http://www.ncbi.nlm.nih.gov/pubmed/19380745>.
- [308] Lee J Sweetlove and Alisdair R Fernie. "The spatial organization of metabolism within the plant cell." In : *Annu. Rev. Plant Biol.* 64 (2013), pp. 723–46. ISSN : 1545-2123. DOI : 10.1146/annurev-arplant-050312-120233. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23330793>.
- [309] Silke S Talsma et al. "Development and in vitro validation of a targeted delivery vehicle for DNA vaccines." In : *J. Control. Release* 112.2 (May 2006), pp. 271–9. ISSN : 0168-3659. DOI : 10.1016/j.jconrel.2006.02.008. URL : <http://www.ncbi.nlm.nih.gov/pubmed/16549219>.
- [310] Brenton Thomas Tan et al. "The cancer stem cell hypothesis : a work in progress." In : *Lab. Invest.* 86.12 (Dec. 2006), pp. 1203–7. ISSN : 0023-6837. DOI : 10.1038/labinvest.3700488. URL : <http://www.nature.com/doi/10.1038/labinvest.3700488><http://www.ncbi.nlm.nih.gov/pubmed/17075578>.
- [311] Dean G Tang. "Understanding cancer stem cell heterogeneity and plasticity." In : *Cell Res.* 22.3 (Mar. 2012), pp. 457–72. ISSN : 1748-7838. DOI : 10.1038/cr.2012.13. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3292302%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [312] Shibu Thomas et al. "CD24 is an effector of HIF-1-driven primary tumor growth and metastasis." In : *Cancer Res.* 72.21 (Nov. 2012), pp. 5600–12. ISSN : 1538-7445. DOI : 10.1158/0008-5472.CAN-11-3666. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3488144%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [313] H S Thompson et al. "Phase I safety and antigenicity of TA-GW : a recombinant HPV6 L2E7 vaccine for the treatment of genital warts." In : *Vaccine* 17.1 (Jan. 1999), pp. 40–9. ISSN : 0264-410X. DOI : 10.1016/S0264-410X(98)00146-7. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0264410X98001467><http://www.ncbi.nlm.nih.gov/pubmed/10078606>.
- [314] F Tian et al. "Prognostic value of serum CA 19-9 levels in pancreatic adenocarcinoma." In : *Ann. Surg.* 215.4 (Apr. 1992), pp. 350–5. ISSN : 0003-4932. URL : <http://www.ncbi.nlm.nih.gov/pubmed/1348409><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1242451>.
- [315] Tianhai Tian et al. "The origins of cancer robustness and evolvability." In : *Integr. Biol. (Camb)*. 3.1 (Jan. 2011), pp. 17–30. ISSN : 1757-9708. DOI : 10.1039/c0ib00046a. URL : <http://xlink.rsc.org/?DOI=C0IB00046A><http://www.ncbi.nlm.nih.gov/pubmed/20944865>.
- [316] Ali Torkamani and Nicholas J. Schork. "Prediction of cancer driver mutations in protein kinases". In : *Cancer Res.* 68.6 (2008), pp. 1675–1682. ISSN : 00085472. DOI : 10.1158/0008-5472.CAN-07-5283.

- [317] Anja Torsvik et al. “Spontaneous malignant transformation of human mesenchymal stem cells reflects cross-contamination : Putting the research field on track - Letter”. In : *Cancer Res.* 70.15 (2010), pp. 6393–6396. ISSN : 00085472. DOI : 10.1158/0008-5472.CAN-10-1305.
- [318] Joseph A. Trapani. “The dual adverse effects of TGF-beta secretion on tumor progression.” In : *Cancer Cell* 8.5 (Nov. 2005), pp. 349–50. ISSN : 1535-6108. DOI : 10.1016/j.ccr.2005.10.018. URL : <http://www.ncbi.nlm.nih.gov/pubmed/16286241>.
- [319] James E Trosko et al. “Ignored hallmarks of carcinogenesis : stem cells and cell-cell communication.” In : *Ann. N. Y. Acad. Sci.* 1028 (Dec. 2004), pp. 192–201. ISSN : 0077-8923. DOI : 10.1196/annals.1322.023. URL : <http://www.ncbi.nlm.nih.gov/pubmed/15650245>.
- [320] M K Tuxen, G Sölétormos, and P Dombernowsky. “Serum tumor marker CA 125 for monitoring ovarian cancer during follow-up.” In : *Scand. J. Clin. Lab. Invest.* 62.3 (2002), pp. 177–88. ISSN : 0036-5513. URL : <http://www.ncbi.nlm.nih.gov/pubmed/12088336>.
- [321] Fyodor D Urnov et al. “Genome editing with engineered zinc finger nucleases.” In : *Nat. Rev. Genet.* 11.9 (Sept. 2010), pp. 636–46. ISSN : 1471-0064. DOI : 10.1038/nrg2842. URL : <http://www.ncbi.nlm.nih.gov/pubmed/20717154>.
- [322] Hadi Valadi et al. “Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells.” In : *Nat. Cell Biol.* 9.6 (June 2007), pp. 654–9. ISSN : 1465-7392. DOI : 10.1038/ncb1596. URL : <http://www.nature.com/doifinder/10.1038/ncb1596%20http://www.ncbi.nlm.nih.gov/pubmed/17486113>.
- [323] Alexandra Van Keymeulen et al. “Distinct stem cells contribute to mammary gland development and maintenance.” In : *Nature* 479.7372 (Nov. 2011), pp. 189–93. ISSN : 1476-4687. DOI : 10.1038/nature10573. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21983963>.
- [324] Rajeshwar P Verma and Corwin Hansch. “Chemical Toxicity on HeLa Cells.” In : *Curr. Med. Chem.* 13.4 (2006), pp. 423–48. ISSN : 0929-8673. URL : <http://www.ncbi.nlm.nih.gov/pubmed/16475931>.
- [325] Jane E Visvader and Geoffrey J Lindeman. “Cancer stem cells in solid tumours : accumulating evidence and unresolved questions.” In : *Nat. Rev. Cancer* 8.10 (Oct. 2008), pp. 755–68. ISSN : 1474-1768. DOI : 10.1038/nrc2499. URL : <http://www.ncbi.nlm.nih.gov/pubmed/18784658>.
- [326] Ilya V. Volodyaev, Elena N. Krasilnikova, and Ruslan N. Ivanovsky. “CO2 Mediated Interaction in Yeast Stimulates Budding and Growth on Minimal Media”. In : *PLoS One* 8.4 (Apr. 2013). Ed. by Alix Therese Coste, e62808. ISSN : 19326203. DOI : 10.1371/journal.pone.0062808. URL : <http://dx.plos.org/10.1371/journal.pone.0062808>.
- [327] Joris Vriens, Bernd Nilius, and Thomas Voets. “Peripheral thermosensation in mammals.” In : *Nat. Rev. Neurosci.* 15.9 (Sept. 2014), pp. 573–89. ISSN : 1471-0048. DOI : 10.1038/nrn3784. URL : <http://www.nature.com/doifinder/10.1038/nrn3784%20http://www.ncbi.nlm.nih.gov/pubmed/25053448>.
- [328] Jonas Wallin, Kerstin Johnsson, and Magnus Fontes. “Latent modeling of flow cytometry cell populations”. In : (2015). arXiv : 1502.04058. URL : <http://arxiv.org/pdf/1502.04058.pdf>.

- [329] Guenther Walther et al. "Automatic clustering of flow cytometry data with density-based merging." In : *Adv. Bioinformatics* 2009 (2009), p. 686759. ISSN : 1687-8027. DOI : 10.1155/2009/686759.
- [330] Lihong Wang et al. "A hypermorphic SP1-binding CD24 variant associates with risk and progression of multiple sclerosis." In : *Am. J. Transl. Res.* 4.3 (Jan. 2012), pp. 347–56. ISSN : 1943-8141. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3426393%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [331] Yue J Wang et al. "Sphere-forming assays for assessment of benign and malignant pancreatic stem cells." In : *Methods Mol. Biol.* 980 (2013), pp. 281–90. ISSN : 1940-6029. DOI : 10.1007/978-1-62703-287-2_15. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23359160>.
- [332] O WARBURG. "On the origin of cancer cells." In : *Science* 123.3191 (Feb. 1956), pp. 309–14. ISSN : 0036-8075. URL : <http://www.ncbi.nlm.nih.gov/pubmed/13298683>.
- [333] Jeremy P T Ward. "Oxygen sensors in context." In : *Biochim. Biophys. Acta* 1777.1 (Jan. 2008), pp. 1–14. ISSN : 0006-3002. DOI : 10.1016/j.bbabi.2007.10.010. URL : <http://www.ncbi.nlm.nih.gov/pubmed/18036551>.
- [334] Alexis B. Webb and Andrew C. Oates. "Timing by rhythms : Daily clocks and developmental rulers." In : *Dev. Growth Differ.* (Nov. 2015), n/a–n/a. ISSN : 1440-169X. DOI : 10.1111/dgd.12242. URL : <http://doi.wiley.com/10.1111/dgd.12242><http://www.ncbi.nlm.nih.gov/pubmed/26542934>.
- [335] Clifford Whatcott et al. "Tumor-stromal interactions in pancreatic cancer." In : *Crit. Rev. Oncog.* 18.1-2 (2013), pp. 135–51. ISSN : 0893-9675. DOI : 10.1016/j.pan.2012.11.311. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23237556><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3632415>.
- [336] Teresa E. Williams et al. "Discovery of RAF265 : A Potent mut-B-RAF Inhibitor for the Treatment of Metastatic Melanoma." In : *ACS Med. Chem. Lett.* 6.9 (Sept. 2015), pp. 961–5. ISSN : 1948-5875. DOI : 10.1021/ml500526p. URL : <http://pubs.acs.org/doi/10.1021/ml500526p><http://www.ncbi.nlm.nih.gov/pubmed/26396681><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4569875>.
- [337] Ch Wong and Kl Goh. "Chronic hepatitis B infection and liver cancer." In : *Biomed. Imaging Interv. J.* 2.3 (July 2006), e7. ISSN : 1823-5530. DOI : 10.2349/biij.2.3.e7. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21614253><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3097640>.
- [338] Tzyy Yue Wong. "Molecular mechanism of extrinsic factors affecting anti-aging of stem cells". In : *World J. Stem Cells* 7.2 (Mar. 2015), p. 512. ISSN : 1948-0210. DOI : 10.4252/wjsc.v7.i2.512. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4369508%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract><http://www.wjgnet.com/1948-0210/full/v7/i2/512.htm>.
- [339] H. Arthur Woods and J. Keaton Wilson. "An information hypothesis for the evolution of homeostasis". In : *Trends Ecol. Evol.* 28.5 (2013), pp. 283–289. ISSN : 01695347. DOI : 10.1016/j.tree.2012.10.021. URL : <http://dx.doi.org/10.1016/j.tree.2012.10.021>.
- [340] World Health Organization. *Classification of tumors*. URL : <http://www.iarc.fr/en/publications/pdfs-online/pat-gen/bb4/index.php>.

- [341] G. D. Wu, W. Wang, and P. G. Traber. "Isolation and characterization of the human sucrase-isomaltase gene and demonstration of intestine-specific transcriptional elements". In : *J. Biol. Chem.* 267.11 (1992), pp. 7863–7870. ISSN : 00219258.
- [342] A H Wyllie. "Apoptosis and the regulation of cell numbers in normal and neoplastic tissues : an overview." In : *Cancer Metastasis Rev.* 11.2 (Sept. 1992), pp. 95–103. ISSN : 0167-7659. URL : <http://www.ncbi.nlm.nih.gov/pubmed/1394797>.
- [343] G Xie et al. "Tumour-initiating capacity is independent of epithelial-mesenchymal transition status in breast cancer cell lines." In : *Br. J. Cancer* 110.10 (May 2014), pp. 2514–23. ISSN : 1532-1827. DOI : 10.1038/bjc.2014.153. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24755887><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4021510%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [344] Murali M Yallapu, Meena Jaggi, and Subhash C Chauhan. "Curcumin nanoformulations : a future nanomedicine for cancer." In : *Drug Discov. Today* 17.1-2 (Jan. 2012), pp. 71–80. DOI : 10.1016/j.drudis.2011.09.009. URL : <http://www.ncbi.nlm.nih.gov/pubmed/21959306><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3259195>.
- [345] Shinya Yamanaka. "Elite and stochastic models for induced pluripotent stem cell generation." In : *Nature* 460.7251 (July 2009), pp. 49–52. ISSN : 0028-0836. DOI : 10.1038/nature08180. URL : <http://www.ncbi.nlm.nih.gov/pubmed/19571877>.
- [346] Shinya Yamanaka and Kazutoshi Takahashi. "Induction of pluripotent stem cells from mouse fibroblast cultures". In : *Tanpakushitsu Kakusan Koso.* 51.15 (2006), pp. 2346–2351. ISSN : 0039-9450.
- [347] Yukiko M. Yamashita et al. "Polarity in stem cell division : asymmetric stem cell division in tissue homeostasis." In : *Cold Spring Harb. Perspect. Biol.* 2.1 (2010), pp. 1–14. ISSN : 19430264. DOI : 10.1101/cshperspect.a001313.
- [348] Ting Yan et al. "Characterization of cancer stem-like cells derived from mouse induced pluripotent stem cells transformed by tumor-derived extracellular vesicles." In : *J. Cancer* 5.7 (Jan. 2014), pp. 572–84. ISSN : 1837-9664. DOI : 10.7150/jca.8865. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4107233%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [349] Ya-Qiong Yang et al. "Epitope imprinted polymer coating CdTe quantum dots for specific recognition and direct fluorescent quantification of the target protein bovine serum albumin." In : *Biosens. Bioelectron.* 54 (2014), pp. 266–72. ISSN : 1873-4235. DOI : 10.1016/j.bios.2013.11.004. URL : <http://www.ncbi.nlm.nih.gov/pubmed/24287415>.
- [350] Yixin Yao and Wei Dai. "Genomic Instability and Cancer." In : *J. Carcinog. Mutagen.* 5.02 (2014). ISSN : 2157-2518. DOI : 10.4172/2157-2518.1000165. URL : <http://www.omicsonline.org/open-access/genomic-instability-and-cancer-2157-2518-5-165.php?aid=24175><http://www.ncbi.nlm.nih.gov/pubmed/25541596><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4274643>.

- [351] Açelya Yilmazer, Irene de Lázaro, and Hadiseh Taheri. “Reprogramming cancer cells : A novel approach for cancer therapy or a tool for disease-modeling?” In : *Cancer Lett.* (2015). ISSN : 1872-7980. DOI : 10.1016/j.canlet.2015.06.027. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0304383515005133><http://www.ncbi.nlm.nih.gov/pubmed/26276716>.
- [352] W. K. A. Yung. “The value of cell line validation”. In : *Neuro. Oncol.* 14.6 (June 2012), pp. 675–675. ISSN : 1522-8517. DOI : 10.1093/neuonc/nos132. URL : <http://neuro-oncology.oxfordjournals.org/cgi/doi/10.1093/neuonc/nos132>.
- [353] Hiba Zahreddine and Katherine L B Borden. “Mechanisms and insights into drug resistance in cancer.” In : *Front. Pharmacol.* 4.March (2013), p. 28. ISSN : 1663-9812. DOI : 10.3389/fphar.2013.00028. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23504227><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3596793>.
- [354] M a Zanta, P Belguise-Valladier, and J P Behr. “Gene delivery : a single nuclear localization signal peptide is sufficient to carry DNA to the cell nucleus.” In : *Proc. Natl. Acad. Sci. U. S. A.* 96.1 (Jan. 1999), pp. 91–6. ISSN : 0027-8424. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=15098%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [355] M Zhang et al. “Epidermal growth factor induces CD44 gene expression through a novel regulatory element in mouse fibroblasts.” In : *J. Biol. Chem.* 272.22 (1997), pp. 14139–14146. ISSN : 0021-9258.
- [356] Weijia Zhang et al. “Microfluidics separation reveals the stem-cell-like deformability of tumor-initiating cells”. In : *Proc. Natl. Acad. Sci.* 109.46 (Nov. 2012), pp. 18707–18712. ISSN : 0027-8424. DOI : 10.1073/pnas.1209893109. URL : <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3503214%7B%5C%7Dtool=pmcentrez%7B%5C%7Drendertype=abstract>.
- [357] Erika Zonari et al. “A role for miR-155 in enabling tumor-infiltrating innate immune cells to mount effective antitumor responses in mice.” In : *Blood* 122.2 (July 2013), pp. 243–52. ISSN : 1528-0020. DOI : 10.1182/blood-2012-08-449306. URL : <http://www.ncbi.nlm.nih.gov/pubmed/23487026>.