



HAL
open science

Caractérisation fonctionnelle des remaniements de structure chromosomique chez des patients atteints de troubles du neurodéveloppement

Julie Masson

► **To cite this version:**

Julie Masson. Caractérisation fonctionnelle des remaniements de structure chromosomique chez des patients atteints de troubles du neurodéveloppement. Génétique humaine. Université de Lyon, 2021. Français. NNT : 2021LYSE1314 . tel-03771703

HAL Id: tel-03771703

<https://theses.hal.science/tel-03771703>

Submitted on 7 Sep 2022

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.



N° d'ordre NNT : 2021LYSE1314

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée au sein de
L'Université Claude Bernard Lyon 1

Ecole Doctorale N°476
Neurosciences et Cognition

Spécialité de doctorat : Sciences de la vie

Discipline : Génétique Humaine

Soutenue publiquement le 13/12/2021, par :

Julie MASSON

Caractérisation fonctionnelle des remaniements de structure chromosomique chez des patients atteints de troubles du neurodéveloppement

Devant le jury composé de :

VINCIGUERRA Christine	PU-PH	Université Claude Bernard Lyon I	Présidente
JAILLARD Sylvie	PU-PH	Université de Rennes	Rapporteure
DUPONT Jean Michel	PU-PH	Université Paris V	Rapporteur
PELLESTOR Franck	PU-PH	Université de Montpellier	Examineur
SCHLUTH-BOLARD Caroline	PU-PH	Université Claude Bernard Lyon I	Directrice
SANLAVILLE Damien	PU-PH	Université Claude Bernard Lyon I	Co-directeur

Université Claude Bernard – LYON 1

Président de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Philippe CHEVALLIER
Vice-Président de la Commission de Recherche	M. Petru MIRONESCU
Directeur Général des Services	M. Pierre ROLLAND

COMPOSANTES SANTE

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud - Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directrice : Mme Christine VINCIGUERRA

COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

Département Génie Electrique et des Procédés (GEP)	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique (CPE Lyon)	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances (ISFA)	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Administrateur Provisoire : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR des Sciences et Techniques des Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE
UFR Faculté des Sciences	Directeur : M. Bruno ANDRIOLETTI

Remerciements

Au professeur Caroline Schluth-Bolard pour ton soutien tout au long de nos différents travaux et particulièrement au cours de cette thèse. Tu m'as permis de venir à bout de ce travail. Bien que j'ai décidé par la suite de changer de voie, j'ai pris beaucoup de plaisir à travailler avec toi sur les remaniements chromosomiques. Je te remercie pour tout ce que tu as pu faire pour moi.

Au Professeur Damien Sanlaville qui m'a accueillie dans son service il y a sept ans déjà avec beaucoup de bienveillance et m'a permis de finaliser ma thèse de sciences. J'ai énormément apprécié de travailler dans ton service et je garderai de très bons souvenirs de toutes ces années. Je te remercie pour ta gentillesse et ta grande écoute.

Au Professeur Jean-Michel Dupont, au Professeur Sylvie Jaillard, et au Professeur Franck Pellestor, je vous remercie chaleureusement d'avoir accepté de juger ce travail de thèse. Au plaisir d'échanger avec vous le jour de la soutenance.

Au Professeur Christine Vinciguerra, je vous remercie d'avoir accepté de juger ce travail et de présider mon jury de soutenance pour la deuxième fois. C'est toujours un honneur de vous avoir dans mon jury.

A toute l'équipe du laboratoire de cytogénétique des Hospices Civils de Lyon pour m'avoir si bien entourée pendant toutes ces années. Je remercie tous les techniciens, les ingénieurs et les biologistes pour tout ce que vous m'avez appris et pour votre bonne humeur. Merci Céline, Flavie et Eudeline pour votre écoute et votre soutien, votre amitié m'est très précieuse.

A toute l'équipe de Julien Courchet, à l'INMG, qui m'a très bien accueillie et soutenue durant la fin de ma thèse.

A l'ensemble des praticiens qui nous ont adressés des patients.

A l'ensemble des praticiens qui sont impliqués dans les publications présentées dans ce travail pour leur collaboration.

Aux patients et leur famille qui ont accepté de participer aux différentes études.

A ma famille qui a été très présente, et d'un grand soutien tout au long de cette thèse.

A mon amie de toujours, Carole, qui est toujours là pour me redonner le sourire pendant les coups durs.

A Sébastien, pour tout.

A Hugo et Agathe, pour toute la joie et la force qu'ils m'apportent au quotidien.

Résumé

Titre : Caractérisation fonctionnelle des remaniements de structure chromosomique chez des patients atteints de troubles du neurodéveloppement

Les remaniements de structure chromosomique apparemment équilibrés (RCAE), qui comprennent les translocations, les insertions, les inversions et les remaniements complexes sont des évènements rares. Sans conséquence sur le phénotype dans la majorité des cas, ils peuvent entraîner des anomalies d'appariement au moment de la méiose et résulter à la formation de gamètes déséquilibrés pouvant mener à des interruptions de grossesse. De plus, 6% des patients avec translocations, 9% des patients avec inversions et 23% des patients avec un remaniement complexe présente un phénotype pouvant résulter d'un déséquilibre cryptique ou d'une pathologie du point de cassure comme une interruption de gène ou une dysrégulation par un effet de position. En effet, outre la séquence génomique, l'expression des gènes est également déterminée par des facteurs épigénétiques. L'ADN est enroulé autour d'histones, formant ainsi la chromatine. Ces histones peuvent subir des modifications post-traductionnelles modulant la compaction de la chromatine. La chromatine est également structurée au sein du noyau, formant une architecture tridimensionnelle hiérarchisée. Ainsi, la chromatine est organisée en territoires chromosomiques, puis en compartiments chromosomiques, en domaines associés topologiquement et en boucles chromatiniennes. L'organisation du génome peut être perturbée par les RCAE, ce qui peut induire une modification de l'expression génique et entraîner l'apparition de troubles du neurodéveloppement. Nous avons étudié l'effet de ces RCAE sur le noyau. Pour cela, nous avons réalisé une étude d'expression du gène *TCF4* afin d'étudier des anomalies d'expression par interruption de différents sites d'initiation de la transcription. Puis nous avons déterminé par hybridation *in situ* en fluorescence en trois dimensions le ratio de position radiale de loci impliqués dans des RCAE. Enfin nous avons caractérisé par caryotype, hybridation *in situ* en fluorescence, analyse chromosomique sur puce à ADN et séquençage du génome entier les chromoanagenesis de six familles dans lesquelles un parent asymptotique a transmis, de manière déséquilibrée, un remaniement chromosomique à sa descendance afin d'appréhender l'appariement des chromosomes anormaux pendant la méiose.

Nous avons montré que les remaniements chromosomiques perturbent les cellules, tant au moment de l'interphase que lors de la méiose. Ils entraînent des répercussions au niveau de la séquence génomique, en interrompant des gènes ou leurs éléments régulateurs. De plus, le phénotype d'un patient peut être différent en fonction du site d'initiation de la transcription qui est impacté par un remaniement chromosomique. En effet les patients présentant un remaniement chromosomique proximal du gène *TCF4*, impliquant les premiers sites d'initiation de la transcription, présentent un phénotype modéré. Les remaniements chromosomiques ont également des répercussions au niveau de l'organisation spatiale de la chromatine. Les ratios de position radiale des loci impliqués dans des translocations ou des inversions, peuvent être modifiés. La relocalisation des dérivés peut alors avoir un effet sur le phénotype des patients en modifiant leur environnement. Lors de la méiose, les chromoanagenesis perturbent l'appariement des chromosomes homologues qui forment alors des figures complexes avec des régions d'asynapsis. Cela peut mener à la formation de gamètes déséquilibrés. Cette recombinaison étant aléatoire et imprévisible, le dépistage prénatal est recommandé pour ces familles.

Abstract

Title: Functional characterization of chromosomal rearrangements in patients with neurodevelopmental disorders

Apparently balanced chromosomal rearrangements (ABCRs) which include translocations, insertions, inversions, and complex rearrangements are rare events. Without affecting the phenotype in most cases, they can cause pairing abnormalities during meiosis and result in the formation of unbalanced gametes which can lead to pregnancy termination. Additionally, 6% of patients with translocations, 9% of patients with inversions, and 23% of patients with complex rearrangement have a phenotype resulting from cryptic imbalance or breakpoint pathology such as gene disruption or dysregulation by a position effect. Indeed, in addition to the genomic sequence, genes expression is also determined by epigenetic factors. DNA is wrapped around histones, forming chromatin. These histones can undergo post-translational modifications modulating chromatin compaction. Chromatin is also structured within the nucleus, in a hierarchical three-dimensional architecture. Indeed, chromatin is organized in chromosomal territories, in chromosomal compartments, in topologically associated domains and in chromatin loops. The organization of the genome can be disrupted by ABCRs, which can induce changes in gene expression and lead to the development of neurodevelopmental disorders. We have studied the effect of these ABCRs on the nucleus. To do this, we performed an expression study of *TCF4* gene to study expression abnormalities by disruption of different transcription start sites. Then we determined by three-dimensional fluorescence *in situ* hybridization the radial position ratio of loci involved in ABCRs. Finally, we characterized by karyotype, fluorescence *in situ* hybridization, chromosomal micro-array analysis and whole genome sequencing the chromoanagenesis of six families in which an asymptomatic parent transmitted, in an unbalanced way, a chromosomal rearrangement to its child, to understand the pairing of abnormal chromosomes during meiosis.

We have shown that chromosomal rearrangements disturbed cells, both during interphase and meiosis. They cause repercussions on the genomic sequence, by disrupting genes or their regulatory elements. Moreover, a patient's phenotype may be different depending on the transcription start site which is impacted by chromosomal rearrangement. In fact, patients with a proximal chromosomal rearrangement of *TCF4* gene, involving the first transcription initiation sites, have a moderate phenotype. Chromosomal rearrangements also have repercussions on the spatial organization of chromatin. The radial position ratios of loci involved in translocations or inversions can be modified. The relocation of the derivatives can then influence the phenotype of the patients by modifying their environment. In meiosis, chromoanagenesis disturbed the pairing of homologous chromosomes which form complex figures with regions of asynapsis. This can lead to the formation of imbalanced gametes. As this recombination is random and unpredictable, prenatal diagnosis is recommended for these families.

Table des matières

Résumé.....	7
Abstract	8
Liste des figures.....	13
Liste des tableaux.....	15
Liste des abréviations.....	16
INTRODUCTION	17
1. Le noyau	18
1.1 La chromatine.....	18
1.2 Méthodes d'études de l'architecture nucléaire.....	23
1.2.1 Méthodes en imagerie	23
1.2.2 Méthodes moléculaires.....	25
1.3 Organisation tridimensionnelle du noyau.....	27
1.3.1 Territoires chromosomiques	27
1.3.2 Compartiments chromosomiques.....	31
1.3.3 Topologically Associated Domains (TAD)	33
1.3.4 Boucles de chromatine.....	35
1.4 Architecture du noyau.....	37
1.4.1 Enveloppe nucléaire	37
1.4.2 Compartiments internes	38
1.4.2.1 Compartiments actifs	38
1.4.2.2 Compartiments répressifs	39
1.5 Fonctionnement du génome en lien avec l'organisation et l'architecture du noyau	41
2. Anomalies chromosomiques.....	42
2.1 Les types d'anomalies chromosomiques.....	42
2.1.1 Les anomalies chromosomiques de nombre.....	42
2.1.2 Les anomalies chromosomiques de structure.....	43
2.1.2.1 Les anomalies déséquilibrées.....	43
2.1.2.2 Les anomalies équilibrées	44
2.2 Les chromoanagenesis	47
2.2.1 Découverte des chromoanagenesis et classification	47
2.2.2 Classification des chromoanagenesis.....	48
2.2.2.1 Les chromothripsis	48
2.2.2.2 Les chromoanasyntésis.....	49

2.2.2.3	Les chromoplaxies	49
2.2.3	Mécanisme de survenue	50
2.3	Mécanismes de réparation des cassures doubles brins.....	51
2.3.1	Recombinaison homologue non allélique.....	51
2.3.2	Appariement simple brin.....	53
2.3.3	Jonction des extrémités non homologues	54
2.3.4	Jonction des extrémités médiée par microhomologie.....	54
2.3.5	Mécanismes réplicatifs.....	56
2.3.5.1	<i>Break Induced Replication (BIR)</i>	56
2.3.5.2	<i>Fork Stalling and Template Switching (FoSTeS)</i>	56
2.4	Identification des remaniements chromosomiques	58
2.4.1	Caryotype et hybridation fluorescente <i>in situ</i> (FISH)	58
2.4.2	Analyse chromosomique sur puce à ADN : ACPA.....	58
2.4.3	Séquençage haut débit.....	59
2.4.3.1	Partie technique : « Wet-lab »	60
2.4.3.2	Détection des SV et CNVs à partir des données de WGS : « <i>Dry-lab</i> »	61
3.	La méiose.....	65
3.1	Principe de la méiose	66
3.2	Recombinaison méiotique.....	69
3.3	Durée de la méiose.....	71
3.4	Conséquences des remaniements chromosomiques sur la méiose	72
3.4.1	Translocations	72
3.4.2	Inversions	74
3.4.3	Insertions.....	75
3.4.4	Remaniements complexes	76
3.4.5	Cas de transmission familiale de chromoanagenesis.....	77
	OBJECTIFS	79
	RESULTATS.....	81
	PARTIE 1 : Etude des remaniements proximaux du gène <i>TCF4</i>	81
	PARTIE 2 : Conséquences des remaniements chromosomiques sur l'architecture nucléaire	109
1.	Matériel et Méthodes	110
1.1	Patients.....	110
1.2	Culture cellulaire	110
1.3	Hybridation <i>in situ</i> en fluorescence en 3D = 3D FISH	112

1.4	Acquisition des images	113
1.5	Analyse des images.....	114
1.6	Analyse statistique	114
2.	Résultats	114
2.1	Choix de la méthode d'analyse	114
2.1.1	Comparaison de deux approches	115
2.1.2	Pipeline d'analyse sur Arivis	117
2.2	Effets des remaniements sur le positionnement radial	118
2.2.1	Patients avec des translocations ayant un effet sur la position radiale.....	119
2.2.2	Patients avec des translocations sans effet de relocalisation	123
2.2.3	Patients avec des inversions.....	127
2.2.4	Les positions radiales sont identiques entre différents individus.....	131
2.2.5	Les chromosomes impliqués dans un remaniement ont des positions radiales identiques.....	132
3.	Discussion	133
	PARTIE 3 : Etude de la transmission de chromoanagenesis familiaux	139
	DISCUSSION	213
	BIBLIOGRAPHIE.....	216

Liste des figures

Figure 1. Visualisation de la chromatine au microscope électronique (adapté de Donald E. Olins et Ada L. Olins, 2003).....	18
Figure 2. Représentation schématique du nucléosome et des différentes modifications post-traductionnelles des histones possibles, adapté de Sawan et Herceg, 2010.....	20
Figure 3. Le chromosome, état de compaction maximal de la chromatine adapté d'Annunziato, 2008	22
Figure 4. Méthodes d'imagerie pour étudier l'architecture tridimensionnelle du noyau, d'après Kempfer et Pombo, 2020	24
Figure 5. Méthode 3C et méthodes dérivées, d'après Kempfer et Pombo, 2020	26
Figure 6. Applications des différentes techniques d'analyse de la chromatine, d'après Umlauf, 2015	27
Figure 7. Visualisation des territoires chromosomiques.....	29
Figure 8. Compartiment interchromatinien, adaptée de Cremer et Cremer, 2001.....	30
Figure 9 Timing de réplication en fonction de l'organisation tridimensionnelle de la chromatine, adapté de Marshall et al., 2019.....	32
Figure 10. Formation des TAD par extrusion de l'ADN, adaptée de Dolgin, 2017	34
Figure 11. Architecture de la chromatine à différentes échelles, avec les méthodes d'étude, adaptée de Kempfer et Pombo, 2020	36
Figure 12. Représentation de la membrane nucléaire, d'après Magistris et al., 2018	37
Figure 13. Usines de transcription, d'après Maland et al., 2009	38
Figure 14. Visualisation des corps d'épissage, adapté de Cremer et Cremer, 2001.	39
Figure 15. Gradient nucléaire d'expression génique, adapté de Chen et al., 2018	42
Figure 16. Représentations d'une translocation réciproque (A) et d'une translocation robertsonienne (B) entre deux chromosomes (Chr) A et B.....	45
Figure 17. Représentation d'une inversion péricentrique du chromosome A (Chr A).....	45
Figure 18. Représentation d'une insertion d'un fragment du chromosome A (Chr A) dans le chromosome B (Chr B).	46
Figure 19. Chromothripsis	49
Figure 20. Chromoanasythesis	50
Figure 21. La recombinaison homologue non allélique est responsable de réarrangements chromosomiques.....	53
Figure 22. Voies de réparation des cassures de l'ADN double brin lors de la méiose	55
Figure 23. Principe du mécanisme d'interruption de la fourche de réplication et de changement de matrice	57
Figure 24. Principe de la méthode d'ACPA.....	59
Figure 25. Principe du séquençage haut débit avec la méthode Illumina	61
Figure 26. Principe du séquençage (WGS) en « <i>paired-end</i> »	61
Figure 27. Remaniements chromosomiques apparemment équilibrés et discordance d'alignements des <i>reads</i> du WGS.....	62
Figure 28. Visualisation sur IGV.....	64
Figure 29. Les différentes étapes de la méiose.....	65
Figure 30. Les différentes étapes de la prophase 1.....	66
Figure 31. Organisation de la chromatine pendant la prophase de la méiose 1.	68

Figure 32. La recombinaison méiotique	70
Figure 33. Possibilités de ségrégation méiotique d'une translocation robertsonienne entre le chromosome A et le chromosome B	72
Figure 34. Possibilités de ségrégation méiotique d'une translocation réciproque entre le chromosome A et le chromosome B	74
Figure 35. Recombinaison méiotique en cas d'inversion péricentrique	75
Figure 36. Possibilités de formation de gamètes en cas d'insertion	75
Figure 37. Formation d'un hexavalent lors d'un remaniement complexe (translocation entre trois chromosomes).....	76
Figure 38. Comparaison des logiciels Imaris et Arivis	116
Figure 39. Pipeline d'analyse sur Arivis	118
Figure 40. Position radiale nucléaire des chromosomes et dérivés du patient AD/1601.....	119
Figure 41. Position radiale nucléaire des chromosomes et dérivés de la patiente LD/0108.....	121
Figure 42. Position radiale nucléaire des chromosomes et dérivés du patient VJ/0601	122
Figure 43. Position radiale nucléaire des chromosomes et dérivés de la patiente CL/2001	124
Figure 44. Position radiale nucléaire des chromosomes et dérivés de la patiente MD/0104	125
Figure 45. Position radiale nucléaire des chromosomes et dérivés du patient CS/1902.....	126
Figure 46. Position radiale nucléaire des chromosomes et dérivés du patient AL/2402.....	128
Figure 47. Position radiale nucléaire des chromosomes et dérivés de la patiente SL/0105	130

Liste des tableaux

Tableau 1 : Caractéristique moléculaire des points de cassure et phénotype des patients.....	111
Tableau 2. Liste des sondes de FISH utilisées.....	113
Tableau 3 : Résultats des tests de comparaison entre Imaris et Arivis sur 30 noyaux	117
Tableau 4. Valeurs des ratios de position radiale des différents loci du patient AD/1601 et résultats des tests statistiques.	120
Tableau 5. Valeurs des ratios de position radiale des différents loci du patient LD/0108 et résultats des tests statistiques.	121
Tableau 6. Valeurs des ratios de position radiale des différents loci du patient VJ/0601 et résultats des tests statistiques.	123
Tableau 7. Valeurs des ratios de position radiale des différents loci de la patiente CL/2001 et résultats des tests statistiques.	124
Tableau 8. Valeurs des ratios de position radiale des différents loci de la patiente MD/0104 et résultats des tests statistiques.	126
Tableau 9. Valeurs des ratios de position radiale des différents loci du patient CS/1902 et résultats des tests statistiques.	127
Tableau 10. Valeurs des ratios de position radiale des différents loci du patient AL/2402 et résultats des tests statistiques.	129
Tableau 11. Valeurs des ratios de position radiale des différents loci de la patiente SL/0105 et résultats des tests statistiques.	130
Tableau 12. Ratios et résultats des tests de comparaisons multiples afin de comparer les positions radiales entre le patient et le témoin.....	131
Tableau 13. Résultats des tests de comparaisons multiples afin de comparer les positions radiales entre deux positions sur un même chromosome	132
Tableau 14. Ratios et résultats des tests de comparaisons multiples afin de comparer les positions radiales entre deux chromosomes impliqués dans une translocation	132

Liste des abréviations

3D-FISH : Hybridation In Situ en Fluorescence en 3 dimensions (*Fluorescent in situ hybridization*)

ACPA : Analyse Chromosomique sur Puce à ADN

ADN : Acide Désoxyribonucléique

ARN : Acide ribonucléique

BIR : Réplication induite par la rupture (*Break Induced Replication*)

Chr : chromosome

CNV : Variation du nombre de copies (*Copy Number Variation*)

DAPI : Diamidino Phénylindole

der: dérivé

DSCR : région critique du syndrome de Down (*Down Syndrome Critical Region*)

FISH : Hybridation In Situ en Fluorescence (*Fluorescent in situ hybridization*)

FoSTeS : Arrêt de la fourche et changement de matrice (*Fork Stalling and Template Swintching*)

HR : Recombinaison homologue (*Homologous recombination*)

kb: kilobases

LAD: domaine associé à la lamine (*Lamina associated domain*)

LCR : duplications segmentaires (*Low Copy Repeat*)

LINE : élément répété de type *Long Interspersed Nuclear Element*

Mb : Mégabases

MMBIR : Réplication induite par la rupture médiée par microhomologie (*Microhomology Mediated Break Induced Replication*)

MMEJ : Jonction des extrémité médiée par microhomologie (*Microhomology Mediated EndJoining*)

NAD : Domaine associé au nucléole (*Nucleolus-Associated Domain*)

NGS : séquençage haut débit (*Next Generation Sequencing*)

NAHR : Recombinaison homologue non allélique (*Non Allelique Homologous Recombination*)

NHEJ : Jonction des extrémités non homologues (*Non Homologous End Joining*)

pb : paires de bases

PCR : Réaction de Polymérisation en Chaîne(*Polymerase Chain Reaction*)

qPCR : PCR quantitative

RCAE : Remaniement Chromosomique Apparemment Equilibré

SSA: Appariement simple brin (*Single Strand Annealing*)

TAD: Domaine associé topologiquement (*Topologically-Associated Domain*)

UCSC : University of California Santa Cruz

WGS : Séquençage du génome entier (*Whole Genome Sequencing*)

INTRODUCTION

Les remaniements chromosomiques apparemment équilibrés (RCAE), regroupant les translocations, les inversions, les insertions et les remaniements complexes sont des événements rares. Ils sont retrouvés dans 1,6 naissances sur 1000. Dans la majorité des cas, ces RCAE sont bénins et n'entraînent pas de phénotype chez le patient. Cependant, ils peuvent entraîner des anomalies au moment de la méiose et être responsables d'interruption de grossesse ou de formation de gamètes déséquilibrés, avec un risque pour la descendance. Il est également rapporté qu'environ 6% des patients avec translocation, 9% des patients avec inversions et 23% des patients avec un remaniement complexe présente un phénotype. Ce phénotype peut résulter de la présence d'un remaniement cryptique, au niveau du point de cassure ou ailleurs dans le génome. Le phénotype peut également être dû à une pathologie du point de cassure, comme l'interruption d'un gène ou un effet de position. En effet, outre la séquence à proprement parlé, l'expression d'un gène est également déterminée par des facteurs épigénétiques, qui sont liés à son environnement. La chromatine n'est pas agencée de manière aléatoire dans le noyau, il existe une architecture tridimensionnelle (3D) hiérarchisée, comprenant territoires, compartiments, domaines et boucles chromatinienne.

Nous allons voir en introduction l'organisation tridimensionnelle et l'architecture du noyau. Puis nous décrirons les anomalies chromosomiques, avec un point particulier sur les remaniements complexes, et les mécanismes de réparation des cassures de l'ADN qui peuvent entraîner ces anomalies. Enfin, nous verrons le déroulement de la méiose, et la répercussion des anomalies chromosomiques sur son déroulement.

1. Le noyau

1.1 La chromatine

Chacune de nos milliards de cellules comporte un noyau renfermant 46 chromosomes : 22 paires d'autosomes et une paire de chromosomes sexuels, les gonosomes. Ainsi une femme a une paire de chromosomes X (46,XX) et un homme un chromosome X et un chromosome Y(46,XY).

Un chromosome est une structure nucléoprotéique composée d'ADN (acide desoxyribonucléique) et de protéines : les histones et les protéines non-histones. L'ADN, qui est le support de notre information génétique, est composé d'un sucre : le désoxy-ribose et de bases nucléotidiques : Adénine, Cytosine, Guanine, Thymine.

Les histones vont permettre de compacter l'ADN. La double hélice d'ADN va en effet s'enrouler autour d'octamères d'histones (composé de quatre paires d'histones : H2A, H2B, H3 et H4). Cela va former des nucléosomes, qui sont la compaction de 146 pb d'ADN autour d'un octamère d'histone. Ces nucléosomes, qui sont l'unité de structure de base ont été observés en microscopie électronique, et donnent l'impression de perles disposées le long d'un fil [1,2]. Il s'agit alors d'une fibre de 11nm, qui est l'état le plus décompacté de la chromatine.

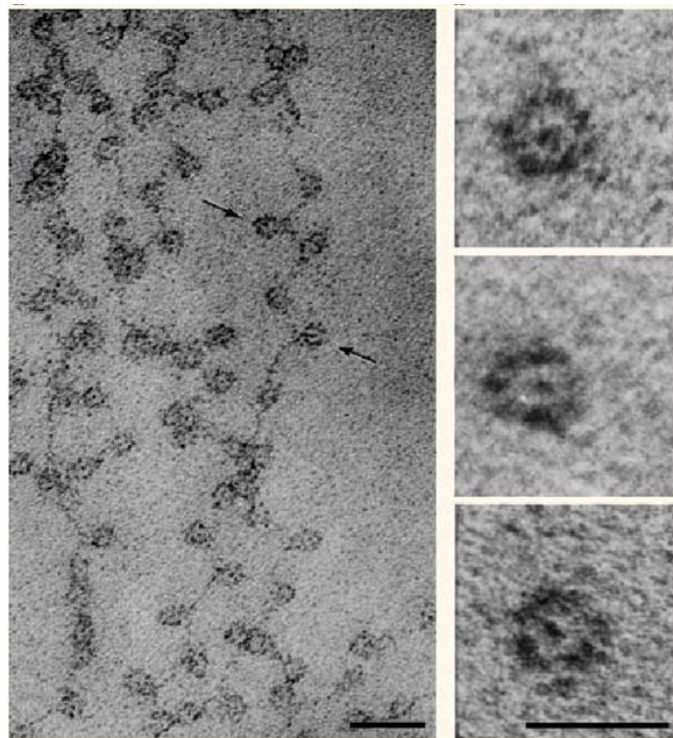


Figure 1. Visualisation de la chromatine au microscope électronique (adapté de Donald E. Olins et Ada L. Olins, 2003)

Sur l'image de gauche le « collier de perle » est visible, avec des nucléosomes pointés par les flèches. Le marqueur de taille représente 30 nm. Sur les trois images de droite, des nucléosomes sont visibles après digestion de la chromatine par des nucléases. Le marqueur de taille représente 10 nm.

Ce « collier de perle » peut encore se sur-enrouler et former la fibre de 30 nm de chromatine (Figure 1) qui est stabilisée par la liaison de l'histone H1 au niveau de l'ADN de liaison entre les nucléosomes. Cette compaction peut être modulée par modification post-traductionnelle des histones, particulièrement l'histone H3. Ces modifications ont principalement lieu au niveau de l'extrémité N-terminale des histones, qui a un rôle dans l'interaction avec les nucléosomes voisins, ce qui peut entraîner une compaction plus ou moins importante. Les principales modifications chimiques post traductionnelles des histones sont : des méthylations des résidus arginine (R), des acétylations, des méthylations, des ubiquitinylations, des ADP-ribosylations et des sumoylations des résidus lysines (K) et des phosphorylations des sérines (S) et thréonines (T). L'acétylation des lysines est réalisée par des acétyltransférases (HAT)[3] et neutralise partiellement les charges positives et donc diminue l'affinité pour l'ADN, ce qui entraîne une diminution de l'interaction histones/ADN. Les régions hyperacétylées sont retrouvées au niveau des gènes exprimés ou des séquences régulatrices [4]. Il s'agit par exemple de l'acétylation de la lysine 27 de l'histone H3 (H3K27ac) pour laquelle est observée un pic au niveau des séquences régulatrices. Contrairement à l'acétylation qui peut être réalisée au niveau de plusieurs résidus lysine par une variété d'HAT, les méthylations, phosphorylations et ubiquitinylations sont en général catalysées par une enzyme spécifique, sur un site spécifique et aboutissent à une fonction spécifique [4]. La méthylation peut apporter 1, 2 ou 3 groupements méthyles sur les fonctions amines des lysines et arginines. En fonction de la position et du nombre de groupements méthyles, les effets sur la chromatine sont différents [4,5]. Certaines marques sont associées à la chromatine active comme la monométhylation, diméthylation et triméthylation des lysines 4 de l'histone H3 (H3K4me1/2/3), de la monométhylation de la lysine 20 de l'histone H4 (H4K20m1) et d'autres marques sont associées à la répression comme la monométhylation, diméthylation et triméthylation de la lysine 9 de l'histone H3 (H3K9me1/2/3), la di et triméthylation de la lysine 27 de l'histone H3 (H3K27me2/3) et la triméthylation de la lysine 20 de l'histone H4 (H4K20me3)[6] (Figure 2). La combinaison de ces différentes modifications permet d'établir le code histone [7].

En dehors de la mitose, deux types de chromatine sont observés en fonction du niveau de compaction : l'euchromatine et l'hétérochromatine [8]. L'euchromatine est décondensée et est généralement riche en gènes, ce sont des régions où la transcription est active. L'hétérochromatine est de la chromatine condensée, ces régions sont moins riches en gènes et sont dites inactives. Le statut de la chromatine

est déterminé par le code histone, les protéines non histones associées et la méthylation de l'ADN. Les protéines non-histones sont les protéines qui restent après l'élimination des histones de la chromatine.

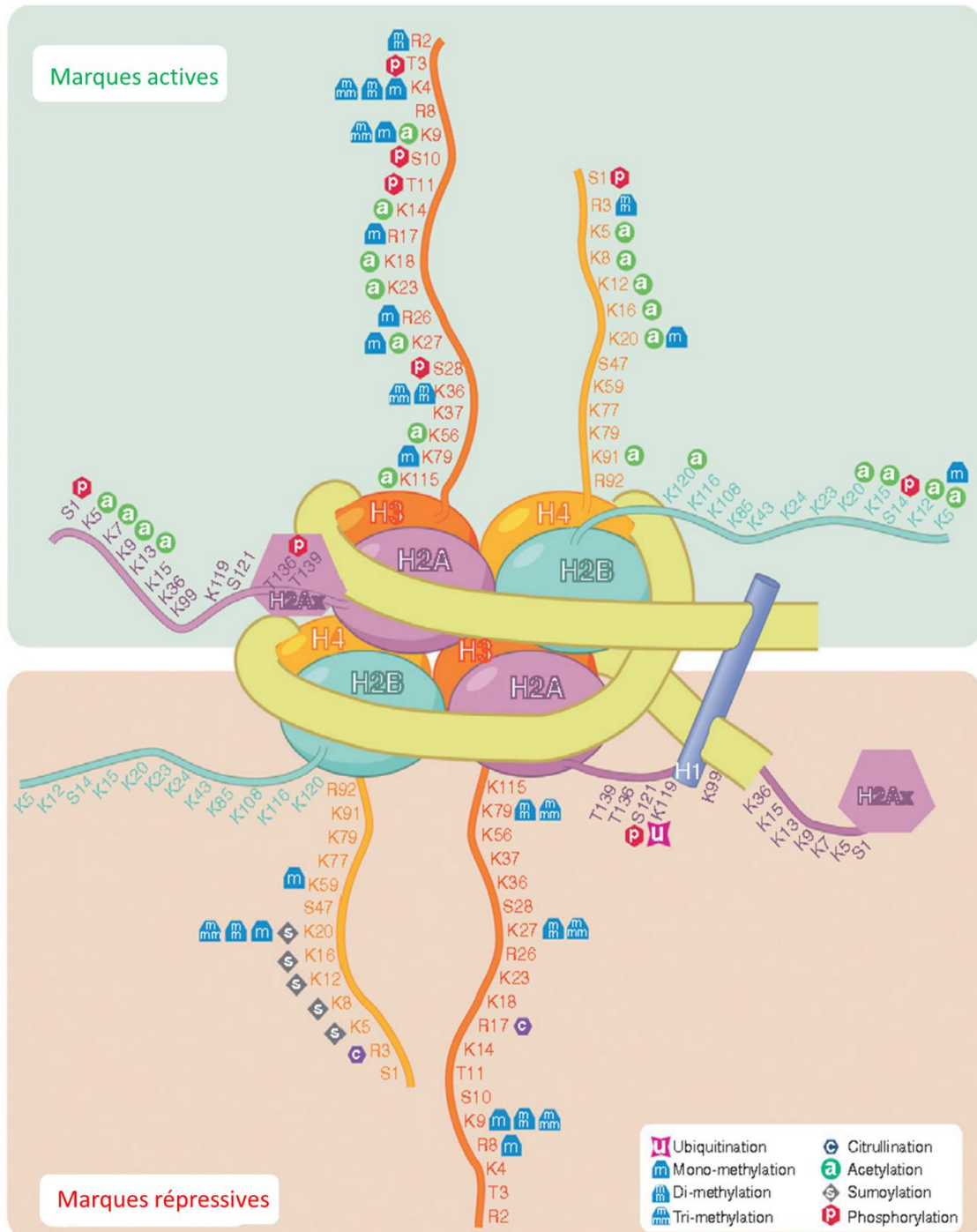


Figure 2. Représentation schématique du nucléosome et des différentes modifications post-traductionnelles des histones possibles, adapté de Sawan et Herczeg, 2010

L'ADN est enroulé autour d'un octamère d'histone. L'histone H1 est liée à l'ADN entre les nucléosomes. Les acides aminés constituant les queues d'histone sont représentés avec les différentes modifications de chaque résidu. Les marques actives sont représentées dans la partie supérieure de la figure et les marques répressives sont représentées dans la partie inférieure de la figure. K : Lysine ; R : arginine ; S : sérine et T : thréonine.

Environ 200 protéines non-histones ont été rapportées [9]. Elles jouent un rôle dans l'organisation et la compaction du chromosome. Elles participent au remodelage des nucléosomes, à la réplication de l'ADN et à la synthèse de l'ARN. Par exemple, les condensines sont impliquées dans la compaction des chromosomes avant la mitose et la cohésine forme un anneau qui encercle les chromatides sœurs pour les maintenir ensemble après la duplication, jusqu'à l'anaphase.

Le chromosome, molécule d'ADN linéaire, est une structure complètement condensée de chromatine, que l'on retrouve au moment de la mitose, et plus précisément au moment de la métaphase (Figure 3). La réalisation d'un caryotype nous permet d'observer les chromosomes. Chaque chromosome est alors composé de deux chromatides sœurs, identiques, qui sont deux molécules d'ADN identiques, reliées par un centromère. Ce centromère va permettre l'assemblage du kinétochore qui permet la liaison des chromatides aux microtubules du fuseau mitotique et va permettre de diviser la cellule mère en deux cellules filles contenant chacune une des deux chromatides sœur. Aux extrémités des chromosomes on retrouve les télomères, qui sont des structures nucléoprotéiques composées de répétitions d'un hexanucléotide TTAGGG et d'un ensemble de six protéines appelé complexe shelterin, qui permettent de protéger les terminaisons chromosomiques. Ceux-ci sont raccourcis à chaque division cellulaire.

Le premier élément structural est le nucléosome, composé d'ADN et de protéines histones. Les protéines histones peuvent subir des modifications post traductionnelles afin de réguler la compaction de la chromatine. En fonction de cette compaction, deux types de chromatines sont distinguées : l'euchromatine qui est décondensée et active et l'hétérochromatine qui est condensée et inactive. Cette compaction de plus en plus importante aboutit au moment de la mitose au chromosome.

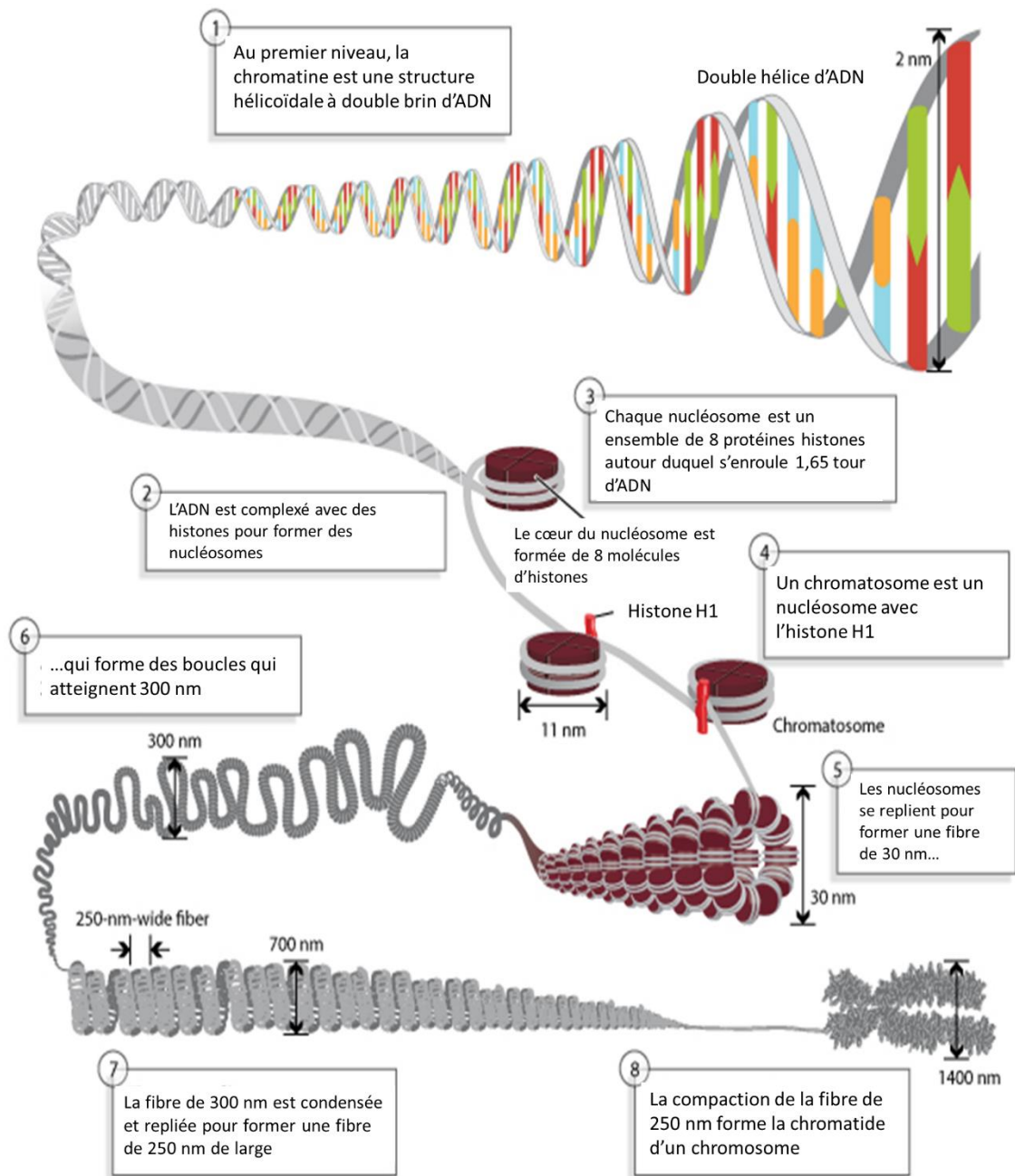


Figure 3. Le chromosome, état de compaction maximal de la chromatine adapté d'Annunziato,2008

L'AND est compacté avec l'aide d'histones. Ce sont des protéines chargées positivement adhérant fortement à l'AND chargé négativement pour former des nucléosomes. Les nucleosomes se surenroulent afin de former une fibre de 30nm qui va se condensée pour aboutir au plus haut niveau de compaction de l'AND : le chromosome.

1.2 Méthodes d'études de l'architecture nucléaire

Nous allons voir, de façon non exhaustive, les principales méthodes d'étude de la chromatine qui permettent d'appréhender l'organisation tridimensionnelle du génome [10,11].

1.2.1 Méthodes en imagerie

Les méthodes d'imagerie permettent de visualiser les différentes structures nucléaires et l'organisation de la chromatine au sein du noyau. Différentes techniques de microscopies, comme la microscopie à contrastes de phases ou la microscopie électronique permettent de visualiser les différents composants du noyau. La technique la plus utilisée pour déterminer les contacts interchromatiniens et la position de régions génomiques est la technique d'hybridation *in situ* en fluorescence [8].

Hybridation *in situ* en fluorescence

Le principe de l'hybridation *in situ* en fluorescence est de cibler une région d'intérêt avec une sonde fluorescente (Figure 4a) [12]. Pour cela il faut que la sonde pénètre dans le noyau de la cellule, qui est donc perméabilisé grâce à l'utilisation de détergents ou de solvants organiques. Suite à ce pré-traitement, les ADNs de la sonde et du génome sont dénaturés par chauffage, puis pendant la période d'hybridation, à 37°C, la sonde va pouvoir se lier à sa région cible. La région d'intérêt, qui est liée à la sonde fluorescente, va pouvoir être visualisée en microscopie à épifluorescence. Les sondes utilisées couvrent des régions de 30kb, jusqu'à la totalité d'un chromosome. Mais il ne s'agit pas d'un grand fragment d'ADN, mais de petits fragments de 100 à 500pb qui sont chevauchants, afin de faciliter la pénétration dans le noyau et l'atteinte de la région cible.

Hybridation *in situ* en fluorescence en 3D

Afin de pouvoir visualiser la région d'intérêt dans le volume du noyau, il faut faire un prétraitement afin de conserver la structure tridimensionnelle du noyau. Le détail de la technique sera traitée dans la partie méthode, et l'analyse des images dans la partie résultats.

Globalement, les cellules vivantes vont être fixées par de la paraformaldehyde. Puis des étapes de perméabilisation du noyau sont réalisées via des bains dans des détergents. Des cycles de congélation/décongélation sont ensuite réalisés en alternant des bains de glycérol et des congélations rapides dans l'azote afin d'augmenter l'accessibilité aux sondes de FISH tout en conservant l'architecture tridimensionnelle du noyau.

L'acquisition des images consiste à faire des coupes optiques du noyau (Z-stack). Puis l'ensemble de ces coupes est traité par un logiciel afin de reconstruire le noyau en trois dimensions.

La technique de 3D-FISH (Figure 4a) est utilisée pour déterminer la distance entre deux loci d'intérêt, pour visualiser la compaction de la chromatine, ou pour déterminer la position radiale d'une région génomique [13].

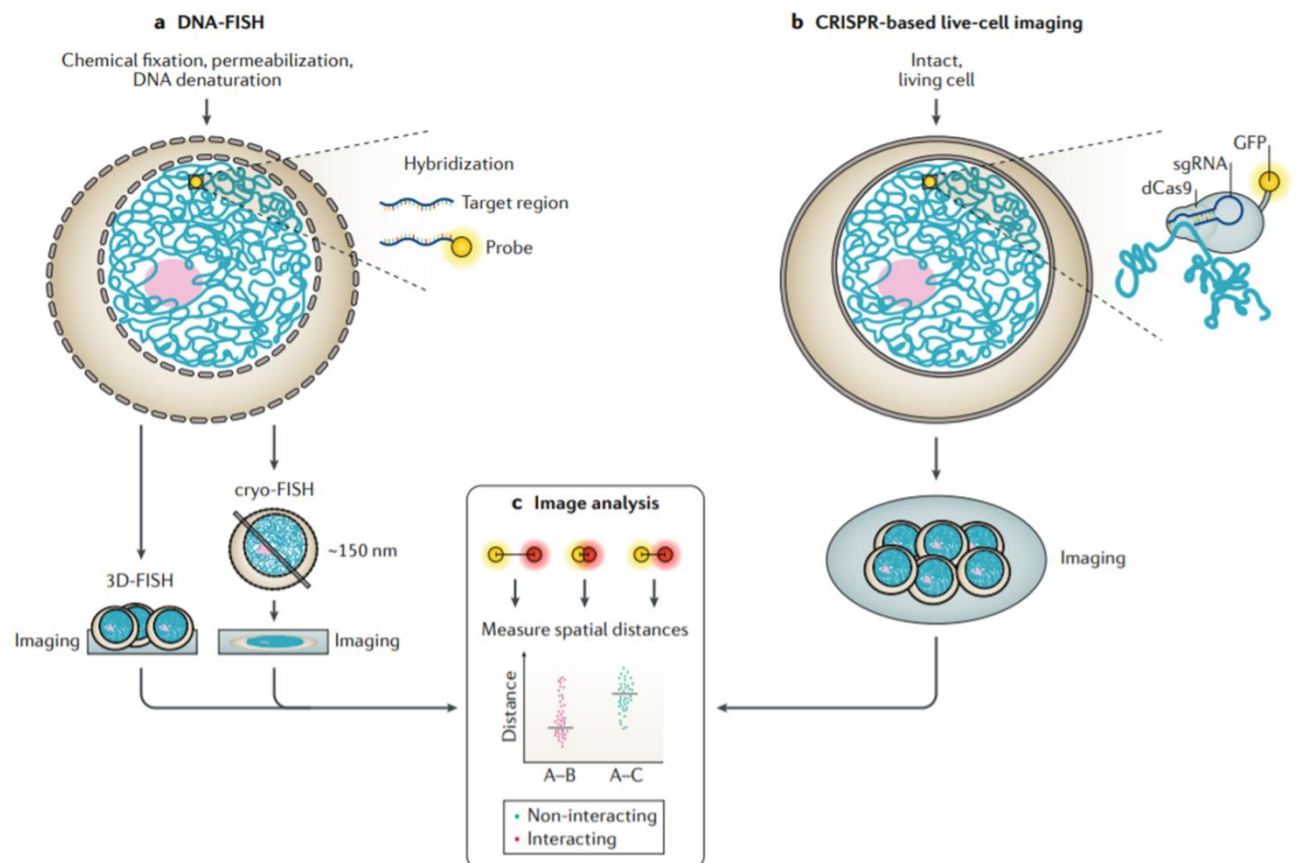


Figure 4. Méthodes d'imagerie pour étudier l'architecture tridimensionnelle du noyau, d'après Kempfer et Pombo, 2020

Cryo-FISH

Il est également possible de réaliser une hybridation sur une fine coupe, de 100 à 200nm, d'une cryosection de cellules fixées [14] (Figure 4a).

Imager des cellules vivantes

Des systèmes, comme le système Lac opérateur/répresseur ou le système CRISPR/Cas9, peuvent permettre de visualiser une région sur des cellules vivantes (Figure 4b). Concernant le système Lac opérateur/répresseur, Il s'agit d'un système constitué de deux composants. Le premier va cibler une

région du génome et est riche en opérateur Lac (LacO). Le deuxième va se lier aux régions LacO via la protéine LacI qui est le répresseur de l'opérateur Lac. Ce deuxième composant est lié à une protéine fluorescente, permettant de visualiser la région d'intérêt ciblée par le premier composant. Quant à l'utilisation du système CRISPR/cas 9, il nécessite l'utilisation d'une cas9 inactivée et fusionnée à une molécule fluorescente comme le GFP (*green fluorescence protein*) qui va être recrutée sur la région cible par l'utilisation d'un ARN guide complémentaire à la région d'intérêt [15].

1.2.2 Méthodes moléculaires

Une famille de méthode, dérivée de la méthode de 3C pour chromosome capture conformation, permet de détecter les fréquences d'interactions chromatinienne afin d'analyser l'organisation spatiale de la chromatine. Il s'agit de détecter et de quantifier les contacts grâce à une technique basée sur la ligation.

Pour cela, les cellules vont être fixées au formaldéhyde, la chromatine est alors figée dans sa conformation 3D, cela va entraîner un « *cross-linking* » des séquences à proximité. Puis la chromatine va être fragmentée par des enzymes de restriction (comme EcoR1 ou HindIII). Ces fragments vont être ligués puis purifiés afin d'inverser le « *cross-linking* ». L'ensemble de ces fragments chimériques (formés par les 2 fragments qui étaient à proximité) forme la librairie 3C (Figure 5) [16].

La fréquence de ligation entre deux loci d'intérêt peut alors être quantifiée par PCR quantitative. La technique de 3C permet donc de connaître les interactions entre 2 loci (un contre un), et implique de connaître les deux loci que l'on veut étudier afin de choisir les amorces spécifiques. Cela peut permettre de tester l'interaction entre un gène et un promoteur.

Cependant, la librairie 3C comprend tous les produits de ligation du génome, d'autres techniques ont donc été développées par la suite (Figure 5).

La technique de 4C, pour « *circular chromosome conformation capture* » [17] ou « *chromosome conformation capture-on-chip* » [18], nécessite une deuxième étape de ligation pour créer des fragments d'ADN circulaires. Les amorces sont choisies dans la séquence connue, permettant de réaliser une PCR-inverse qui va détecter les produits de ligation. La banque ainsi réalisée peut être séquencée ou hybridée sur puce à ADN. Cela permet donc de connaître les interactions entre un locus et le reste du génome (un contre tous).

La technique de 5C, pour « *chromosome conformation capture carbon copy* » [19], est réalisée en utilisant un mélange de couples d'amorces (non spécifiques). Cela détecte les interactions entre tous les fragments dans une région donnée, allant jusqu'à plusieurs mégabases (plusieurs contre plusieurs). Cela implique de bien choisir tous les couples d'amorces nécessaires à l'étude d'une région donnée.

La technique la plus utilisée aujourd’hui afin d’étudier les contacts chromatinien est la technique Hi-C, pour « *high throughput 3C* » [20]. Dans cette technique, l’étape de ligation est précédée d’une étape de biotinylation des extrémités. Puis l’activité exonucléase de l’ADN polymérase T4 est utilisée afin d’enlever la biotine des extrémités qui ne sont pas liguées. Les fragments ligués, qui sont biotinylés, sont ensuite capturés en utilisant des billes de streptavidine. Cela permet de ne conserver que les fragments qui résultent d’une ligation entre deux régions différentes. Cette librairie peut alors être séquencée. Cela permet de détecter toutes les interactions du génome (tous contre tous). Les données de Hi-C permettent de générer des matrices d’interactions à l’échelle du génome. Ces matrices montrent des fragments de chromatine enrichis ou déplétés en interactions, grâce à un code couleur. Les données de Hi-C vont permettre d’élucider l’organisation du génome humain à une fine échelle (Figure 6).

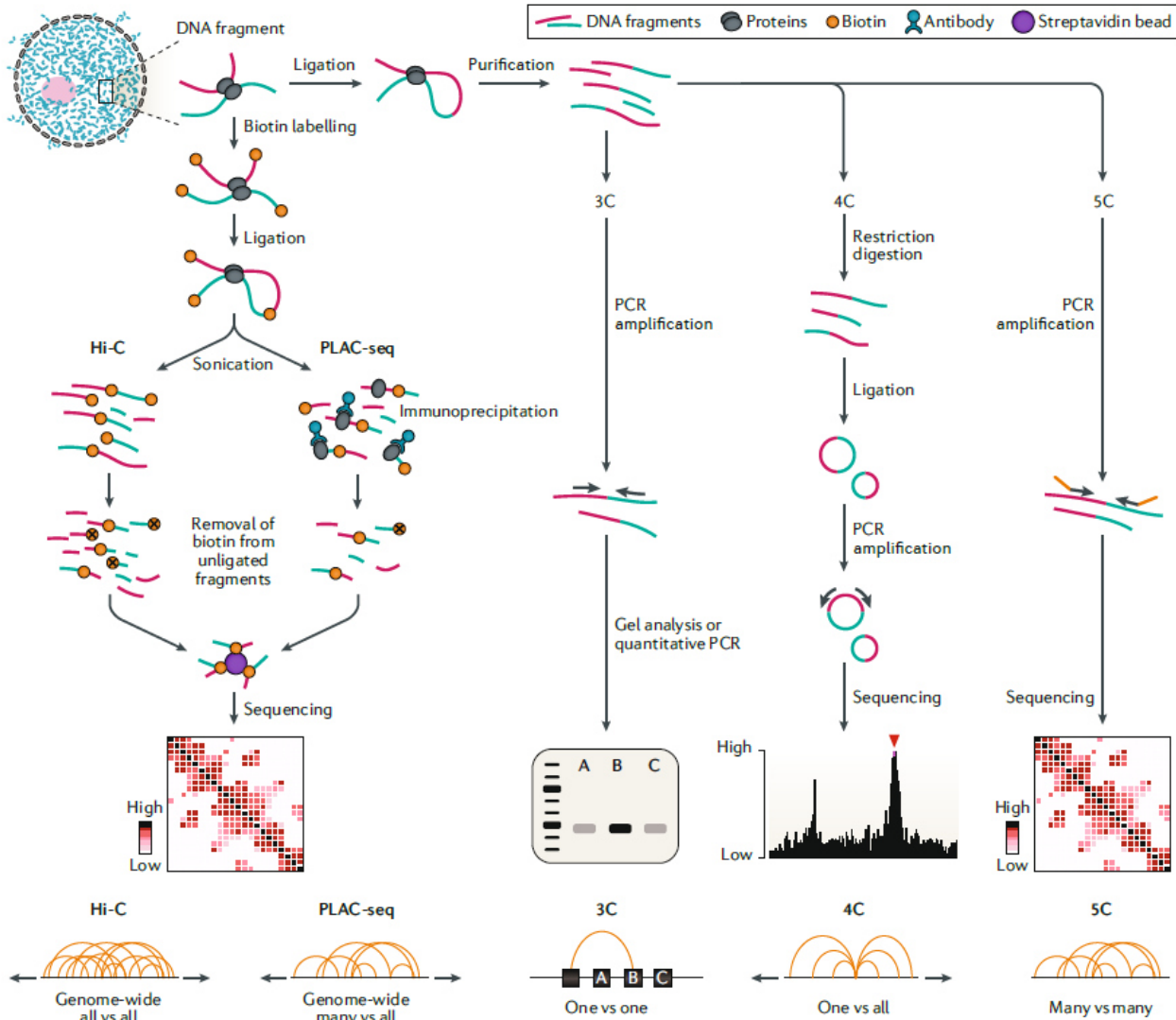


Figure 5. Méthode 3C et méthodes dérivées, d’après Kempfer et Pombo, 2020


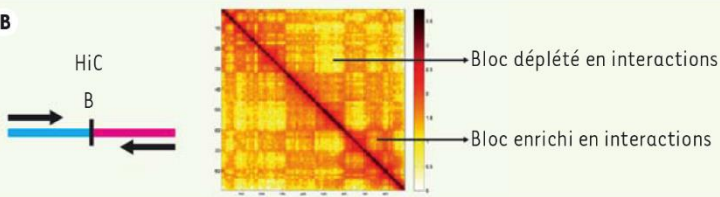
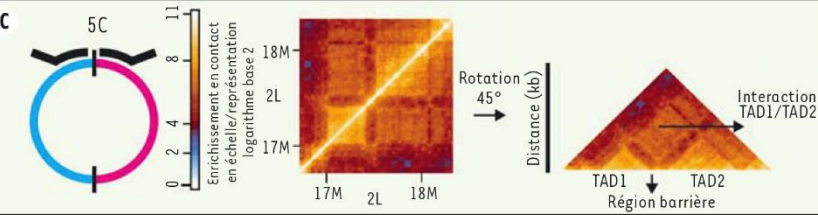
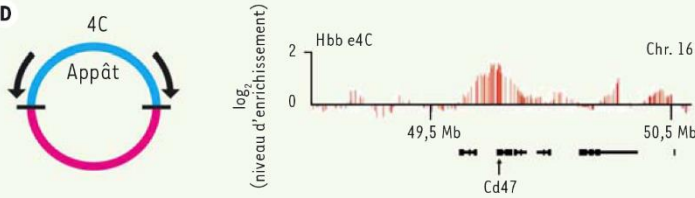
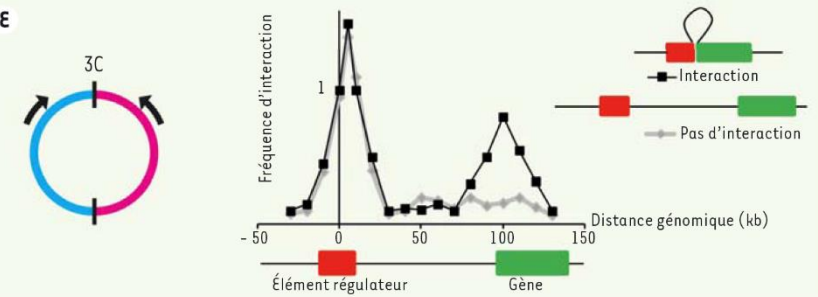
Techniques		Applications
A	ADN FISH 	Localisation intranucléaire des gènes, organisation des territoires chromosomiques
B	HiC  Bloc déplété en interactions Bloc enrichi en interactions	Caractérisation des compartiments A et B Caractérisation des TAD Organisation 3D des chromosomes interphasiques ou métaphasiques
C	5C  Enrichissement en contact en échelle/représentation logarithme base 2 Rotation 45° Distance (kb) Interaction TAD1/TAD2 Région barrière	Caractérisation des TAD Organisation 3D des chromosomes Détection des boucles chromatinienne Organisation des interactions éléments régulateurs/promoteurs à très large échelle
D	4C  Appât log ₂ (niveau d'enrichissement) Hbb e4C Chr. 16 49,5 Mb 50,5 Mb Cd47	Détection des boucles chromatinienne Réseaux d'interactions fonctionnelles et/ou transcriptionnelles entre l'appât et le reste du génome
E	3C  Fréquence d'interaction Distance génomique (kb) Élément régulateur Gène Interaction Pas d'interaction	Détection des boucles chromatinienne Interactions élément régulateur/promoteur en cis et en trans

Figure 6. Applications des différentes techniques d'analyse de la chromatine, d'après Umlauf, 2015

1.3 Organisation tridimensionnelle du noyau

Grâce à l'utilisation des méthodes de visualisation en imagerie, puis grâce aux cartographies obtenues par les méthodes moléculaires, l'organisation tridimensionnelle du noyau s'est dévoilée peu à peu.

1.3.1 Territoires chromosomiques

C'est en 1885 que Carl Rabl suggère l'organisation en territoires de la chromatine pendant l'interphase. Cependant, durant les années 1950-1970 la croyance est que la chromatine est entremêlée, à l'image d'une assiette de spaghettis. La première observation aura été faite en 1977, sur des cellules ovariennes de hamsters traitées à l'acide acétique, séchées et colorées au Giemsa, permettant l'observation de « taches » de chromatines : les territoires chromosomiques [21].

Une autre expérience, de micro-lésions de l'ADN, a permis de confirmer l'existence des territoires chromosomiques. En effet, des micro-lésions de l'ADN ont été réalisées à l'aide d'un microfaisceau laser au niveau d'une petite région du noyau de cellules de hamsters. Puis de la thymidine tritiée a été ajoutée à la culture, permettant de marquer les portions de chromatine où une réparation a eu lieu. L'observation des chromosomes pendant la mitose permet de mettre en évidence de l'ADN marqué restreint à une région sur un ou quelques chromosomes, alors que l'on s'attendrait à avoir de nombreux chromosomes marqués dans le cas d'un modèle d'organisation aléatoire [22–24] (Figure 7 A et B).

Le développement de l'hybridation *in situ* et des sondes de peintures chromosomiques a permis la première visualisation des territoires chromosomiques dans le noyaux de cellules fixées : en effet l'espace occupé en interphase par un chromosome peut être visualisé grâce à une sonde de peinture (ensemble de sondes couvrant l'ensemble d'un chromosome) le marquant spécifiquement [25–27] (Figure 7C).

Les territoires chromosomiques ont été ensuite visualisés dans les cellules vivantes de hamster par incorporation de nucléotides fluorescents pendant la réplication de l'ADN [28–30]. Cela a permis de visualiser qu'il n'y avait pas de chevauchement entre les territoires chromosomiques.

L'hybridation *in situ* en fluorescence a ensuite permis de visualiser plusieurs chromosomes simultanément et de confirmer que chaque chromosome occupe un territoire distinct et ne chevauche pas les autres territoires chromosomiques [31,32] (Figure 7D, E, F et G).

L'existence des territoires chromosomiques a également pu être confirmée récemment par l'utilisation d'une technique moléculaire, le séquençage Hi-C [20].

La position des territoires chromosomiques dans le noyau n'est pas aléatoire. En effet il a été montré que les petits chromosomes (16, 17,18, 19, 20, 21, 22) ont une position plus centrale que les grands chromosomes [33,34]. Les petits chromosomes sont plus riches en gènes que les grands chromosomes. La densité en gènes semble contrôler le positionnement radial des territoires chromosomiques. En effet, en comparant les positions radiales des chromosomes 18 et 19 qui sont semblables en taille, il est observé que le chromosome 19, qui est trois fois plus denses en gènes que le chromosome 18, est plus central [12,35].

Cependant, l'agencement des territoires chromosomiques varie d'un type cellulaire à l'autre [36]. Il a également été montré que la position radiale est corrélée à la densité génique dans le cas de cellules sphériques [35]. Dans le cas contraire, la position radiale est plus corrélée à la taille des chromosomes, avec les petits chromosomes au centre et les grands chromosomes en périphérie [32,37].

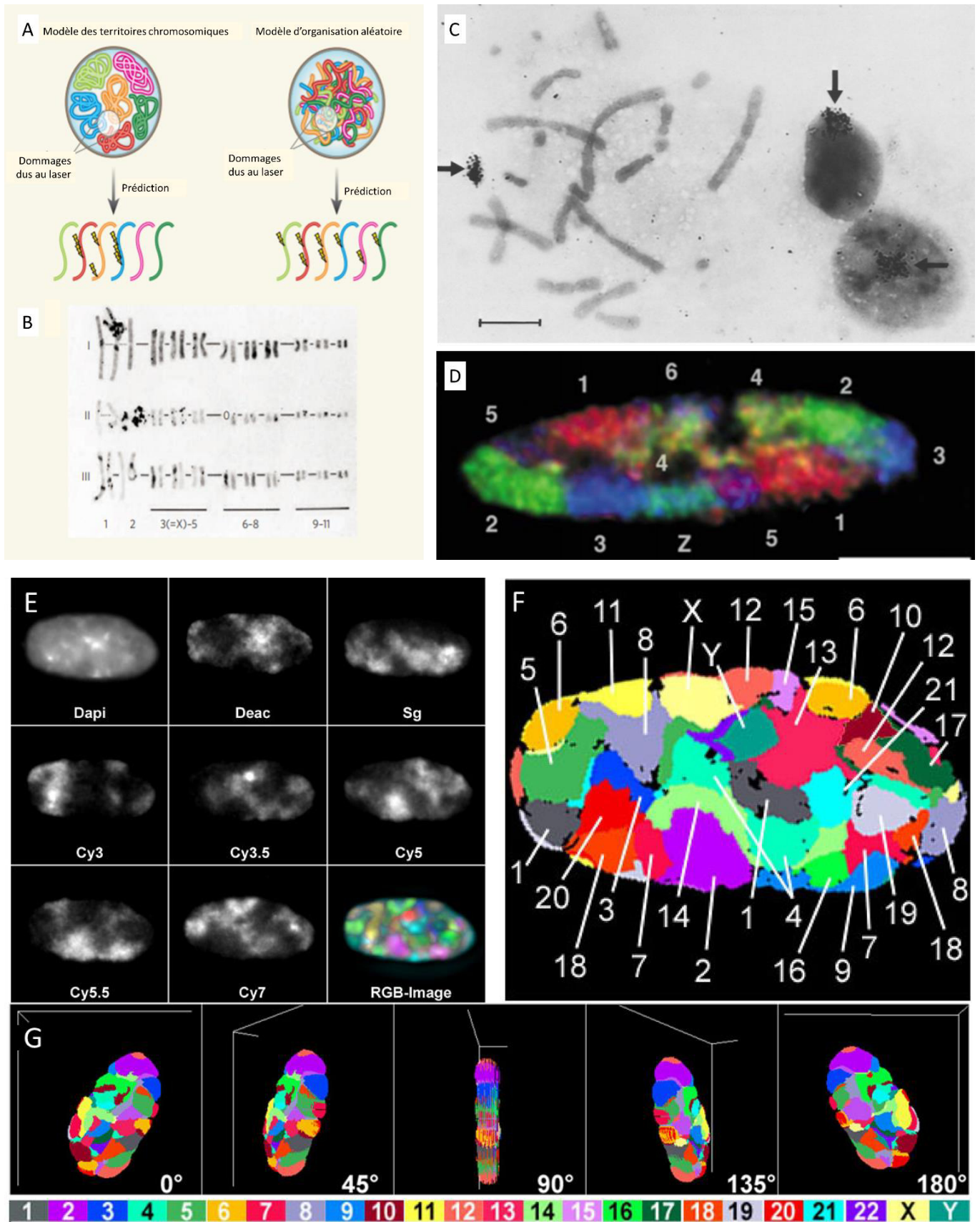


Figure 7. Visualisation des territoires chromosomiques

A et B. Des dommages localisés sont réalisés à l'aide d'un microfaisceau laser, adapté de Meaburn et Misteli, 2007, l'image B est reprise de Zorn et al., 1979. A. Dans le cas d'un modèle avec des territoires chromosomiques, cela n'atteint qu'un petit groupe de chromosomes, alors que dans le cas d'un modèle d'organisation aléatoire,

cela concernerait tous les chromosomes. B. Les trois caryotypes de cellules de hamster ayant subi des microlésions montrent que seul un petit nombre de chromosomes a été endommagé. Cela confirme donc l'existence de territoires chromosomiques.

C. Autoradiographie montrant une plaque en métaphase et deux noyaux en interphase de la lignée cellulaire hybride 29-11B (un seul chromosome humain, le chromosome X, dans une lignée de hamster chinois) après hybridation *in situ* avec de l'ADN génomique humain marqué au tritium. Le chromosome X humain est fortement marqué en métaphase (indiqué par la flèche). Les noyaux en interphase présentent une région (indiquée par les flèches) fortement marquée. Le chromosome X est donc organisé au sein d'un territoire chromosomique. La barre représente 10 μm . Adapté de Schardin et al., 1985.

D. Visualisation des territoires chromosomiques dans un noyau de fibroblaste de poulet (contenant 7 paires de chromosomes), en hybridation *in situ* en fluorescence. Adapté de Harbermann, 2001.

E, F, G. Visualisation et représentation des territoires chromosomiques, d'après Bolzer et al., 2005. E. Plan d'une image de 3D-FISH avec la fluorescence du DAPI et de 7 fluorochromes. F. Représentation des territoires chromosomiques. G. Reconstruction 3D des territoires chromosomiques vue sous plusieurs angles.

L'espace libre de chromatine entre les territoires chromosomiques est appelé compartiment interchromatinien (Figure 8). Les forces électromagnétiques font que les fibres chromatiniennes se repoussent [38]. Cet espace, allant de quelques nanomètres à plusieurs micromètres, permet la décondensation des boucles chromatiniennes permettant ainsi la transcription et le passage des molécules d'ARN produites vers les pores de la membrane nucléaire [39]. L'espace interchromatinien comprend des domaines non chromatiniens, les corps nucléaires comme les corps de Cajal, les corps d'épissage, les corps polycombs, les foyers de transcription et les corps PML (que nous détaillerons par la suite). La régulation et la transcription des gènes n'est possible que si la chromatine est en contact avec l'espace interchromatinien. Il a été observé que la transcription et l'épissage ont lieu à périphérie mais également à l'intérieur des territoires chromosomiques, grâce à des canaux de compartiment interchromatinien passant à l'intérieur des territoires chromosomiques [37].

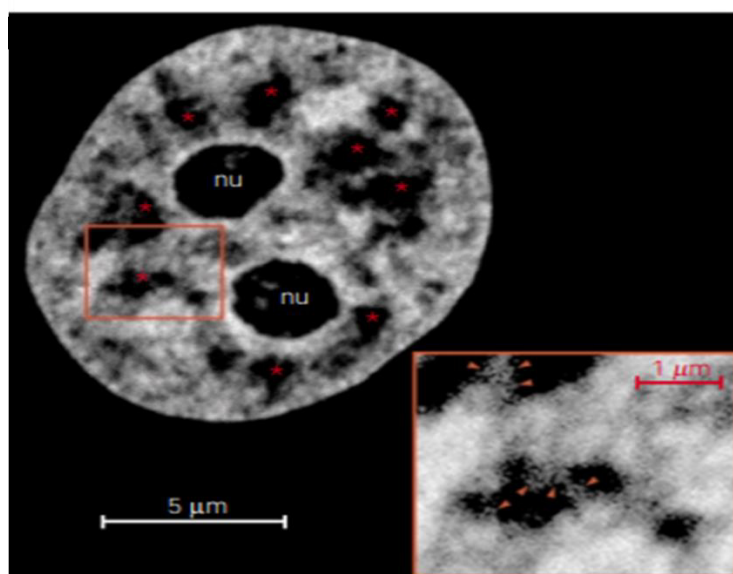


Figure 8. Compartiment interchromatinien, adaptée de Cremer et Cremer, 2001.

La chromatine couplée à la GFP (green fluorescent protéine) est visualisée en blanc pour les zones de haute densité, en gris pour les zones de faible densité et en noire pour les deux nucléoles et le compartiment interchromatinien.

La chromatine de chaque chromosome occupe un territoire nucléaire assez délimité, que l'on appelle territoire chromosomique. Ceux-ci ne sont pas définis aléatoirement. Dans les cellules sphériques, les chromosomes riches en gènes ont tendance à avoir une position radiale plus centrale que les chromosomes pauvres en gènes. L'agencement des territoires chromosomiques varie d'un type cellulaire à l'autre. L'espace interchromatinien laissé libre entre ces territoires chromosomiques permet la régulation et la transcription des gènes, ainsi que l'acheminement de l'ARNm au niveau des pores nucléaires.

1.3.2 Compartiments chromosomiques

Les territoires chromosomiques sont connus depuis de nombreuses années, mais l'organisation plus fine de la chromatine est explorée depuis une dizaine d'années grâce aux techniques d'étude de la conformation de la chromatine comme le séquençage Hi-C.

Les premiers résultats de technique Hi-C, avec une résolution de l'ordre du mégabase, sépare le génome en deux compartiments appelés compartiments A et B représentant respectivement l'euchromatine et l'hétérochromatine [20].

Le compartiment A, composé d'euchromatine, est plus riche en gènes et a un niveau d'expression de ces gènes plus importants.

Les interactions ont lieu préférentiellement entre compartiments du même type, A avec A et B avec B.

On peut distinguer deux types d'hétérochromatine :

- L'hétérochromatine constitutive composée de l'ADN répété impliqué dans les télomères et centromères.
- L'hétérochromatine facultative où les gènes sont réprimés par désacétylation des histones. Il s'agit d'une structure dynamique et réversible. Cette hétérochromatine peut se décompacter et devenir transcriptionnellement active.

La compartimentalisation varie en fonction du type cellulaire, et également au cours du développement[40]. Il est également observé une différence au niveau de la réplication, avec une réplication plus précoce pour le compartiment A et plus tardive pour le compartiment B, avec des zones de transitions entre les compartiments[40,41] (Figure 9).

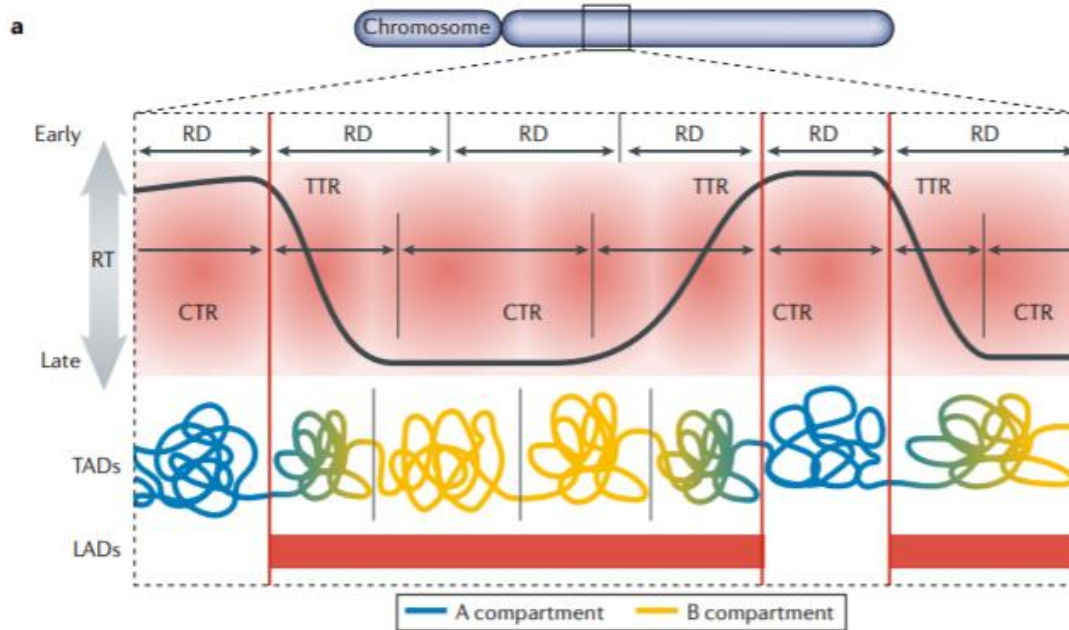


Figure 9 Timing de réplication en fonction de l'organisation tridimensionnelle de la chromatine, adapté de Marshall et al., 2019.

RD : domaine de réplication ; TTR : région de timing de transition ; CTR : région de timing constante ; TAD : domaines associés topologiquement ; LAD : domaines associés à la lamine.

Les régions de réplication précoce (early CTR) correspondent au compartiment A, les régions de réplifications tardives (late CTR) correspondent au compartiment B, et les régions TTR correspondent aux transitions entre les compartiments A et B. Les régions de réplication tardive et les TTR correspondent aux LAD.

La technique de Hi-C devient de plus en plus résolutive, et la technique de Hi-C haute résolution avec une résolution de 25kb permet de définir six sous compartiments avec des marques d'histones différentes[42]. Il est défini deux sous compartiments pour le compartiment A (A1 et A2) et quatre pour le compartiment B (B1, B2, B3 et B4).

Les deux sous compartiments A1 et A2 sont denses en gènes, fortement exprimés, ont les mêmes marques d'histones associées à l'activation (comme H3K27ac et H3Kame1) et sont déplétés en domaines associés à la lamine (LAD) et domaines associés au nucléole (NAD). Leur différence concerne la réplication, bien que précoce pour les deux sous compartiments, elle finit plus précocement pour le sous compartiment A1 que pour le sous compartiment A2. Le sous compartiment A2 contient également des gènes plus grands.

Le sous compartiment B1 correspond à de l'hétérochromatine facultative. Le sous compartiment B2 inclus majoritairement de l'hétérochromatine centromérique et est enrichi en LAD et NAD. Le sous compartiment B3 est enrichi en LAD mais fortement déplété en NAD. Le sous-compartiment B4

représente seulement 11Mb du chromosome 19 et comprend un groupe de régions contenant des gènes de la superfamille KRAB-ZNF (130 sur les 278 au total).

Le génome est séparé en deux compartiments distincts : les compartiments A et B représentant respectivement l'euchromatine et l'hétérochromatine. Les interactions sont privilégiées entre deux régions appartenant au même compartiment. La technique de Hi-C haute résolution a permis de retrouver six sous compartiments.

1.3.3 Topologically Associated Domains (TAD)

Grâce à l'augmentation de la résolution de la technique de séquençage Hi-C, des domaines à l'échelle du mégabase sont observés. Ces domaines sont appelés domaines associés topologiquement ou TAD pour *topologically associated domains* [43]. Il s'agit d'un domaine de contact préférentiel, les régions à l'intérieur d'un TAD interagissent préférentiellement avec d'autres régions du même TAD plutôt qu'avec d'autres régions du génome. Ces régions sont représentées par un triangle sur les matrices d'interaction de Hi-C. Les interactions chromatiniennes s'arrêtent brutalement aux deux extrémités qui représentent des régions frontières du génome. Ces frontières sont enrichies en CTCF, qui est un élément insulateur. Les bordures ont une forte interaction que l'on peut voir sur les cartes de Hi-C (« *corner peaks* », au sommet du triangle). Cela suggère la formation de boucles entre les différents sites de CTCF. Ces boucles sont presque toujours entre des sites CTCF avec une orientation convergente. Un modèle de « *loop extrusion* » permet d'expliquer la formation de ces TAD[44–47] : la chromatine est engagée dans un anneau de cohésine formant une boucle qui est extrudée jusqu'à ce que l'anneau de cohésine rencontre deux sites CTCF convergents (Figure 10). La suppression ou la modification d'orientation d'un site CTCF est suffisant pour supprimer ou décaler une frontière de TAD [48,49].

Une partie des TAD identifiés a des frontières communes avec celles des régions associées à la lamine (LAD) et des régions non LAD, et avec les compartiments A et B [43]. Les contacts entre enhancer et promoteur d'un gène sont restreints au niveau des TAD [50].

Initialement, l'étude des TAD a été réalisée sur des cellules souches embryonnaires murines et humaines, et sur des fibroblastes IMR90 par Dixon et al. Ils ont identifié 2200 TAD d'une taille moyenne de 880 kb occupant 91 % du génome de la souris. Avec les données de Hi-C haute résolution, des sous-TAD, d'une taille médiane de 185kb ont été déterminés dans 9 types cellulaires, humains et murins[42]. La taille et le nombre de TAD dépendent fortement de la résolution des données de Hi-C et de la méthode d'analyse[51]. Ces TAD sont stables entre différents types cellulaires [42,43] et fortement conservés entre les espèces[43].

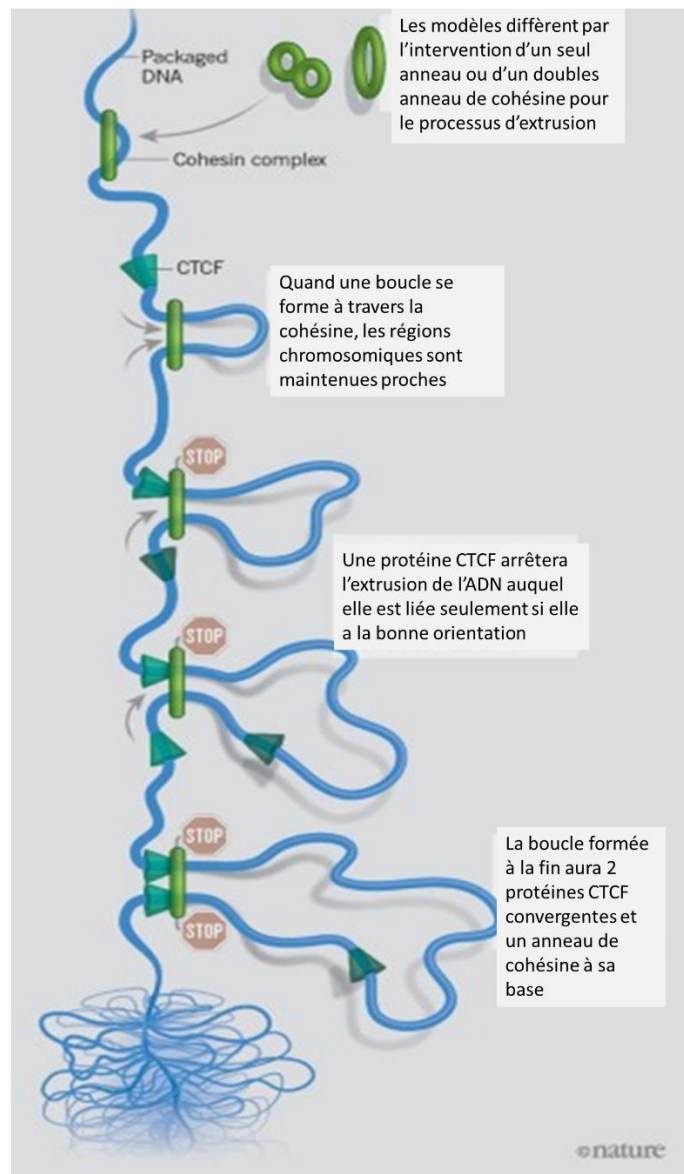


Figure 10. Formation des TAD par extrusion de l'ADN, adaptée de Dolgin, 2017

Les TAD sont formés par extrusion de l'ADN dans un anneau de cohésine, jusqu'à la rencontre de deux sites CTCF convergents.

Les TAD sont des régions de contact préférentiel délimitées par des insulateurs : des motifs CTCF convergents. La chromatine forme une boucle entre les motifs CTCF par l'action conjointe des CTCF et des anneaux de cohésine. Ils ont une taille moyenne de l'ordre du mégabase.

1.3.4 Boucles de chromatine

Les éléments de cis-régulation, comme les *enhancers* sont parfois à une distance relativement importante (plusieurs dizaines de kilobases) de leur gène cible. Une boucle de chromatine peut alors permettre de rapprocher spatialement cet enhancer du promoteur de son gène cible.

Un exemple connu est celui du locus de la β -globine pour lequel des contacts préférentiels sont observés grâce à la technique de 3C entre le LCR (Locus Control Region) et le gène *HBB* dans des cellules érythroïdes [52]. La formation de ces boucles est spécifique et dépendante du type cellulaire, les autres cellules (cellules non érythroïdes) n'exprimant pas de β -globine ne présentent pas cette interaction[53].

Il peut également y avoir la formation de boucles entre un gène et un *silencer*.

Les boucles peuvent également permettre les contacts entre régions régulatrices, comme par exemple le locus *IGF2/H19* où il existe un regroupement des éléments régulateurs. Une méthylation différentielle est mise en place en fonction de l'origine parentale de l'allèle afin que la conformation entraîne l'éloignement de l'*enhancer* du gène *IGF2* sur l'allèle maternel et du gène *H19* sur l'allèle paternel.

Les éléments de réponses au complexe Polycomb (PRE), où se fixent les complexes polycomb, se regroupent dans le noyau afin de former un environnement répressif ce qui entraînent la répression des gènes liés [54,55].

La plupart des boucles sont petites, elles rapprochent des régions espacées de moins de 2Mb, et fortement conservées entre les différents types cellulaires et entre les espèces[42].

Les boucles chromatiniennes permettent de rapprocher des régions génomiques espacées de quelques dizaines voire centaines de kb afin de permettre la régulation de la transcription des gènes. Les contacts promoteur/enhancer sont associés à l'activation des gènes, et les PRE se regroupent afin de créer un micro-environnement répressif.

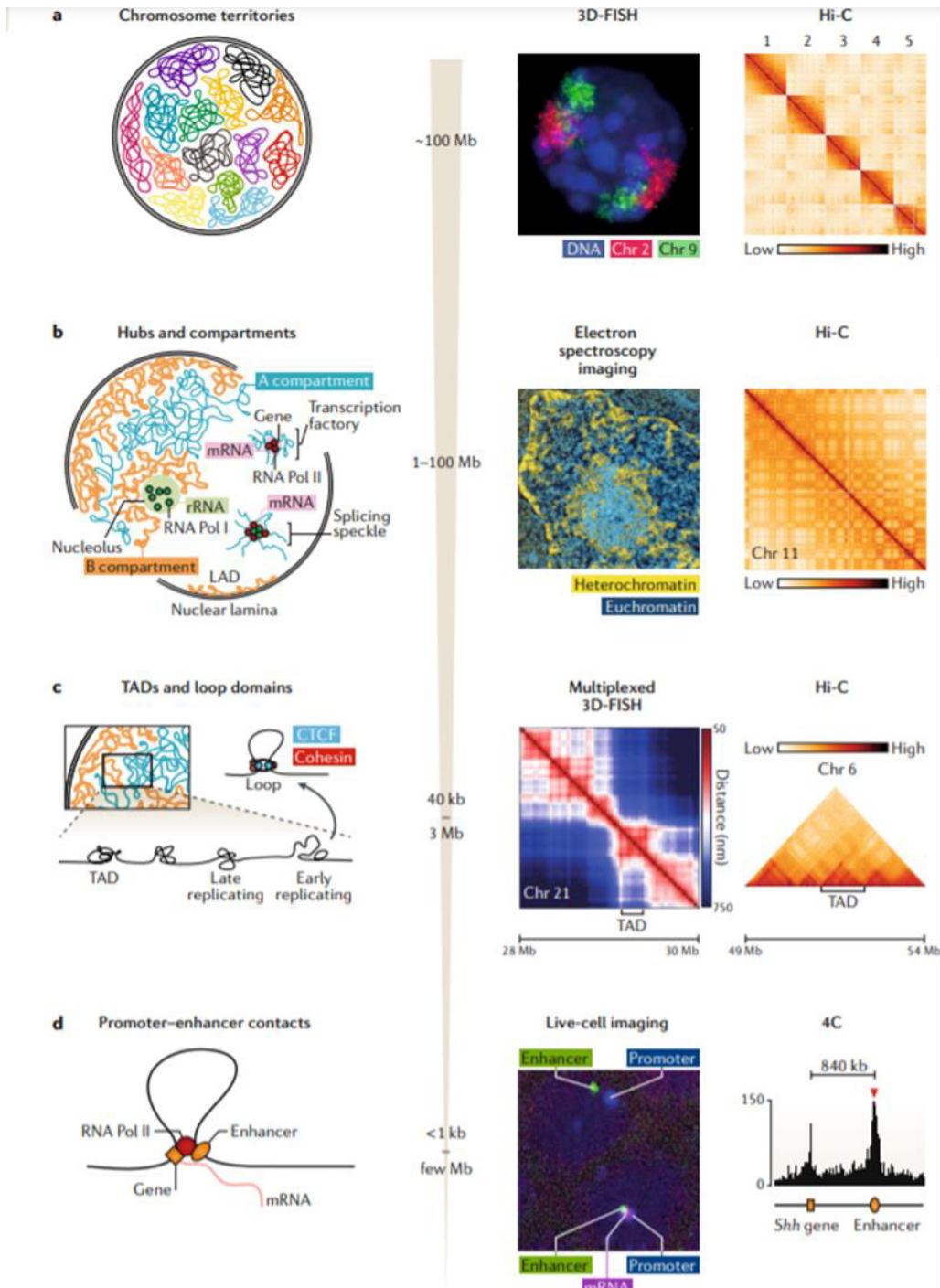


Figure 11. Architecture de la chromatine à différentes échelles, avec les méthodes d'étude, adaptée de Kempfer et Pombo, 2020

En résumé, le développement des techniques moléculaires, notamment le Hi-C, a permis de déterminer l'architecture de la chromatine et les interactions ayant lieu à différents niveaux d'organisation (Figure 11). Les territoires chromosomiques, qui avaient déjà été mis en évidence grâce aux techniques d'imagerie ont pu être confirmés. Puis les compartiments A et B ont pu être mis en évidence, suivi des TAD, grâce à la progression des techniques et l'augmentation de leur résolution.

1.4 Architecture du noyau

Le noyau, lieu de la réplication et de la transcription, est un organe cellulaire structuré.

L'expression des gènes est soumise à différentes régulations. Tout d'abord au niveau du génome lui-même, avec la présence de séquences promotrices. Puis au niveau épigénétique avec par exemple les modifications des histones ou le positionnement nucléaire. En effet il existe de nombreux domaines nucléaires qui peuvent interagir sur l'expression génique.

1.4.1 Enveloppe nucléaire

Le génome humain est contenu dans un compartiment séparé : le noyau, qui est protégé par l'enveloppe nucléaire (Figure 12). L'enveloppe nucléaire est composée d'une double bicouche phospholipidique. La membrane externe, en contact avec le cytoplasme, est en continuité avec la membrane du réticulum endoplasmique rugueux. La membrane interne, en contact avec le nucléoplasme, est liée aux lamines, un réseau de filaments intermédiaires. Entre ces deux membranes, il y a l'espace périnucléaire. Ces deux membranes se rejoignent autour des pores nucléaires, qui sont des gros complexes transmembranaires composés de plus de 500 nucléoporines[56]. Ces canaux aqueux permettent le passage de petites molécules entre les compartiments cytoplasmique et nucléaire. Les lamines sont également connectées avec les pores nucléaires. La chromatine est liée aux lamines et à la membrane interne.

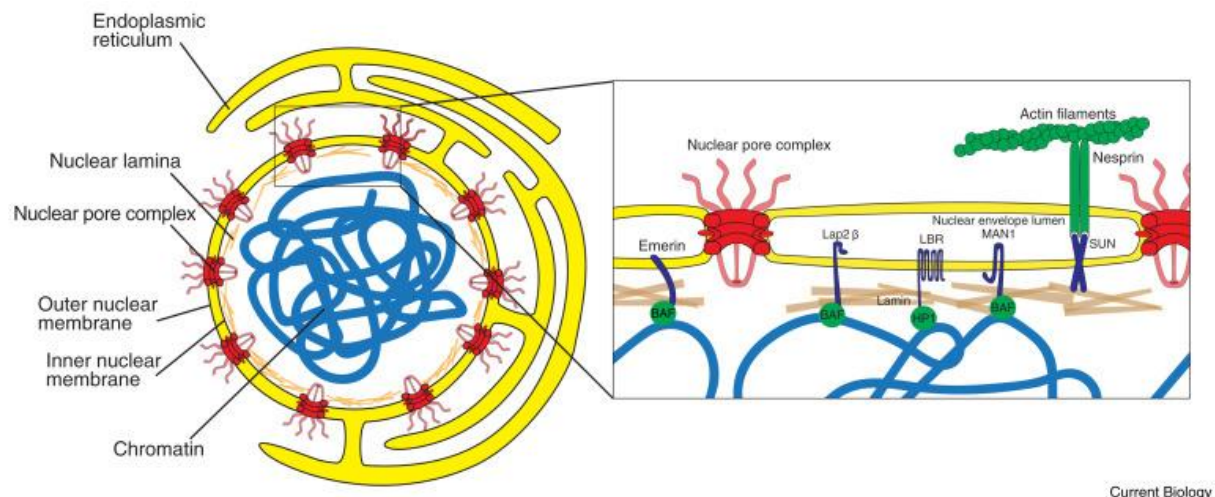


Figure 12. Représentation de la membrane nucléaire, d'après Magistris et al., 2018

1.4.2 Compartiments internes

Le noyau contient des corps nucléaires, avec des rôles plus ou moins connus. Ces corps nucléaires ne sont pas entourés par une membrane, ils sont définis par l'accumulation de protéines, voire d'acides nucléiques. Ils compartimentent le noyau afin de créer des sites distincts de concentration de réactifs nécessaires à certaines réactions. Seul le nucléole est visible en microscopie électronique, les autres corps nucléaires sont identifiés grâce à des techniques d'immunofluorescence. Certains compartiments sont associés à l'expression des gènes et d'autres à leur répression.

Les corps de Cajals

D'un diamètre maximum de 1,5 μ m, ils peuvent être au nombre de 10 dans le noyau. Ils sont impliqués dans l'assemblage et la maturation des petites particules ribonucléo-protéiques nucléaires (snRNP). Ils sont également impliqués dans le contrôle de la longueur des télomères[57].

1.4.2.1 Compartiments actifs

Usines de transcription

Ce sont des sites discrets, présentant une accumulation locale d'ARN polymérase II (Figure 13). Ils contiennent l'ensemble de la machinerie moléculaire nécessaire pour l'expression des gènes. L'ADN passe à travers l'usine et est transcrit grâce aux polymérases[58]

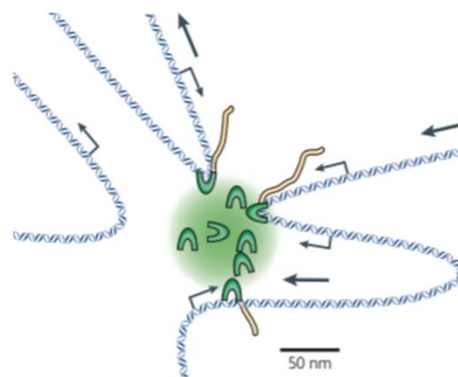


Figure 13. Usines de transcription, d'après Maland et al., 2009

Des régions de transcription active (d'un même chromosome ou de chromosomes différents) sont rassemblées au niveau d'une usine de transcription, riche en ARN polymérase II (croissants verts). Les filaments jaunes représentent l'ARN transcrit.

Corps d'épissage ou speckls

Il s'agit d'un des corps nucléaires les plus nombreux, jusqu'à plusieurs dizaines par noyau[59]. Ce sont les deuxièmes corps nucléaires les plus volumineux après le nucléole. Ils sont présents au niveau

central du noyau et exclus de la périphérie[60]. Ils sont riches en facteurs d'épissage (snRNP, protéines SR)[59,61] (Figure 14). Mais ils contiennent également de nombreux facteurs (p300/CBP, pTEF-beta) impliqués dans la production d'ARNm par l'ARN polymérase II[59,61]. De plus, les régions chromosomiques situées proches des speckls ont un nombre plus importants de gènes, avec une expression génique plus importante[60]. La périphérie des corps d'épissage agirait comme une plaque tournante centrale pour coordonner plusieurs étapes de l'expression des gènes[61].

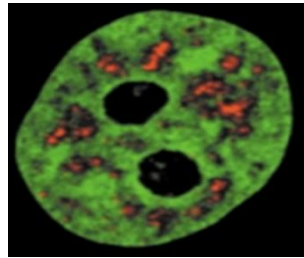


Figure 14. Visualisation des corps d'épissage, adapté de Cremer et Cremer, 2001.

Visualisation de chromatine couplée à la GFP (en vert) et des corps d'épissage marqués avec des anticorps dirigés contre les non-snRNP (en rouge). Les corps d'épissage sont au niveau du compartiment interchromatinien.

Les corps PML

Les corps PML (protéine proleucémie myéloïde) sont des structures sphériques de 0,1 à 1µm de diamètre, au nombre de 5 à 20 par noyau, organisées par la protéine PML[62]. La transcription du gène PML est contrôlée par les interférons α , β et γ , ainsi que par la protéine p53. La protéine PML subit des modifications post traductionnelles, dont des phosphorylations et sumoylations, impliquées dans la biogénèse des corps PML. Les dommages à l'ADN et les kinases induites par le stress phosphorylent les PML régulant leur stabilité. Ils sont impliqués dans le recrutement de nombreuses protéines, intervenant dans la transcription, la sénescence, l'apoptose et la défense anti-virale.

1.4.2.2 Compartiments répressifs

Nucléole

Le nucléole fut le premier domaine nucléaire identifié dans le noyau (Fontanan 1781), étant le plus gros et le seul visible en microscopie électronique. Il a une taille allant jusqu'à plusieurs micromètres. Il est le lieu de synthèse des ARN ribosomiques et d'assemblage des ribosomes. Il s'agit du regroupement des régions organisatrices du nucléole (NOR) qui contiennent des répétitions en tandem des gènes ribosomiaux. Ces NOR sont situés sur les bras courts des chromosomes acrocentriques (chromosomes 13, 14, 15, 21 et 22).

La périphérie du nucléole est enrichie en hétérochromatine constitutive. La chromatine associée au nucléole, NAD pour nucleolus associated domains, a des caractéristiques semblables aux LAD.

Corps polycomb

Les interactions chromatinienne à longue distance ne sont pas limitées à l'activation de l'expression des gènes, des régions se regroupent également autour des corps polycomb qui entraînent la répression des gènes. Les protéines polycomb sont des régulateurs épigénétiques qui forment des complexes, Polycomb repressive complexe 1 (PRC1) et 2 (PRC2), qui sont responsables de la méthylation de la lysine 27 de l'histone H3 (H3K27me3) entraînant la compaction de la chromatine et donc la répression des gènes. Les protéines polycomb se fixent au niveau des éléments PRE (PcG responsive element) des gènes cibles.

Lamines nucléaires

Les lamines, qui sont des filaments intermédiaires, sont situés juste sous l'enveloppe nucléaire. Elles sont divisées en deux classes, les lamines A et les lamines B. Les lamines A comprennent les lamines de type A et C qui sont issues du gène *LMNA* par épissage alternatif. Les lamines B comprennent les lamines de types B1 et B2 qui sont codées respectivement par les gènes *LMNB1* et *LMNB2*.

Les lamines A, B et C s'assemblent pour former la lamina nucléaire qui est un réseau protéique fibreux qui tapisse la membrane interne du noyau. Elle forme une couche de 10 à 20nm et est interrompue par les pores nucléaires. La lamina confère une stabilité à l'enveloppe nucléaire et lui donne sa forme. Elle se dépolymérise lors de la mitose et se reforme en fin de mitose. Elle entre également en interaction avec la chromatine, déterminant alors une position périphérique pour celle-ci. Les régions associées à la lamina ou LAD (lamina associated domains) sont peu exprimées. Une expérience de Dam ID consistant en l'introduction dans les cellules en culture d'une protéine chimérique composée d'une adénine méthyl transférase (Dam pour DNA adénine méthylase) et de la protéine d'intérêt, la lamine B1, montre que la lamine correspond à un environnement répressif[63]. En effet les portions du génome devenant méthylées et donc associées à la lamine B1 sont pauvres en gènes, ne possèdent pas d'ARN polymérase II et ont la signature épigénétique d'une chromatine répressive (beaucoup de marques H3K9me2/3 et de H3K27me3 et peu de marques H3K4me2). La détection de ces zones de méthylation permet d'identifier les LAD et de réaliser des cartes de LAD. Chez l'homme, il a été retrouvé 1300 LAD de 0,1 à 10 MB, couvrant environ 40 % du génome [63]. Comme les TAD, les LAD sont délimités par des sites de liaison CTCF [63]. La densité en gènes dans les LAD est deux fois plus faibles que dans le reste du génome et les gènes présents sont 5 à 10 fois moins exprimés que les autres[63]. Certains LAD sont spécifiques d'un type cellulaire, ce sont des LAD facultatifs, tandis que

d'autres sont constitutivement associés à la lamine comme l'hétérochromatine centromérique et certaines régions télomériques, ce sont les LAD constitutifs (cLADs)[64]. Nous pouvons prendre comme exemple les loci *IgH* (chaîne lourde d'immunoglobuline) et *IgK* (chaîne légère kappa de l'immunoglobuline). Ils ont une localisation périphérique dans les cellules n'appartenant pas à la lignée B et une localisation plus centrale dans les lignées B[65]. La taille et la position génomique des LAD constitutives sont fortement conservées entre l'homme et la souris[64].

1.5 Fonctionnement du génome en lien avec l'organisation et l'architecture du noyau

Nous avons pu voir que la chromatine est organisée en trois dimensions dans l'espace nucléaire afin de pouvoir permettre l'expression des gènes et la réplication de l'ADN. La position radiale des territoires chromosomiques n'est pas aléatoire [66–69] ce qui suggère que les interactions entre plusieurs territoires chromosomiques ne sont pas aléatoires [70].

La position radiale est corrélée à la densité en gène dans les cellules sphériques, et à la taille des chromosomes pour les cellules non sphériques [37,70]. D'autres facteurs que la taille et la densité génique sont impliqués dans l'organisation fonctionnelle du génome. Il peut s'agir d'interaction avec des compartiments nucléaires comme les régions NOR des chromosomes acrocentriques avec le nucléole [70,71]. La position périphérique dans le noyau est associée à une répression de la transcription des gènes. Nous pouvons prendre comme exemple le chromosome X inactif. Celui-ci a une position plus périphérique que le chromosome X actif [65].

Plusieurs expériences de relocalisation à la périphérie par ancrage artificiel à la lamine ont été réalisées afin d'explorer le lien entre le positionnement en périphérie du noyau et la diminution de l'expression due à l'environnement répressif [72–74]. Il s'agit d'expérience utilisant un système inductible pour relocaliser un fragment chromosomique de l'intérieur du noyau vers la périphérie. Un transgène contenant de nombreux sites de l'opérateur lac (*lacO*) est intégré au génome, dans une région cible. Il est relocalisé par l'interaction avec une construction comprenant la protéine *lacI* (répresseur lac, qui se fixe au niveau des sites *lacO*) fusionnée à une protéine de la membrane interne du noyau comme l'emerine[74], la lamine B1[72] ou *lap2β*[73]. Les différentes publications ne vont pas dans le même sens. Si certaines équipes ont démontré, après relocalisation périphérique, que l'expression de certains gènes à proximité des sites *LacO* était diminuée[73,74], avec même une répression de certains gènes plus éloignés des sites *LacO*[73], certaines ne trouvent pas de différence au niveau du profil d'expression entre la localisation interne et la relocalisation en périphérie[72].

Il a été montré qu'il existe un modèle dans lequel un gradient d'expression génique s'étend de la périphérie vers le centre du noyau [66]. Ce modèle a été affiné par la suite et serait plus complexe, dû au positionnement du génome en fonction des compartiments nucléaires [60] (figure 15).

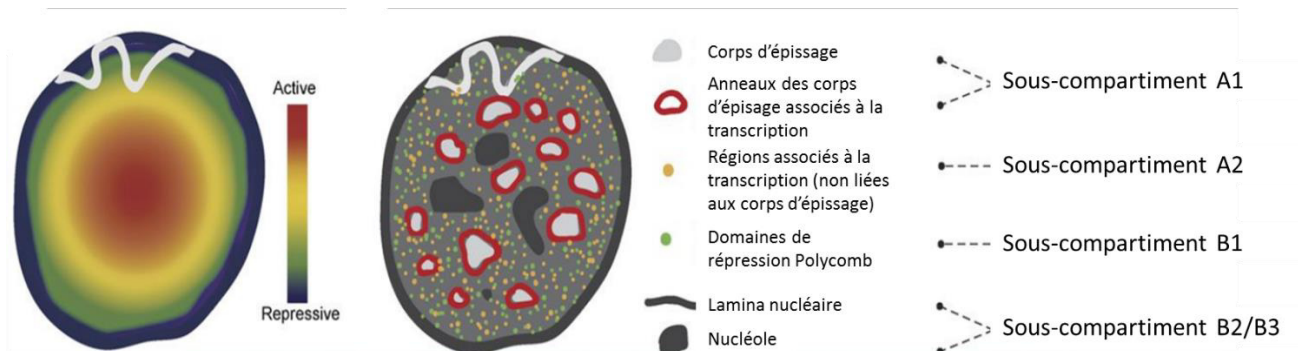


Figure 15. Gradient nucléaire d'expression génique, adapté de Chen et al., 2018

A gauche l'image représente le modèle dans lequel l'activité transcriptionnelle augmente radialement de faible près de la périphérie nucléaire à élevée vers le centre nucléaire. La représentation à droite représente le nouveau modèle affiné : les régions chromosomiques transcriptionnellement inactives sont proches de la périphérie nucléaire et nucléolaire (noir). A l'intérieur du noyau se trouvent des « hot spot » de transcription, qui se localisent à la périphérie des corps d'épissage et à proximité des usines de transcription.

2. Anomalies chromosomiques

Notre étude portant sur l'impact des remaniements chromosomiques sur l'organisation 3D du génome, l'expression génique et la méiose, nous allons présenter dans cette partie les différents types d'anomalies chromosomiques, en se concentrant sur les anomalies de structures. Puis nous parlerons de remaniements de structure extrêmement complexes, les chromoanagenesis. Dans un troisième chapitre, nous aborderons les mécanismes de réparation des cassures doubles brins, lesquels peuvent être à l'origine de remaniements de structure. Enfin nous verrons les méthodes d'identification de ces remaniements chromosomiques.

2.1 Les types d'anomalies chromosomiques

Il existe deux grands types d'anomalies chromosomiques, les anomalies de nombres qui sont dues à des problèmes de non-disjonction méiotique et les anomalies de structure qui sont les conséquences de cassures doubles brins mal réparées par les mécanismes de réparation.

2.1.1 Les anomalies chromosomiques de nombre

Les anomalies chromosomiques de nombre affectent le nombre de chromosome. Les aneuploïdies sont définies par un chromosome surnuméraire, on parle alors de trisomie (les trisomies viables à l'état homogène étant les trisomies 13, 18 et 21) ou la perte d'un chromosome, on parle alors de monosomie

(non viable, sauf la monosomie X chez la femme). Ces anomalies sont souvent dues à des problèmes de non-disjonction méiotique, c'est-à-dire lorsque les deux chromosomes homologues (lors de la première division) ou les deux chromatides sœurs (lors de la deuxième division) ne se séparent pas dans les deux cellules filles générant alors un gamète avec deux chromosomes homologues, qui après fécondation donnera un zygote avec trois chromosomes (trisomie) et un gamète sans chromosome qui après fécondation donnera un zygote avec un seul chromosome (monosomie).

On peut observer aussi des polyploïdies, qui correspondent à un nombre anormal de lots haploïdes entiers. Par exemple la triploïdie avec trois lots de chromosomes donc 69 chromosomes. La triploïdie peut être liée à la fécondation d'un ovocyte par deux spermatozoïdes ou à la non-expulsion du deuxième globule polaire.

Ces anomalies aboutissent à des morts *in utero* quand cela n'est pas viable, ou à des syndromes génétiques avec des déficiences intellectuelles et/ou des malformations.

2.1.2 Les anomalies chromosomiques de structure

Les anomalies chromosomiques de structure sont des remaniements qui peuvent affecter un seul ou plusieurs chromosomes. Elles sont la conséquence de cassures double brin suivies d'un recollement anormal dû à une mauvaise réparation. Cela peut aboutir ou non à la perte de matériel chromosomique. Le réarrangement est dit équilibré s'il ne s'accompagne ni de perte ni de gain de matériel chromosomique. Dans le cas contraire il est déséquilibré, on parle alors de variation de nombre de copies (*Copy Number Variation* ou CNV).

2.1.2.1 Les anomalies déséquilibrées

Les anomalies chromosomiques déséquilibrées se manifestent comme des CNV. Ces anomalies peuvent être *de novo*, ou peuvent résulter d'un déséquilibre d'un remaniement parental équilibré (voir chapitre 3.4 Conséquences des remaniements chromosomiques sur la méiose).

Certaines de ces anomalies peuvent être la conséquence d'une recombinaison homologue non allélique (NAHR, *non allelic homologous recombination*) au moment de la mitose ou de la méiose. Il y a alors recombinaison entre deux répétitions segmentaires (appelés *low copy repeats* LCR) entraînant une duplication ou une délétion du fragment d'ADN compris entre les deux LCR. Il s'agit du mécanisme de survenue des remaniements récurrents (duplication ou délétion récurrentes), dont le plus fréquent est la délétion 22q11.2.

Pour les CNV avec perte de matériel, on peut observer des délétions terminales, impliquant alors un seul point de cassure avec perte du fragment terminal ; des délétions interstitielles, impliquant deux

points de cassures avec perte du fragment situé entre ces deux points de cassures ; des chromosomes en anneaux résultant d'une cassure à chaque extrémité du chromosome avec pertes des fragments terminaux et recollement du chromosome formant un anneau.

Pour les CNV avec gain de matériel, on peut observer des duplications en tandem (fragment dupliqué orienté dans le même sens) ou en miroir (fragment dupliqué orienté en sens inverse), et des marqueurs (petit chromosome surnuméraire).

On peut également observer des isochromosomes qui sont des chromosomes anormaux formés de deux bras longs ou deux bras courts. On observe alors une perte des bras courts associé à un gain des bras longs (ou inversement).

Ces CNV sont visibles au caryotype s'ils sont supérieurs à 5 à 10Mb. Pour les CNV plus petits on parle de microremaniement. La méthode de prédilection pour les retrouver est l'ACPA (Analyse chromosomique sur puce à ADN). Ils peuvent également être recherchés spécifiquement en FISH.

Ces anomalies sont le plus souvent responsables d'un phénotype, qui dépend de la fonction des gènes qu'elles contiennent.

2.1.2.2 Les anomalies équilibrées

Les remaniements chromosomiques apparemment équilibrés (RCAE) regroupent les translocations, les inversions, les insertions et les remaniements complexes sans perte ni gain de matériel chromosomique. Concernant les remaniements visibles au caryotype, dont la résolution est limitée, le terme apparemment équilibré est utilisé car un déséquilibre cryptique associé ne peut pas être éliminé.

Les translocations

On peut distinguer deux sortes de translocations : les translocations réciproques et les translocations robertsonniennes. Ces translocations peuvent être héritées d'un parent ou survenir *de novo*.

Les translocations réciproques impliquent deux points de cassure (un sur chaque chromosome) et résultent en un échange des fragments terminaux entre deux chromosomes non homologues (Figure 16.1)

Les translocations robertsonniennes impliquent la totalité des bras longs des chromosomes acrocentriques, à savoir les chromosomes 13, 14, 15, 21, 22. Il s'agit le plus souvent de cassure juxtacentromérique dans les deux bras courts, avec accolement des bras longs et perte des bras courts de deux acrocentriques, induisant un caryotype équilibré à 45 chromosomes. (Figure 16.2)

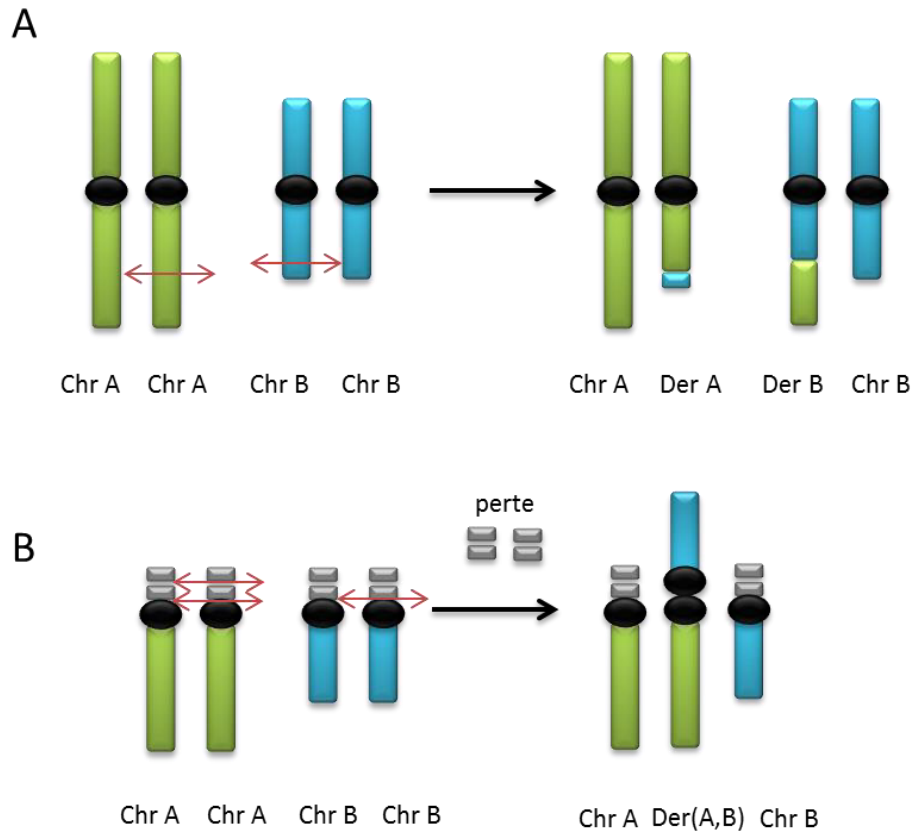


Figure 16. Représentations d'une translocation réciproque (A) et d'une translocation robertsonienne (B) entre deux chromosomes (Chr) A et B

Cette translocation aboutit à la formation de dérivés (Der). La translocation robertsonienne implique la perte des bras courts des chromosomes acrocentriques (séquences répétées). Les doubles flèches rouges indiquent les points de cassures.

Les inversions

Les inversions impliquent deux points de cassure au niveau du même chromosome avec retournement du fragment intermédiaire. Les inversions sont péricentriques si les points de cassure sont de part et d'autre du centromère. Dans le cas où les deux points de cassures sont sur le même bras chromosomique on parle d'inversion paracentrique (Figure 17).

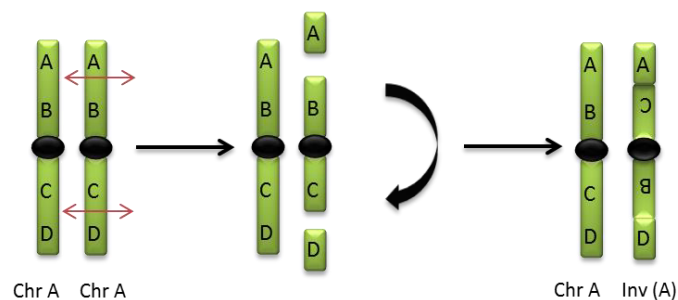


Figure 17. Représentation d'une inversion péricentrique du chromosome A (Chr A)

On obtient un dérivé d'inversion : inv(A). Les doubles flèches rouges indiquent les points de cassures.

Les insertions

Les insertions impliquent trois points de cassure, deux points de cassure sur un chromosome donneur avec excision d'un fragment chromosomique et un point de cassure sur un autre ou sur le même chromosome (accepteur) aboutissant à l'insertion de ce fragment (Figure 18).

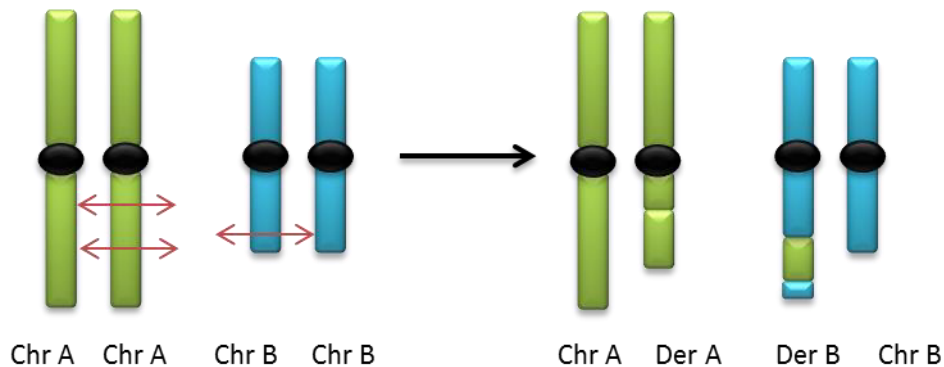


Figure 18. Représentation d'une insertion d'un fragment du chromosome A (Chr A) dans le chromosome B (Chr B).

On obtient un dérivé du chromosome B (der B) contenant le fragment du chromosome et un dérivé du chromosome A (der A) qui n'a plus ce fragment. Les doubles flèches rouges indiquent les points de cassures.

Les remaniements complexes

Historiquement, les remaniements complexes étaient les remaniements comportant plus de trois points de cassures et impliquant au moins deux chromosomes (insertion exclue)[75]. En 2011, Stephens et al., décrivent un patient présentant un réarrangement extrêmement complexe avec 42 points de cassure regroupés sur le bras long du chromosome 4 d'un patient atteint de leucémie lymphoïde chronique [76]. La définition du remaniement complexe a alors été revue. Un remaniement complexe est défini comme un remaniement pour lequel le conseil génétique est complexe. Il s'agit donc des remaniements qui ne sont pas des translocations, des inversions ou des insertions. Les remaniements complexes, et plus précisément les chromoanagenesis, seront détaillés davantage dans le chapitre « 2.2 Les chromoanagenesis ».

Conséquences

Ces RCAE sont retrouvés dans 1,54 naissances pour 1000 [77]. En général ils sont associés à un phénotype normal, cependant on observe un phénotype anormal chez 9% des personnes avec une inversion et 6% des personnes avec une translocation [78]. En cas de remaniement complexe, un phénotype anormal est présent dans 23% des cas [79]. En 2018, une étude a été menée afin de suivre

des patients présentant des RCAE en moyenne sur 17 ans. Cela leur a permis de réévaluer le risque de phénotypes neurodéveloppementaux anormaux à 27% [80].

Pour expliquer cela, trois hypothèses ont été soulevées : 1. un déséquilibre cryptique associé au remaniement (40% des cas) [81], 2. une pathologie du point de cassure [82] ou 3. une association fortuite. En ce qui concerne les pathologies du point de cassure, on peut observer une interruption du gène [83,84], une dérégulation de son expression par un effet de position [85] ou la formation d'un gène de fusion entraînant la formation d'une protéine de fusion (très décrit en cancérologie).

Les RCAE peuvent également entraîner des anomalies d'appariement des chromosomes homologues au moment de la méiose (voir chapitre 3.4 Conséquences des remaniements chromosomiques) sur la méiose, à l'origine de troubles de la reproduction ou de déséquilibre dans la descendance.

2.2 Les chromoanagenesis

2.2.1 Découverte des chromoanagenesis et classification

Le terme chromoanagenesis (« chromo » pour chromosome et « anagenesis » pour renaissance) regroupe trois types de remaniements : les chromothripsis, les chromoanasythesis et les chromoplexies, qui ont été découverts séquentiellement, grâce à l'apport des stratégies de séquençage haut débit.

En 2011 a été utilisé pour la première fois le terme « chromothripsis », venant du grec, « chromo » pour chromosomes et « thripsis » pour pulvérisation, concernant le remaniement extrêmement complexe retrouvé chez un patient atteint de leucémie lymphoïde chronique [76]. Ce patient avait un remaniement impliquant 42 points de cassure, concentrés sur le bras long du chromosome 4. Suite à plusieurs observations, les auteurs ont fait l'hypothèse d'un évènement unique, pendant lequel le chromosome 4 aurait éclaté, puis les fragments se seraient recollés entre eux mais certains auraient été perdus lors du processus. Cet évènement catastrophique unique est contraire à la théorie de l'accumulation de remaniements expliquant jusque-là le processus de cancérogenèse/leucémogénèse [86]. En effet, chez ce patient, la concentration des remaniements sur un bras chromosomique, et même parfois sur des régions très courtes au sein de ce bras chromosomique (7 réarrangements regroupés sur 30 kb), un nombre de copies égal soit à 2 (normal) soit à 1 (délétion), des fragments inversés et non inversés, des jonctions entre deux fragments éloignés, sont en faveur d'un évènement unique. Les études des fragments de jonction ont montré qu'il s'agissait de réparation de type jonction des extrémités non homologues (ou Non-Homologous End Joining, NHEJ) ou jonction des extrémités médiée par microhomologie (ou Microhomology Mediated End Joining, MMEJ). Par la suite, plus de 700 lignées cancéreuses et 2700 tumeurs ont été analysées, ce qui a permis d'évaluer la fréquence de

ce type de remaniement à 2 à 3% des cancers. La fréquence est encore plus élevée pour les tumeurs osseuses, avec une fréquence d'environ 25%.

Par la suite, un autre type de remaniement extrêmement complexe mais avec des gains de matériel a été décrit [87], dénommés « chromoanasythesis », « chromo » pour chromosome et « anasythesis » pour reconstitution. Cette fois-ci les études de fragments de jonction sont plutôt en faveur d'un mécanisme réplcatif, de type interruption de la fourche de réplication et changement de matrice (ou Fork Stalling and Template Switching, FoSTeS) ou réplication induite par la cassure médiée par microhomologie (Microhomology-mediated break-induced replication) sans pulvérisation.

En 2013, apparaît le terme de « chromoplexie » pour « pleko » qui signifie tisser ou tresser [88] , qui correspond à la formation de chromosomes chimériques. Ce mécanisme implique souvent la formation de gènes de fusion.

Tout d'abord décrits en cancérologie, des chromothripsis sont très rapidement décrits en constitutionnel [89]. Kloosterman et al., rapportent le cas d'un enfant avec une déficience intellectuelle sévère, un retard de langage et une cyphoscoliose chez lequel ont été trouvés par séquençage haut débit des remaniements chromosomiques, avec 12 points de cassures, impliquant trois chromosomes (1, 4 et 10).

2.2.2 Classification des chromoanagenesis

2.2.2.1 Les chromothripsis

Le chromothripsis résulte d'un éclatement, en général localisé, d'un ou plusieurs chromosomes en des dizaines voire des centaines de fragments [76,90] (Figure 19). Après les cassures doubles brins, les fragments sont réassemblés grâce à la jonction des extrémités non homologues (NHEJ) en majorité, ou grâce à la jonction des extrémités médiée par micro homologie (MMEJ). En effet, les fragments de jonction ont été beaucoup étudiés grâce à la réalisation de séquençage haut débit, ce qui a permis de retrouver dans une majorité des cas des petites délétions ou insertions, et parfois des microhomologies de quelques nucléotides au niveau des points de cassure. Ceci est en faveur d'une réparation par NHEJ ou MMEJ.

Lors du réassemblage, certains fragments peuvent être perdus. Les fragments sont recollés dans le sens normal ou inversé, sans préférence. On observe donc au final un chromosome remanié, avec des fragments qui sont réassemblés de manière aléatoire. Les chromothripsis peuvent être équilibrés ou non, avec la présence de délétions.

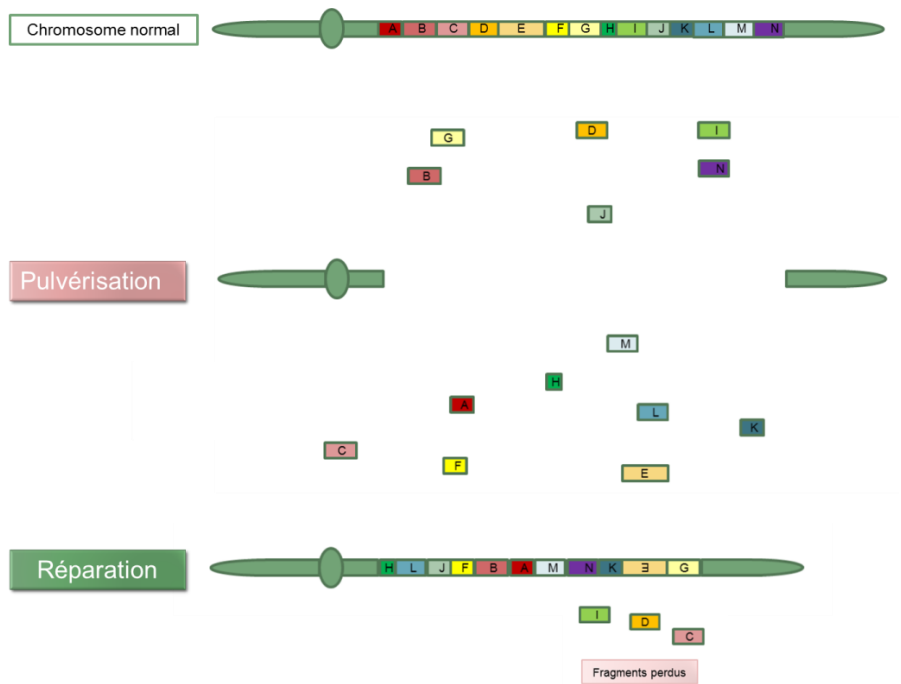


Figure 19. Chromothripsis

Pulvérisation du chromosome suivi d'un recollement aléatoire avec des fragments en orientation normale ou inversés. Certains fragments sont perdus.

2.2.2.2 Les chromoanasthésis

Les chromoanasthésis sont des remaniements qui intéressent en général un seul chromosome qui présente au moins cinq points de cassures [90]. Ces remaniements ne sont pas équilibrés et sont caractérisés par la présence de gains (duplications, triplications) et de pertes. On observe également au niveau des fragments de jonction l'insertion de courtes séquences homologues à une autre séquence du génome (sur le chromosome impliqué ou sur un autre) appelée *templated insertion*. Ces observations permettent d'impliquer le mécanisme de réparation par interruption de la fourche de réplication et de changement de matrice (FoSTes) (Figure 20). Et MMBIR ?

2.2.2.3 Les chromoplaxies

Les chromoplaxies ont été décrites uniquement en cancérologie [88]. Ce remaniement implique deux chromosomes ou plus. Il s'agit d'un remaniement en général équilibré, avec formation de chromosomes chimériques par jonction de plusieurs segments d'ADN de plusieurs chromosomes. Des délétions au niveau des fragments de jonction, après réparation par NHEJ ou MMEJ, ont pu être observées.

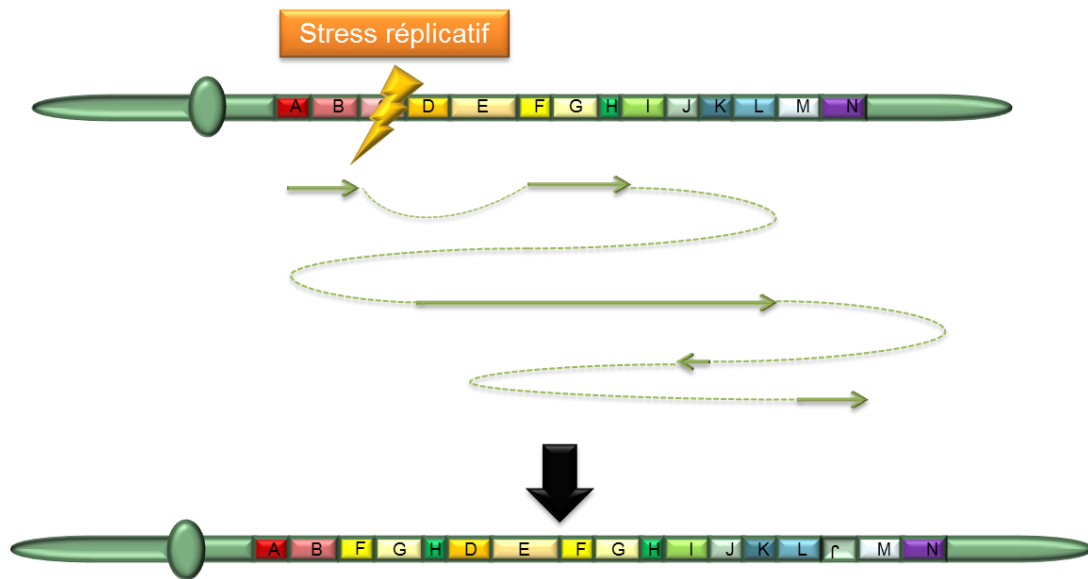


Figure 20. Chromoanasythesis

Un stress réplcatif (éclair jaune) induit un changement de matrice de réplcation. Ces changements peuvent être répétés et aboutir à un chromosome remanié.

2.2.3 Mécanisme de survenue

Les nombreuses cassures de l'ADN peuvent être causées par des facteurs environnementaux comme les radiations ionisantes et les radicaux libres, ou des facteurs endogènes comme le stress réplcatif qui peut être dû à des anomalies de l'activité de l'ADN polymérase et de la topoisomérase [91].

Les chromothripsis constitutionnels observés sont préférentiellement d'origine paternelle [92], comme c'est le cas pour les anomalies de structure [75]. On observe en effet une grande vulnérabilité aux cassures de l'ADN et une réparation limitée lors de la spermiogénèse. De plus les spermatogonies subissent des centaines de divisions mitotiques avant d'entrer en méiose. Le regroupement des points de cassure observé dans une majorité des chromothripsis permet de poser comme hypothèse que la pulvérisation a lieu au moment où les chromosomes sont condensés. L'initiation du chromothripsis pourrait alors avoir lieu au cours de l'une des trois étapes : lors d'une division pré-méiotique de la spermatogonie, lors de la première division méiotique, ou lors de la spermiogénèse (différenciation des spermatides ronds en spermatozoïdes) [92].

Différentes hypothèses ont été soulevées pour expliquer la survenue des chromoanagenesis comme l'apoptose incomplète [93] ou l'attrition télomérique [94]. L'hypothèse la plus probable aujourd'hui est la formation d'un micronoyau où un ou plusieurs chromosomes se retrouvent encapsulés dans une enveloppe nucléaire et pulvérisés suite à l'échec de la réplcation complète de l'ADN précédant la mitose [95,96].

La formation du micronoyau fait suite à un défaut de captation d'un centromère par le fuseau mitotique au moment de la division cellulaire [95]. Cette hypothèse a été exploré par l'équipe de Ly en 2017, en inactivant le centromère du chromosome Y [97]. Celui-ci a alors été séquestré dans un micronoyau et a donné lieu à un remaniement complexe du chromosome Y. L'environnement de ce micronoyau n'est pas propice au bon déroulement de la division cellulaire, car il est déficient en machinerie de réplication et de réparation de l'ADN [98]. Il y a alors deux hypothèses. Dans un premier cas, cet environnement va entraîner l'explosion du chromosome, qui après réintégration dans le noyau principal sera réassemblé. La deuxième hypothèse est que l'asynchronie entre le micronoyau et le noyau principal entraîne l'explosion chromosomique du chromosome séquestré [96,99]. En effet au moment de l'entrée en phase M de la mitose, la membrane du micronoyau se rompt alors que le chromosome n'est pas encore répliqué. Des fourches de réplication sont encore actives, le chromosome se retrouve alors fragmenté. Les mécanismes de réparation sont alors mis en route, mais la recombinaison homologue est rapidement saturée, et des mécanismes moins conservatifs sont alors mis en place, comme la réparation par jonction des extrémités non homologues.

2.3 Mécanismes de réparation des cassures doubles brins

Suite aux nombreuses cassures doubles brins générées pendant le cycle cellulaire, des mécanismes de réparation sont mis en jeu. Il existe deux voies de réparation des cassures doubles brins : la réparation par recombinaison homologue et la réparation par jonction des extrémités. Ces deux voies sont en compétition pour réparer les cassures doubles brins. Lors de la méiose, la réparation par recombinaison homologue (précédemment décrite) est la voie préférentiellement utilisée. Chez la souris, il a été montré que la protéine Ku70 (nécessaire à la voie de réparation par jonction des extrémités) est régulée négativement pendant la prophase de la première méiose, afin de défavoriser la réparation par jonction des extrémités non homologues [100]. Cependant, bien que rare, une mauvaise réparation des cassures doubles brins peut amener à des remaniements chromosomiques de structure [101,102].

2.3.1 Recombinaison homologue non allélique

La recombinaison homologue non allélique (NAHR pour *Non Allelic Homologous Recombination*) est une forme de recombinaison homologue qui peut avoir lieu au moment de la mitose ou de la méiose. Dans cette voie, la réparation se fait également en se servant du chromosome homologue comme matrice mais à cause de séquences répétées, cette réparation est trompée. Elle fait intervenir ou elle survient entre deux répétitions segmentaires ou LCR (low copy repeats), qui sont des éléments répétés de grandes tailles : 1 à 400kb partageant une homologie supérieure à 90 %. Les séquences répétées

représentent une part importante du génome (environ 50% du génome), les LCR représentent quant à elles environ 5 % du génome [103]. La forte homologie entre ces LCR peut créer un mésappariement avec un alignement rapprochant deux séquences répétées identiques mais non alléliques. Un *crossing over* au niveau de ces régions répétées va aboutir à des remaniements chromosomiques chez les cellules filles.

Si les deux régions répétées sont en orientation directe, la recombinaison homologue non allélique entre elles va aboutir à des délétions et des gains des régions comprises entre ces deux éléments répétés. Elle peut avoir lieu entre deux chromosomes homologues ou deux chromatides sœurs aboutissant à des délétions et des duplications. Elle peut également avoir lieu au sein d'une même chromatide, entraînant uniquement des délétions. La recombinaison homologue non allélique entre deux LCR inversés peut aussi aboutir à la formation d'un isochromosome (Figure 21).

Si les deux régions répétées sont en orientation inverse, la recombinaison homologue non allélique va aboutir à une inversion du fragment compris entre les éléments répétés.

Les recombinaisons entre deux LCR aboutissent à des remaniements récurrents (délétion, gain, inversion). Une trentaine de maladies avec des phénotypes différents (déficience intellectuelle, maladies sanguines, infertilité masculine) sont liés à des remaniements récurrents médiés par recombinaison homologue non allélique [91,104]. Certains remaniements récurrents ne sont pas pathogéniques.

Plus les LCR sont grands et plus leur homologie est importante, plus y a de risque de recombinaison homologue non allélique. Des études sur la levure ont également montré que plus la distance est courte entre les deux LCR plus il y a de risque de recombinaison homologue non allélique [105].

La recombinaison homologue non allélique peut également avoir lieu entre les autres éléments répétés du génome comme les séquences Alu et les séquences LINEs (long interspersed repetitive elements). Ces événements sont plus rares et sont responsables de remaniements non récurrents. Une cassure double brin qui a lieu dans une séquence répétée induit donc le risque de mauvaise sélection de la matrice pour la réparation [102,104].

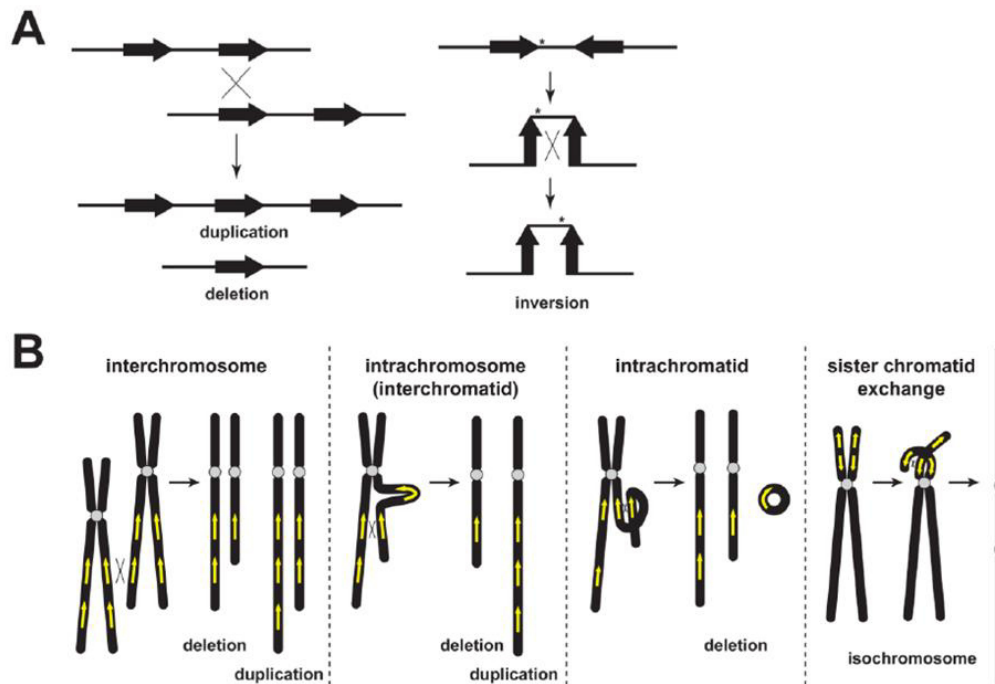


Figure 21. La recombinaison homologue non allélique est responsable de réarrangements chromosomiques
 D'après Liu et al., 2012[91].

A. Différents réarrangements à la suite d'une recombinaison homologue non allélique entre deux LCR. Les recombinaisons homologues non alléliques au niveau : de deux LCR ayant la même orientation aboutissent à des délétions et duplication, de deux LCR ayant une orientation inverse aboutissent à des inversions.

B. La recombinaison homologue non allélique peut avoir lieu entre deux chromosomes homologues ou deux chromatides sœurs aboutissant à des délétions et des duplications. Elle peut également avoir lieu au sein d'une chromatide, entraînant uniquement des délétions. La recombinaison homologue non allélique entre deux LCR inversés peut aussi aboutir à la formation d'un isochromosome.

2.3.2 Appariement simple brin

L'appariement simple brin ou SSA pour *Single Strand Annealing* est un mécanisme de recombinaison homologue non conservatif qui se met en place quand une cassure double brin intervient entre deux séquences répétées orientées dans la même direction[106] (Figure 22). Cette voie est indépendante de rad51. La résection 5'3' aboutissant à des extrémités 3' d'ADN simple brins s'étend jusqu'aux régions répétées, les brins complémentaires s'hybrident via l'homologie entre ces séquences répétées et les brins d'ADN simples brins dépassant sont digérés. Cela entraîne la délétion d'une copie de la séquence répétée, et de la séquence d'ADN comprise entre les deux répétitions.

2.3.3 Jonction des extrémités non homologues

La jonction des extrémités non homologues ou NHEJ (non homologous end joining) est un mécanisme de réparation dans lequel deux fragments d'ADN sont recollés entre eux, indépendamment de leur homologie [102] (Figure 22). Cette réparation nécessite la liaison de l'hétérodimère Ku70-Ku80 aux extrémités de la cassure double brin. Cet hétérodimère se fixe aux extrémités de la cassure et les protège ainsi d'une résection. Puis il va recruter l'ADN ligase IV et ses protéines accessoires, qui rejoignent les deux extrémités d'ADN [107]. Cette réparation est génératrice d'erreur et peut entraîner la perte ou l'insertion de quelques nucléotides. Cela peut également entraîner des remaniements de grandes tailles, comme des translocations, en réparant deux extrémités de cassures doubles brins appartenant à deux chromosomes différents [108,109].

2.3.4 Jonction des extrémités médiée par microhomologie

Le mécanisme de jonction des extrémités médiée par microhomologie ou MMEJ (*microhomology mediated end joining*), est également appelé NHEJ alternatif. Ce mécanisme de réparation utilise souvent une microhomologie de 5 à 25 paires de base. Il est mis en place après une résection 5'3' et crée toujours des délétions qui encadrent le point de cassure original [110] (Figure 22). Il est fréquemment associé à des remaniements chromosomiques: translocations, inversions, remaniements complexes [111].

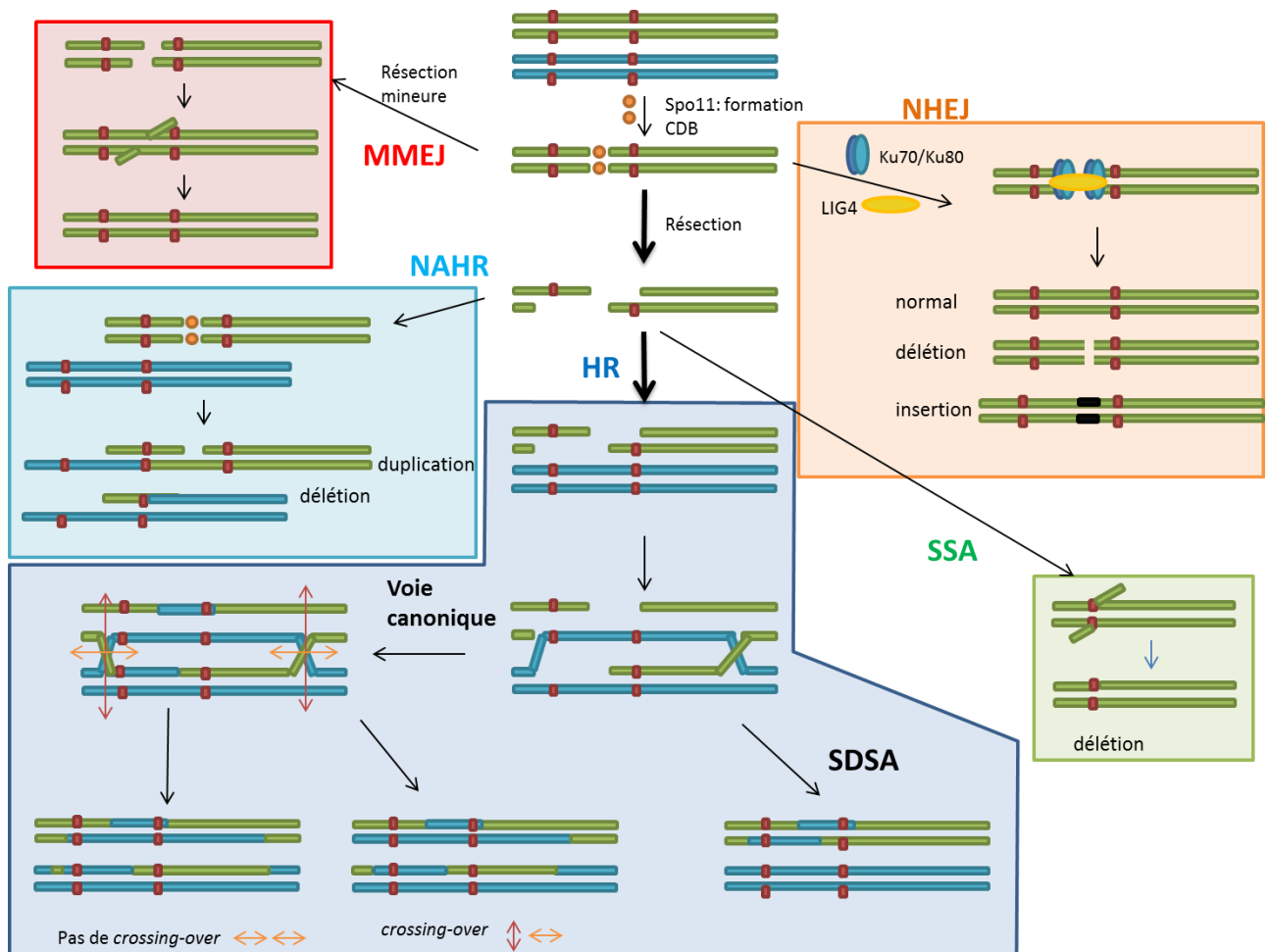


Figure 22. Voies de réparation des cassures de l'ADN double brin lors de la méiose

Les cassures doubles brins (CDB) programmées sont induites par le dimère de protéines Spo11. Lors de la méiose la voie préférentiellement utilisée est celle de la recombinaison homologue (HR pour homologous recombinaison). Cette voie à deux mécanismes de résolution : la voie canonique qui peut amener à un *crossing over* et la voie de synthèse dépendante de l'hybridation de brin (SDSA). En cas de CDB au niveau de séquences répétées (carrés rouges), la recombinaison homologue peut être perturbée. Les chromosomes peuvent mal s'aligner, il s'ensuivra alors une recombinaison homologue non allélique (NAHR) aboutissant à des duplications et délétions. Les fragments cassés peuvent également être recollés via jonction des extrémités : la voie NHEJ (jonction des extrémités non homologue) est mise en place suite au recrutement de l'hétérodimère Ku70/Ku80. Il existe également une voie alternative : la jonction des extrémités médiée par microhomologie (MMEJ) qui implique le plus souvent une microhomologie au niveau des extrémités cassées et qui est mise en place après une petite résection des extrémités. Ces réparations par jonction des extrémités peuvent alors amener à des petits remaniements au niveau du fragment de jonction (insertion, délétion) ou à des remaniements de grandes tailles (translocation, inversion). La voie SSA (single strand annealing) permet une réparation lorsque la CDB est comprise entre deux zones répétées, mais elle implique toujours une délétion de la séquence comprise entre ces deux répétitions.

2.3.5 Mécanismes réplicatifs

2.3.5.1 *Break Induced Replication (BIR)*

Le mécanisme BIR est mis en place lors de la réplication de l'ADN, quand il y a une cassure du brin qui sert de trame à la fourche de réplication [112]. La fourche est alors brisée. Une résection du brin 5' va générer une extrémité 3' simple brin qui peut envahir une région de microhomologie sur une matrice différente pour établir une nouvelle fourche de réplication et redémarrer la synthèse. Cette invasion est dépendante de RAD51 et nécessite une homologie de 70 à 100 nucléotides. Il peut y avoir de multiples envahissements de nouvelles matrices ce qui peut générer un remaniement complexe.

Le MMBIR (*micro-homology break_induced replication*) est indépendant de RAD51 et utilise des séquences d'homologies plus courtes ce qui génère un risque de remaniement chromosomique plus important.

2.3.5.2 *Fork Stalling and Template Switching (FoSTeS)*

Un autre mécanisme qui peut être à l'origine de remaniements chromosomiques est celui d'interruption de la fourche de réplication et de changement de matrice appelé FoSTeS. La fourche de réplication se bloque à cause d'une lésion du brin d'ADN, la réplication est alors interrompue [112]. Le brin en retard se désengage et envahit une autre fourche de réplication active par l'appariement de séquences de microhomologie. La synthèse du brin retardé se poursuit brièvement au niveau de la fourche envahie, avant qu'il se retire à nouveau. Ce brin peut envahir séquentiellement d'autres fourches de réplication actives avant de reprendre finalement au niveau de la matrice d'origine. Cela peut générer des gains ou des pertes en fonction du positionnement des fourches de réplication envahies (Figure 23).

Le terme MMBIR/FoSTeS est souvent utilisé pour parler des mécanismes de réparation mis en jeu lors des anomalies de la fourche de réplication.

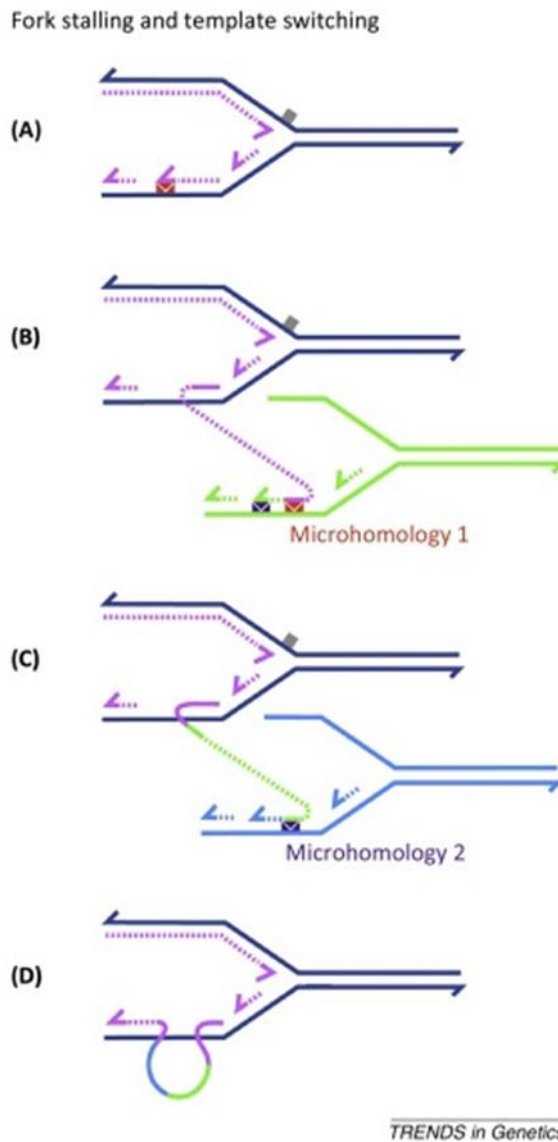


Figure 23. Principe du mécanisme d'interruption de la fourche de réplication et de changement de matrice

D'après Ottaviani et al., 2014 [112]. (A) La fourche de réplication est bloquée. (B) Le brin retard envahit donc une autre fourche de réplication grâce à des zones de microhomologie. (C) Ce brin peut envahir d'autres fourches de réplication. (D) Le brin reprend sa place au niveau de la matrice d'origine, un remaniement a été intégré au chromosome.

En résumé, lors de cassures doubles brins, la réparation par recombinaison homologue est la voie de prédilection. Mais d'autres mécanismes de réparation peuvent entrer en jeu au lieu de cette voie canonique. Ces mécanismes, non conservatifs, peuvent être à l'origine de remaniements chromosomiques (Figure 22).

2.4 Identification des remaniements chromosomiques

Les remaniements chromosomiques peuvent être identifiés grâce à différentes techniques, qui peuvent être ciblées ou pangénomiques ; avec une résolution faible (de l'ordre de 10Mb) ou une résolution à la base près et qui peuvent nous permettre de mettre en évidence des remaniements de structures équilibrés ou non.

2.4.1 Caryotype et hybridation fluorescente *in situ* (FISH)

Le caryotype est une analyse pangénomique. Ce fut la première technique permettant de détecter les anomalies de nombres et les remaniements de structure supérieur à 5 à 10 Mb. Pour cela les cellules sanguines sont mises en culture, puis sont synchronisé afin d'avoir le maximum de cellules au même stade du cycle cellulaire. Les cellules sont bloquées en mitose, en métaphase, par l'ajout de colchicine. Puis un choc hypotonique est réalisé afin de faire éclater les cellules, suivi d'une étape de fixation. Après étalement sur une lame, les chromosomes subissent une dénaturation en bandes (par la chaleur pour les bandes R et enzymatique pour les bandes G). Les lames peuvent enfin être visualisées au microscope. Les chromosomes sont classés en fonction de leur taille, et de leur banding.

La FISH permet une analyse ciblée, de la structure et du nombre des chromosomes grâce à l'utilisation de sondes marquées correspondant à un locus du génome humain et pouvant atteindre une résolution d'environ 150 à 200 kb en fonction du type de sondes utilisées (sondes locus spécifique, sondes centromériques, sondes subtélomériques et sondes de peinture). Elle repose sur la complémentarité des brins d'ADN et peut être appliquée sur chromosomes métaphasiques et noyaux interphasiques. La FISH permet de retrouver des anomalies chromosomiques non visibles au caryotype et permet de caractériser les remaniements identifiés au caryotype. La technique de FISH a également été utilisée pour déterminer les points de cassure des remaniements de structure. Pour cela une démarche par progression pas à pas le long du chromosome est réalisée jusqu'à ce qu'une sonde soit interrompue.

2.4.2 Analyse chromosomique sur puce à ADN : ACPA

L'ACPA permet une analyse pangénomique des CNV mais ne permet pas de détecter des remaniements équilibrés.

Le principe de cette technique (Figure 24) est de mettre en compétition deux ADN marqués par 2 fluorochromes différents : l'ADN du patient en rouge par la cyanine 5 (Cy5) et l'ADN du témoin en vert par la cyanine 3 (Cy3). Après normalisation, les ADN sont mélangés puis sont déposés sur une lame, appelée puce à ADN. Des centaines de milliers d'oligonucléotides représentant des régions du génome sont fixés sur cette lame, permettant une analyse de tout le génome à une résolution d'environ 30kb.

Les fragments d'ADN du patient et du témoin entrent alors en compétition pour s'hybrider sur les oligonucléotides. Lorsque qu'il n'y a pas de CNV, autant d'ADN du patient que du témoin va s'hybrider sur la lame et le spot émettra alors une lumière jaune. Lorsque que le patient à une perte alors il y aura moins d'ADN du patient qui sera hybridé et le signal apparaîtra vert. Pour un gain, il y aura un excès d'ADN du patient et le signal sera rouge. L'analyse de ces signaux se fait informatiquement, les signaux lumineux sont transformés en résultats logarithmiques, puis l'analyse des profils se fait sur un logiciel de visualisation. Lorsqu'il n'y a pas de déséquilibres la courbe se trouve alors sur le 0. En cas de délétion hétérozygote la courbe est déviée vers le -1 et en cas de duplication la courbe est déviée vers +0.58. Lorsqu'un CNV est retrouvé chez un patient il doit être confirmé par une deuxième technique : une PCR quantitative (qPCR) ou une technique d'hybridation *in situ* en fluorescence (FISH) en fonction de la taille et du type de remaniement.

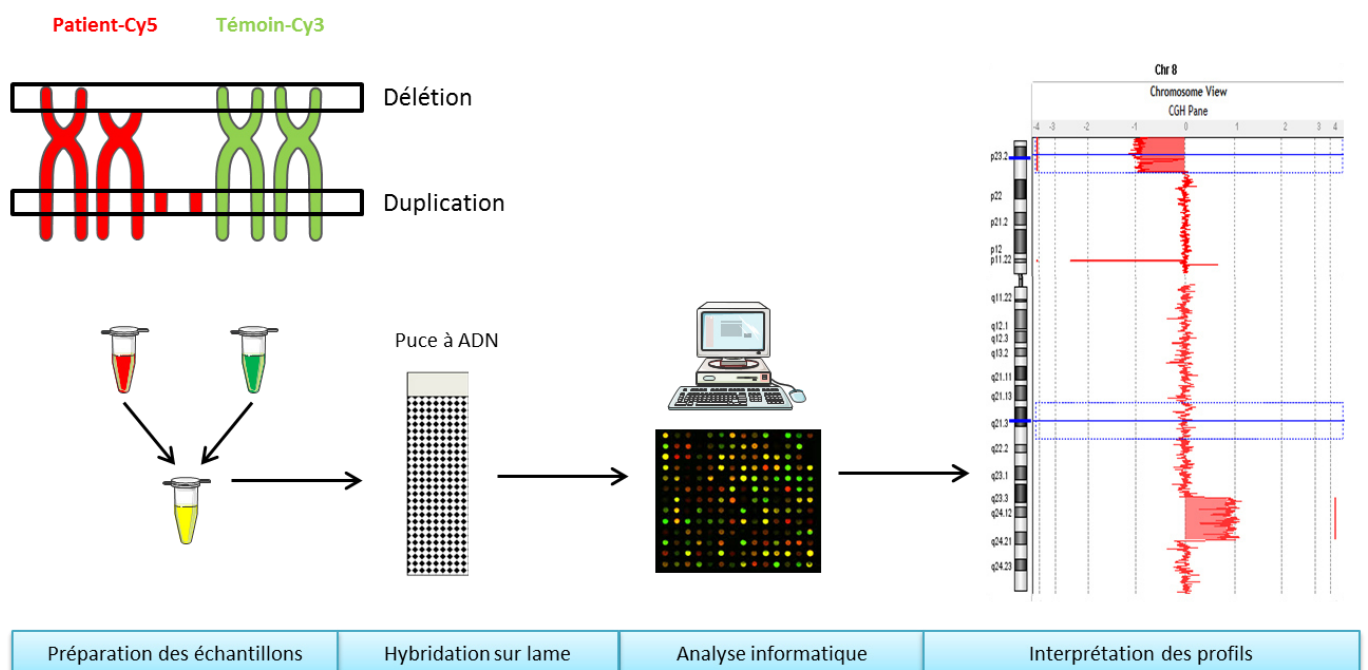


Figure 24. Principe de la méthode d'ACPA

On peut observer sur le profil à droite une perte de la région terminale du bras court du chromosome, et un gain au niveau du bras long du chromosome.

2.4.3 Séquençage haut débit

Le séquençage haut débit permet de séquencer des centaines de fragments d'ADN en parallèle. En fonction du type de librairie réalisé, il permet d'étudier un panel de gènes, l'exome ou le génome. Dans le cas d'une analyse pangénomique, les CNV, SNV (*single nucleotide variant*) et les variants de structure équilibrés peuvent être détectés.

La technique de séquençage est composée d'une partie technique en laboratoire appelée « wet-lab » et une partie bio-informatique appelée « dry-lab ». En wet-lab, il y a trois grandes étapes : 1. la préparation de la librairie qui consiste à créer une collection de fragments d'ADN ; 2. L'amplification clonale qui consiste à amplifier chacun de ces fragments d'ADN par PCR ; 3. Le séquençage qui permet d'obtenir la séquence de chacun de ses fragments.

2.4.3.1 Partie technique : « Wet-lab »

Nous allons ici décrire la technologie Illumina qui est celle que nous utilisons. Il s'agit d'une technique de séquençage « short-read » c'est-à-dire que la taille des lectures/*reads* est petite (environ 75-150 pb).

Préparation de la librairie

L'ADN est tout d'abord extrait à partir d'un échantillon de sang. Il est ensuite fragmenté, et constitue une librairie. Des adaptateurs (séquences nucléotidiques qui vont permettre aux fragments d'ADN de se fixer sur la *FlowCell*) permettant le séquençage et des index (ou codes-barres, qui sont une séquence de six nucléotides, propre à chaque patient) permettant le multiplexage de l'ADN de plusieurs individus sont ensuite ligués à ces fragments. L'ADN est dénaturé en simple brin et déposé sur la *FlowCell*, où les fragments vont se fixer grâce aux adaptateurs.

Amplification clonale

L'amplification par PCR en pont des banques capturées va aboutir à la formation de clusters qui sont des régions contenant un grand nombre de séquences identiques. Le fragment se fixe grâce à ses adaptateurs (Figure 25).

Séquençage

Le principe du séquençage repose sur la synthèse d'ADN couplée à l'émission de fluorescence. A chaque cycle, les 4 bases marquées et bloquantes sont ajoutées, celle correspondant au fragment est insérée par une ADN polymérase et après excitation par un laser, la fluorescence émise par chaque cluster est récupérée par une caméra et la base est lue. Après excision du groupement bloquant, un nouveau cycle commence. Les cycles de séquences sont répétés pour lire chaque base les unes après les autres (Figure 20). On réalise un séquençage en « paired-end » en 2x100pb de façon à obtenir des « lectures » (communément appelées « *reads* ») de 100pb pour chacune des extrémités des fragments. Après avoir séquencé les 100-150 bases d'une extrémité du fragment, il est retourné, et

l'autre extrémité est séquencée de la même manière. Les deux lectures sont donc appariées, il s'agit d'un séquençage dit *paired-end*

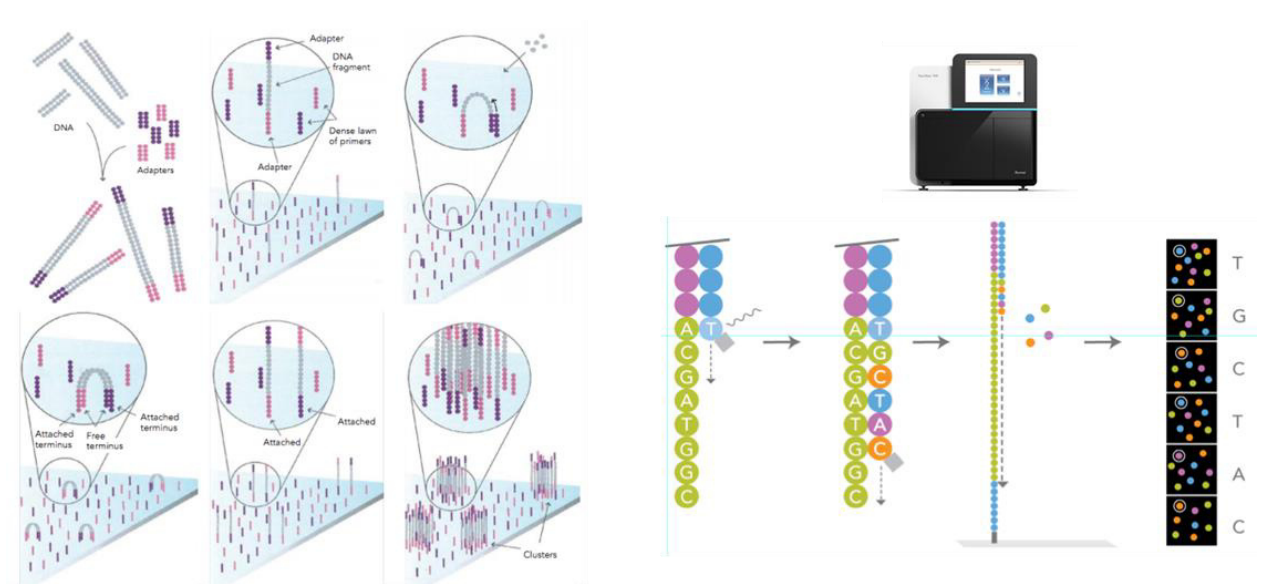


Figure 25. Principe du séquençage haut débit avec la méthode Illumina

2.4.3.2 Détection des SV et CNVs à partir des données de WGS : « Dry-lab »

Les lectures sont alors alignées sur le génome de référence (Figure 26). L'analyse de ces alignements par des logiciels de détection de variants permet d'identifier les remaniements de structure équilibrés et déséquilibrés grâce à trois types d'analyse : l'analyse de profondeur, l'analyse des paires de « reads » aberrantes et l'analyse des « split reads ».

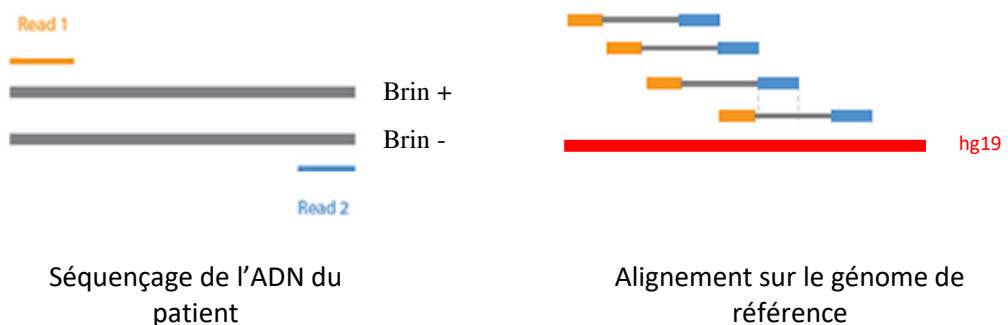


Figure 26. Principe du séquençage (WGS) en « *paired-end* »

Les extrémités des fragments d'ADN génomique d'environ 350 pb sont séquencées sur 101 pb: la lecture 1 (*Read 1*) est séquencée sur le brin positif du fragment et son complémentaire (*Read2*) est séquencé sur le brin négatif. Alignement sur le génome de référence (hg19) : en cas d'orientation normale, le *Read 1* est en amont du *Read 2*.

Analyse des paires de « reads » aberrantes

Théoriquement, les deux lectures d'une même paire s'alignent sur le même chromosome et sont séparées par une distance dépendant de la longueur des fragments séquencés (environ 350 pb). De plus, une des deux lectures est théoriquement alignée sur le brin positif pendant que son partenaire l'est sur le brin négatif. Si une paire se trouve de part et d'autre d'un point de cassure, ces considérations ne seront plus respectées. Par exemple, dans le cas d'une translocation, les deux lectures d'une même paire seront alignées sur deux chromosomes différents. Dans le cas d'une inversion, les deux lectures d'une même paire seront séparées par une distance supérieure à celle attendue (souvent plusieurs Mb), avec une orientation anormale (elles seront toutes deux alignées sur le même brin) (Figure 27).

Sur IGV les lectures ne présentant aucune anomalie sont figurées en gris. Lorsque les lectures apparaissent dans une autre couleur, cela signifie qu'il y a un défaut d'alignement, de distance ou d'orientation entre les deux lectures de la paire (Figure 28).

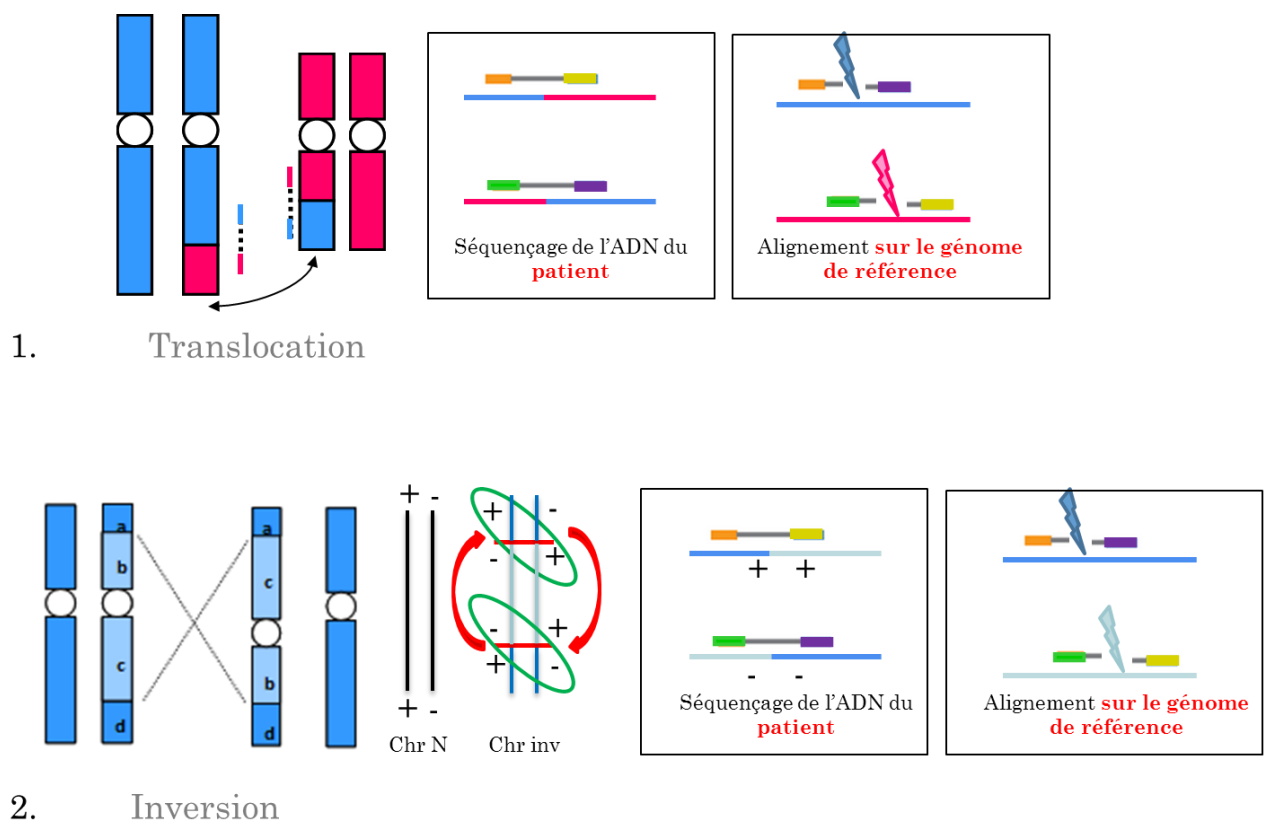


Figure 27. Remaniements chromosomiques apparemment équilibrés et discordance d'alignements des *reads* du WGS

1. Translocation : les lectures de part et d'autre du point de cassure s'alignent sur 2 chromosomes différents
2. Inversion : les lectures de part et d'autre du point de cassure présentent une orientation anormale (les deux lectures s'alignent sur le même brin). Légende : Chr N : chromosome normal, Chr inv : chromosome inversé

Analyse de profondeur

L'analyse de profondeur permet de détecter les CNV, la quantité de séquence alignées à un locus est proportionnelle à la quantité d'ADN. La profondeur correspond au nombre de fois où une base a été séquencée. En cas de délétion, la profondeur de lecture sera diminuée d'un facteur 2 et en cas de gain, elle sera multipliée par 1,5. Cette profondeur de lecture peut être visualisée sur IGV (Figure 28).

Analyse des « *split-reads* »

Lors de la visualisation des remaniements sur IGV des « *splits reads* » sont apparents (figure 28). Ce sont des *reads* tronqués. C'est-à-dire qu'ils ne s'alignent qu'en partie sur la région du génome que l'on visualise. Ces *reads* « chevauchent » les points de cassure, ils sont à la fois sur les deux fragments chromosomiques s'étant recollés ensemble. En récupérant la séquence de ces « *splits reads* », on peut, grâce à l'outil *Blat* du site Genome Bioinformatics de l'UCSC, connaître à la base près où se situe le point de cassure. La séquence du fragment de jonction permet de mettre en évidence des microhomologies au niveau des points de cassure, des insertions ou des délétions.

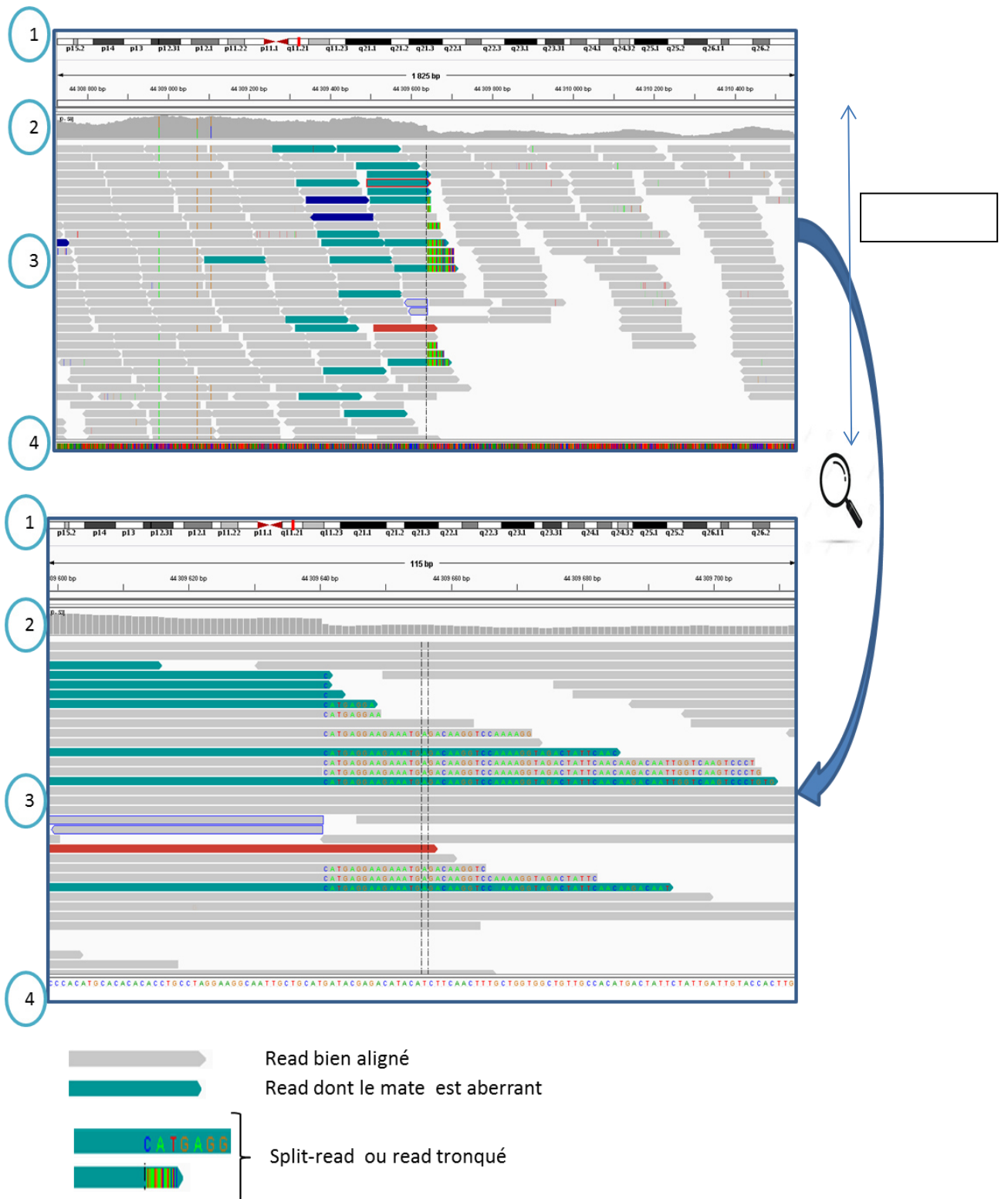


Figure 28. Visualisation sur IGV

En 1 : la représentation du chromosome sur lequel on visualise des « reads », avec en rouge la région d'intérêt. En 2 : la profondeur de lecture. Sur cette image une différence de profondeur de facteur 2 de part et d'autre du point de cassure est observée. En 3 : cette zone permet de visualiser tous les « reads ». Les « reads » sont gris quand ils sont bien alignés, et de couleur quand il y a une aberration de position, de distance ou d'orientation entre les 2 « reads » de la paire. Les *splits read*, qui sont des *reads* tronqués donc alignés qu'en partie sur la région visualisée, sont représentés en bariolé ou avec la séquence de la partie non alignée. En 4 : séquence de référence.

3. La méiose

Lors de la méiose, qui consiste en une double division cellulaire, une réparation des cassures double brin par les mécanismes alternatifs peut aboutir à la formation de gamète avec des remaniements chromosomiques. De plus, la méiose peut être perturbée par les remaniements chromosomiques, et cela peut aboutir à la formation de gamètes déséquilibrés. Nous verrons dans cette partie le déroulement de la méiose en portant une attention particulière sur la phase de recombinaison méiotique. Puis nous verrons les conséquences des remaniements chromosomiques.

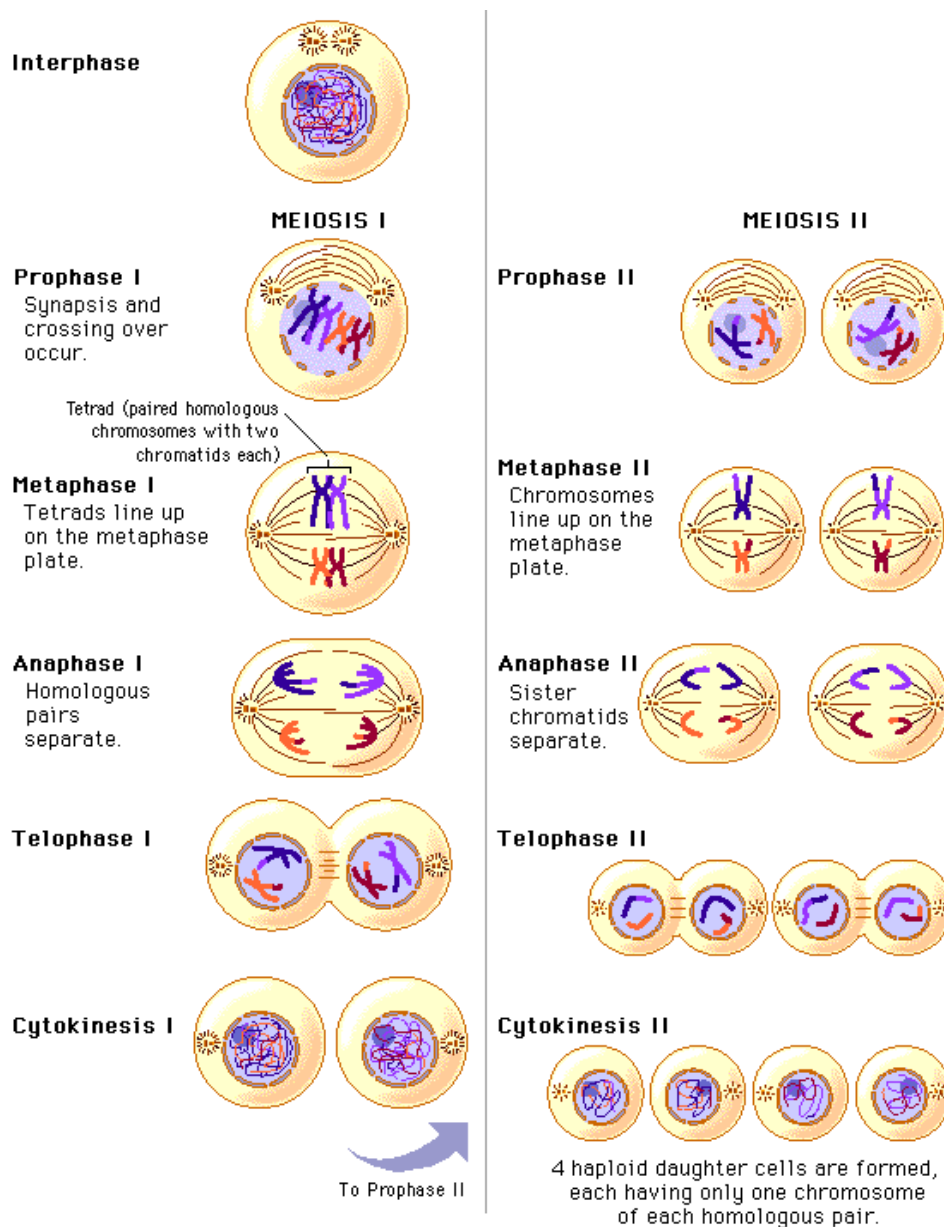


Figure 29. Les différentes étapes de la méiose.

La réplication de l'ADN est suivie par deux divisions, aboutissant à la formation de quatre gamètes haploïdes contenant un seul chromosome de chaque paire de chromosomes homologues.

(D'après http://www.phschool.com/science/biology_place/labbench/lab3/concepts2.html)

3.1 Principe de la méiose

La méiose est le processus de formation des gamètes grâce à une double division cellulaire. A partir d'une cellule mère diploïde, qui contient deux exemplaires de chaque chromosome, quatre cellules filles haploïdes différentes vont être formées, contenant chacune la moitié du patrimoine génétique de la cellule mère (figure 29).

Avant la première division, il y a une réplication de l'ADN. La cellule qui est diploïde va dupliquer son ADN qui sera ensuite condensé en deux chromatides sœurs. Cette réplication est suivie de deux ségrégations.

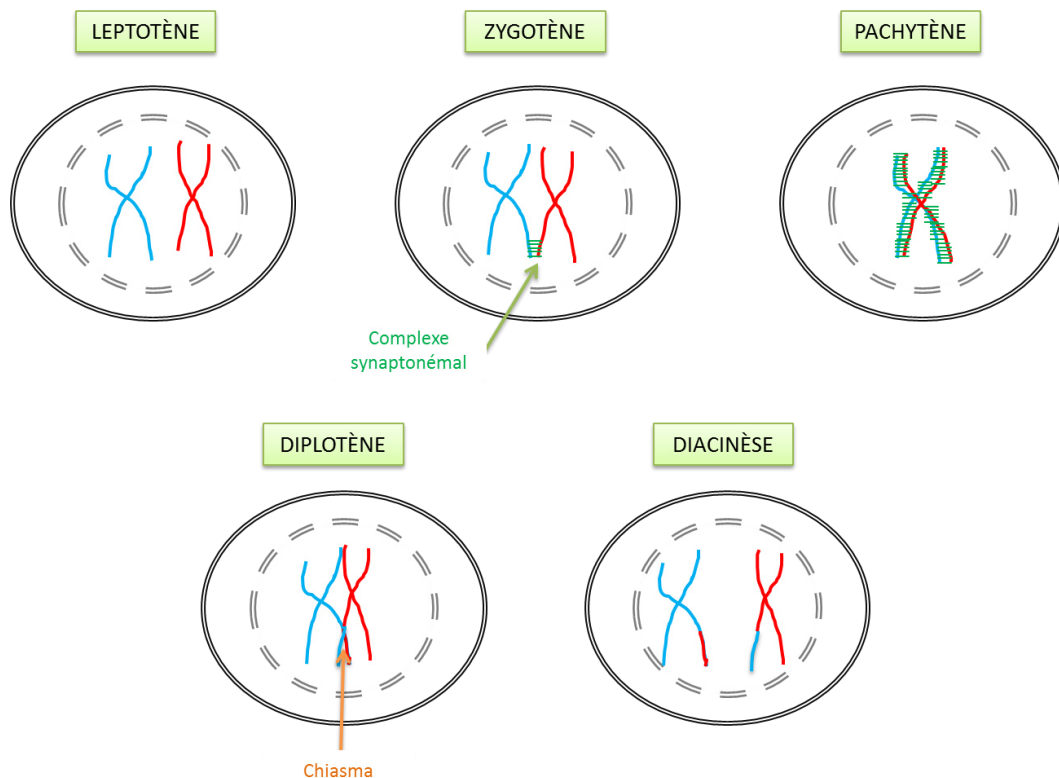


Figure 30. Les différentes étapes de la prophase 1.

Les différentes étapes de la prophase 1. Au cours du zygotène le complexe synaptonémal (traits verts) commence à se former, celui-ci est complet au passage au pachytène. Lors du pachytène il y a formation de chiasmata, et lorsque le complexe synaptonémal se désassemble au diplotène, on observe des bivalents reliés par des chiasmata.

La première division de la méiose (méiose 1) est une division réductionnelle consistant à la séparation des chromosomes homologues. On obtient alors deux cellules haploïdes.

La méiose 1 est divisée en quatre étapes : La prophase, la métaphase, l'anaphase et la télophase. On peut sous-diviser la prophase en plusieurs stades : le leptotène, le zygotène, le pachytène, le diplotène et la diacinèse (figure 30). Lors de la prophase 1, au leptotène, chaque chromosome doit localiser et reconnaître son chromosome homologue afin de s'aligner avec lui. Lors de cette phase, des cassures doubles brins sont programmées, permettant de maintenir l'appariement des chromosomes homologues. Puis, au zygotène, une protéine de structure, le complexe synaptonémal, va s'assembler entre les deux chromosomes homologues (Figure 31). Lorsque le synapsis, qui est l'union des chromosomes homologues, est complet, le stade de pachytène commence.

Au stade du pachytène, les chromosomes vont s'enjamber (on parle de *crossing over*). Ces enjambements sont essentiels pour la formation de chiasmata (connexions entre chromosomes homologues).

Au diplotène, le complexe synaptonémal se désassemble. On observe alors des bivalents connectés par des chiasmata [113].

En fin de prophase 1, les chromosomes homologues sont orientés chacun de leur côté, ce qui permet à la métaphase 1 un attachement bipolaire des chromosomes homologues au fuseau de microtubules via le kinétochore. Puis lors de l'anaphase 1, les deux chromosomes se séparent en étant attirés aux pôles opposés par la contraction du réseau de fibres et la répulsion des chromosomes entre eux. Lors de la télophase 1, les membranes nucléaires sont reformées et les deux cellules se séparent.

La deuxième division (méiose 2) est une division équationnelle, identique à une mitose classique, consistant en la séparation des chromatides afin d'aboutir à quatre cellules différentes : les gamètes.

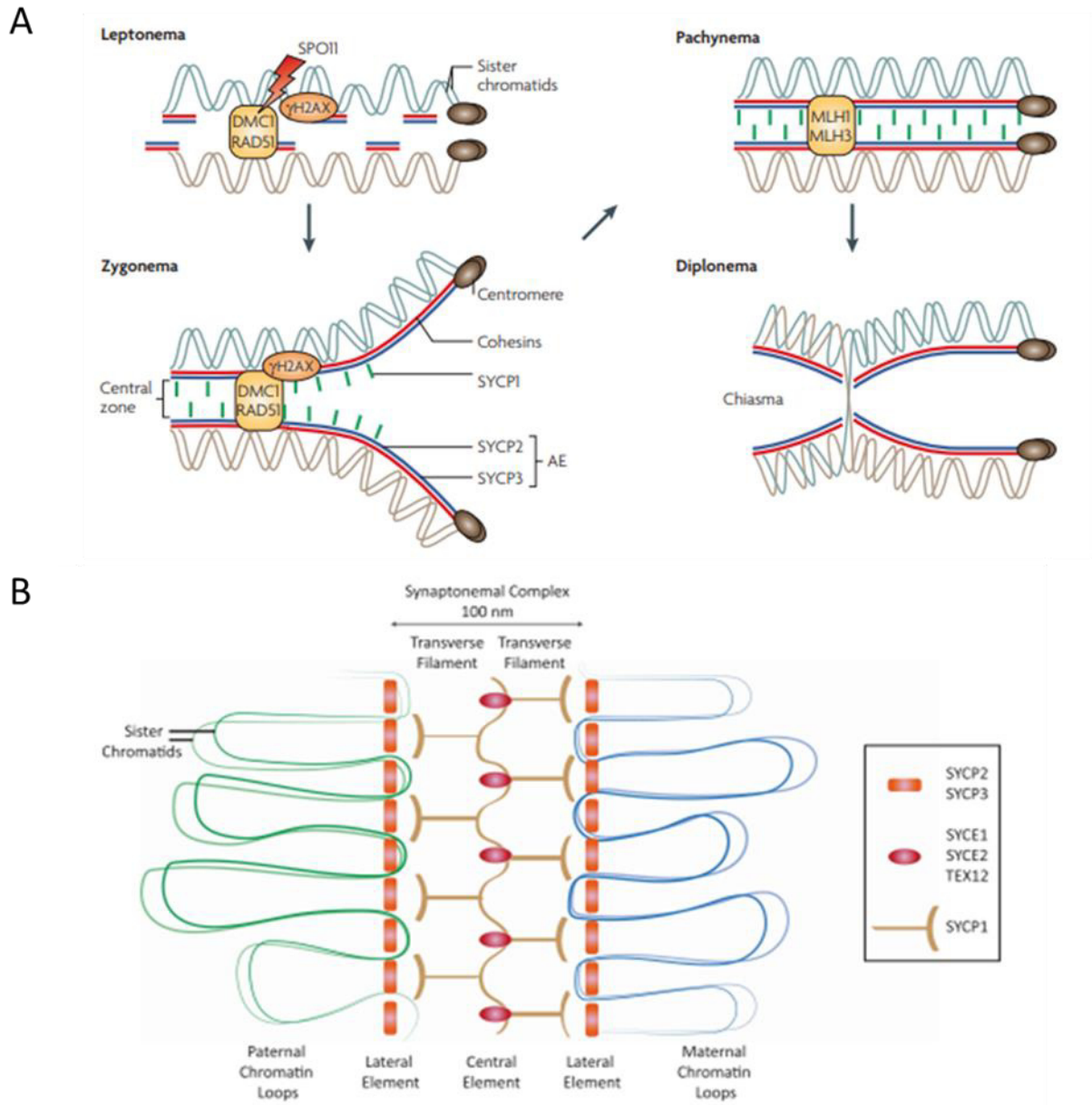


Figure 31. Organisation de la chromatine pendant la prophase de la méiose 1.

A. D'après Handel et al., 2010[114] | Au cours du leptotène les chromosomes homologues commencent à s'aligner mais ne sont pas appariés, les éléments du complexe synaptonémal commencent à s'assembler (SYCP3 et SYCP2). Les cassures doubles brins programmées sont induites par la protéine Spo11. Les cassures sont reconnues par la machinerie de réparation par recombinaison homologue et réséquées, ce qui déclenche la liaison des protéines DMC1 et RAD51 qui colocalisent avec des structures appelées nodules de recombinaison. Au zygotène, l'appariement entre les homologues continue, le complexe synaptonémal est formé et le synapsis est initié. Le pachytène correspond au moment où le synapsis est complet. Le stade pachytène est long et inclut la maturation d'un sous-ensemble de sites de recombinaison méiotique (<10%) en crossing-over. Quand la recombinaison est terminée (diplotène), le complexe synaptonémal se désassemble et les chromosomes se compactent. Les chromosomes homologues sont alors maintenus ensemble par les *crossing over*, ce sont des bivalents.

B. D'après MacGregor et al., 2019 [115]. Les chromosomes sont organisés en un réseau de boucles chromatiniennes dont la base s'associe au complexe synaptonémal qui est composé de trois parties : les éléments latéraux (SYCP2 et SYCP3) qui sont

le long de chaque chromosomes, les filaments transverses (SYCP1) qui maintiennent les chromosomes homologues ensemble et les éléments centraux (SYCE1, SYCE2 et TEX12) qui stabilisent le complexe synaptonémal.

3.2 Recombinaison méiotique

Lors du pachytène, les chromosomes forment des chiasmata qui sont initiés par la recombinaison méiotique. La recombinaison joue un rôle crucial pendant la méiose. Elle contribue tout d'abord au brassage chromosomique. Mais elle joue également un rôle dans l'assemblage des chromosomes par paires, le processus de synapsis et la formation des chiasmata.

La recombinaison méiotique est initiée par des centaines de cassures doubles brins programmées (contrairement à la recombinaison somatique qui a lieu suite à des cassures doubles brins fortuites)[115]. Cela va entraîner une cascade de recombinaisons homologues, utilisant comme matrice l'ADN situé à la même position allélique sur le chromosome homologue. Ces réparations des cassures doubles brins vont aboutir ou non à un *crossing over* (échange réciproque de bras chromosomiques). Ces cassures doubles brins sont formées par un homodimère de la protéine SPO11 par réaction de transestérification de type topoisomérase [113,116,117]. Un monomère de SPO11 se lie de manière covalente à chaque extrémité 5' du brin d'ADN. Cette liaison covalente de SPO11 est clivée par le complexe MRN (Mre11-Rad50-Nbs1) puis l'exonucléase Exo1 va faire une résection exonucléotidique 5'3' aboutissant à des extrémités 3' d'ADN simple brin. Les protéines DMC1 et RAD51, qui sont des protéines d'échange de brins, vont se fixer aux extrémités, former des filaments nucléoprotéiques puis vont catalyser l'invasion du chromosome homologue par une extrémité 3' d'ADN simple brin, ce qui va créer un hétéroduplex d'ADN. Il y a ensuite séparation des brins de l'ADN envahi, ce qui va entraîner la formation d'un intermédiaire : la boucle de déplacement (D-loop)[104]. La resynthèse des brins d'ADN est alors amorcée par une polymérase à chacune des extrémités 3' simple brin en utilisant le brin homologue comme matrice. Il existe deux modèles de résolution. Dans la voie de la double jonction de Holliday (voie canonique), il y a invasion du chromosome homologue par la deuxième extrémité 3' d'ADN simple brin, il y a donc deux jonctions Holliday (jonction mobile entre quatre brins d'ADN) par les deux extrémités 3'. La résolution est alors obtenue par clivage des deux jonctions de Holliday, ce qui peut résulter ou non en la formation d'un *crossing over* selon l'orientation de la coupure. La deuxième voie, la voie de synthèse dépendante de l'hybridation de brin (SDSA : *Synthesis Dependent Strand Annealing*), [118] ne peut aboutir qu'à des non *crossing over*. Dans cette voie, le brin est déplacé de la boucle de déplacement (D-loop). Il s'hybride alors avec la seconde extrémité de la cassure qu'il utilise comme matrice pour s'allonger. Il n'y aura pas de formation de chiasmata mais on peut aboutir à un phénomène appelé conversion génique (modification de la

séquence d'ADN, qui est celle du chromosome homologue), qui prend l'apparence d'un double *crossing over* d'une petite zone restreinte [114] (Figure 32).

Lors de la méiose, le choix de la matrice de réparation est biaisé ; le chromosome homologue va être préféré à la chromatide sœur pour la réparation. Une machinerie empêche l'utilisation de la chromatide sœur pour réparer la cassure [116].

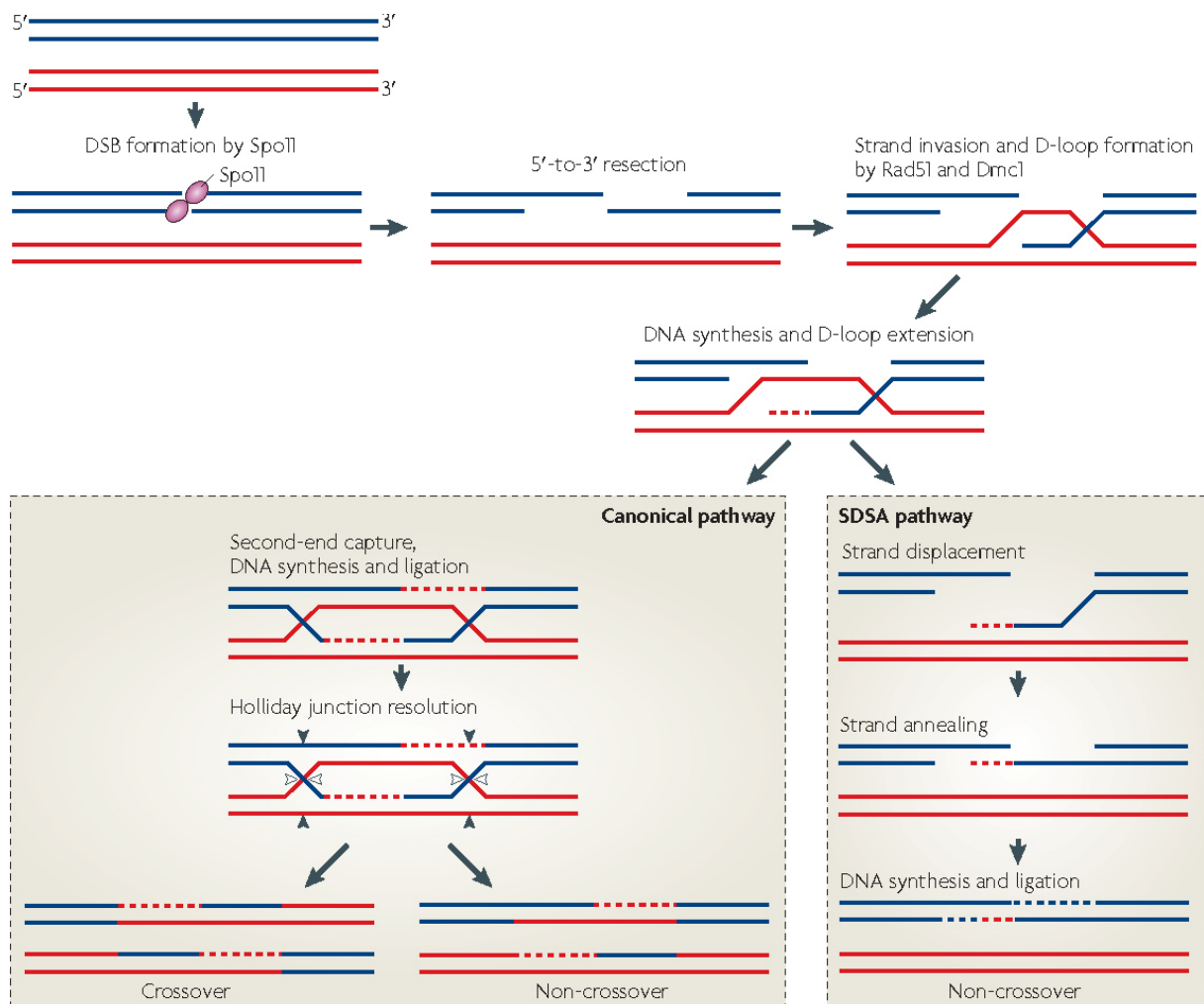


Figure 32. La recombinaison méiotique

D'après Sasaki et al., 2010[104]. Les cassures doubles brins sont induites par la protéine Spo11 pendant la prophase de la méiose 1.

L'exonucléase Exo1 va faire une résection exonucléotidique 5'3'aboutissant à des extrémités d'ADN simple brin qui vont chercher puis envahir le chromosome homologue en formant une boucle de déplacement (D-loop) grâce aux protéines d'échange de brins : Dmc1 et Rad51. La synthèse d'ADN est alors amorcée aux deux extrémités. Puis il existe deux mécanismes de résolution.

Dans la voie canonique (voie des doubles jonctions de Holliday), la capture de la seconde extrémité, va générer une seconde jonction de Holliday. Les jonctions de Holliday peuvent être résolues en coupant l'extérieur des brins (pointes de flèche remplies) ou brins croisés (pointes de flèche ouvertes) de chaque jonction (deux des quatre résolutions possibles sont montrées ici). La résolution peut conduire ou non à un *crossing over*, bien que les *crossing over* soient plus fréquents. La voie de synthèse dépendante de l'hybridation de brin (SDSA) conduit à des non *crossing-over*.

La recombinaison homologue crée un lien physique entre les deux chromosomes homologues, ce qui permet leur bonne orientation sur le fuseau méiotique et leur bonne séparation lors de l'anaphase. La recombinaison méiotique est donc indispensable à une bonne stabilité du génome et la bonne distribution des chromosomes dans les gamètes [104].

Malgré le grand nombre de cassure doubles brins (plusieurs centaines par noyau) [119] la recombinaison va générer des *crossing over* dans une petite minorité des cas : on en observe en moyenne entre 1 par chromosome et 1 à 3 par bras chromosomique [120]. Les *crossing over* correspondent à un échange réciproque de la fin des bras chromosomiques des deux chromosomes homologues, à partir de la cassure double brin générée. Dans la grande majorité des cas, la recombinaison homologue génère des « non-crossing-overs », qui sont des réparations locales de l'ADN.

Un mécanisme de surveillance s'assure que chaque bivalent a réalisé un *crossing over*. En absence de *crossing over*, il y a défaut de ségrégation.

La distribution des cassures doubles brins n'est pas aléatoire, elles ont préférentiellement lieu dans certaines régions, on parle de « *hot spot* de recombinaison » [114,121]. Chez la levure, il a été montré que les sites de coupures sont marqués en amont par la triméthylation de la lysine 4 des histones H3 (H3K4me3) [122]. Chez l'humain, le motif de 13 nucléotides CCNCCNTNNCCNC est retrouvé dans plus de 40% des *crossing-overs* méiotiques [123].

3.3 Durée de la méiose

La grande différence entre la méiose chez les hommes et les femmes est sa durée. La méiose 1 peut durer des années (en moyenne entre 13 et 51 ans) chez la femme mais seulement 23 jours chez l'homme [124]. A 5 semaines de vie fœtale, les ovogonies subissent de nombreuses mitoses afin de générer un stock qui ne sera pas renouvelable (environ 700 000 à la naissance et 400 000 à la puberté). Ces ovogonies rentrent en méiose à 12 semaines de vie fœtale et sont bloquées en prophase 1 au stade diplotène, ces cellules sont devenues des ovocytes primaires. La maturation reprendra au moment du pic pré-ovulatoire de LH, soit environ 13 ans plus tard pour les premières. Les ovocytes primaires finissant leur maturation avant la ménopause sont alors restés au stade de prophase 1 pendant en moyenne 51 ans. La spermatogenèse commence au moment de la puberté et est un processus continu qui dure environ 64 jours. L'étape de méiose 1 dure seulement 23 jours.

3.4 Conséquences des remaniements chromosomiques sur la méiose

Pour tous ces remaniements équilibrés, le risque est celui de la formation de gamètes déséquilibrés entraînant alors après la fécondation des zygotes anormaux ce qui peut amener soit à des avortements spontanés soit à la naissance d'un enfant présentant des anomalies congénitales.

3.4.1 Translocations

Dans le cas des translocations, les chromosomes homologues ne formeront pas de bivalents au moment de la prophase de la méiose I, mais se regrouperont afin de réunir tous les chromosomes impliqués dans la translocation et permettre l'appariement de toutes les séquences/régions homologues.

Ainsi en cas de translocations robertsoniennes, il va y avoir formation d'un trivalent. Il existe alors 2 types de ségrégations possibles : la ségrégation alterne qui produit des gamètes pouvant être normaux ou équilibrés (avec la même translocation que le parent) et la ségrégation adjacente qui produit des gamètes trisomique ou monosomique pour un des deux chromosomes impliqués. Dans le cas de translocations robertsoniennes entre deux chromosomes acrocentriques homologues, il n'y a que des gamètes déséquilibrés qui peuvent être formés (Figure 33).

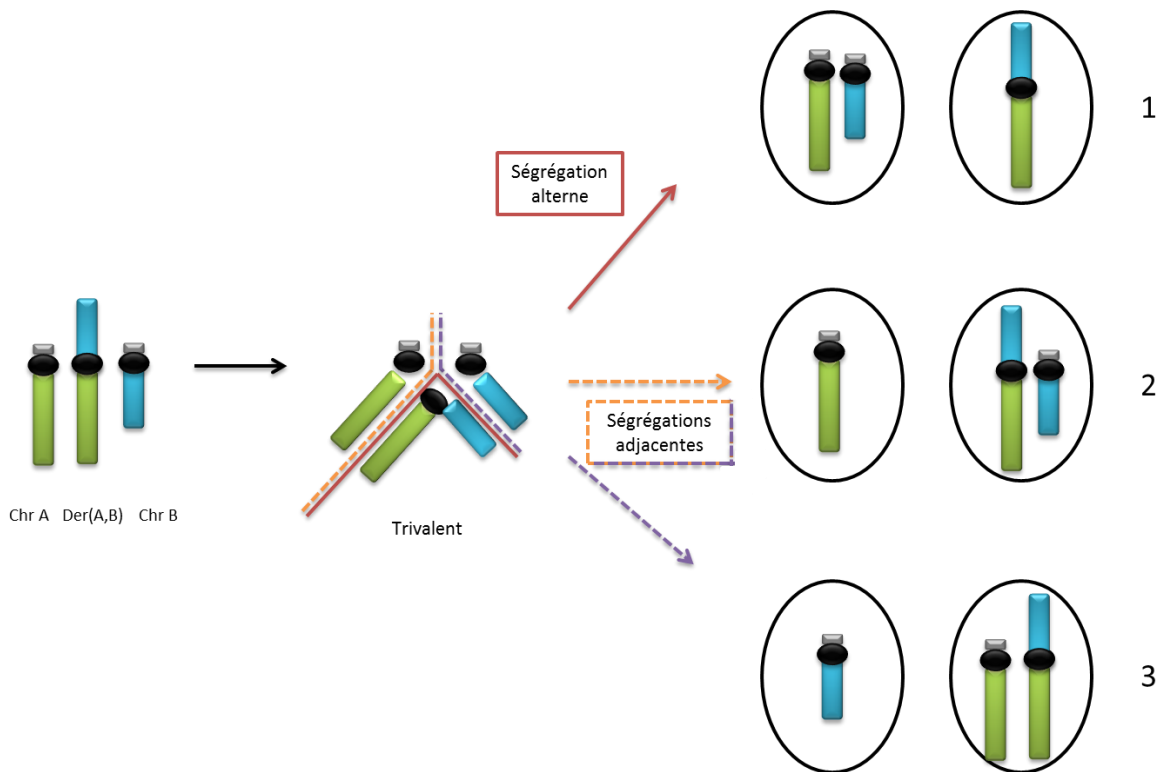


Figure 33. Possibilités de ségrégation méiotique d'une translocation robertsonienne entre le chromosome A et le chromosome B

Les chromosomes vont former un trivalent afin de s'aligner. La ségrégation alterne (1) aboutit à la formation de gamètes normaux ou équilibrés. Les ségrégations adjacentes aboutissent à la formation de gamètes déséquilibrés. Après fécondation, les zygotes en (2) auront une monosomie B pour le premier et une trisomie B pour le deuxième ; et les zygotes en (3) auront une monosomie A et une trisomie A.

Dans le cas des translocations réciproques, il y a formation d'un tétravalent. Il existe alors trois grands types de ségrégation : 1. Les ségrégation de type 2:2 où deux chromosomes sont transmis dans chaque gamètes ; 2. Les ségrégations de type 3:1 où trois chromosomes sont transmis dans un des gamètes, et un chromosome dans le deuxième ; 3. Les ségrégations 4:0 où les quatre chromosomes sont dans un gamète et le deuxième n'a pas de chromosome.

Dans le cas des ségrégations 2:2, la ségrégation peut être alterne ou adjacente. La ségrégation alterne produit des gamètes normaux ou équilibrés alors que la ségrégation adjacente produit des gamètes anormaux. Il existe deux types de ségrégation adjacente : le type 1 et le type 2.

Dans la ségrégation adjacente 1, il y a formation d'un gamète avec un chromosome dérivé (der A ou réciproquement der B) et l'autre chromosome normal (chr B ou réciproquement chr A). Cela entraîne alors la formation d'un zygote avec une duplication d'un fragment transloqué et une délétion de l'autre fragment transloqué (duplication du fragment du chromosome B et délétion du fragment du chromosome A ou réciproquement).

Dans la ségrégation adjacente 2, il y a formation d'un gamète avec un chromosome dérivé (der A ou réciproquement der B) et son propre homologue (chr A ou réciproquement chr B). Cela entraîne alors la formation d'un zygote avec une trisomie d'un chromosome transloqué excepté le fragment transloqué et une monosomie de l'autre chromosome excepté le fragment transloqué (trisomie du chromosome A excepté le fragment transloqué et monosomie du chromosome B excepté le fragment transloqué ; ou réciproquement) (Figure 34).

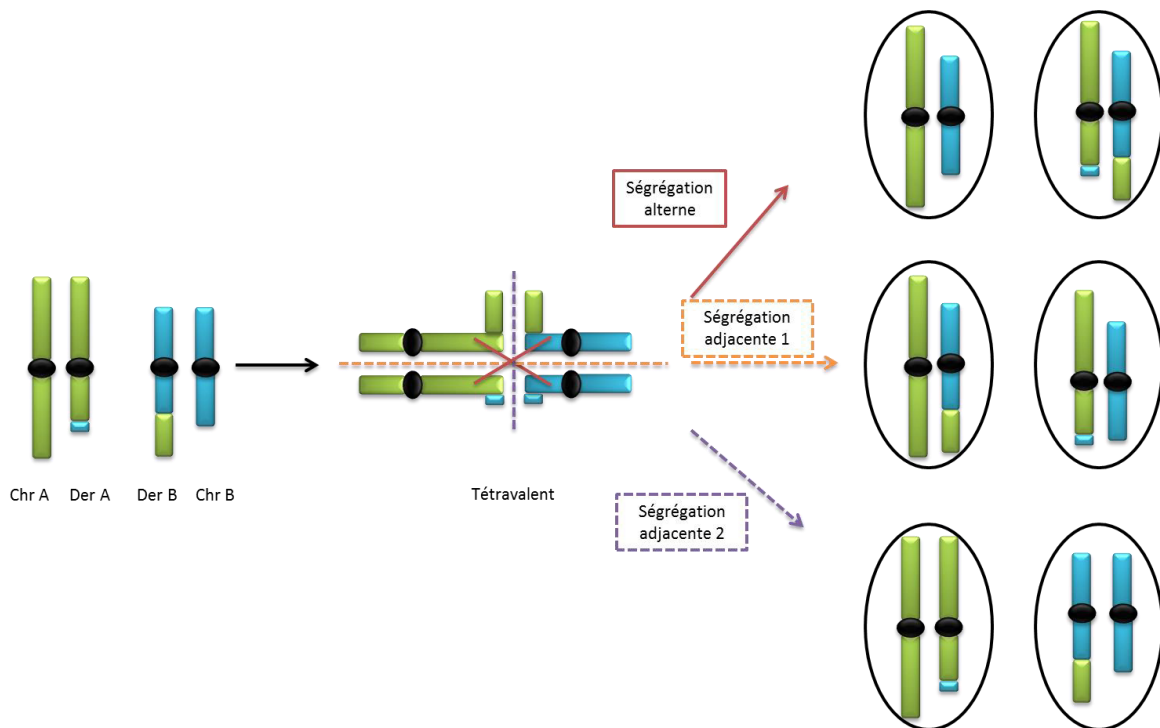


Figure 34. Possibilités de ségrégation méiotique d'une translocation réciproque entre le chromosome A et le chromosome B

Les chromosomes vont former un tétravalent afin de s'aligner. La ségrégation alterne aboutit à la formation de gamètes normaux ou équilibrés. Les ségrégations adjacentes aboutissent à la formation de gamètes déséquilibrés. La ségrégation adjacente 1 donnera des zygotes avec une délétion d'un fragment du chromosome B et une trisomie d'un fragment du chromosome A (et inversement). La ségrégation adjacente 2 donnera des zygotes avec une trisomie du chromosome A à l'exception du fragment transloqué qui sera en nombre normal et une monosomie du chromosome B à l'exception du fragment transloqué qui sera en nombre normal (et inversement).

3.4.2 Inversions

Les inversions entraînent des difficultés d'appariement au moment de la méiose, avec formation d'une boucle d'appariement, qui permet de replacer les fragments d'ADN dans la même orientation.

Pour les inversions péricentriques, s'il y a une recombinaison avec *crossing over* dans cette boucle d'appariement, il y a alors formation de 2 gamètes avec une perte d'un fragment pour l'un et un gain pour l'autre (appelé duplication-déficiences). Il y a perte ou gain des fragments distaux par rapport aux points de cassure.

Pour les inversions paracentriques, le segment en gain ou perte inclut le centromère ce qui entraîne la formation de gamètes avec des chromosomes acentriques ou dicentriques, qui sont soit perdus, soit instables (Figure 35).

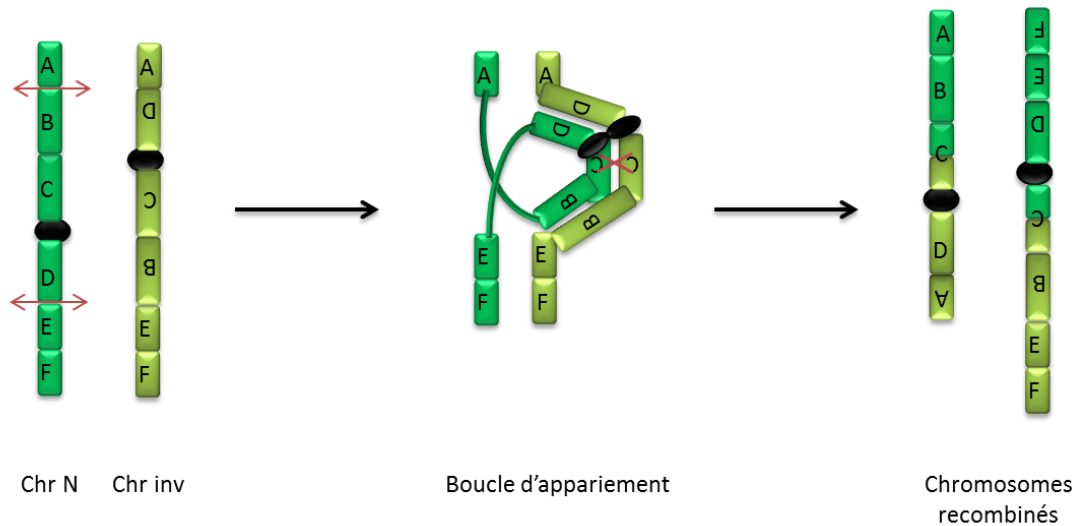


Figure 35. Recombinaison méiotique en cas d'inversion péricentrique

Le chromosome normal (Chr N) va former une boucle d'appariement afin de s'aligner correctement avec le chromosome inversé (Chr inv). Un *crossing over* (croix rouge) est représenté dans la boucle d'inversion au niveau du fragment C. Cela aboutit à la formation de deux chromosomes recombinés déséquilibrés : le premier avec le fragment A dupliqué et les fragments E et F délétés et le deuxième avec le fragment A délété et les fragments E et F dupliqués.

3.4.3 Insertions

Les insertions peuvent amener à la formation de gamètes trisomiques ou monosomiques pour le segment inséré en fonction de la ségrégation des chromosomes (Figure 36).

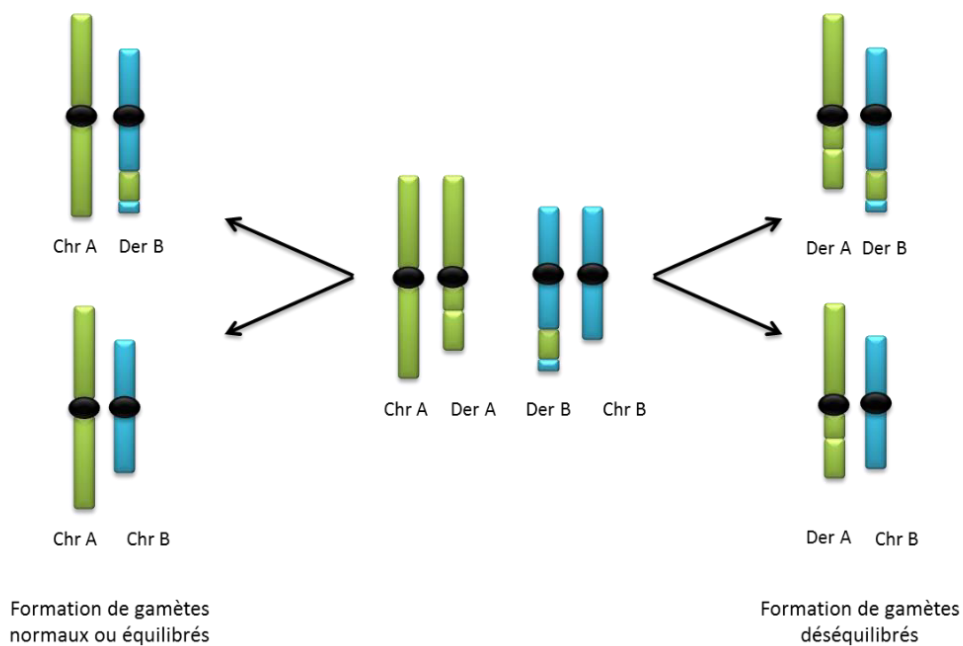


Figure 36. Possibilités de formation de gamètes en cas d'insertion

3.4.4 Remaniements complexes

Les principales études concernant les remaniements complexes ont porté sur ceux qui impliquent un échange entre trois chromosomes. Il y a alors formation d'un hexavalent lors de la méiose (Figure 137). Il existe quatre grands types de ségrégation : 1. Les ségrégations de type 3:3 où trois chromosomes sont transmis dans chaque gamète ; 2. Les ségrégations de type 4:2 où quatre chromosomes sont transmis dans un des gamètes, et deux chromosomes dans le deuxième ; 3. Les ségrégations 5:1 où cinq chromosomes sont transmis dans un des gamètes, et un chromosome dans le deuxième ; 4. Les ségrégations 6:0 où les six chromosomes sont dans un gamète et le deuxième n'a pas de chromosome. Dans le cas des ségrégations 3:3, la ségrégation peut être alterne ou adjacente. La ségrégation alterne produit des gamètes normaux ou équilibrés alors que la ségrégation adjacente produit des gamètes anormaux. Les cas rapportés de naissance dans des familles dont un parent est porteur d'un remaniement complexe montre que la ségrégation 3:3 adjacente 1 est la plus fréquente, suivie de la ségrégation 4:2 [125]. Saadalaah and Hulten ont observé un nombre réduit de chiasmata entraînant donc rarement des recombinaisons [126]. Plus les réarrangements sont complexes, plus le risque d'avoir des gamètes anormaux est important.

L'analyse du sperme dans des remaniements complexes de type échange entre trois chromosomes montre une faible proportion de gamètes normaux ou équilibrés (13-27%). De plus tous les patients présentaient une oligospermie, traduisant l'effet délétère de ces remaniements sur la spermatogenèse [75].

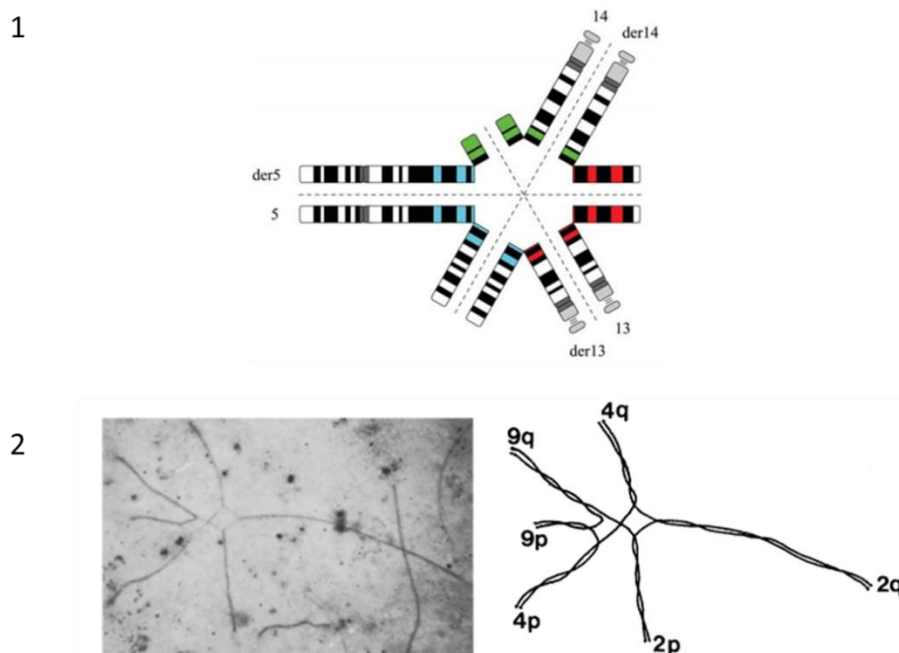


Figure 37. Formation d'un hexavalent lors d'un remaniement complexe (translocation entre trois chromosomes).

1. D'après Pellestor et al., 2011. Formation d'un hexasome dans le cas d'une translocation t(5;13;14). 2. D'après Saadallah and Hulten, 1985. Microscopie électronique montrant un hexasome dans le cas d'une translocation t(2;4;9).

3.4.5 Cas de transmission familiale de chromoanagenesis

Douze familles avec transmission de chromoanagenesis sont décrites à ce jour [127–134], avec trois cas de chromoanasythesis [127,133,134] et neuf cas de chromothripsis [128–132]. Dans la majorité des familles, le parent porteur est asymptomatique. La transmission du chromoanagenesis est décrite comme stable dans seulement quatre familles [127,129,133,134], dans les autres cas elle est déséquilibrée. Une recombinaison méiotique ou une mauvaise ségrégation des chromosomes dérivés lors de la méiose sont responsables d'un nouveau remaniement chez l'enfant, responsable du phénotype.

OBJECTIFS

L'organisation du génome peut être perturbée par les RCAE, ce qui peut induire une modification de l'expression génique et entraîner l'apparition de troubles du neurodéveloppement. Pour aborder les mécanismes mis en jeu, nous allons réaliser un travail en trois parties afin de répondre aux questions suivantes : Les RCAE peuvent-ils modifier l'expression d'un gène différemment en fonction de la localisation d'un point de cassure par rapport au site d'initiation de la transcription ? Peuvent-ils perturber l'architecture 3D du génome, et quels en sont les effets au niveau du phénotype ? Des remaniements très complexes peuvent-ils être transmis à la descendance ?

La première partie portera sur la caractérisation de remaniements proximaux du gène *TCF4*. Ce gène complexe, avec de nombreux transcrits de différentes tailles est connu pour être responsable du syndrome de Pitt-Hopkins associant déficience intellectuelle sévère, absence de langage et une dysmorphie faciale caractéristique. Cependant certains patients présentent un phénotype plus modéré avec une déficience intellectuelle légère ou modérée et un retard de langage.

La deuxième partie de cette thèse explorera les conséquences des remaniements chromosomiques sur l'architecture tridimensionnelle du noyau. Pour cela nous avons réalisé des techniques d'hybridation *in situ* en fluorescence en 3D, afin de déterminer la position radiale des loci impliqués dans les remaniements étudiés.

La troisième partie de ce travail a pour objectif de caractériser par caryotype, hybridation *in situ* en fluorescence, analyse chromosomique sur puce à ADN et séquençage du génome entier les chromoanagenesis de six familles dans lesquelles un parent asymptotique a transmis, de manière déséquilibrée, un remaniement chromosomique à sa descendance selon un mécanisme inhabituel.

Ce projet de thèse permettra d'explorer différentes répercussions des remaniements chromosomiques, au niveau de la séquence génomique du patient, de l'organisation spatiale de la chromatine et de la descendance par déséquilibre lors de la méiose.

RESULTATS

PARTIE 1 : Etude des remaniements proximaux du gène *TCF4*

Le syndrome de Pitt Hopkins (PTHS) associe une déficience intellectuelle sévère et une dysmorphie caractéristique. Il est souvent accompagné de crises d'épilepsie sévères et de troubles de la respiration à type d'accès d'hyperventilation.

Ce syndrome est dû à des anomalies du gène *TCF4*, comprenant 20 exons et ayant de nombreuses isoformes de taille courte (exons 9 -20), intermédiaire (exons 7-20) ou longue (exons 1-20), codant pour une protéine « basic helix-loop-helix » (bHLH) qui est un facteur de transcription. La majorité des cas décrits sont des mutations hétérozygotes de la région critique du gène (exons 7-19) ou des délétions plus larges.

Nous rapportons ici le cas de trois patients présentant un remaniement en début de gène. La première patiente est une fille de six ans porteuse d'une inversion péricentrique du chromosome 18, 46,XX,inv(18)(p11.2q21.1). Le séquençage du génome entier a caractérisé le point de cassure à la base près, aux position chr18:1262334_1262336 et chr18:53254747_53254751 (hg19). Ce point de cassure a perturbé le site proximal d'initiation de la transcription (TSS) de *TCF4*, dans le premier intron du gène. Les deuxièmes et troisièmes patients sont un fils et sa mère, porteurs d'une délétion de 39kb caractérisée par une micro-puce chromosomique à haute résolution (chr18:53253833_53282302, hg19) englobant les deux premiers exons du gène *TCF4* et incluant le TSS proximal. Ces patients n'ont pas un phénotype classique mais présentent une déficience intellectuelle modérée, ils n'ont pas de dysmorphie particulière, ni de crises d'épilepsie ou d'accès d'hyperventilation.

Nous avons réalisé une étude d'expression par RT-qPCR (PCR quantitative réalisée après une reverse transcription) sur des lymphocytes sanguins chez deux de ces patients ainsi que chez un patient avec un syndrome de Pitt Hopkins typique et chez des témoins sains. Plusieurs couples d'amorces, situées en partie proximale ou distale du gène, ont été utilisés afin de différencier les formes longues de l'ensemble des isoformes (forme totale).

Nous observons une diminution d'environ 50% de l'expression de la forme totale du gène *TCF4* chez l'ensemble des patients par rapport aux témoins. L'expression des formes courtes est également diminuée chez l'ensemble des patients, mais plus fortement chez les patients avec un phénotype classique. Les patients avec un remaniement en début de gène présentent donc une augmentation du taux d'ARNm pour les isoformes les plus courtes par rapport aux patients présentant un syndrome de Pitt Hopkins typique.

Les patients décrits ici, ainsi que les patients précédemment rapportés présentant des altérations structurelles proximales du gène *TCF4*, permettent de délimiter un phénotype de déficience intellectuelle légère avec une dysmorphie faciale non spécifique sans caractéristiques de PTHS. Il suggère également un gradient de sévérité phénotypique inversement corrélé au nombre de sites d'initiation de la transcription intacts. L'expression d'isoformes courtes compense en partie la perte d'isoformes plus longues.

European Journal of Medical Genetics
Disruption of TCF4 proximal promoter is associated with mild intellectual disability
 –Manuscript Draft–

Manuscript Number:	EJMG-D-21-00462
Article Type:	Research Paper
Section/Category:	Clinical Report
Keywords:	TCF4; Pitt Hopkins syndrome; Intellectual Disability; Next Generation Sequencing
Corresponding Author:	Julie Masson Hospices Civils de Lyon Centre de pathologie et biologie Est FRANCE
First Author:	Julie Masson
Order of Authors:	Julie Masson Linda Pons Tiffany Busa Chantal Missirian Matthew Lines Hélène Tevissen Flavie Diguët Pierre-Antoine Rollat-Farnier Damien Sanlaville Caroline Schluth-Bolard
Abstract:	<p>TCF4 gene (18q21.1) encodes for a transcription factor with multiple isoforms playing a critical role during neurodevelopment. Molecular alterations of this gene are associated with Pitt-Hopkins syndrome, a severe condition characterized by intellectual disability, specific facial features and autonomic nervous system dysfunction.</p> <p>We report here three patients presenting with structural variations of the proximal part of TCF4 associated with a mild phenotype. The first patient is a six-years-old girl carrier of a pericentric inversion of chromosome 18, 46,XX,inv(18)(p11.2q21.1). Whole genome sequencing characterized the breakpoint at the base-pair level at chr18:1262334_1262336 and chr18:53254747_53254751 (hg19). This latter breakpoint disrupted the proximal transcription start site (TSS) of TCF4 in the first intron of the gene. The second and third patients are a son and his mother, carrier of a 39kb deletion characterized by high-resolution chromosomal micro-array (chr18:53253833_53282302, hg19) encompassing the two first exon of TCF4 gene and including the proximal TSS. Expression studies on blood lymphocytes in these patients showed a marked decrease of mRNA level for long isoforms of TCF4 and an increased level for shorter isoforms.</p> <p>The patients described here, together with previously reported patients with proximal structural alterations of TCF4, help to delineate a phenotype of mild ID with non-specific facial dysmorphism without characteristic features of PTHS. It also suggests a gradient of phenotypic severity inversely correlated with the number of intact TCF4 TSS, with expression of short isoforms compensating in part the loss of longer isoforms.</p>
Suggested Reviewers:	<p>Irina Giurgea irina.giurgea@inserm.fr Clinical expertise in Pitt Hopkins Syndrome</p> <p>Niels Tommerup ntommerup@sund.ku.dk Expertise in balanced chromosomal rearrangement characterization</p>

1 **Disruption of TCF4 proximal promoter is associated with mild intellectual**
2 **disability**

3

4 Julie Masson^{1,2,3*}, Linda Pons^{1*}, Tiffany Busa⁴, Chantal Missirian⁴, Matthew Lines⁵, H el ene Tevissen⁶,
5 Flavie Digu et¹, Pierre-Antoine Rollat-Farnier¹, Damien Sanlaville^{1,2,3}, Caroline Schluth-Bolard^{1,2,3}

6 ¹Service de G en etique, Laboratoire de Cytog en etique Constitutionnelle, Centre de Biologie et de Pathologie Est, Hospices
7 Civils de Lyon, Bron, France

8 ²Lyon Neuroscience Research Center, GENDEV Team, CNRS UMR 5292, INSERM U1028, UCBL1, Bron, France

9 ³Universit e Claude Bernard Lyon 1, France

10 ⁴Service de G en etique, H opital de la Timone Enfant, APHM, Marseille, France

11 ⁵Division of Metabolics, Children’s Hospital of Eastern Ontario, Ottawa, Ontario, Canada

12 ⁶Centre Hospitalier de Valence, Service de p diatrie, Valence, France

13 * These authors should be regarded as joint first authors

14

15 **Corresponding author:** Dr Caroline Schluth-Bolard

16 Service de G en etique, Laboratoire de Cytog en etique Constitutionnelle - Centre de Biologie et de Pathologie Est
17 - 59, boulevard Pinel, BRON 69677 Cedex, France - caroline.schluth@yahoo.fr

18

19 **Manuscript word count:** 2433

20 **ABSTRACT**

21 *TCF4* gene (18q21.1) encodes for a transcription factor with multiple isoforms playing a critical role
22 during neurodevelopment. Molecular alterations of this gene are associated with Pitt-Hopkins
23 syndrome, a severe condition characterized by intellectual disability, specific facial features and
24 autonomic nervous system dysfunction.

25 We report here three patients presenting with structural variations of the proximal part of *TCF4*
26 associated with a mild phenotype. The first patient is a six-years-old girl carrier of a pericentric
27 inversion of chromosome 18, 46,XX,inv(18)(p11.2q21.1). Whole genome sequencing characterized
28 the breakpoint at the base-pair level at chr18:1262334_1262336 and chr18:53254747_53254751
29 (hg19). This latter breakpoint disrupted the proximal transcription start site (TSS) of *TCF4* in the first
30 intron of the gene. The second and third patients are a son and his mother, carrier of a 39kb deletion
31 characterized by high-resolution chromosomal micro-array (chr18:53253833_53282302, hg19)
32 encompassing the two first exon of *TCF4* gene and including the proximal TSS. Expression studies on
33 blood lymphocytes in these patients showed a marked decrease of mRNA level for long isoforms of
34 *TCF4* and an increased level for shorter isoforms.

35 The patients described here, together with previously reported patients with proximal structural
36 alterations of *TCF4*, help to delineate a phenotype of mild ID with non-specific facial dysmorphism
37 without characteristic features of PTHS. It also suggests a gradient of phenotypic severity inversely
38 correlated with the number of intact *TCF4* TSS, with expression of short isoforms compensating in
39 part the loss of longer isoforms.

40 **Word count: 237**

41 **KEY WORDS:** *TCF4*, Pitt Hopkins syndrome, intellectual disability, next generation sequencing

42 INTRODUCTION

43 TCF4 is a transcription factor that belongs to the class I Helix-loop-Helix protein family (E proteins). It
44 binds DNA as a dimer at Ephrussi (E) box sequences (CANNTG) to regulate expression of target genes.
45 TCF4 protein is composed by several domains including activation domains (AD1, AD2, AD3), C-
46 terminal bHLH-domain, two regulatory domains, CE and Rep, and nuclear localization signal (NLS-1,
47 NLS-2) and nuclear export signal (NES-1, NES-2) (Fig. 1A). The gene is located on the long arm of
48 chromosome 18, at 18q21.2. It shows a complex structure including many alternative exons and
49 alternative transcription start sites (TSS), producing 18 N-terminally distinct isoforms that differ in
50 tissue expression, subcellular localization and transcriptional activity (Sepp et al., 2011). All TCF4
51 transcripts have exons 9-19 in common (NM_001083962.1). All TCF4 isoforms have AD2, AD3 and
52 bHLH-domain in common. Small transcripts encode proteins without NLS-1, but the bHLH domains
53 contains an NLS (NLS-2) and two NESs allowing TCF4 exchange between nuclei and cytoplasm (Greb-
54 Markiewicz et al., 2019). *TCF4* is expressed in multiple tissue during the development with a high
55 level expression in the embryonic central nervous system. TCF4 plays a particular role during
56 neurodevelopment, including neurogenesis and neuronal differentiation but also in post-mitotic
57 neurons function (Schoof et al., 2020). It has been associated to various neurological conditions
58 through different mechanisms, some of them being not fully characterized. *TCF4* polymorphisms
59 have been associated to schizophrenia and other psychiatric disorders such as bipolar troubles,
60 depression and post-traumatic stress (Teixeira et al., 2021). Expansion of CTG trinucleotides in intron
61 1 of the gene is responsible for Fuchs corneal dystrophy through the disturbance of MBNL1 splicing
62 pathway (Wieben et al., 2012). *TCF4* is also involved in Pitt-Hopkins syndrome (PTHS), a rare
63 neurodevelopmental disorder, presenting with severe intellectual disability, seizures, microcephaly,
64 hypotonia, autistic features, characteristic facial appearance (deep set eyes, midface protrusion,
65 large mouth) and typical hyperventilation access. Molecular alterations responsible for PTHS are
66 various and are responsible for haplo-insufficiency. They include missense, nonsense and splice site
67 point mutations, small insertions or deletion leading to frameshift, as well as large deletion involving

68 all or part of the gene (Amiel et al., 2007; Marangi et al., 2011; Whalen et al., 2012; Zweier et al.,
69 2008). In some rare cases, disruption by translocation breakpoint has been described (Schluth-Bolard
70 et al., 2013; Talkowski et al., 2012). All these alterations are mainly located between exon 7 and 20.
71 More recently, milder phenotypes associated with *TCF4* alterations have emerged, including isolated
72 mild to moderate intellectual disability (Kalscheuer et al., 2008; Kharbanda et al., 2016; Maduro et
73 al., 2016). They involved variants in the proximal part of the genes (exon 1 to 7), including small
74 deletions and gene disruption. A genotype-phenotype correlation attempt was proposed relying on
75 *TCF4* structure, suggesting that variants within exons 1-4 and exons 4-6 give rise to a recurrent
76 phenotype with mild intellectual disability; variants within exons 7-8 cause a severe phenotype
77 without typical facial features; variants within exons 9-19 cause typical Pitt-Hopkins syndrome
78 (Bedeschi et al., 2017; Zollino et al., 2019).

79 We describe herein alterations of the proximal promoter of *TCF4* in three patients (two unrelated
80 families) associated with mild intellectual disability without typical features of PTHS, expanding the
81 phenotype and molecular mechanism spectrum of *TCF4* proximal alterations.

82 **Clinical report**

83 **Patient 1**

84 The patient 1 is a 6 years-old Caucasian girl born to healthy unrelated parents. She was born at full
85 term, (weight: 4225g (+2 SDS), no length and head circumference data available), without any
86 distress. She was diagnosed with constipation and gastro-oesophageal reflux at 3 years of age. She
87 sat independently at 10 months of age, could walk unaided from 18 months of age, and could run
88 and climb from 6 years of age. She had difficulties in fine motor skills and required for example,
89 assistance with buttons and fasteners. She developed some speech from about 3 years of age and is
90 now able to write her name. Toilet training is still incomplete at night. She presents some
91 behavioural troubles with hyperactivity, anxiety, frequent temper tantrums and sleep disorders.

92 Hearing tests were normal. She had glasses for strabismus. Psychometric tests concluded to mild
93 intellectual disability without autistic troubles. Growth was normal. Mild dysmorphism was noticed
94 (Fig. 2A and 2) with deeply set eyes, prominent nasal bridge, slightly low-hanging columella and
95 everted vermilion of the lower lip. No respiratory abnormalities were described. Brain MRI and EEG
96 were not performed. Clinical scoring for PTSH according to Whalen et al. was 6/20. Standard blood
97 karyotype showed a balanced pericentric *de novo* inversion of chromosome 18
98 46,XX,inv(18)(p11.2;q21.1) (Fig. 3A). Microarray comparative genomic hybridization showed no
99 cryptic imbalance. Breakpoints were further characterized by whole genome sequencing using the
100 Illumina TruSeq protocol (Illumina, San Diego, California, USA) and 101 bp-reads paired-end
101 sequencing on Illumina NextSeq 500. Alignment on reference genome hg19 and breakpoint calling
102 were performed as previously described (Schluth-Bolard et al., 2019). WGS yielded a 16.7X mean
103 sequencing depth. It made it possible to delineate the breakpoints on chromosome 18 in a 147 bp
104 region at 18p11.32 (chr18:1262295_1262441) and in a 25 bp region at 18q21.2
105 (chr18:53,254,733_53,254,758). After PCR amplification, Sanger sequencing of junction fragments,
106 identified the breakpoints on chromosome 18 at the positions chr18:1262334_1262336 at 18p11.32
107 and chr18:53254747_53254751 at 18q21.2 (Fig. 3B and supplemental figure 1). These breakpoints
108 were further verified by FISH using RP11-372E2 at 18p11.32 and RP11-569E16 at 18q21.2 BAC clones
109 (BlueFish, BlueGnome, Cambridge, UK) (supplemental figure 1). For each probe it showed a split
110 signal on either derivative chromosome. The breakpoint on 18p11.32 disrupted a long non-coding
111 RNA sequence *LINC00470* and the breakpoint on 18q21.2 disrupted the *TCF4* gene in the first intron
112 (NM_001083962.1) in a transcription start site according to GeneHancer and Fantom5 databases.

113 **Patient 2**

114 The patient 2 is a 3 years-old Caucasian boy born to unrelated parents. He was born at full term,
115 (weight: 3670g (+0.5 SDS), length: 50.5cm (+0.5 SDS), head circumference: 34cm (-1 SDS)), without
116 any distress. He underwent left hip dysplasia surgery at 3 years of age. He sat independently at 10

117 months of age, could walk unaided from 24 months of age. He developed some speech from about 3
118 years and started to associate few words. Behaviour was agitated but very friendly. At 30 months of
119 age, weight was 13kg (-0.5 SDS), length was 83cm (-2 SDS), and head circumference was 46cm (-2
120 SDS). Mild dysmorphism was noticed (Fig. 2C and 2D) with deeply set eyes, horizontal eyebrows, thin
121 vermilion of the upper lip and brachycephaly. No respiratory abnormalities were described. EEG was
122 normal. Brain MRI was not performed. Clinical scoring for PTSH according to Whalen et al. was 4/20.
123 Standard karyotype was normal. CMA was performed using a 180,000-oligonucleotide microarray
124 (SurePrint G3 Human CGH Microarray Kit, 4 × 180 K, Agilent Technologies, Santa Clara, CA, USA) and
125 then a 1 × 1M-oligonucleotide microarray (SurePrint G3 Human CGH Microarray Kit, 1x1M, Agilent
126 Technologies) to further refine the CNV, according to the manufacturer's instructions. It showed a
127 18q21.2 deletion with a minimal size of 39 kb extending from position chr18:53253833 to
128 chr18:53282302 and a maximal size extending from position chr18:53242214 to chr18:53288196
129 (Fig. 3C). The deletion was confirmed by quantitative PCR using primers in intron 3 of *TCF4*
130 (NM_001083962.1). This deletion contained the first three exons of *TCF4* (NM_001083962.1).

131 **Patient 3**

132 The patient 3 is the mother of patient 2. She was born from unrelated parents. We lack information
133 about initial psychomotor development. She underwent right hip dysplasia surgery at 12 months of
134 age. She attended ordinary primary schooling and then required special education since 10 years of
135 age. She is 36 years-old and independent for daily life. She can read, write and count. She works as a
136 maid. Head circumference is 54cm (-1 SDS). Mild dysmorphism was also noticed (Fig. 2E and 2F) with
137 deep set eyes, broad nasal bridge, low-set ears and brachycephaly. Standard karyotype was normal.
138 CMA shows the same CNV as in her son. Quantitative PCR using primers in intron 3 confirmed the
139 deletion.

140 **Expression study of TCF4 isoforms**

141 RNA was extracted from whole blood with the PAXgene Blood RNA kit (Qiagen, Venlo, Netherlands)
142 from patients 1 and 3. As controls, we used RNA from two healthy women without chromosomal
143 rearrangements and RNA from a patient presenting with a typical PTHS due to *TCF4* disruption in
144 intron 6 by a translocation described by Schluth-Bolard et al. (Schluth-Bolard et al., 2013), and
145 recently clinically evaluated (supplemental table 1). Reverse transcription into cDNA was performed
146 with the SuperScript II Reverse Transcriptase kit (Invitrogen, Carlsbad, California, USA) from 0.5 µg of
147 RNA with random primers. Five primer pairs (supplemental table 2) were selected on *TCF4*
148 (NM_1083962.1) with exons 4-5 and exons 5-6 representing the longer isoforms, exons 7-8
149 representing the intermediary isoform and exons 9-10 and exons 14-15 representing long and short
150 isoforms. cDNA were amplified using the QuantiTect SYBR Green PCR kit (Qiagen, Venlo,
151 Netherlands) on the LightCycler 480 (Roche, Bale, Switzerland) according to standard protocol.
152 Results were then normalised on the expression of *ACTB* gene and relative mRNA ratio calculated
153 according to the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). Statistical analysis were performed
154 using the R software. Patients and PTHS control mRNA ratio means were compared to normal control
155 (fixed value, $\mu=1$) using a student t-test with Bonferroni corrections. Patients and PTHS control
156 mRNA ratio means were compared using one-way analysis of variance (ANOVA) followed by post-hoc
157 testing with Tukey's honest significant difference (HSD) test for pairwise comparisons. We set the
158 significance level at 5%.

159 Expression level of all *TCF4* transcripts was globally decreased in patients 1 and 3 and in PTHS control
160 compared to healthy controls (all corrected p-values < 0,01, except for exons 7-8 with p-value <0,5)
161 (Fig. 4). When comparing the patients to PTHS control, the decrease of expression level was greater
162 for long isoforms (p-value <0.001 for exons 4-5 and exons 5-6) while it was similar for the
163 intermediary isoform (p-value non-significant (NS) for exons 7-8) and the total amount of isoforms

164 (p-value NS for exons 10-11 and exons 15-16). This suggested that expression level of short isoforms
165 might be increased in patients 1 and 3 to compensate the decrease of long isoforms expression.

166 **DISCUSSION**

167 We characterised herein three patients (one sporadic and two related) presenting alterations of the
168 first TSS of *TCF4* associated with a mild phenotype. *TCF4* is associated with Pitt-Hopkins syndrome, a
169 severe autosomal dominant condition characterized by severe intellectual disability with absent or
170 limited speech, specific facial features including narrow forehead, thin lateral eyebrows, wide nasal
171 bridge, flared nasal alae, full cheeks, wide mouth and thickened helices, marked autonomic nervous
172 system dysfunction with breathing regulation anomalies (Zollino et al., 2019) . Other features include
173 myopia, constipation, unstable gait, slender fingers and abnormal palmar creases. This phenotype is
174 due to variants, including nonsense, frameshift, splice site and missense, affecting exons 7 to 19 of
175 *TCF4* (Whalen et al., 2012; Zweier et al., 2008). Recently, atypical phenotypes related to alterations of
176 the proximal part of the gene were additionally described and attempts for genotype-phenotype
177 correlations were proposed (Bedeschi et al., 2017; Zollino et al., 2019).

178 We reviewed structural balanced and unbalanced variants affecting the proximal part of *TCF4* (exons
179 1 to 8) (Fig 1B). We could identify three groups according to the severity of the phenotype. Then we
180 correlated these groups to the 6 TSS identified for *TCF4* according to Fantom5 atlas and Genhancer.
181 The first group is composed of eight patients (Kalscheuer et al., 2008; Kharbanda et al., 2016;
182 Maduro et al., 2016), including the three patients described herein, with variants in the three first
183 exons of *TCF4* (Table 1). Their phenotype is characterized by mild intellectual disability. Walking was
184 delayed but acquired between 18 to 24 months. Acquisition of speech was also delayed but present
185 and patients were able to learn writing. There may be some behavioural troubles, in particular
186 hyperactivity. Adults were all autonomous for daily life. Facial dysmorphism is discrete and non-
187 specific, high forehead and pointed chin being the most common features. There were no feature
188 evocative of PTHS, in particular no autism, no stereotypy, no breathing anomalies nor seizures.

189 Alterations included either disruption by translocation breakpoints or microdeletion, but no single
190 nucleotide variant was described so far in this part of gene (Fig. 1B). These variants affected TSS1
191 and/or TSS2 of *TCF4*. In patient 1, the only disruption of TSS1 was sufficient for the phenotype,
192 whereas in the family described by Maduro et al., breakpoints were located between TSS1 and TSS2.
193 Thus, only the longest isoforms seem to be affected in this group. Expression studies performed for
194 other patients (Kharbanda et al., 2016; Maduro et al., 2016) showed results concordant with our
195 findings, with a decreased expression of long isoforms and an increase of the shorter isoforms.

196 In the second group, patients showed alterations of exons 4 to 8 (Bedeschi et al., 2017; Goodspeed et
197 al., 2018; Whalen et al., 2012). They presented with a more severe phenotype, with mild to
198 moderate intellectual disability. Walking was acquired later between 18 months to 5 years. Language
199 was severely delayed and most of patients were non-verbal. Stereotypies, constipation, strabismus
200 and myopia were also present. Although accurate genomic coordinates were not available for all
201 cases, these deletions have affected TSS 3 and 4, but may have left intact TSS5, at the origin of the
202 shorter isoforms.

203 Finally, the patients of the third group presented structural variants involving exons 6 to 8 (Rosenfeld
204 et al., 2009; Schluth-Bolard et al., 2013; Whalen et al., 2012). They showed a severe phenotype, with
205 severe ID, walking acquired after 3 years, absence of language, breathing anomalies, stereotypies
206 and facial features characteristic of PTHS. The deletions encompassed TSS5, affecting transcription of
207 shorter isoforms. For patient 1 of Schluth-Bolard et al., (2013) this TSS is left intact by the
208 translocation breakpoint, but a position effect inactivating this TSS can be hypothesized to explain
209 the severe phenotype of the patient.

210 This suggests that shorter isoforms could compensate at least for a part the function of longer
211 isoform. These isoforms (from exon 9) contain the bHLH domain and NLS signals, important for the
212 protein function. Moreover, the longest isoforms of *TCF4* did not seem to be expressed in brain
213 structures (Sepp et al., 2011).

214 In conclusion, we further delineate the phenotype associated with structural variants involving the
215 first three exons of *TCF4*. Patients showed mild ID with a specific developmental pattern and some
216 facial features. This work also suggests a gradient of phenotypic severity associated to structural
217 variants of the proximal part of *TCF4*, with a correlation of the phenotype to the number of affected
218 TSS of *TCF4* and subsequently the number of affected isoforms.

219 **ACKNOWLEDGMENTS:**

220 The authors wish to thank patients and their families for their kind cooperation, all their colleagues
221 who provided help for this article and the Hospices Civils de Lyon for their grant support.

222 **AUTHOR CONTRIBUTION:**

223 Caroline Schluth-Bolard and Damien Sanlaville designed the study. Linda Pons, Julie Masson, Flavie
224 Diguët, Pierre-Antoine Rollat-Farnier, Caroline Schluth-Bolard and Damien Sanlaville analysed the
225 data. Matthews Lines and Tyffanie Busa were involved in the patients' care and evaluation. Linda
226 Pons and Julie Masson wrote the first draft of the manuscript and all other authors contributed to
227 the final form of manuscript.

228 **DECLARATION OF CONFLICTING INTERESTS:**

229 The authors have no conflicts of interest to declare.

230 **FUNDING:**

231 This work was supported by a Programme de Recherche Translationnelle en Santé of the Agence
232 Nationale de Recherche and of the Direction Générale de l'Offre de Soins.

233 **FIGURE TITLES AND LEGENDS**

234 **Figure 1** : Genotype/phenotype correlation

235 A: Representation of main proteins isoforms according to Sepp et al., 2011, leading to long (TCF4-B),
236 intermediate (TCF4-D) and short (TCF4-A) length protein. Coloured shaded area represented
237 functional domains. All isoforms included transcriptional activation domain 2 (AD2) and basic helix-
238 loop-helix (bHLH) domain. Short length forms lacked nuclear localisation signal 1 (NLS-1) domain and
239 transcriptional activation domain 1 (AD1) but had NLS-2. Intermediate length form included the NLS-
240 1 domain but lacked the AD1 domain. Exon numbering referred to transcript NM_001083962.1.

241 B: Review of proximal translocations and deletion (before exon 9) already reported. Vertical lines
242 represented translocations and horizontal lines represented deletions. Green lines represented
243 patient with mild phenotype (mild ID), yellow lines represented patients with intermediary
244 phenotype (mild to moderate ID, language severely delayed) and red lines patients with severe
245 phenotype (severe ID, no language). Green arrows represented the principal transcription start sites
246 (TSS) according to GeneHancer and Fantom5.

247 **Figure 2:** facial characteristics of patients

248 A-B: patient 1 at 6 years-old; C-D: patient 2 at 3 years-old; E-F: patient 3 at 36 years old (mother of
249 patient 2).

250 **Figure 3:** Characterization of chromosomal structural variants in patients 1 and 2.

251 A and B: Patient 1. a: partial G-band karyogram and ideogram showing chromosome 18 pericentric
252 inversion between 18p11.3 and 18q21.2 bands (400 bands resolution). b: sequences of junction
253 fragments for short arm junction (on top) and for long arm junction (bottom) (hg 19). Reference
254 sequences for the short arm (18p) are in red and for the long arm (18q) are in blue. +/- indicates
255 forward/reverse strand. The junction sequences are in red/blue (inv 18). C. CMA profile for patient 2
256 (1M microarray): arr[GRCh37] 18q21.2(53243833_53282302)x1 mat.

257 **Figure 4:** Expression of *TCF4* mRNA in blood for patients 1 and 3 (dark grey bar), as well as healthy
258 controls (white bar) and PTHS control (dark bar).

259 The results of statistical analyses performed using ANOVA to compared mRNA ratio means of the two
260 patients and the PTHS control, are shown as mean value +/- standard error of the mean (SEM). ***:
261 $p < 0.001$, NS: non-significant. . Exon numbering referred to transcript NM_001083962.1.

262 **TABLES (WITH A TITLE AND LEGEND FOR EACH)**

263 **Table 1:** Clinical characteristics of patients with exons 1-4 rearrangements.

	Kalscheuer et al., 2008	Kharbanda et al., 2016	P1 Maduro et al., 2016	P2 Maduro et al., 2016	P3 Maduro et al., 2016	P1	P2	P3	Total
Sex	F	M	M	M	F	F	M	F	5M/3F
Age (years)	15	11	4	31	53	6	3	36	
Karyotype	46,XX,t(18;20) (q21.1;q11.2)	?	46,XY,t(14;18) (q22;q21)	46,XY,t(14;18) (q22;q21)	46,XX,t(14;18) (q22;q21)	46,XX,inv(18) (p11.2;q21.1)	46,XY	46,XY	
CMA	arr(X,Y)x1,(1-22)x2	18q21.2 ?	arr(X,Y)x1,(1-22)x2	arr(X,Y)x1,(1-22)x2	arr(X,1-22)x2	arr(X,1-22)x2	18q21.2(53253833_5 3282302)x1	18q21.2(53253833_5 3282302)x1	
WGS (hg19)	chr18:53161764_ 53280103	chr18:53231903_ 53421483	chr18: 53200018_ 53219411	chr18: 53200018_ 53219411	chr18: 53200018_ 53219411	chr18:53254747_ 53254750	chr18:53253833_ 53282302	chr18:53253833_ 53282302	
Mechanism	disruption intron 3	deletion 189,5 kb exons 1-3	deletion 19kb intron 3	deletion 19kb intron 3	deletion 19kb intron 3	Disruption intron 1	deletion 39kb exon 1-3	deletion 39kb exon 1-3	6 deletions/ 2 disruptions
Growth	N	N	N	N	ns	N	-2SD	ns	
Development									
ID	moderate	mild	mild	mild	mild	mild	mild	mild	8/8
Age at walking (months)	17	delayed	18	delayed	ns	18	24	ns	6/6
Speech acquisition	delayed (> 20 m)	delayed	delayed (> 2 y)	delayed	ns	delayed (> 3y)	delayed (> 3y)	delayed	7/7
Writing	+	+	na	+	ns	+	na	+	5/5
unstable gait	-	-	+	-	-	-	-	-	1/8
Behavioral troubles	-	+	-	-	-	+	+	-	3/8
Facial dysmorphism									
brachycephaly	-	-	+	-	-	-	+	+	3/8
high forehead	+	-	+	+	+	+	+	+	7/8
deep set eyes	-	-	-	-	-	+	+	+	
horizontal eyebrows	-	-	-	-	-	+	+	-	3/8
upslanting palpebral fissures	-	-	+	+	-	-	-	-	2/8
broad nasal bridge	+	-	-	-	-	+	-	+	3/8
low-set /dysplastic ears	+	-	+	-	-	-	-	+	3/8
thin upper lip	-	-	-	-	-	+	+	-	2/8
everted lower-lip	-	-	-	-	-	+	+	+	3/8
pointed chin	-	-	+	+	-	+	+	+	5/8
high arched palate	-	-	+	+	-	-	-	-	2/8
Skeletal findings									
hip dysplasia	-	-	-	-	-	-	+	+	2/8
limb asymmetry	+	-	-	-	-	-	-	-	1/8
pectus carinatum	-	-	+	-	-	-	-	-	1/8
finger/toe clinodactyly	-	-	+	-	-	-	-	-	1/8
Other findings									
strabismus	-	-	-	-	-	+	-	-	1/8
hearing loss	+	-	-	-	-	-	-	-	1/8
constipation	-	-	-	-	-	+	-	-	1/8
GER	-	-	-	-	-	+	-	-	1/8
Whalen score			0	0	0	6	4	4	

264 N: normal; ns: not specified ; na : not applicable

REFERENCES

- Amiel, J., Rio, M., de Pontual, L., Redon, R., Malan, V., Boddaert, N., Plouin, P., Carter, N.P., Lyonnet, S., Munnich, A., Colleaux, L., 2007. Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *Am. J. Hum. Genet.* 80, 988–993. <https://doi.org/10.1086/515582>
- Bedeschi, M.F., Marangi, G., Calvello, M.R., Ricciardi, S., Leone, F.P.C., Baccarin, M., Gueneri, S., Orteschi, D., Murdolo, M., Lattante, S., Frangella, S., Keena, B., Harr, M.H., Zackai, E., Zollino, M., 2017. Impairment of different protein domains causes variable clinical presentation within Pitt-Hopkins syndrome and suggests intragenic molecular syndromology of TCF4. *Eur. J. Med. Genet.* 60, 565–571. <https://doi.org/10.1016/j.ejmg.2017.08.004>
- Goodspeed, K., Newsom, C., Morris, M.A., Powell, C., Evans, P., Golla, S., 2018. Pitt-Hopkins Syndrome: A Review of Current Literature, Clinical Approach, and 23-Patient Case Series. *J. Child Neurol.* 33, 233–244. <https://doi.org/10.1177/0883073817750490>
- Greb-Markiewicz, B., Kazana, W., Zarębski, M., Ożyhar, A., 2019. The subcellular localization of bHLH transcription factor TCF4 is mediated by multiple nuclear localization and nuclear export signals. *Sci. Rep.* 9, 15629. <https://doi.org/10.1038/s41598-019-52239-w>
- Kalscheuer, V.M., Feenstra, I., Ravenswaaij-Arts, C.M.A.V., Smeets, D.F.C.M., Menzel, C., Ullmann, R., Musante, L., Ropers, H.-H., 2008. Disruption of the TCF4 gene in a girl with mental retardation but without the classical Pitt–Hopkins syndrome. *Am. J. Med. Genet. A.* 146A, 2053–2059. <https://doi.org/10.1002/ajmg.a.32419>
- Kharbanda, M., Kannike, K., Lampe, A., Berg, J., Timmusk, T., Sepp, M., 2016. Partial deletion of TCF4 in three generation family with non-syndromic intellectual disability, without features of Pitt-Hopkins syndrome. *Eur. J. Med. Genet.* 59, 310–314. <https://doi.org/10.1016/j.ejmg.2016.04.003>
- Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} Method. *Methods San Diego Calif* 25, 402–408. <https://doi.org/10.1006/meth.2001.1262>
- Maduro, V., Pusey, B.N., Cherukuri, P.F., Atkins, P., du Souich, C., Rupps, R., Limbos, M., Adams, D.R., Bhatt, S.S., Eydoux, P., Links, A.E., Lehman, A., Malicdan, M.C., Mason, C.E., Morimoto, M., Mullikin, J.C., Sear, A., Van Karnebeek, C., Stankiewicz, P., Gahl, W.A., Toro, C., Boerkoel, C.F., 2016. Complex translocation disrupting TCF4 and altering TCF4 isoform expression segregates as mild autosomal dominant intellectual disability. *Orphanet J. Rare Dis.* 11, 62. <https://doi.org/10.1186/s13023-016-0439-6>
- Marangi, G., Ricciardi, S., Orteschi, D., Lattante, S., Murdolo, M., Dallapiccola, B., Biscione, C., Lecce, R., Chiurazzi, P., Romano, C., Greco, D., Pettinato, R., Sorge, G., Pantaleoni, C., Alfei, E., Toldo, I., Magnani, C., Bonanni, P., Martinez, F., Serra, G., Battaglia, D., Lettori, D., Vasco, G., Baroncini, A., Daolio, C., Zollino, M., 2011. The Pitt-Hopkins syndrome: Report of 16 new patients and clinical diagnostic criteria. *Am. J. Med. Genet. A.* 155, 1536–1545. <https://doi.org/10.1002/ajmg.a.34070>
- Rosenfeld, J.A., Leppig, K., Ballif, B.C., Thiese, H., Erdie-Lalena, C., Bawle, E., Sastry, S., Spence, J.E., Bandholz, A., Surti, U., Zonana, J., Keller, K., Meschino, W., Bejjani, B.A., Torchia, B.S., Shaffer, L.G., 2009. Genotype-phenotype analysis of TCF4 mutations causing Pitt-Hopkins syndrome shows increased seizure activity with missense mutations. *Genet. Med. Off. J. Am. Coll. Med. Genet.* 11, 797–805. <https://doi.org/10.1097/GIM.0b013e3181bd38a9>
- Schluth-Bolard, C., Diguët, F., Chatron, N., Rollat-Farnier, P.-A., Bardel, C., Afenjar, A., Amblard, F., Amiel, J., Blesson, S., Callier, P., Capri, Y., Collignon, P., Cordier, M.-P., Coubes, C., Demeer, B., Chaussonot, A., Demurger, F., Devillard, F., Doco-Fenzy, M., Dupont, C., Dupont, J.-M., Dupuis-Girod, S., Faivre, L., Gilbert-Dussardier, B., Guerrot, A.-M., Houlier, M., Isidor, B., Jaillard, S., Joly-Hélas, G., Kremer, V., Lacombe, D., Caignec, C.L., Lebbard, A., Lebrun, M., Lesca, G., Lespinasse, J., Levy, J., Malan, V., Mathieu-Dramard, M., Masson, J., Masurel-Paulet, A., Mignot, C., Missirian, C., Morice-Picard, F., Moutton, S., Nadeau, G., Pebrel-

- Richard, C., Odent, S., Paquis-Flucklinger, V., Pasquier, L., Philip, N., Plutino, M., Pons, L., Portnoï, M.-F., Prieur, F., Puechberty, J., Putoux, A., Rio, M., Rooryck-Thambo, C., Rossi, M., Sarret, C., Satre, V., Siffroi, J.-P., Till, M., Touraine, R., Toutain, A., Toutain, J., Valence, S., Verloes, A., Whalen, S., Edery, P., Tabet, A.-C., Sanlaville, D., 2019. Whole genome paired-end sequencing elucidates functional and phenotypic consequences of balanced chromosomal rearrangement in patients with developmental disorders. *J. Med. Genet.* 56, 526–535. <https://doi.org/10.1136/jmedgenet-2018-105778>
- Schluth-Bolard, C., Labalme, A., Cordier, M.-P., Till, M., Nadeau, G., Tevissen, H., Lesca, G., Boutry-Kryza, N., Rossignol, S., Rocas, D., Dubruc, E., Edery, P., Sanlaville, D., 2013. Breakpoint mapping by next generation sequencing reveals causative gene disruption in patients carrying apparently balanced chromosome rearrangements with intellectual deficiency and/or congenital malformations. *J. Med. Genet.* 50, 144–150. <https://doi.org/10.1136/jmedgenet-2012-101351>
- Schoof, M., Hellwig, M., Harrison, L., Holdhof, D., Lauffer, M.C., Niesen, J., Viridi, S., Indenbirken, D., Schüller, U., 2020. The basic helix-loop-helix transcription factor TCF4 impacts brain architecture as well as neuronal morphology and differentiation. *Eur. J. Neurosci.* 51, 2219–2235. <https://doi.org/10.1111/ejn.14674>
- Sepp, M., Kannike, K., Eesmaa, A., Urb, M., Timmusk, T., 2011. Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5' exon usage and splicing. *PLoS One* 6, e22138. <https://doi.org/10.1371/journal.pone.0022138>
- Talkowski, M.E., Rosenfeld, J.A., Blumenthal, I., Pillalamarri, V., Chiang, C., Heilbut, A., Ernst, C., Hanscom, C., Rossin, E., Lindgren, A.M., Pereira, S., Ruderfer, D., Kirby, A., Ripke, S., Harris, D.J., Lee, J.-H., Ha, K., Kim, H.-G., Solomon, B.D., Gropman, A.L., Lucente, D., Sims, K., Ohsumi, T.K., Borowsky, M.L., Loranger, S., Quade, B., Lage, K., Miles, J., Wu, B.-L., Shen, Y., Neale, B., Shaffer, L.G., Daly, M.J., Morton, C.C., Gusella, J.F., 2012. Sequencing Chromosomal Abnormalities Reveals Neurodevelopmental Loci that Confer Risk across Diagnostic Boundaries. *Cell* 149, 525–537. <https://doi.org/10.1016/j.cell.2012.03.028>
- Teixeira, J.R., Szeto, R.A., Carvalho, V.M.A., Muotri, A.R., Papes, F., 2021. Transcription factor 4 and its association with psychiatric disorders. *Transl. Psychiatry* 11, 1–12. <https://doi.org/10.1038/s41398-020-01138-0>
- Whalen, S., Héron, D., Gaillon, T., Moldovan, O., Rossi, M., Devillard, F., Giuliano, F., Soares, G., Mathieu-Dramard, M., Afenjar, A., Charles, P., Mignot, C., Burglen, L., Van Maldergem, L., Piard, J., Aftimos, S., Mancini, G., Dias, P., Philip, N., Goldenberg, A., Le Merrer, M., Rio, M., Josifova, D., Van Hagen, J.M., Lacombe, D., Edery, P., Dupuis-Girod, S., Putoux, A., Sanlaville, D., Fischer, R., Drévilion, L., Briand-Suleau, A., Metay, C., Goossens, M., Amiel, J., Jacqueline, A., Giurgea, I., 2012. Novel comprehensive diagnostic strategy in Pitt-Hopkins syndrome: clinical score and further delineation of the TCF4 mutational spectrum. *Hum. Mutat.* 33, 64–72. <https://doi.org/10.1002/humu.21639>
- Wieben, E.D., Aleff, R.A., Tosakulwong, N., Butz, M.L., Highsmith, W.E., Edwards, A.O., Baratz, K.H., 2012. A common trinucleotide repeat expansion within the transcription factor 4 (TCF4, E2-2) gene predicts Fuchs corneal dystrophy. *PLoS One* 7, e49083. <https://doi.org/10.1371/journal.pone.0049083>
- Zollino, M., Zweier, C., Balkom, I.D.V., Sweetser, D.A., Alaimo, J., Bijlsma, E.K., Cody, J., Elsea, S.H., Giurgea, I., Macchiaiolo, M., Smigiel, R., Thibert, R.L., Benoist, I., Clayton-Smith, J., Winter, C.F.D., Deckers, S., Gandhi, A., Huisman, S., Kempink, D., Kruisinga, F., Lamacchia, V., Marangi, G., Menke, L., Mulder, P., Nordgren, A., Renieri, A., Routledge, S., Saunders, C.J., Stembalska, A., Balkom, H.V., Whalen, S., Hennekam, R.C., 2019. Diagnosis and management in Pitt-Hopkins syndrome: First international consensus statement. *Clin. Genet.* 95, 462–478. <https://doi.org/10.1111/cge.13506>
- Zweier, C., Sticht, H., Bijlsma, E.K., Clayton-Smith, J., Boonen, S.E., Fryer, A., Grealley, M.T., Hoffmann, L., den Hollander, N.S., Jongmans, M., Kant, S.G., King, M.D., Lynch, S.A., McKee, S., Midro, A.T., Park, S.-M., Ricotti, V., Tarantino, E., Wessels, M., Peippo, M., Rauch, A., 2008. Further

delineation of Pitt-Hopkins syndrome: phenotypic and genotypic description of 16 novel patients. *J. Med. Genet.* 45, 738–744. <https://doi.org/10.1136/jmg.2008.060129>

SUPPLEMENTAL DATA

Supplemental table 1: Clinical characteristics of patients with exons 4-8 rearrangements

Supplemental table 2: RT-qPCR primers

Supplemental figure 1: Characterization of chromosome 18 inversion in patient 1

Figure 1

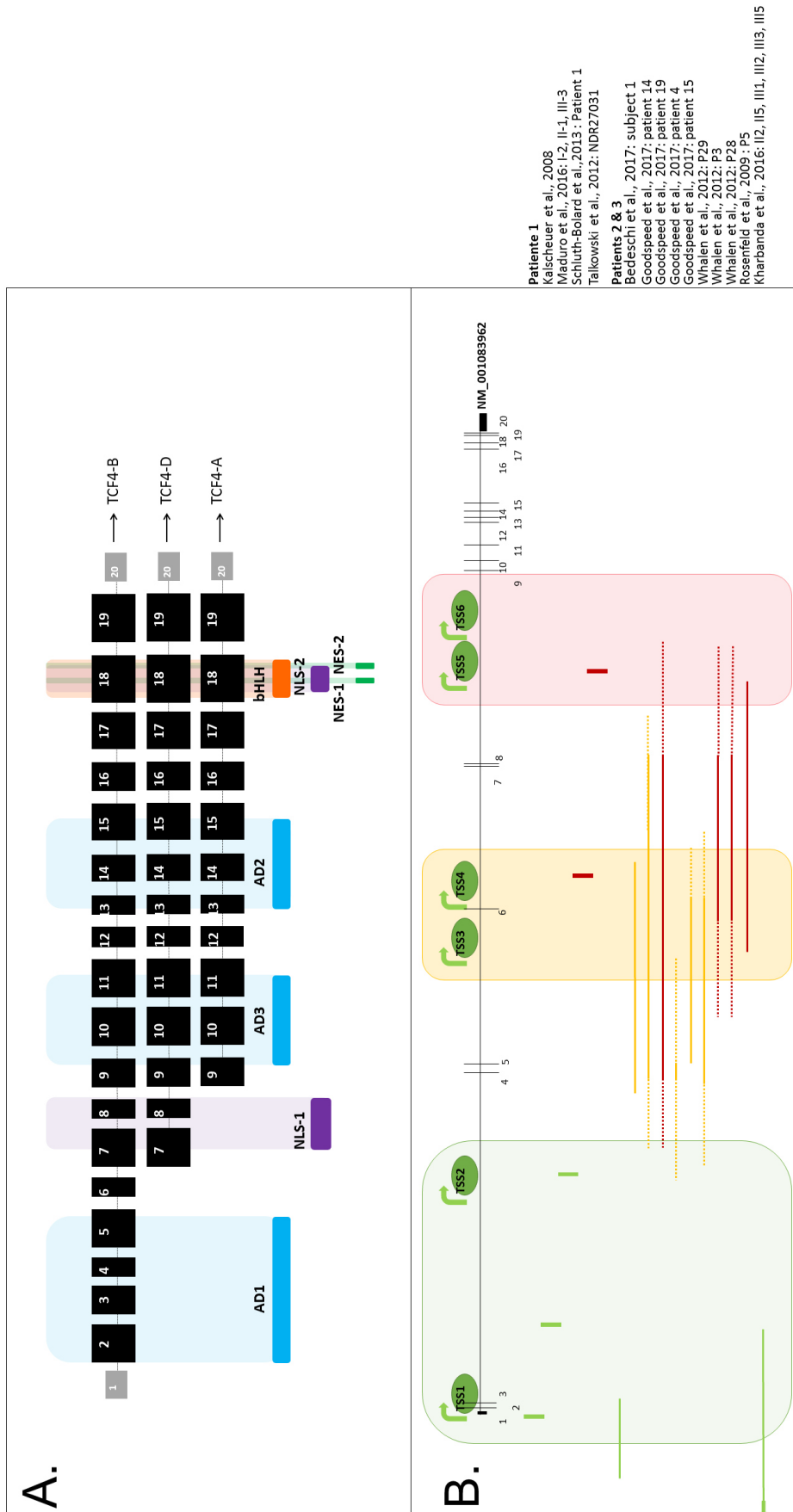
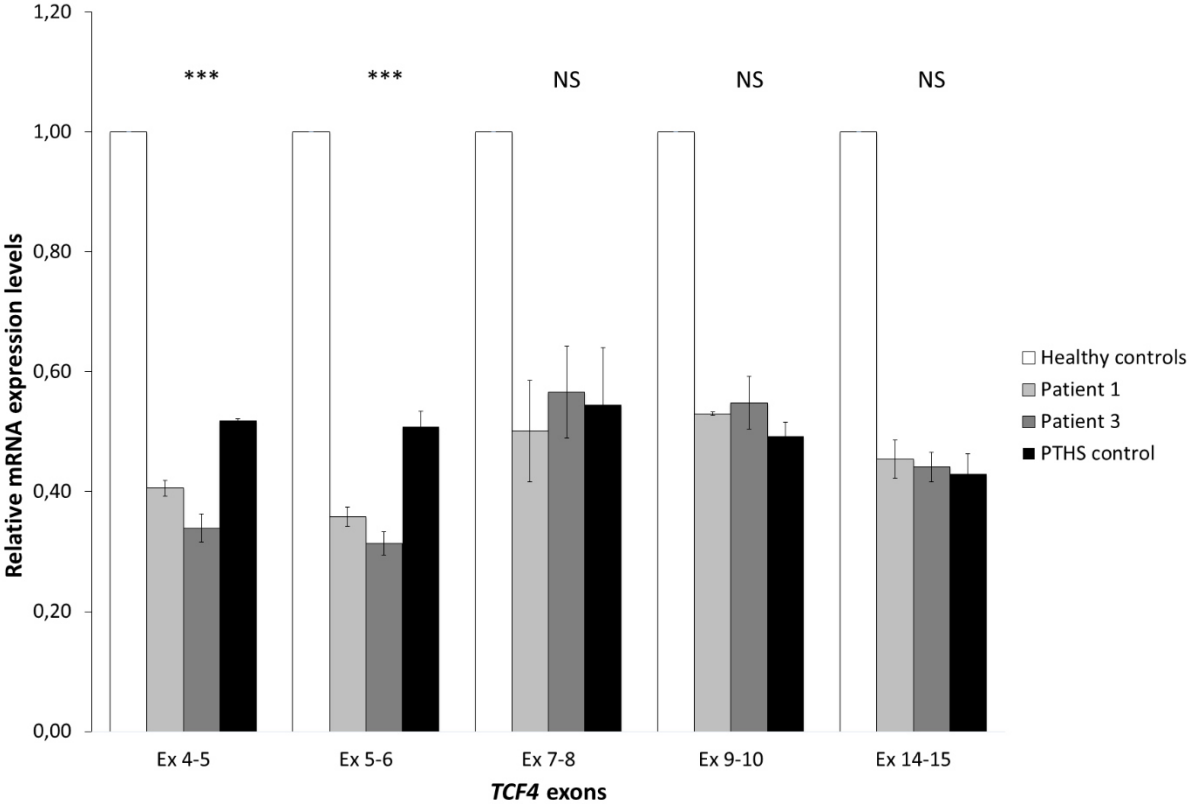
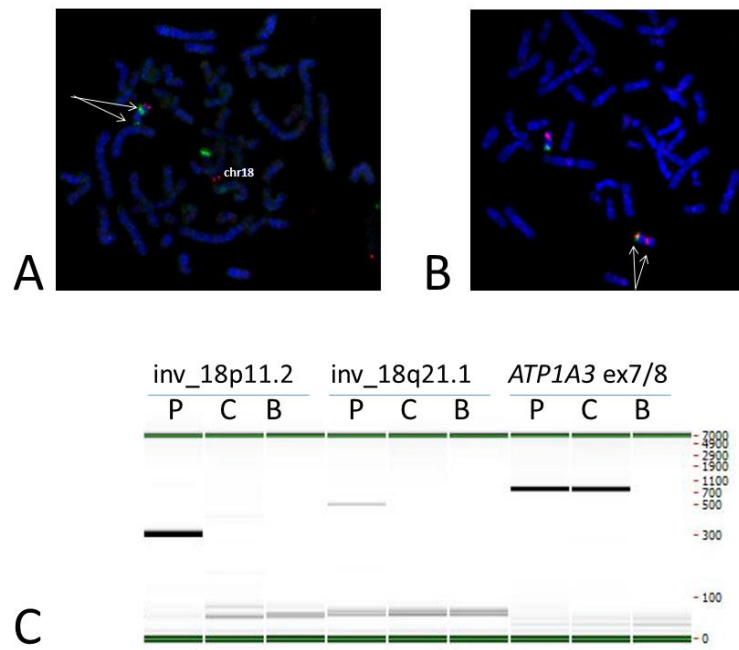


Figure 4



Supplementary data



Supplemental figure 1: Characterization of chromosome 18 inversion in patient 1. A: FISH using BAC probe RP11-372E2 (18q11.32) labelled in green and 18qter subtelomeric probe (CytoCell) labelled in red. The double arrow shows RP11-372E2 probe disruption on the derivative 18. B: FISH using BAC probe RP11-569E16 en 18q21.1labelled in red and 18pter subtelomeric probe (CytoCell) labelled in green. The double arrow shows RP11-569E16 probe disruption on the derivative 18. C: PCR amplification of junction fragment (Labchip view) in patient 1 (P) and a healthy female control (C) without chromosomal rearrangement. The junction fragments are specific for the patient. Inv_18p11.2: junction on the short arm; inv_18q21.1: junction on the long arm. ATP1A3 ex7/8 : control amplification for DNA quality. B: Blank.

Supplemental table 1: Clinical characteristics of patients with exons 4-8 rearrangements

	P29 Whalen et al., 2012	P1 Bedeschi et al., 2017	P4 Goodspeed et al., 2018	P15 Goodspeed et al., 2018	P14 Goodspeed et al., 2018	P19 Goodspeed et al., 2018	P5 Rosenfeld et al., 2009	P3 Whalen et al., 2012	P28 Whalen et al., 2012	P4 Schluth-Bolard et al., 2013
Sex	F	F	?	?	?	?	F	F	M	F
Age (years)	3	55	5	3	3	6	11	4	21	10
Mechanism	deletion 159 kb (exons 4-6)	deletion 98kb (exons 4-6)	deletion 20 kb (exons 4-5)	deletion 138kb (exons 5-6)	deletion 188kb (exons 4-8)	deletion 100 kb (exons 4-8)	deletion 101 kb (exons 6-8)	deletion (exons 6-8)	deletion (exons 6-8)	disruption (intron 6)
Development										
ID	+	mild	+	+	+	+	+	Severe	severe	severe
Age at walking (years)	<3	1.5	5	2.5	2.5	3.5	>3 with assistance	>3	>3	>3
Speech acquisition	absent	delayed	babbling	babbling	babbling	absent	absent	Absent	absent	absent
Writing	-	+	-	-	-	-	-	-	-	-
Behavioral troubles	+	-	-	+	+	-	+	+	+	+
stereotypies*	+	-	+	+	+	+	+	+	+	+
Seizures	-	-	+	+	-	-	ne	-	-	-
Facial dysmorphism										
high forehead	-	+	ns	ns	ns	ns	ns	-	-	-
deep set eyes*	-	+	ns	ns	ns	ns	ns	+	+	+
broad nasal bridge	+	+	ns	ns	ns	ns	ns	+	+	+
tented upper lip	+	-	ns	ns	ns	ns	ns	+	+	+
everted lower-lip	+	+	ns	ns	ns	ns	ns	+	+	+
large mouth	+	-	ns	ns	ns	ns	ns	+	+	+
protrusion of mid/lower face		-	ns	ns	ns	ns	ns	+	+	-
Other findings										
strabismus	+	+	+	na	+	+	-	+	-	-
myopia astigmatism	-	+	+	na	+	+	+	-	-	-
Gastro-intestinal findings										
Constipation	-	-	-	+	+	-	-	-	-	-
Breathing anomalies	-					+	+	-	+	+
Whalen score	12	ns	17	14	16	20	ns	16	17	15

ns: not specified ; na : not applicable ; ne : not evaluated.

Supplementary table 2. RT-qPCR primers

Gene	Primers	
	Reverse	Forward
<i>ACTB</i>	ATTGGCAATGAGCGGTTTC	CGTGGATGCCACAGGACT
<i>TCF4</i> Ex 4/5	ACAGAAGTAGCTCAGGGTCCT	GGTGGAGAGAGATTGTCATGTGA
<i>TCF4</i> Ex 5/6	CTCCACCTTTTGTCAATTCCAG	GGCAACCCTGTAAGTTTGATTCC
<i>TCFA</i> Ex 7/8	CTTCCTCCAAACCAGCAACC	CCCAACATTCTGCATAGCC
<i>TCF4</i> Ex 9/10	ACCAGCACTTTCCCTAGCTC	GCCCAACATTCTGCATAGC
<i>TCF4</i> Ex14/15	ACAGGCCTCATCGTCTCCTA	CACTGCATGGTTCCGGAGAA

PARTIE 2 : Conséquences des remaniements chromosomiques sur l'architecture nucléaire

L'organisation tridimensionnelle du génome n'est pas aléatoire et joue un rôle dans la régulation de l'expression des gènes. Il a été démontré que la perturbation de cette organisation par des remaniements de structure pouvait entraîner des dérégulations et aboutir à des phénotypes variés. Ceci a particulièrement été étudié ces dernières années à l'échelle des TADs [135], dont les disruptions peuvent modifier les interactions entre les gènes et leurs éléments régulateurs. L'étude de l'architecture 3D du génome est donc d'une importance cruciale pour comprendre les mécanismes des pathologies.

Nous avons donc entrepris d'explorer les conséquences des RCAE sur l'architecture 3D du génome, chez huit patients atteints de déficience intellectuelle et/ou malformation congénitale. Ces patients sont porteurs d'un RCAE qui a été préalablement caractérisé sur le plan moléculaire par WGS. Pour ces patients nous étudions les conséquences sur l'organisation nucléaire (3D-FISH), l'expression des gènes (RNA-seq), les marques épigénétiques (ChIP-seq) et les interactions chromatiniennes (Hi-C). La compilation des différentes données permettra d'avoir une image globale de la répercussion des remaniements chromosomiques.

Dans cette partie, seuls les résultats de l'étude en 3D-FISH seront présentés, les autres données étant en cours d'analyse ou d'acquisition. Ce travail d'étude en 3D-FISH a été présenté lors du mini symposium VEGAS de Lausanne (16-17 septembre 2021).

Avec la technique de 3D FISH nous avons voulu déterminer si les remaniements chromosomiques sont responsables d'une relocalisation de la chromatine impliquée dans le noyau. Pour cela nous avons déterminé pour chacun des remaniements la position du locus « point de cassure », qui représente le locus aux coordonnées du point de cassure, pour le chromosome dérivé du patient, pour le chromosome normal du patient (contrôle interne) et pour le chromosome d'un témoin (contrôle externe).

1. Matériel et Méthodes

1.1 Patients

Les patients étudiés dans ce projet ont été recrutés dans le cadre d'un autre projet de l'équipe, le projet ANI [136]. Le projet ANI visait à caractériser des remaniements chromosomiques apparemment équilibrés (RCAE) par séquençage haut débit chez 55 patients atteints de déficience intellectuelle et/ou de malformations congénitales.

Pour cette étude, huit patients issus du projet ANI ont été inclus. Ils présentent un phénotype anormal avec une déficience intellectuelle et/ou des malformations congénitales. Six d'entre eux sont porteurs d'une translocation réciproque et deux d'une inversion péricentrique, toutes survenues *de novo*. Les RCAE ont été caractérisés préalablement par un caryotype. De plus une ACPA a été réalisée afin de vérifier l'absence de CNV pathogène. Les points de cassure des remaniements ont été caractérisés au niveau moléculaire par WGS et des données sur l'expression des gènes au niveau du point de cassure sont également disponibles [136].

Le détail du phénotype de chaque patient inclus dans cette étude ainsi que ses points de cassures sont résumés dans le tableau 1.

L'ensemble des analyses a été réalisé après l'obtention du consentement éclairé des patients ou de leur représentant légal.

1.2 Culture cellulaire

Les lignées lymphoblastoïdes sont obtenues par immortalisation des lymphocytes B par infection par le virus Epstein Barr (EBV). Ces lignées sont cultivées dans un milieu contenant du RPMI et 20% de sérum de veau foetal, auquel on ajoute deux antibiotiques et un antifongique. Ces cellules en suspension sont cultivées dans des flacons de 25 cm² placés dans un incubateur à 37° dont l'atmosphère est contrôlée, avec une humidité de 90% et un taux de CO₂ de 5%. Elles sont divisées quand les rosettes sont bien développées (environ 2 divisions par semaine).

Le comptage des cellules est réalisé grâce à une cellule de Malassez, en déposant 15µL de cellules en suspension diluées au demi dans le bleu de trypan.

Tableau 1 : Caractéristique moléculaire des points de cassure et phénotype des patients

Patient	Age (ans)	Sexe	Caryotype	Type de RCAE	Points de cassure (hg19)	Phénotype
(AD/1601)	18	M	46,XY,t(9;17)(q33;q22)	Translocation	chr9:129,869,403(/401)-129,869,408 chr17:45,807,432-45,807,437(/435)	Acro-ostéolyse, dysplasie spondylo-épiphysaire
(VJ/0601)	11	M	46,XY,t(4;14)(p15;q32.1)	Translocation	chr4:24,332,309-24,332,315 chr14:98,758,653-98,758,657	DI modérée, dysmorphie faciale, TSA, Retard langage, Allergies multiples
(AL/2402)	11	M	46,XY,inv(3)(p24q25.2)	Inversion	chr3:16,652,378(/375)-16,652,421 chr3:16,657,768(/771)-16,657,884(/886) chr3:162,351,577-162,351,610(/608)	DI, apnée du sommeil, asthme, encoprésie, constipation, macrocranie mineure, surpoids, agénésie du corps calleux, brachydactylie
(CS/1902)	29	M	46,XY,t(4;14)(q31.1;q32.2)	Translocation	chr4:148,252,852-148,252,854 chr14:99,097,976-99,097,980	Difficultés relationnelles, troubles psychiatriques, peu de dextérité manuelle, difficultés des apprentissages, retard de langage modéré, dysmorphie faciale
(SL/0105)	13	F	46,XX,inv(2)(p12q22)	Inversion	chr2:74,789,593(/592)- 74,789,727 chr2:148,888,435(/436)-148,888,476	Microcéphalie, DI sévère, épilepsie pharmacorésistante, ataxie, surdité, constipation, difficultés d'alimentation, troubles du comportement (anxiété, agressivité, troubles du sommeil), strabisme
(CL/2001)	19	F	46,XX,t(11;15)(p14;q14)	Translocation	chr11:20,219,439-20,219-438 chr15:37,377,784-37,377,785	Dysmorphie faciale, microcéphalie, DI sévère, absence de langage, épilepsie, communication inter auriculaire et interventriculaire, retard staturo-pondéral, hernie hiatale, malposition laryngée, retard de myélinisation du corps calleux
(MD/0104)	15	F	46,X,t(X;7)(p11.1;q36.1)	Translocation	chr7:149,451,597(/94)-149,451,594 chrX:58,568,914 (/911)/58,568,915	Épilepsie pharmacorésistante (absences atypiques, crises tonico-cloniques, cloniques et atoniques avec chutes), ataxie, tremblements d'action, troubles de déficit de l'attention avec hyperactivité, retard psychomoteur, hépatite auto-immune
(LD/0108)	4	F	46,X,t(X;1)(p11.2;p36.1)	Translocation	chr1:7,130,530(/532)-7,130,531(/533) chrX:49,826,908(/906)-49,826,909(/907)	Retard psychomoteur, DI, convulsions, hypotonie axiale, dysmorphie faciale, strabisme, dyschromie cutanée (lignes de Blaschko)

F : féminin ; M : masculin ; DI : déficience intellectuelle ; TSA : trouble du spectre autistique

1.3 Hybridation *in situ* en fluorescence en 3D = 3D FISH

La technique de 3D-FISH a été réalisée sur les lignées lymphoblastoïdes de huit patients présentant un remaniement de structure de type translocation (six patients) ou inversion (deux patients) avec deux points de cassures. Le marquage fluorescent a été réalisé en utilisant deux sondes de couleurs différentes (FITC et TRITC) de part et d'autre de chaque point de cassures sur une lame du patient et sur une lame d'un témoin. Dans le cas des translocations $t(X;1)$ et $t(X;7)$, une seule sonde a pu être trouvée à proximité du point de cassure du chromosome X. De plus le témoin de ces deux patientes avec des translocations impliquant le chromosome X est un garçon, afin de s'affranchir de la différence de position entre les chromosomes X actif et inactif.

La 3D FISH consiste à réaliser une hybridation *in situ* en fluorescence sur des cellules avec une organisation tridimensionnelle conservée grâce à un pré-traitement afin de déterminer le positionnement d'un locus dans le noyau ou la distance entre deux loci.

Pour cela, la suspension de cellules de lignée lymphoblastoïde va être centrifugée afin d'éliminer le surnageant et de reprendre les cellules dans du milieu de culture (RPMI avec 20% SVF) avec une concentration d'environ 2×10^6 cellules/mL. Puis 400 μ L vont être déposés sur des lames avec traitement de surface à la polylysine, qui par ses propriétés électrostatiques, va permettre la fixation des cellules (Lames polysine, Thermoscientific). Ces lames sont ensuite placées dans l'incubateur à 37°C, avec une humidité de 90% et un taux de CO₂ de 5%, pendant 1h afin de permettre la meilleure adhérence des cellules. Les lames sont alors rincées dans le PBS1X avant d'être fixées dans un bain de 4% paraformaldehyde/PBS1X pendant 10min. Après 3 lavages de 5min dans le 0,05% Triton X-100/PBS1X, les lames sont incubées pendant 20 minutes dans un bain de 0,5% triton X100/PBS1X qui permet de perméabiliser le noyau des cellules. Les lames sont ensuite placées environ 60 min dans un bain de glycérol 20% avant de subir 6 cycles de congélation/décongélation. Pour cela les lames sont plongées dans l'azote liquide, puis après décongélation sont de nouveau placées dans le bain de glycérol. Ces cycles de congélation/décongélation dans l'azote liquide permettent d'augmenter l'accessibilité du noyau aux sondes de FISH sans affecter l'architecture 3D de la chromatine. Les lames sont ensuite déprotéinisées dans un bain de 0,1N HCl pendant 10 minutes. Puis elles sont rincées dans un bain de 2XSSC pendant 5 min avant d'être placées dans un bain de 50%formamide/2XSSC à pH 7 pendant au moins 5 jours. A ce stade les lames peuvent être conservées plusieurs semaines.

L'ADN est ensuite dénaturé en présence d'une sonde (Empire Genomics) à 75°C pendant 3 à 5 minutes en fonction des sondes (Tableau 2). Puis l'hybridation a lieu sur la nuit, dans une chambre noire et humide (Slide Moat BOEKEL Scientific). Les lames sont ensuite lavées dans un bain de 0,5XSS/0,1% SDS à 70°C pendant 5min, puis dans un bain de PBS1X/0,2% Tween pendant 5 min, puis dans un bain de

PBS1X pendant 10min. Les lames sont ensuite montées en utilisant un milieu de montage avec du DAPI.

Tableau 2. Liste des sondes de FISH utilisées

Patient	Chromosome		BAC
AD/1601	17	1	RP11-115L18
		2	RP11-433J24
	9	3	RP11-142P11
		4	RP11-689B7
VJ/0601	14	1	RP11-42D11
		2	Rp11-122F9
	4	3	RP11-1151C21
		4	RP11-886F19
AL/2402	3p	1	RP11-15E222
		2	Rp11-357E4
	3q	3	RP11-482K5
		4	RP11-1111F23
CS/1902	14	1	RP11-42D11
		2	RP11-751B15
	4	3	RP11-160J16
		4	RP11-481I20
SL/0105	2p	1	RP11-203F24
		2	RP11-259N19
	2q	3	RP11-970F18
		4	RP11-947L11
CL/2001	15	1	RP11-715F2
		2	RP11-597G23
	11	3	RP11-698J9
		4	RP11-6K5
MD/0104	X	1	RP11-348G24
	7	3	RP11-767K21
		4	RP11-481B16
LD/0108	X	1	RP11-58H17
	1	3	RP11-54K8
		4	RP11-438L4

1.4 Acquisition des images

Les lames sont ensuite imagées grâce à un microscope confocal Nikon C2 plus avec un objectif à immersion de 60X. Le logiciel relié est le celui de Nikon (Nis-Element). Les images sont imagées avec une résolution de 1024x1024 pixels. Nous réalisons une série de coupe optique (Z-stack) avec un pas 0,3 µm entre deux coupes optiques.

1.5 Analyse des images

Les images sont analysées grâce au logiciel Arivis (version 3.4.0). Ce logiciel permet de visualiser les images en 3D. La création d'un pipeline permet de modéliser le noyau ainsi que les différents spots. Puis il permet de calculer la distance entre chaque spot et la périphérie du noyau (a) et entre chaque spot et le centre du noyau (b). Avec ces données nous pouvons calculer par la suite un ratio de position radiale grâce à la formule $b/(a+b)$. Ces ratios sont calculés pour 50 à 100 noyaux pour chacune des sondes (tableau supplémentaire 1).

L'analyse des images 3D permet d'obtenir pour chaque sonde utilisée le ratio de position radiale. Pour une même sonde, ce ratio est déterminé pour le chromosome normal (dénommé par la suite chrA avec A représentant le numéro du chromosome), le chromosome remanié (dénommé derA) et le chromosome normal d'un témoin (chrAT).

Une étude de comparaison de méthodes d'analyse a été réalisée en utilisant le logiciel Imaris (version 7.7.2). Le ratio de position radiale est déterminé de manière relative par variation du volume d'un objet « noyau ».

1.6 Analyse statistique

Puis un test ANOVA avec un risque $\alpha = 5\%$ est réalisé grâce au logiciel Prism (version 7.0) afin de comparer les moyennes des ratios de position radiale des deux dérivés et des deux chromosomes normaux du patient et du témoin. Ce test est suivi d'un test post hoc, le test des comparaisons multiples de Tuckey, afin de comparer les moyennes deux à deux.

2. Résultats

Nous avons déterminé la position relative des loci impliqués dans le RCAE au sein du noyau, que l'on a appelé le ratio de position radiale. Ce ratio est calculé à partir des distances absolues du locus au centre du noyau et du locus à la périphérie. Pour réaliser cela, nous avons mis au point un pipeline sur le logiciel Arivis.

2.1 Choix de la méthode d'analyse

Afin de pouvoir analyser les images acquises au microscope, il faut modéliser les différents volumes : c'est-à-dire le noyau et les différents spots. Pour cela, nous avons créé des objets qui sont définis par l'intensité de fluorescence du DAPI pour le noyau et des sondes locus spécifiques pour les différents spots. Nous avons donc obtenu un objet « noyau » et quatre objets « sondes ». Nous souhaitons déterminer la position relative des objets « sondes » au sein de l'objet « noyau ».

2.1.1 Comparaison de deux approches

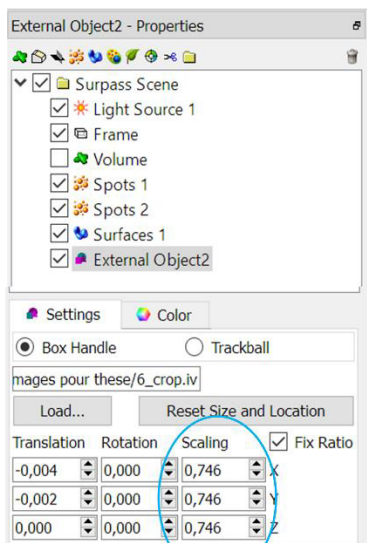
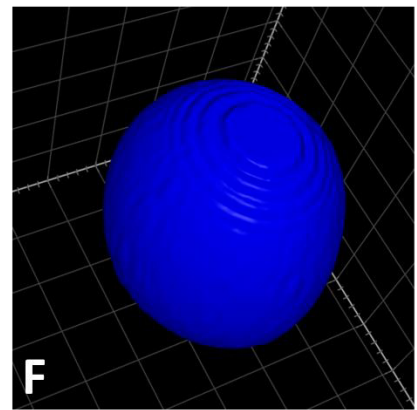
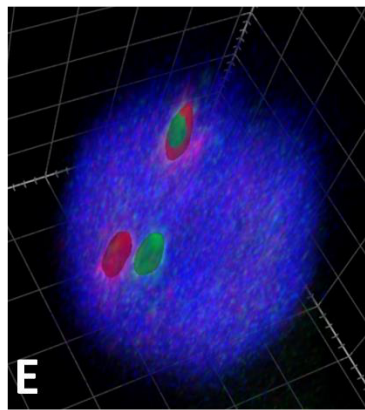
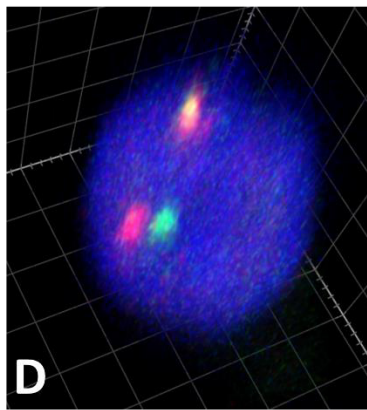
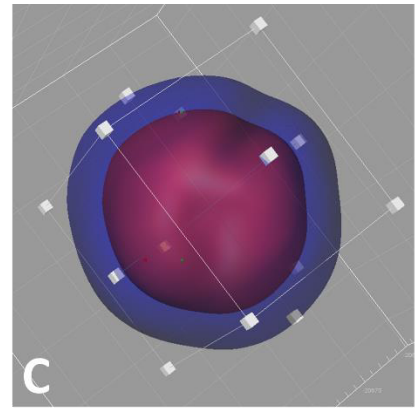
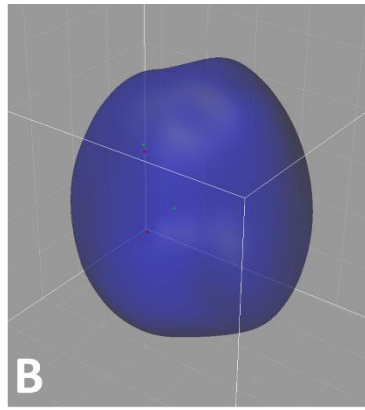
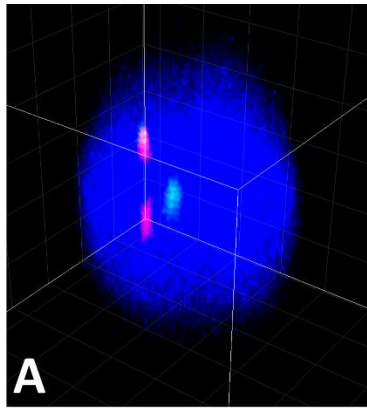
Pour cela, nous pouvons utiliser plusieurs techniques. La première, qui semble la plus juste car les noyaux ne sont pas des sphères parfaites mais de forme elliptique, est la détermination de la position relative directement par variation du volume de l'objet noyau. Nous pouvons également faire la mesure de distances absolues entre les centres des différents objets, et entre les centres et la périphérie.

Nous avons utilisé ces deux techniques sur 30 noyaux afin de les comparer, et de vérifier que la technique par mesure de distances est correcte (Figure 38).

La méthode par position relative est la méthode de référence qui a été initialement mise en place avec le logiciel Imaris (version 7.7.2) [137]. Des objets « noyau » et « sondes » sont créés grâce à la fluorescence du DAPI et des sondes locus spécifique. Nous considérons que la surface de l'objet « noyau » représente la limite de la périphérie du noyau (100%) et le centre de l'objet « noyau » le centre du noyau (0%). Un nouvel objet est inséré, qui est une copie de l'objet « noyau » que l'on nommera « noyau 2 ». La position des objets « sonde » (donc la position des signaux de FISH) est déterminée en diminuant le volume du « noyau2 », jusqu'à ce qu'il chevauche le signal de FISH. Les objets « noyau » et « noyau 2 » conservent bien sûr le même centre. Le pourcentage de régression du « noyau 2 » par rapport à l'objet « noyau » correspond au ratio de position radiale pour la sonde qui était chevauchante.

La méthode par mesure des distances absolues a été développée sur le logiciel Arivis (version 3.4.0). A partir des images acquises au microscope, le noyau et les différents spots au sein de celui-ci sont modélisés grâce à la fluorescence du DAPI et des sondes locus spécifique utilisées. Puis le logiciel détermine le centre de chacun des volumes et mesure la distance entre le centre de chaque objet « sonde » et la périphérie de l'objet « noyau » (distance a) et entre le centre de chaque objet « sonde » et le centre de l'objet « noyau » (distance b). Puis le ratio de position radiale est calculé grâce à la formule $b/(a+b)$.

Ce test a été réalisé avec les sondes RP11-433J24 et RP11-142P11. Pour chacune des sondes, 30 mesures pour le spot du chromosome normal (chr) et le spot du chromosome dérivé (der) ont été réalisées par les deux techniques. Les résultats pour les deux techniques ont été comparés par un test de Student, qui n'a montré aucune différence significative (Tableau 3). La méthode par mesure des distances absolues peut donc être utilisée de la même manière que la méthode par position relative.



Type	Name	Distance, D-Center (µm)	Distance, D-Edge (µm)
	Green	1.951	-1.614
	Green T	2.933	-0.892
	Red	2.198	-1.235
	Red T	2.935	-0.893
	Nucleus	-	-

G

H

Figure 38. Comparaison des logiciels Imaris et Arivis

Comparaison sur un même noyau des logiciels Imaris (A,B, C et G) et Arivis (D, E, F, H). A et D: Visualisation en 3D du noyau sur Imaris et Arivis respectivement. Le bleu correspond à la fluorescence du DAPI et le vert et le rouge aux sondes locus spécifique RP11-433J24 et RP11-142P11 respectivement. B : Modélisation du noyau en bleu et des sondes en vert et rouge. On visualise le centre des objets pour les sondes. C : technique d'insertion d'une copie de l'objet « noyau » (External Object2) afin de pouvoir varier son volume jusqu'à ce qu'il chevauche un signal d'un objet « sonde », ici le signal vert qui correspond au chromosome vert transloqué. G : visualisation du résultat, ici le ratio de position radiale est de 0,746 pour la sonde verte du chromosome dérivé. E : modélisation des sondes vertes et rouges. F : modélisation du noyau. H : visualisation des résultats pour les 4 signaux avec

Green et Red pour les signaux vert et rouge du chromosome normal et Green T et Red T pour les signaux vert et rouge du chromosome dérivé. Le calcul du ratio de position radiale pour le signal vert du chromosome dérivé donne un résultat de 0,767.

Tableau 3 : Résultats des tests de comparaison entre Imaris et Arivis sur 30 noyaux

Spot analysé	Moyenne sur Imaris	Moyenne sur Arivis	p-value (T-test)
Spot Vert chr	0,54913333	0,59931124	0,2364047
Spot Vert der	0,5909	0,63287861	0,28597571
Spot rouge chr	0,60813333	0,64923538	0,35526173
Spot rouge der	0,60043333	0,64736149	0,25737758

2.1.2 Pipeline d'analyse sur Arivis

Nous avons donc par la suite utilisé la technique de mesures des distances absolues sur Arivis. Nous avons développé un pipeline afin que l'analyse soit semi-automatique (Figure 39). En effet il faut toujours vérifier la modélisation car les seuils de détection peuvent varier en fonction de l'intensité de fluorescence et du bruit de fond.

A partir d'une image acquise au microscope confocal contenant plusieurs noyaux, nous pouvons les analyser un par un, en déterminant une région d'intérêt (ROI pour *Region of interest*) en 2D (Figure 39A). Puis, dans cette région d'intérêt, l'objet « noyau » est créé grâce à la fluorescence du DAPI (Figure 39B). Nous indiquons pour cela des paramètres de taille et des seuils de détection. Ensuite les objets « sondes » sont modélisés, sonde par sonde (Figure 2C et 2D). Il faut également indiquer une taille et régler le seuil de détection. La sensibilité de séparation (*split sensitivity*) des signaux peut également être modifiée si deux signaux sont trop proches pour être distingués ou au contraire si un signal est séparé en deux alors qu'il ne devrait pas. Il faut ensuite distinguer les différents loci : celui du chromosome normal (1 signal) et celui du chromosome dérivé (2 signaux, 1 vert et 1 rouge qui se superposent).

Nous créons l'objet « Foci » qui regroupe les deux sondes (Figure 2E), afin de pouvoir ajouter les modules qui permettent de calculer la distance entre le centre de nos objets « sondes » et le centre du « noyau » (Figure 2F) et entre le centre des « sondes » et la périphérie du « noyau » (Figure 2G).

Nous nous servons ensuite de ces données pour calculer les ratios de position radiale (figure 2H et 2I).

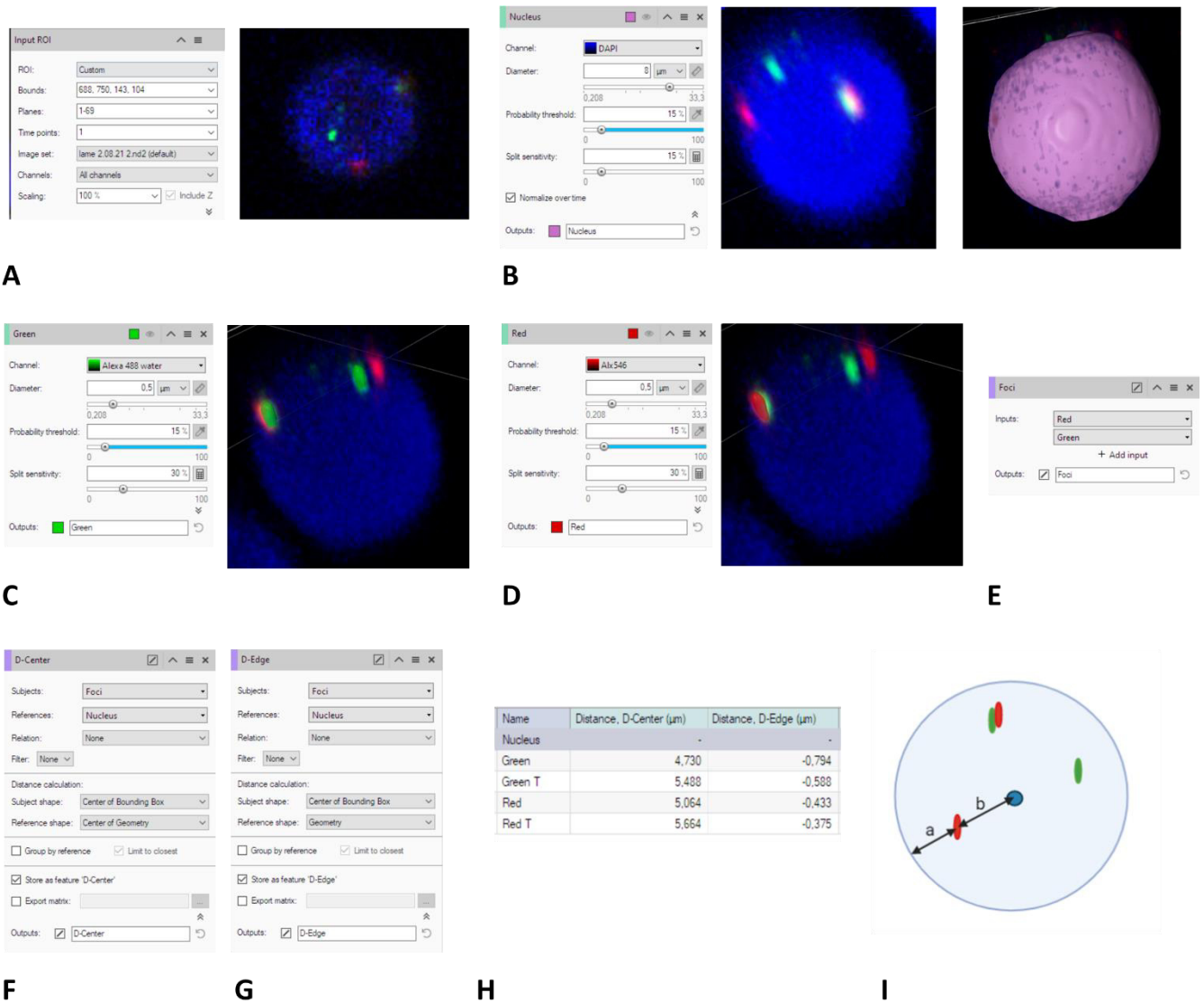


Figure 39. Pipeline d'analyse sur Arivis

A : Le choix d'une région d'intérêt se réalise en 2D. B : Le noyau est visualisé en 3D puis modélisé, ici en violet. C et D : Le logiciel modélise ensuite les sondes locus spécifique. E : l'objet Foci est créé afin de regrouper les deux objets sondes (Green et Red). F et G : des modules de calcul permettent de mesurer les distances entre les Foci et le centre de l'objet « noyau » (F) et entre les Foci et la surface de l'objet « noyau » (G). H : visualisation des résultats pour les quatre signaux (Green et Red pour le chromosome normal et Green T et Red T pour le chromosome dérivé). I : représentation du noyau, des quatre signaux et des distances mesurées. $\text{Ratio} = b/(a+b)$.

2.2 Effets des remaniements sur le positionnement radial

Nous avons regardé pour chaque patient si les loci « point de cassure » des chromosomes dérivés ont une position différente par rapport aux chromosomes normaux.

Pour plus de facilité de compréhension, les chromosomes dérivés des patients seront notés « derA » avec A représentant le chromosome impliqué, les chromosomes normaux des patients seront notés « chrA » et ceux des témoins « chrAT »

2.2.1 Patients avec des translocations ayant un effet sur la position radiale

2.2.1.1 AD/1601

Le patient AD/1601 présente une translocation entre un chromosome 9 et un chromosome 17, au niveau des bras longs.

Afin de comparer les positions des dérivés 9 (der9) et 17 (der17) par rapport aux positions des chromosomes 9 et 17 du patient (chr9 et chr17) et du témoin (chr9T et chr17T) nous avons regroupés les données de ratio de position radiale de différentes sondes. Ainsi les dérivés 9 et 17 sont représentés par l'ensemble des sondes qui sont de part et d'autre du point de cassure, soit les sondes 2 et 3 pour le der9 et 1 et 4 pour le der17. Les chromosomes 9 normaux sont représentés par les sondes 3 et 4 et les chromosomes 17 normaux par les sondes 1 et 2 (Figure 40).

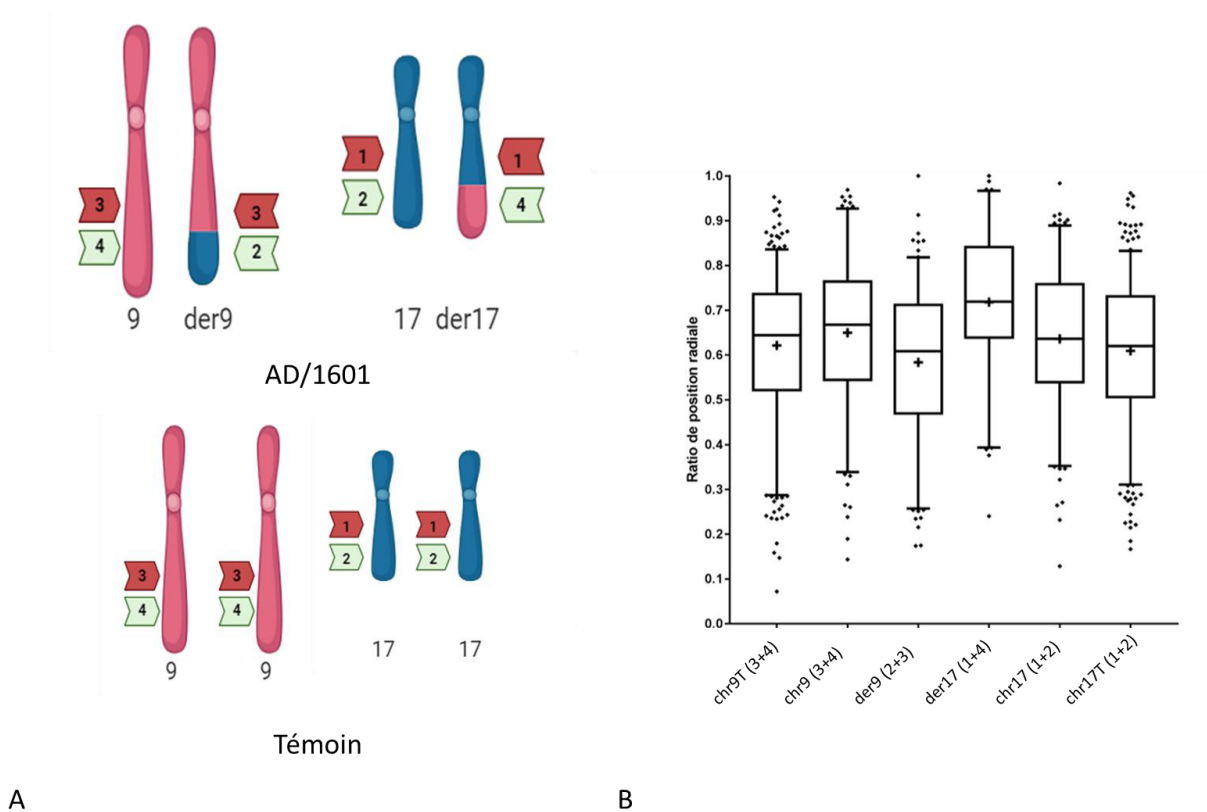


Figure 40. Position radiale nucléaire des chromosomes et dérivés du patient AD/1601

A : représentation du remaniement du patients et placements des sondes utilisées (1,2,3,4). B : représentation des boîtes à moustaches des ratios de position radiale pour les chromosomes 9 et 17 du patient et du témoin et pour les dérivés 9 et 17 du patient. Ne sont comprises que les valeurs de 5 à 95%. Les points représentent les valeurs extrêmes. La croix représente la moyenne.

Le test ANOVA est significatif ($F=9,632$; $p\text{-value} < 0.001$) avec le ratio de position radial du der17 qui est égal à 0,72 et qui est significativement différent du chr17 (ratio à 0,64), du chr17T (ratio à 0,61), du

chr9 (ratio à 0,65) et du chr9T (ratio à 0,65). Le ratio de position radiale du der9, qui est égal à 0,58, n'est significativement différent que du chr17 et du chr9. Les valeurs des p-values ajustées sont présentées dans le tableau 4.

Tableau 4. Valeurs des ratios de position radiale des différents loci du patient AD/1601 et résultats des tests statistiques.

Locus A	Ratio du locus A	Locus B	Ratio du locus B	p-value ajustée
der9	0,5839	chr9	0,6498	0,0029
		chr9T	0,6215	0,1276
		chr17	0,6364	0,0347
		chr17T	0,6094	0,539
der17	0,7181	chr9	0,6498	0,0171
		chr9T	0,6215	<0,0001
		chr17	0,6364	0,0018
		chr17T	0,6094	<0,0001

Le der17 a donc une position significativement plus périphérique par rapport aux chromosomes 9 et 17. Le der9 semble avoir une position plus centrale par rapport aux chromosomes 9 et 17, mais cela n'est significatif que par rapport aux chromosomes normaux du patient.

2.2.1.2 LD/0108

La patiente LD/0108 présente une translocation entre les bras courts d'un chromosome 1 et les bras courts d'un chromosome X.

De la même manière que le patient précédent, nous avons comparés les moyennes de ratios de position radiale des der1, chr1, chr1T, derX, chrX et chrXT. Pour cela, nous avons regroupé les données des sondes 1 et 4 pour le der1 et 3 et 4 pour les chr1 et chr1T. Le derX est représenté par la sonde 3 et les chrX et chrXT par la sonde 1 (Figure 41).

Le test ANOVA est significatif ($F=20,53$; $p\text{-value} < 0,0001$). Le ratio de position radiale du der1, qui est de 0,66 est significativement différent de celui du chr1 (ratio à 0,57). Le derX a un ratio de position radiale égal à 0,73, ce qui est significativement différent du ratio du chr1. Les valeurs des p-values ajustées sont présentées dans le tableau 5.

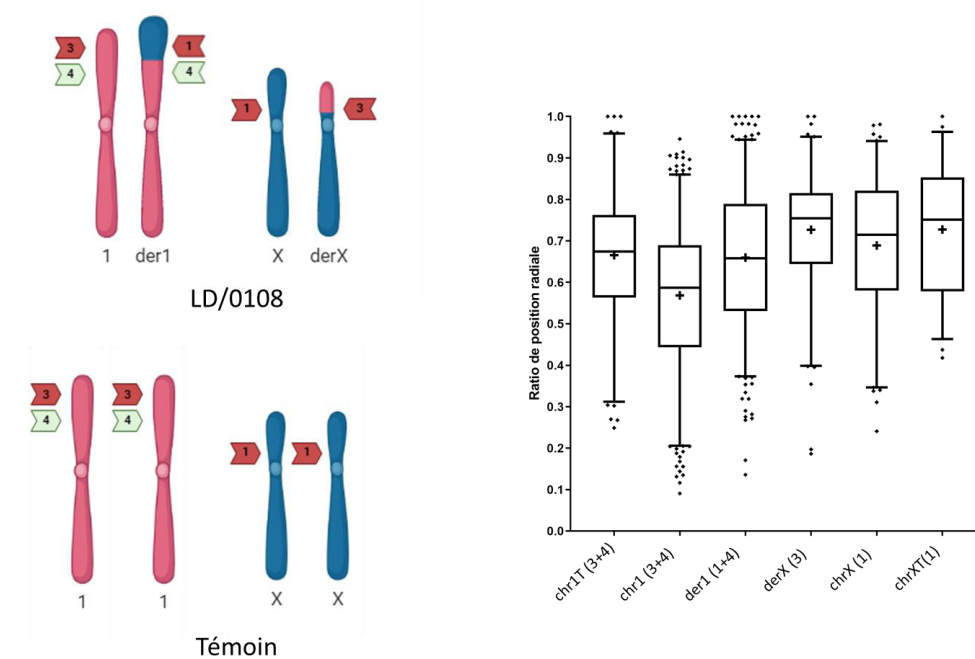


Figure 41. Position radiale nucléaire des chromosomes et dérivés de la patiente LD/0108

A : représentation du remaniement du patients et placement des sondes utilisées (1,2,3,4). B : représentation des boîtes à moustaches des ratios de position radiale pour les chromosomes 1 et X de la patiente et du témoin et pour les dérivés 1 et X de la patiente. Ne sont comprises que les valeurs de 5 à 95%. Les points représentent les valeurs extrêmes. La croix représente la moyenne.

Tableau 5. Valeurs des ratios de position radiale des différents loci du patient LD/0108 et résultats des tests statistiques.

Locus A	Ratio du locus A	Locus B	Ratio du locus B	p-value ajustée
der1	0,6598	chr1	0,5684	<0,0001
		chr1T	0,6654	0,9998
		chrX	0,6890	0,6892
		chrXT	0,7275	0,1030
derX	0,7267	chr1	0,5684	<0,0001
		chr1T	0,6654	0,1208
		chrX	0,6890	0,6386
		chrXT	0,7275	>0,9999

Les der1 et derX sont plus périphériques par rapport au chromosome 1 (chr1). Les fragments du bras court et du bras long du chromosome 1 sont alors dans une position plus périphérique que leur position normale, avec un positionnement plus périphérique pour le derX que pour le der1 (p-value ajustée=0,0110).

2.2.1.3 VJ/0601

Le patient VJ/0601 présente une translocation entre les bras courts d'un chromosome 4 et les bras longs d'un chromosome 14.

De la même manière que les patients précédents, nous avons comparés les moyennes des ratios de position radiale des der4, chr4, chr4T, der14, chr14 et chr14T. Pour cela, nous avons regroupé les données des sondes 1 et 3 pour le der4, 3 et 4 pour le chr4 et chr4T, 2 et 4 pour le der14 et 1 et 2 pour le chr14 et chr14T (Figure 42).

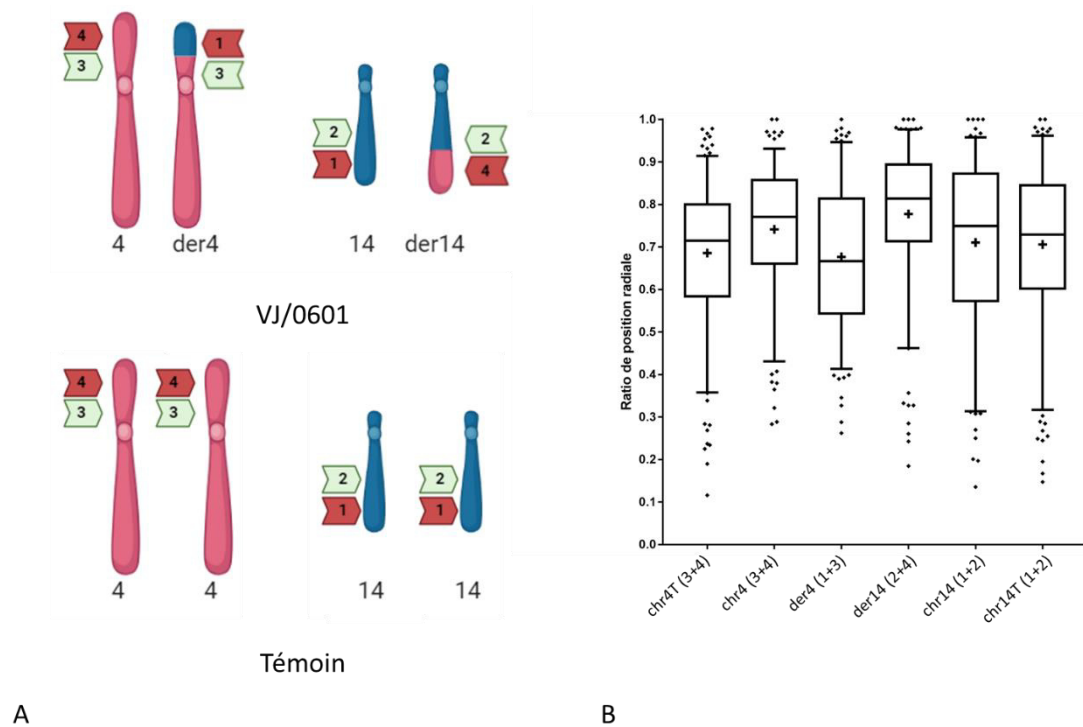


Figure 42. Position radiale nucléaire des chromosomes et dérivés du patient VJ/0601

A : représentation du remaniement du patient et placement des sondes utilisées (1,2,3,4). B : représentation des boîtes à moustaches des ratios de position radiale pour les chromosomes 4 et 14 du patient et du témoin et pour les dérivés 4 et 14 du patient. Ne sont comprises que les valeurs de 5 à 95%. Les points représentent les valeurs extrêmes. La croix représente la moyenne.

Le test ANOVA est significatif ($F=8,647$; $p\text{-value}<0,0001$). Le ratio de position radiale du der14, qui est de 0,78 est significativement différent de ceux du chr4 (ratio à 0,74), du chr4T (ratio à 0,69), du chr14 (ratio à 0,71) et du chr14T (ratio à 0,71).

Le der4 a un ratio de position radiale égal à 0,68, ce qui est significativement différent du ratio du chr4. Cependant il n'y a pas de différence significative avec les ratios de positions radiales des chr4T, chr14 et chr14T. Les valeurs des p-values ajustées sont présentées dans le tableau 6.

Tableau 6. Valeurs des ratios de position radiale des différents loci du patient VJ/0601 et résultats des tests statistiques.

Locus A	Ratio du locus A	Locus B	Ratio du locus B	p-value ajustée
der4	0,6766	chr4	0,7413	0,0087
		chr4T	0,6857	0,9994
		chr14	0,7106	0,4712
		chr14T	0,7059	0,7358
der14	0,7774	chr4	0,7413	<0,0001
		chr4T	0,6857	<0,0001
		chr14	0,7106	0,0047
		chr14T	0,7059	0,0003

Le der14 est donc statistiquement plus périphérique que les chromosomes 4 et 14 normaux. Il n'y a pas de différence significative entre les positions du der4 et des chr4T, chr14 et chr14T. Le dérivé 4 semble plus central que le chr4.

2.2.2 Patients avec des translocations sans effet de relocalisation

2.2.2.1 CL/2001

La patiente CL/2001 présente une translocation entre les bras courts d'un chromosome 11 et les bras longs d'un chromosome 15 (Figure 42).

Les ratios de positions radiales ont été comparés pour les der11, chr11, chr11T, der15, chr15, chr15T. Le test ANOVA est significatif ($F=2,898$; $p\text{-value}=0,0131$) mais aucun des tests de comparaisons multiples n'est significatif (Tableau 7).

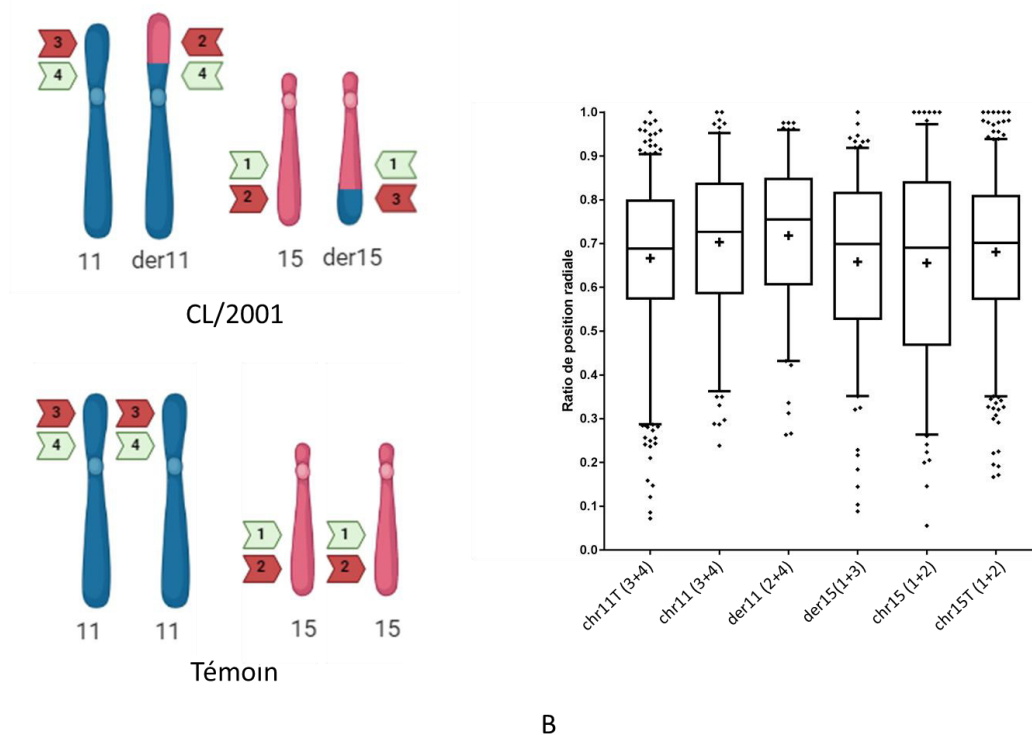


Figure 43. Position radiale nucléaire des chromosomes et dérivés de la patiente CL/2001

A : représentation du remaniement du patient et placement des sondes utilisées (1,2,3,4). B : représentation des boîtes à moustaches des ratios de position radiale pour les chromosomes 11 et 15 de la patiente et du témoin et pour les dérivés 11 et 15 de la patiente. Ne sont comprises que les valeurs de 5 à 95%. Les points représentent les valeurs extrêmes. La croix représente la moyenne.

Tableau 7. Valeurs des ratios de position radiale des différents loci de la patiente CL/2001 et résultats des tests statistiques.

Locus A	Ratio du locus A	Locus B	Ratio du locus B	p-value ajustée
der11	0,7183	chr11	0,7032	0,9848
		chr11T	0,6664	0,0795
		chr15	0,6558	0,0594
		chr15T	0,6810	0,3848
der15	0,6583	chr11	0,7032	0,2241
		chr11T	0,6664	0,9965
		chr15	0,6558	>0,9999
		chr15T	0,6810	0,7488

Les ratios des positions radiales des dérivés 11 et 15 ne sont donc pas différents des ratios de positions radiales des chromosomes 11 et 15 normaux du patient et du témoin. Le remaniement n'a pas entraîné de relocalisation des dérivés par rapport aux chromosomes normaux.

2.2.2.2 MD/0104

La patiente MD/0104 présente une translocation entre les bras longs d'un chromosome 7 et les bras longs d'un chromosome X (Figure 44).

Les ratios de positions radiales ont été comparés pour les der7, chr7, chr7T, derX, chrX, chrXT.

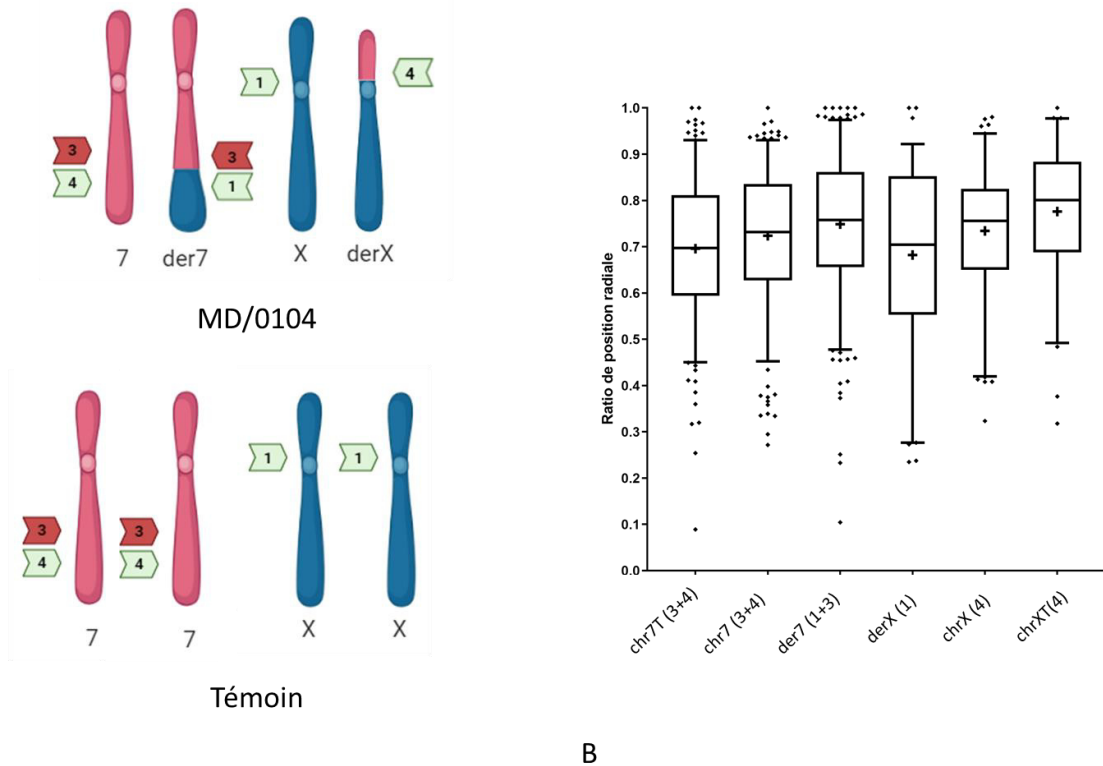


Figure 44. Position radiale nucléaire des chromosomes et dérivés de la patiente MD/0104

A : représentation du remaniement du patient et placement des sondes utilisées (1,2,3,4). B : représentation des boîtes à moustaches des ratios de position radiale pour les chromosomes 7 et X de la patiente et du témoin et pour les dérivés 7 et X de la patiente. Ne sont comprises que les valeurs de 5 à 95%. Les points représentent les valeurs extrêmes. La croix représente la moyenne.

Le test ANOVA est significatif ($F=5,794$; $p\text{-value}<0,0001$) mais les seuls tests de comparaisons multiples significatifs concernant les comparaisons entre les positions des chromosomes dérivés par rapport aux chromosomes normaux sont entre le der7, qui a un ratio de position radiale égal à 0,75 ; et le chr7T (ratio à 0,70) et le derX, qui a un ratio de position radiale égal à 0,68 ; et le chrXT (ratio à 0,78). Il n'y a pas de différence significative entre les positions radiales des dérivés et des chromosomes normaux du patient. Les valeurs des p-values ajustées sont présentées dans le tableau 8.

Le remaniement n'a pas entraîné de relocalisation des dérivés par rapport aux chromosomes normaux du patient.

Tableau 8. Valeurs des ratios de position radiale des différents loci de la patiente MD/0104 et résultats des tests statistiques.

Locus A	Ratio du locus A	Locus B	Ratio du locus B	p-value ajustée
der7	0,7488	chr7	0,7237	0,3988
		chr7T	0,6959	0,0017
		chrX	0,7343	0,9650
		chrXT	0,7761	0,7779
derX	0,6821	chr7	0,7237	0,2786
		chr7T	0,6959	0,9832
		chrX	0,7343	0,2062
		chrXT	0,7761	0,0030

2.2.2.3 CS/1902

Le patient CS/1902 présente une translocation entre les bras longs d'un chromosome 4 et les bras longs d'un chromosome 14 (Figure 45). Le point de cassure sur le chromosome 14 est très proche de celui du patient VJ/0601.

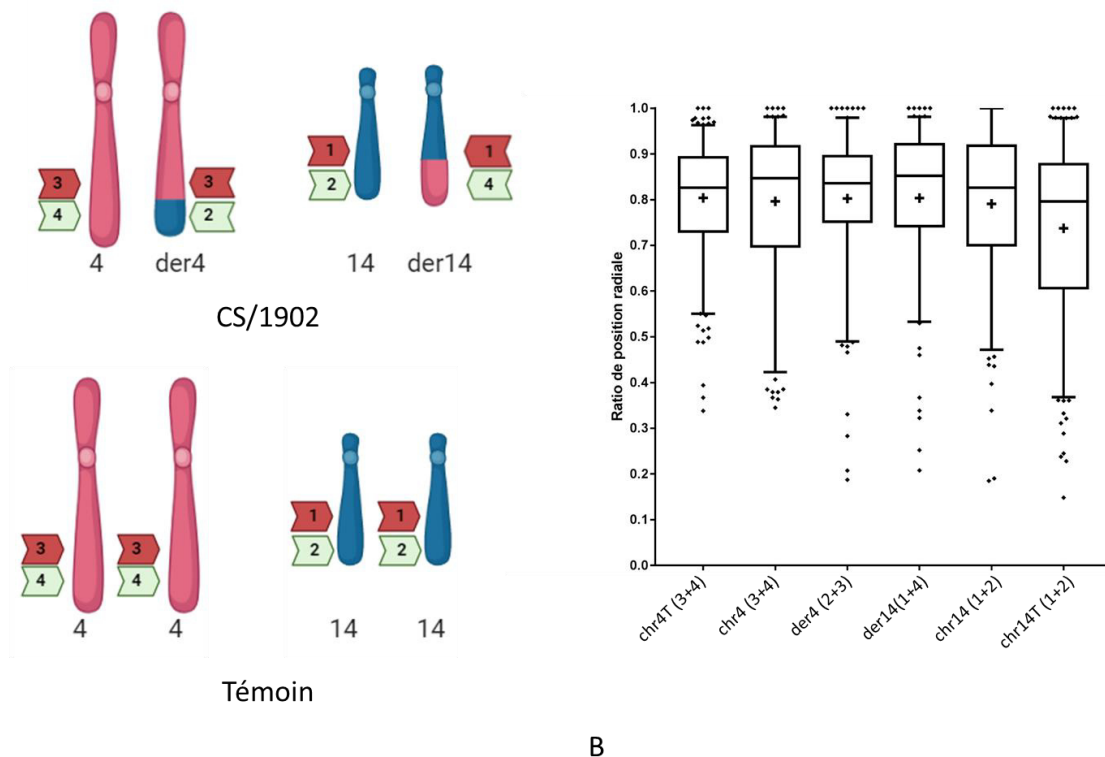


Figure 45. Position radiale nucléaire des chromosomes et dérivés du patient CS/1902

A : représentation du remaniement du patient et placement des sondes utilisées (1,2,3,4). B : représentation des boîtes à moustaches des ratios de position radiale pour les chromosomes 4 et 14 du patient et du témoin et pour les dérivés 4 et 14 du patient. Ne sont comprises que les valeurs de 5 à 95%. Les points représentent les valeurs extrêmes. La croix représente la moyenne.

Les ratios de positions radiales ont été comparés pour les der4, chr4, chr4T, der14, chr14, chr14T. Le test ANOVA est significatif ($F=6,045$; $p\text{-value}<0,0001$). Seule la position du chromosome 14 chez le témoin est statistiquement différente des autres ratios. Les positions radiales des dérivés 4 et 14 ne sont pas différentes des positions des chromosomes 4 et 14 normaux du patient. Les valeurs des p-values ajustées sont présentées dans le tableau 9.

Le remaniement n'a pas entraîné de relocalisation des dérivés.

Tableau 9. Valeurs des ratios de position radiale des différents loci du patient CS/1902 et résultats des tests statistiques.

Locus A	Ratio du locus A	Locus B	Ratio du locus B	p-value ajustée
der4	0,8025	chr4	0,7964	0,9992
		chr4T	0,8040	>0,9999
		chr14	0,7908	0,9837
		chr14T	0,7374	0,0006
der14	0,8037	chr4	0,7964	0,9983
		chr4T	0,8040	0,9999
		chr14	0,7908	0,9765
		chr14T	0,7374	0,0005

2.2.3 Patients avec des inversions

Pour les patients avec des inversions, les quatre sondes sont sur le même chromosome. Pour les deux patients, il s'agit d'une inversion péricentrique, donc il y a deux sondes au niveau du point de cassure sur le bras court et deux sondes au niveau du point de cassure sur le bras long. Nous appellerons derAp (avec A faisant référence au chromosome) le locus au niveau du point de cassure sur le bras court, et derAq le locus au niveau du point de cassure sur le bras long. Pour les positions des chromosomes de référence, nous appellerons chrAp/chrApT et chrAq/chrAqT les positions des loci représentant le point de cassures sur les chromosomes normaux du patient et du témoin, respectivement pour le bras court et le bras long du chromosome impliqué.

2.2.3.1 AL/2402

Le patient AL/2402 présente une inversion péricentrique du chromosome 3. Les moyennes des ratios de positions radiales ont été comparées pour les der3p, chr3p, chr3pT, der3q, chr3q, chr3qT. Pour cela nous avons regroupé les données des sondes 1 et 2 pour le chr3p et chr3pT, des sondes 3 et 4 pour le chr3q et chr3qT, des sondes 1 et 3 pour le der3p et des sondes 2 et 4 pour le der3q (Figure 46).

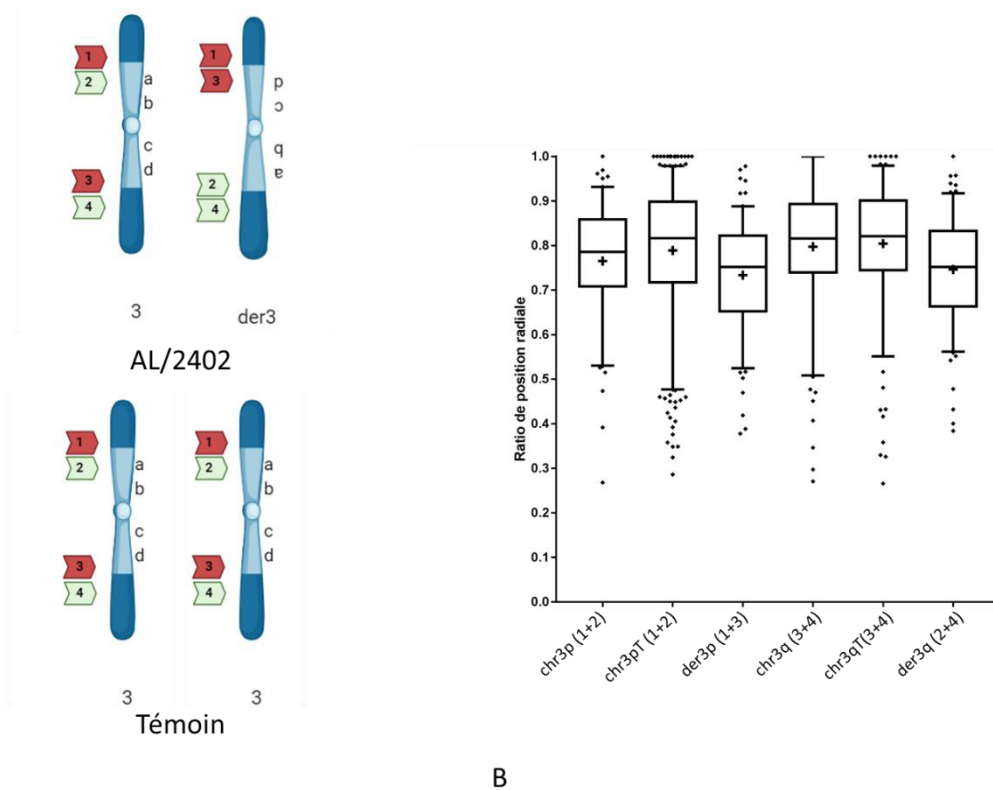


Figure 46. Position radiale nucléaire des chromosomes et dérivés du patient AL/2402

A : représentation du remaniement du patient et placement des sondes utilisées (1,2,3,4). B : représentation des boîtes à moustaches des ratios de position radiale pour les loci « point de cassure » du bras court et du bras long des chromosomes 3 normaux pour le patient et le témoin et du dérivé 3 du patient. Ne sont comprises que les valeurs de 5 à 95%. Les points représentent les valeurs extrêmes. La croix représente la moyenne.

Le test ANOVA est significatif ($F=7.558$, $p\text{-value}<0,0001$). Les tests de comparaisons multiples de Tuckey montrent des différences significatives entre la moyenne du ratio de position radiale du der3p, qui a un ratio égal à 0,73 ; et du ratio du chr3pT (ratio à 0,79), du chr3q (ratio à 0,80), du chr3qT (ratio à 0,80). Le ratio de position radiale du der3q, qui est de 0,75, est significativement différent du ratio de position radiale du chr3pT, du chr3q et du chr3qT. Il n'y a pas de différence significative entre le der3p et le chr3p et entre le der3q et le chr3p. Les valeurs de p-values ajustées des tests de comparaison multiples sont mentionnées dans le tableau 10.

Tableau 10. Valeurs des ratios de position radiale des différents loci du patient AL/2402 et résultats des tests statistiques.

Locus A	Ratio du locus A	Locus B	Ratio du locus B	p-value ajustée
der3p	0,7336	chr3p	0,7652	0,3894
		chr3pT	0,7889	0,0004
		chr3q	0,7973	0,0006
		chr3qT	0,8043	<0,0001
der3q	0,7464	chr3p	0,7652	0,8615
		chr3pT	0,7889	0,0159
		chr3q	0,7973	0,0137
		chr3qT	0,8043	0,0014

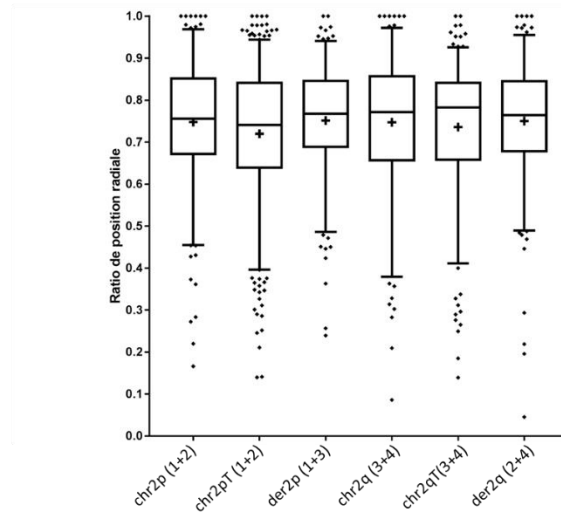
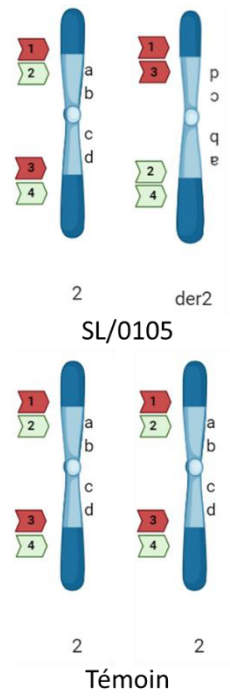
Les positions des loci « point de cassure » sur les bras court et long sont donc significativement plus centrales par rapport aux mêmes positions sur les chromosomes 3 normaux. Cependant les différences ne sont pas significatives par rapport au contrôle interne (le locus 3p du chromosome normal du patient).

2.2.3.2 SL/0105

La patient SL/0105 présente une inversion péricentrique du chromosome 2. Les moyennes des ratios de positions radiales ont été comparées pour les der2p, chr2p, chr2pT, der2q, chr2q, chr2qT. Pour cela nous avons regroupé les données des sondes 1 et 2 pour le chr2p et chr2pT, des sondes 3 et 4 pour le chr2q et chr2qT, des sondes 1 et 3 pour le der2p et des sondes 2 et 4 pour le der3q.

Le test ANOVA n'est pas significatif (F=1,897, p-value=0,0915).

L'inversion du chromosome 2 n'a pas modifié la position radiale des loci au niveau des deux points de cassures.



A

B

Figure 47. Position radiale nucléaire des chromosomes et dérivés de la patiente SL/0105

A : représentation du remaniement du patient et placement des sondes utilisées (1,2,3,4). B : représentation des boîtes à moustaches des ratios de position radiale pour les loci « point de cassure » du bras court et du bras long des chromosomes 2 normaux pour la patiente et le témoin et du dérivé 2 de la patiente. Ne sont comprises que les valeurs de 5 à 95%. Les points représentent les valeurs extrêmes. La croix représente la moyenne.

Tableau 11. Valeurs des ratios de position radiale des différents loci de la patiente SL/0105 et résultats des tests statistiques.

Locus A	Ratio du locus A	Locus B	Ratio du locus B	p-value ajustée
der2p	0,7514	chr2p	0,7479	>0,9999
		chr2pT	0,7199	0,1936
		chr2q	0,7472	0,9998
		chr2qT	0,7360	0,9118
der2q	0,7501	chr2p	0,7479	>0,9999
		chr2pT	0,7199	0.234
		chr2q	0,7472	>0,9999
		chr2qT	0,7360	0,9376

2.2.4 Les positions radiales sont identiques entre différents individus

Nous pouvons également nous servir des tests de comparaisons multiples précédents pour comparer les moyennes des ratios de positions radiales des loci « point de cassure » entre les chromosomes normaux du patient et les chromosomes du témoin (Tableau 12).

Nous pouvons observer une différence significative dans seulement trois cas : la comparaison entre le chr4 du patient VJ/0601 et le chr4T (p-value ajustée=0,0219), entre le chr14 du patient CS/1902 et le chr14T (p-value ajustée=0,0107) et entre le chr1 et le chr1T de la patiente LD/0108 (p-value ajustée <0,0001).

Tableau 12. Ratios et résultats des tests de comparaisons multiples afin de comparer les positions radiales entre le patient et le témoin

Patient	Locus 1	Ratio du	Locus 2	Ratio du	p-value ajustée
AD/1601	chr17	0,6364	chr17T	0,6094	0,4663
	chr9	0,6498	chr9T	0,6215	0,4083
VJ/0601	chr4	0,7413	chr4T	0,6857	0,0219 (*)
	chr14	0,7106	chr14T	0,7059	0,9998
CS/1902	chr4	0,7964	chr4T	0,8040	0,9968
	chr14	0,7908	chr14T	0,7374	0,0107 (*)
CL/2001	chr11	0,7032	chr11T	0,6664	0,3027
	chr15	0,6558	chr15T	0,6810	0,7134
MD/0104	chr7	0,7237	chr7T	0,6959	0,3493
	chrX	0,7343	chrXT	0,7761	0,5105
LD/0108	chr1	0,5684	chr1T	0,6654	<0,0001 (****)
	chrX	0,6890	chrXT	0,7275	0,7900
AL/2402	chr3p	0,7652	chr3pT	0,7889	0,5172
	chr3q	0,7973	chr3qT	0,8043	0,9966
SL/0105	chr2p	0,7479	chr2pT	0,7199	0,2849
	chr2q	0,7472	chr2qT	0,7360	0,9791

En conclusion, nous observons peu de variabilité inter-individus.

Il est également intéressant de regarder les tests de comparaison entre les positions des loci des bras longs et des bras courts dans le cas des inversions (Tableau 13). Il n'y a aucune différence significative entre les ratios de position radiale des loci « points de cassure » des bras courts et des bras longs des chromosomes 2 et 3. Il n'y a pas de différence significative entre les positions du locus au niveau du bras long et du locus au niveau du bras court pour les chromosomes 2 et 3.

Tableau 13. Résultats des tests de comparaisons multiples afin de comparer les positions radiales entre deux positions sur un même chromosome

Patient	Locus 1	Ratio du locus 1	Locus 2	Ratio du locus 2	p-value ajustée
AL/2402	chr3p	0,7652	chr3q	0,7973	0,3421
	chr3pT	0,7652	chr3qT	0,8043	0,7785
SL/0105	chr2p	0,7479	chr2q	0,7472	>0,9999
	chr2pT	0,7199	chr2qT	0,7360	0,8056

2.2.5 Les chromosomes impliqués dans un remaniement ont des positions radiales identiques

Nous avons comparé les moyennes des ratios de positions radiales des loci « points de cassure » des deux chromosomes impliqués dans les translocations. Toutes ses comparaisons sont non significatives sauf les comparaisons entre le chr7T et le chrXT (MD/0104), entre le chr4T et chr14T (CS/1902) et entre le chr1 et le chrX (LD/01/08) qui donnent des différences significatives (Tableau 14). Nous avons vu dans les résultats de CS/1902 que le chr14T était significativement différent des chr14, der4 et der14. Les chromosomes impliqués dans une translocation ont donc majoritairement la même position radiale dans le noyau.

Tableau 14. Ratios et résultats des tests de comparaisons multiples afin de comparer les positions radiales entre deux chromosomes impliqués dans une translocation

Patient	Locus 1	Ratio du locus 1	Locus 2	Ratio du locus 2	p-value ajustée
AD/1601	chr9	0,6498	chr17	0,6364	0,9727
	chr9T	0,6215	chr17T	0,6094	0,9137
VJ/0601	chr4	0,7413	chr14	0,7106	0,5628
	chr4T	0,6857	chr14T	0,7059	0,8392
CS/1902	chr4	0,7964	chr14	0,7908	0,9995
	chr4T	0,8040	chr14T	0,7374	<0,0001 (****)
CL/2001	chr11	0,7032	chr15	0,6558	0,2138
	chr11T	0,6664	chr15T	0,6810	0,8953
MD/0104	chr7	0,7237	chrX	0,7343	0,9916
	chr7T	0,6959	chrXT	0,7761	0,0024 (**)
LD/0108	chr1	0,5694	chrX	0,6890	<0,0001 (****)
	chr1T	0,6654	chrXT	0,7275	0,2936

3. Discussion

Afin d'étudier l'impact des remaniements chromosomiques sur l'architecture tridimensionnelle du noyau, nous avons déterminé les ratios de position radiale au niveau de chacun des loci « point de cassure » pour nos huit patients. Pour chaque locus il existe une grande variabilité intra-individuelle pour la position nucléaire. L'analyse d'un grand nombre de noyaux nous permet de déterminer un ratio de position radiale moyen.

Nous observons très peu de variabilité interindividuelle. En effet sur les 16 loci analysés, seulement trois ratios de position radiale présentent une différence significative entre les patients et les témoins. La disposition des territoires chromosomiques semble donc stable d'un individu à l'autre. Il pourrait être intéressant de réaliser ces études avec plusieurs témoins afin de lisser les petites différences inter-individuelles et permettre de comparer plus finement les ratios de position radiale des patients avec ceux du pool de témoins.

Nous avons également observé que les ratios de position radiale des deux loci impliqués dans un remaniement chromosomique sont identiques dans treize cas sur seize. La proximité de ces loci a dû favoriser une erreur de réparation lors de la survenue de cassures doubles brins à ces niveaux pendant la méiose. Cette proximité a déjà été observée pour les translocations récurrentes impliquées en pathologie cancéreuse [138]. Cela a également été reproduit dans des cellules primaires pré lymphocytes B de souris qui ont été soumises à de fortes doses de radiations ionisantes afin d'induire des cassures doubles brins [139]. Cela a permis de conclure que la probabilité de translocations entre deux chromosomes est proportionnelle à la distance entre les deux territoires chromosomiques de ces chromosomes. Les translocations interviennent entre deux chromosomes qui ont des territoires chromosomiques voisins dans l'architecture nucléaire.

La position radiale est corrélée à la densité génique dans le cas de cellules sphériques [33]. Nous pouvons observer dans nos données une tendance. Les loci des chromosomes denses en gènes ont une position radiale plus centrale que les chromosomes moins denses en gènes (Tableau supplémentaire 2). Cependant nous n'avons pas déterminé la localisation d'un territoire chromosomique mais simplement celle d'un locus précis appartenant à un territoire chromosomique.

Nous avons observé une relocalisation des dérivés par rapport aux chromosomes normaux pour trois patients avec une translocation et un patient avec une inversion. Pour les autres patients, le remaniement n'a pas entraîné de modification de la position radiale des chromosomes au sein du noyau. Cette relocalisation, bien que significative, est généralement légère. Les remaniements chromosomiques n'entraînent pas de forte modification de l'organisation tridimensionnelle du noyau.

Les chromosomes dérivés ne sont pas fortement relocalisés de la périphérie vers le centre du noyau ou inversement. Nous allons analyser plus précisément les corrélations génotype-phénotype pour trois patients.

Pour le patient AD/1601, le dérivé 17 adopte une position plus périphérique et le dérivé 9 une position plus centrale par rapport aux chromosomes 9 et 17 normaux du patient. Le point de cassure sur le chromosome 17 est proche du cluster de gènes *HOXB*, en amont de ce cluster. Ce cluster se retrouve situé sur le dérivé 9 qui a été légèrement relocalisé vers le centre du noyau. Un cas de translocation $t(12;17)(p13.3;q21.3)$ a été rapporté chez un enfant présentant un retard du développement et des malformations squelettiques. Le point de cassure au niveau du chromosome 17 est également proche du cluster *HOXB*, mais en aval du cluster. Chez la souris, les gènes *Hoxb* sont connus pour jouer des fonctions importantes dans le développement du système nerveux central et dans le développement squelettique [140]. Des études complémentaires seront nécessaires pour préciser si cette relocalisation centrale est associée à une modification d'expression des gènes du cluster *HOXB* et si elle peut jouer un rôle dans le phénotype du patient.

Les patientes LD/0108 et MD/0104 sont porteuses d'une translocation X-autosome. Chez les femmes, un des deux chromosomes X est inactivé, en dehors de ses régions pseudo-autosomiques (région PAR). L'inactivation du chromosome X est régulée par le centre d'inactivation de l'X (XIC) qui orchestre la transcription d'un ARN non codant, XIST. Ce dernier s'associe avec le chromosome par lequel il est transcrit et modifie la conformation de la chromatine qui devient de l'hétérochromatine. Dans le cas d'une translocation entre un chromosome X et un autosome, l'inactivation peut être étendue sur la portion de l'autosome reliée au chromosome X, résultant en une monosomie des gènes proches du point de cassure. Mais il a été montré que cette inactivation est moins efficace sur les autosomes et ne concerne alors pas la totalité de l'autosome transloqué sur le chromosome X mais seulement une courte distance depuis le point de cassure relié au chromosome X [141,142]. En théorie, l'inactivation de l'X est aléatoire dans les cellules. Cependant il peut y avoir un biais d'inactivation de l'X, c'est-à-dire que l'un des chromosomes X est préférentiellement inactivé, en faveur de la situation la plus favorable à la croissance cellulaire. Dans le cas d'une translocation X autosome, le biais d'inactivation est en faveur du chromosome X normal, afin de ne pas avoir de monosomie partielle pour l'autosome impliqué, ni de disomie fonctionnelle du fragment du chromosome X présent sur l'autosome.

Il a été également rapporté que le chromosome X inactivé se localise préférentiellement à la périphérie du noyau [34,67,69]. Pour les deux patientes LD/0108 et MD/0104, il n'y a pas de différence significative entre les positions des loci étudiés entre le chromosome X et le dérivé X. Une étude d'inactivation de l'X a été réalisée sur un prélèvement sanguin pour ces deux patientes et il existe un

biais d'inactivation en faveur du chromosome X normal. Pour la patiente LD/0108, l'étude d'inactivation a également été réalisée sur les deux types de peau, pigmentée et hypopigmentée. Sur peau pigmentée il existe un biais d'inactivation en faveur du chromosome X normal. Le profil d'inactivation en peau hypopigmentée n'est pas biaisé. Nous pouvons supposer que le profil d'inactivation n'est pas le même pour les lignées et qu'il n'y a pas de biais, vu que les positions des chromosomes X sont identiques.

Teller et al. ont rapporté, dans des fibroblastes, que le territoire chromosomique du chromosome X inactivé était finalement plus central que celui du chromosome X activé [143]. Ce qui différencie surtout les territoires chromosomiques des chromosomes X activé et inactivé était leur forme : allongé pour le X activé et sphérique pour le X inactivé. Cette forme allongée plate pour le chromosome X activé le rendrait plus proche de la périphérie nucléaire [143].

Pour la patiente LD/0108, les deux loci de part et d'autre du point de cassure du chromosome 1 sont relocalisés dans une position plus périphérique. Le dérivé X, qui comprend le fragment distal du bras court du chromosome 1, a une position encore plus périphérique que le dérivé 1. Nous pouvons alors faire deux hypothèses. D'un côté, la relocalisation du fragment 1p36 de la patiente peut expliquer son phénotype, dû à une diminution de l'expression des gènes contenus dans ce fragment. Cela correspond alors à un phénotype de délétion 1p36 caractérisé par un retard de développement, une déficience intellectuelle, une hypotonie, une épilepsie, une détérioration de l'ouïe, une dysmorphie faciale caractéristique et des malformations cardiaques congénitales [144]. D'un autre côté, si le dérivé X est inactivé, l'hétérochromatinisation de la chromatine du chromosome 1 par l'ARN XIST pourrait entraîner une inactivation du fragment du chromosome 1 transloqué et sa relocalisation en périphérie.

Le patient VJ/0601 présente un dérivé 14 qui a été relocalisé en position plus périphérique que les chromosomes 4 et 14 normaux et un dérivé 4 qui a une position plus centrale que le chromosome 4 normal. Donc les loci en amont du point de cassure sur les chromosomes 4 et 14 ont une position plus périphérique par rapport aux loci sur les chromosomes 4 et 14 normaux respectivement. Le locus en aval du point de cassure sur le chromosome 4 a une position plus centrale que le locus sur le chromosome 4 normal. Ce patient présente un point de cassure sur le chromosome 14 proche de celui du patient CS/1902, pour qui aucun effet de relocalisation n'a été mis en évidence. Pour ces deux patients, un diagnostic par effet de position a été posé. En effet leur point de cassure sépare le gène *BCL11B* de sa séquence régulatrice (*enhancer*) [145]. Ce gène est situé sur le dérivé 4, dont le ratio de position radiale du locus « point de cassure » n'a pas été significativement modifié. Dans cet exemple, c'est la séparation du gène de son enhancer qui joue un rôle majeur dans la survenue du phénotype.

L'étude réalisée en 3D-FISH permet de déterminer les positions des loci impliqués dans les remaniements de patients. Cela a permis de mettre en évidence des relocalisations de certains loci, pouvant participer au phénotype des patients. Cependant cette étude doit être complétée afin d'appréhender au mieux le retentissement global des remaniements chromosomiques.

Des études d'expression vont être réalisées par une technique de RNA-seq, qui est une technique qui utilise le séquençage à haut débit pour séquencer l'ARN. Cela permet de mettre en évidence des différences d'expression entre un patient et un pool de témoin.

Une étude des interactions chromatiniennes sera réalisée par séquençage Hi-C afin d'observer si les remaniements chromosomiques entraînent une modification des interactions chromatiniennes à courte ou à longue distance.

Enfin, une étude des marques épigénétiques sera réalisée par une technique de Chip-seq (*Chromatin Immunoprecipitation Assay*). Cette technique permet d'identifier les interactions entre les protéines et l'ADN à l'échelle du génome. Pour cela les protéines liées à l'ADN sont fixées de manière covalente par le formaldéhyde, il s'agit d'un « cross-linkage ». Puis la chromatine est fragmentée et les complexes ADN-protéines sont isolés par immunoprécipitation en présence de l'anticorps d'intérêt. Les protéines sont ensuite éliminées et les fragments d'ADN obtenus sont séquencés.

La compilation des données de ces différentes techniques permettra d'avoir une image globale de la répercussion des remaniements chromosomiques.

Données supplémentaires de la partie 2

Tableau supplémentaire 1 : Nombre de spots analysés en 3D-FISH

Patient	Locus	Nombre de Spot analysés	Patient	Locus	Nombre de Spot analysés
AD/1601	chr17	174	MD/0104	chr7	258
	chr17T	374		chr7T	222
	der17	178		der7	279
	der9	170		derX	79
	chr9	174		chrX	100
	chr9T	374		chrXT	67
VJ/0601	chr4	176	LD/0108	chr1	300
	chr4T	214		chr1T	102
	der4	170		der1	300
	der14	182		derX	100
	chr14	176		chrX	100
	chr14T	214		chrXT	51
CS/1902	chr4	167	AL/2402	chr3p	128
	chr4T	235		chr3pT	400
	der4	170		der3p	144
	der14	164		der3q	144
	chr14	167		chr3q	160
	chr14T	234		chr3qT	196
CL/2001	chr11	152	SL/0105	chr2p	206
	chr11T	220		chr2pT	240
	der11	120		der2p	190
	der15	183		der2q	190
	chr15	152		chr2q	174
	chr15T	219		chr2qT	236

Tableau supplémentaire 2 : Données de densité génique et de ratios de position radiale

Chromosome	Nombre de gènes	Nombre de bases	Densité génique (nb de gènes/base)	Ratio de position radiale patient (P)	Ratio de position radiale témoin (T)	Moyenne ratio de position radiale P+T
19	1454	63790860	2,2793E-05			
17	1576	81691216	1,9292E-05	0,6364	0,6094	0,6229
11	2093	134978784	1,5506E-05	0,7032	0,6664	0,6848
20	927	63644868	1,4565E-05			
12	1652	133464434	1,2378E-05			
16	1098	89995999	1,2201E-05			
1	2968	245203898	1,2104E-05	0,5684	0,6654	0,6169
6	1963	170740541	1,1497E-05			
15	1122	100114055	1,1207E-05	0,6558	0,681	0,6684
10	1440	135480874	1,0629E-05			
14	1098	105311216	1,0426E-05	0,7507	0,7217	0,7362
3	2032	199411731	1,019E-05	0,7813	0,7966	0,7889
18	766	77753510	9,8516E-06			
9	1299	134505819	9,6576E-06	0,6498	0,6094	0,6357
2	2288	243315028	9,4034E-06	0,7476	0,7280	0,7378
7	1443	158431299	9,108E-06			
5	1643	180967295	9,079E-06			
X	1184	152634166	7,7571E-06	0,7117	0,7518	0,7317
8	1127	145908738	7,724E-06			
4	1297	191610523	6,7689E-06	0,7689	0,7449	0,7569
13	748	114151656	6,5527E-06			
21	303	46976537	6,45E-06			
22	288	49476972	5,8209E-06			
Y	231	50961097	4,5329E-06			

PARTIE 3 : Etude de la transmission de chromoanagenesis familiaux

Les chromoanagenesis sont des remaniements de structure survenant au cours d'un unique évènement catastrophique. Ils sont classés en trois catégories : les chromothripsis, les chromoanasyntesis et les chromoplexies. Seuls les chromothripsis et chromoanasyntesis sont décrits en constitutionnel. Les chromothripsis sont caractérisés par l'éclatement d'un (ou plusieurs) chromosome au cours d'un seul évènement unique suivi par le recollement aléatoire des morceaux par un mécanisme de jonction des extrémités non homologues. Cela peut aboutir à des pertes de fragments chromosomiques. Les chromoanasyntesis sont eux des remaniements dus à un stress réplicatif où le mécanisme d'interruption de la fourche de réplication et de changement de matrice (FoSTeS) entre en jeu. Ce remaniement est caractérisé par la présence de gains multiples et de pertes. Ces chromoanagenesis peuvent être responsables de différents phénotypes en fonction des CNVs générés et des gènes impliqués. La plupart des chromoanagenesis sont *de novo*. Cependant douze familles ont été décrites dans la littérature avec un chromoanagenesis hérité d'un parent. Dans la majorité des cas, ce remaniement n'a pas été transmis de façon stable, et les enfants ont un phénotype dû à ce déséquilibre.

Nous avons étudié six familles dans lesquelles un parent sain porteur d'un remaniement chromosomique complexe l'a transmis à sa descendance selon un mécanisme inhabituel. Ces remaniements ont été caractérisés par caryotype, FISH, ACPA et séquençage haut débit du génome entier (WGS).

Nous avons ainsi caractérisé au niveau moléculaire les chromosomes remaniés du parent et du/des enfants et montré que dans les six familles le parent est porteur d'un remaniement de type chromoanagenesis. Ces remaniements ont été transmis de façon variable à la descendance, en fonction des recombinaisons méiotiques et de la ségrégation des dérivés, pouvant aboutir à la formation d'un nouveau remaniement.

Le fœtus 1 de la famille 1 (Fi-fe1) présente alors une trisomie 21 partielle contenant la région critique du syndrome de Down. L'enfant de la famille 2 (F2-C) présente une délétion de l'exon 6 du gène du facteur IX *F9* responsable d'une hémophilie. L'enfant de la famille 3 (F3-C) présente onze gains dont un gain de 11,1 Mb en 10p15p13 responsable du phénotype de déficience intellectuelle. L'enfant de la famille 4 (F4-C) présente quatre gains de régions du chromosome 1 insérés dans un chromosome 2 dont un gain de 15,9Mb responsable de son phénotype de déficience intellectuelle et de retard du développement. Lors de sa deuxième grossesse, la mère porte un fœtus présentant une fente labiale isolée et un remaniement déséquilibré avec deux délétions de 18,2 et 3,4 Mb du chromosome 1 considérées comme pathologiques. Les deux enfants de la famille 5 (F5-C1 et F5-C2) présentent

plusieurs CNVs en commun, mais le premier enfant (F5-C1) présente une délétion de 4,1 Mb probablement responsable de son phénotype de retard de développement psychomoteur. L'enfant de la famille 6 présente plusieurs CNVs : une délétion de 1.3Mb en 1p31.1, un gain de 3.4Mb en 1q24.2.q24.3 et un gain de 1 Mb en 3q13.11 tous classés en variants de signification indéterminé (VOUS).

Lors de la méiose, l'appariement des chromosomes anormaux entraîne la formation de figures complexes, avec la présence de boucle d'asynapsis. Les crossing-over ayant lieu suite aux jonctions de Holliday peuvent entraîner une recombinaison aléatoire aboutissant à la formation de nouveaux remaniements déséquilibrés. Ce travail montre que les chromoanagenesis peuvent être associés à un phénotype normal, qu'ils peuvent être transmis à l'identique mais que la recombinaison homologue lors de la méiose est imprévisible. Le conseil génétique ainsi que le choix de la technique la plus adaptée pour le diagnostic prénatal sont donc complexes. Ce travail souligne l'importance du WGS pour le diagnostic des chromoanagenesis. Enfin, il questionne plus fondamentalement les mécanismes d'appariement méiotiques des remaniements complexes, encore très mal appréhendés.

Familial transmission of chromoanagenesis leads to unpredictable unbalanced rearrangements through meiotic recombination

Julie Masson^{1,2,3}, Celine Pebrel-Richard⁴, Gwenael le Gudayer⁵, Mathilde Fretigny⁶, Marion Beaumont^{1,7}, Kevin Uguen^{1,8}, Pierre-Antoine Rollat-Farnier^{1,9}, Flavie Diguët¹, Isabelle Perthus⁴, Matthieu Egloff¹⁰, Damien Haye¹¹, Marie-Noëlle Bonnet Dupeyron¹², Audrey Putoux¹¹, Fabienne Raskin-Champion¹³, Marianne Till¹, Nicolas Chatron^{1,2,3}, Bérénice Doray¹⁴, Claire Bardel^{1,3,9}, Christine Vinciguerra^{3,6}, Damien Sanlaville^{1,2,3}, Caroline Schluth-Bolard^{1,2,3}

¹Service de Génétique, Laboratoire de Cytogénétique Constitutionnelle, Centre de Biologie et de Pathologie Est, Hospices Civils de Lyon, Bron, France

²Lyon Neuroscience Research Center, GENDEV Team, CNRS UMR 5292, INSERM U1028, UCBL1, Bron, France

³ Université Claude Bernard Lyon 1, Lyon, France

⁹ Cellule Bioinformatique Plateforme NGS, Hospices Civils de Lyon, Bron, France

⁷ Laboratoire Eylau-Unilabs, Neuilly-sur-Seine, France

⁸ Service de Génétique, CHU Brest, Brest, France

⁴ Service de génétique, CHU Clermont Ferrand, Clermont Ferrand, France

¹¹ Service de génétique, Hospices Civils de Lyon, Bron, France

¹² Service de génétique de Valence, Valence, France

¹⁰ Service de génétique de Poitiers, Poitiers, France

⁶ Service d'hématologie, Laboratoire d'hématologie, Centre de Biologie et de Pathologie Est, Hospices Civils de Lyon, Bron, France

¹³ Service de gynécologie médicale et obstétrique, Groupement Hospitalier Sud, Hospices Civils de Lyon, Pierre-Bénite, France

¹⁴ Service de génétique, CHU de la Réunion - Hôpital Félix Guyon, Saint-Denis, France

Corresponding author: Pr Caroline Schluth-Bolard

Laboratoires de Diagnostic Génétique
Nouvel Hôpital Civil
Hôpitaux Universitaires de Strasbourg
1 place de l'Hôpital
67000 STRASBOURG France
Tel : 33 3 3 69
Email : caroline.schluth-bolard@chru-strasbourg.fr

Manuscript word count:

Abstract

Chromoanagenesis occur during a single catastrophic event and can result in loss or gain of material, responsible for different phenotypes depending on the CNVs generated and the genes involved. We studied six families in which a healthy parent carrying a chromoanagenesis transmitted it in an unbalanced way to the descent. These rearrangements were characterized by karyotyping, fluorescent in situ hybridization, chromosomal micro-array analysis and whole genome sequencing.

We characterized at the molecular level the rearranged chromosomes of the parent and the offspring and hypothesized a meiosis-pairing figure of parental chromosomes. Crossing-over resulted in meiotic recombination between normal and abnormal homologous chromosomes, leading in the formation of new unbalanced rearrangements.

This work underlines that chromoanagenesis can be associated with a normal phenotype and can be transmitted in an unbalanced way because of meiotic recombination leading to unpredictable rearrangements. Despite of the importance of WGS for diagnosis of chromoanagenesis, genetic counseling and technical choice for prenatal diagnosis are therefore complex. Finally this work questions the meiotic pairing mechanisms of complex rearrangements, still very poorly understood.

Abstract word count: 173

Introduction

Complex chromosomal rearrangements (CCRs) were initially described as structural anomalies with at least three breakpoints and implicating at least two chromosomes (Pellestor et al., 2011). The definition was questioned after the characterization of a highly complex rearrangement involving only one chromosome, presenting with 42 breakpoints (Stephens et al., 2011). CCRs were then defined as rearrangements for which genetic counseling is complex. Therefore, these rearrangements are not translocations, inversions or insertions. The advent of genome sequencing technologies enabled the characterization of increasingly complex rearrangements. These rearrangements caused by a unique event and characterized by multiple breakpoints implicating one or a few chromosomes and possible copy number variations (CNV) are named chromoanagenesis (Holland and Cleveland, 2012): “chromo” for chromosome and “agenesis” for rebirth. This term refers to three different types of rearrangements: chromothripsis, chromoanasythesis and chromoplexy depending on the mechanism, with only the two first described in constitutional disorders. Chromothripsis was first described in cancer (Stephens et al., 2011) and then in congenital disorders (Kloosterman et al., 2011) and results from a shattering, usually localized in one or a few chromosomes, into tens or even hundreds of fragments, followed by a reassembling of these fragments by non-homologous end joining (NHEJ) repair. Chromothripsis may be balanced or unbalanced with the presence of deletions predominantly. Chromoanasythesis was described just after (Liu et al., 2011), and results of a defective DNA replication and involves repair mechanisms like fork stalling and template switching (FoSTes) and micro-homology mediated break induced replication (MMBIR). This result in a rearranged chromosome including many copy number variations (CNV): duplication, triplications and deletions. Constitutional chromoanagenesis are rare events and occur de novo in main cases.

Given the complexity of these rearrangements, one should expect that meiotic segregation would be disturbed dramatically. Indeed, first meiotic division requires chromosome pairing and crossing-over to proceed properly. The high number of breakpoints may give rise to complex figures to manage this

pairing step. Nevertheless, stable and unstable transmissions of chromoanagenesis have already been described (Bertelsen et al., 2016; de Pagter et al., 2015; Gu et al., 2013; Gudipati et al., 2019; Kloosterman et al., 2012; Kurtas et al., 2019; Sabatini et al., 2018; Weckselblatt et al., 2015).

Here we report six families with one parent presenting chromoanagenesis with an unstable transmission to the descent, resulting in congenital disorders. Abnormal chromosomes were characterized using fluorescent in-situ hybridization (FISH), chromosomal micro-array analysis (CMA) and whole genome sequencing (WGS) to further explore the complexity of these chromosomal rearrangements and hypothesize the meiotic recombination mechanisms.

Material and method

Patients

Informed consent was obtained from the patients and/or their parents for genetic analyses.

Patients will be named according to their family (F1 for family 1 for example) and then their place (M for mother, F for father, C for child and fe for fetus). In case of several children or fetuses, a number will be added after.

Family 1

The proband was a healthy 42-year-old woman, gravida 3, para 1 (F1-M). Her first pregnancy was marked by high risk combined screening test for trisomy 21 ($>1/10$). An amniocentesis was performed at 14 weeks of gestation (WG), revealing a derivative chromosome 21 on karyotype interpreted as a possible 21q isochromosome (F1-fe1). This pregnancy was terminated. The parental study revealed in the proband an abnormal chromosome 21, presenting a different pattern from the der(21) of the fetus (F1-fe1). The second pregnancy was also marked by a high risk of Down's syndrome ($> 1/10$). A chorionic villus sample performed at 11th WG showed a normal karyotype. Unfortunately, *in utero* fetal death occurred at 12th WG (F1-fe2). An amniocentesis was also performed during her third pregnancy.

The fetal karyotype showed the same altered chromosome 21 as in the proband (F1-fe3). The baby was born at full term without any noticeable problem.

Family 2

The proband was a 15-year-old boy (F2-C). At the age of 8 he was referred for the diagnosis of hemophilia B. MLPA (Multiplex Ligation-dependent Probe Amplification) was performed and revealed an atypical profile with a deletion of exon 6 of the *F9* gene surrounded by a duplication of the other exons of the gene. This patient also presented with developmental delay and language delay. Proband's mother was in good health.

Family 3

The proband was a 4.5-year-old boy (F3-C) followed for psychomotor retardation with axial hypotonia and dysmorphism. He was the only child of healthy non consanguineous parents. At 1 year of age, axial hypotonia with hyperlaxity and sagging of the arches of the feet and a macroglossia were noted. He presented prominent forehead. At 18 months, he crawled and walked with help, and he presented fatigability. His hands have very marked folds, he presented short, curled fingers, low implanted thumbs and hypoplastic nails. Proband's father (F3-F) was in good health, with a normal development.

Family 4

The proband (F4-C) was referred at the age of 2.5 years for intellectual disability, developmental delay, hypotonia and facial dysmorphism. He was the first child of a 27-year-old woman (F4-M) who presented with mild neurodevelopmental features. Karyotype revealed an insertion of a part of the long arm of chromosome 1 into the long arm of chromosome 2. FISH confirmed the insertion and showed a deletion of the associated 2q33.3 region.

During the second pregnancy of F4-M, the fetus (F4-fe) presented an isolated bilateral cleft lip. As part of the family survey, an amniocentesis was performed at 23 WG and the pregnancy was terminated at 25 WG due to genomic imbalances.

Family 5

The proband (F5-C1) was a 3.5-year-old boy from a family of two children. He was born at 37 WG with a height of 48.5 cm (22th centile) and a weight of 2.7 kg (5th centile). He grew proportionately, following a -2SD curve for height and a curve below the third centile for weight. At the age of 3.5 years, OFC was (occipito-frontal circumference) at -4SD. His mother had a height of 55 cm and OFC at -2SD and his father had a height of 170 cm and OFC at -0.5SD. He had a delayed bone maturation. He presented delayed psychomotor development, he started walking at the age of 19 months and he said his first words at 2.5 years. Attention disorders were noted. He had plagiocephaly, moderate hyperlaxity. Clinical examination revealed a facial dysmorphism with bitemporal narrowing, sparse eyebrows, deeply set eyes, depressed nasal bridge, tubular nose, short philtrum, microstomy, dental anomalies and protruding ears with overfolded helix.

The proband's brother (F5-C2), who was a 6-year-old boy, presented attention disorders and had language difficulties. Proband's father (F5-F) was in good health, with a normal development.

Family 6

The proband was a 13-year-old boy (F6-C) followed for delayed language acquisition and attention disorders. He walked at 16 months. He presented overgrowth with a height of 187 cm (+4SD). He had a neuropsychological evaluation at 7 years, showing an IQ of 90. His father (F6-F) was 187 cm tall and had no neurodevelopmental issue.

Karyotype

Karyotype was performed according to standard procedures, from blood, amniotic fluid or chorionic villus sampling with a 400-550 bands resolution after RHG and GTG banding.

DNA extraction

DNA was extracted from EDTA blood sample using the QIAmp DNA Blood Midikit kit from Qiagen (Qiagen, Vento, Netherlands) or from amniotic fluid (F1-fe1, F1-fe3, F4-fe) using the Complete DNA

and RNA Purification MasterPure kit (Epicentre Biotechnologie, Madison, WI, USA) or from chorionic villus culture (F1-fe2) by the QIAamp Blood Mini Kit (Qiagen), according to the manufacturer's instructions.

Chromosomal micro-array Analysis (CMA)

CMA was performed from 1µg genomic DNA using a 60,000 (F3-F and F3-C) or 180,000 (F1-M, F1-fe1, F1-fe2, F1-fe3, F2-M, F2-C, F4-M, F4-C, F4-fe, F5-F, F5-C, F6-C) oligonucleotide microarray (SurePrint G3 Human CGH Microarray Kit, 4 × 180 K or SurePrint G3 Human CGH Microarray 8x60K, Agilent Technologies, Santa Clara, California, USA), according to the manufacturer's instruction. Slides were scanned on Agilent's scanner and images were extracted using Feature Extraction software. Analyzes were carried out according to hg19 reference genome with Genomic Workbench 5.0 software (family 2) using the following parameters: ADM-2: 5.0, z-score threshold: 2.0, window: 0.2 Mb, cutoff: 0.25; or Agilent Cytogenomics 4.0.3.12 software (other families) using the following parameters: ADM-2, threshold: 5.0, window: 0.5 Mb. A copy number variation (CNV) was validated if the log₂ ratio is abnormal for at least three contiguous oligonucleotides. CNVs were interpreted according to the ACMG (Riggs et al., 2020) and Achropuce network recommendations (<https://www.acpa-achropuce.com/wp-content/uploads/2021/01/Guide-interpretation-des-CNV-2020.pdf>).

Whole Genome sequencing

Whole genome library were prepared from 1 to 3 µg genomic DNA using the TruSeq[®] DNA PCR-free kit (Illumina, San Diego, CA, USA) following the manufacturer's protocol in order to obtain fragments of 350 bp as previously described (Jourdy et al., 2021; Schluth-Bolard et al., 2019). Libraries were sequenced according to a 2x101 paired-end protocol with a 15X to 30X mean sequencing depth.

The demultiplexing was performed using the bcl2fastq software (v 2.16). The reads (in FASTQ format) obtained were aligned against the hg19 (F1-M, F1-fe1, F1-fe2, F2-M, F2-C, F3-F, F3-C) or hg38 (F4-M, F5-F, F5-C, F6-F) version of the human genome by the bwa software (MEM algorithm, version 0.7.13), for each of the patients. The alignments were analyzed by two variant callers: ERDS (v1.1) (Zhu et al.,

2012) and BreakDancer (v1.4.5) (Chen et al., 2009). The variants obtained were compared to an internal database listing the breakpoints identified in a multitude of patients in order to remove recurrent variants and annotated using an in-house program (Svagger). The selected rearrangements were visualized and validated on the Broad Institute v2.3 Integrative Genomics Viewer (IGV) software. Alignment of the split reads was carried out using the Blat tool of the Genome Bioinformatics site of the UCSC (University of California Santa Cruz; <http://genome.ucsc.edu/>). In the results part, all the coordinates have been converted to hg38 version of genome in order to be homogeneous.

PCR amplification of junction fragment and Sanger sequencing

Breakpoints not characterized by split reads analysis were confirmed by PCR and Sanger sequencing, as previously described (Schluth-Bolard et al., 2019). In order to determine the location of the breakpoint, the sequence of the junction fragment was aligned of the reference genome (hg19) using the Blat tool of the Genome Bioinformatics site of the UCSC (University of California Santa Cruz; <http://genome.ucsc.edu/>). (See primers in supplemental table 1).

Quantitative PCR (q-PCR)

Quantitative PCR was performed to confirm CNVs associated to the rearrangement using QuantiTect SYBR Green PCR kit (QIAGEN, Courtaboeuf, France) according to manufacturer's recommendations. The qPCR reaction was performed with specific primers (supplemental table 2) for the CNV to be confirmed and primers of a reference gene (*ADORA2B*) using the LightCycler LC480 (Roche), in both patient and control samples.

Fluorescent in situ hybridization (FISH)

Fluorescence *In Situ* Hybridization (FISH) was performed on interphase nuclei or lymphocytes blocked on metaphases of patients and siblings to verify the CNVs found with CMA or WGS and to verify the structure of the rearrangements. See probes in supplemental table 3.

Results

Family 1

F1-M karyotype showed a rearranged chromosome 21 with additional material on its short arm (Figure 1a). To characterize this rearrangement, CMA was first performed. It identified a complex pattern along the long arm of chromosome 21, including 11 gains ranging from 197kb to 610kb and one deletion of 302kb (Figure 1b, supplemental table 4). These CNVs were considered as likely benign. The rearrangement was further refined owing to WGS (12.1X) that revealed 28 breakpoints with 12 one-copy gains ranging from 26kb to 617kb, one two-copy gain of 15kb and one loss of 309 kb (supplemental table 4). All the gained segments were clustered on its short arm as it was further supported by FISH studies (supplemental figure 1). All junctions could be delineated at the base-pair level, however the position of the first breakpoint could not be accurately map on the reference genome since it involved alpha-satellite sequences of the short arm of chromosome 21 (supplemental table 5). In four out of 13 junctions, microhomology (from 3 to 9 bp) was observed and nucleotide insertions were noticed (10, 88, 57, 24, and 140 bp) in five junctions, including templated insertions. All these features were suggestive of a replication-based mechanism.

F1-fe1 karyotype displayed a rearranged chromosome 21 that appeared as an isodicentric of the long arm (Figure 1a). CMA identified two one-copy gains (Figure 1b). This result was confirmed by WGS (15.19X) showing a 322kb gain chr21:26176397-26498731 (fragment F) and a large 20.6Mb involving half of the long arm of chromosome 21 and including the Down syndrome critical region (fragments H to Z). It has to be noticed that the region chr21:41508869-41769255 corresponding to the deleted “R” fragment in F1-M was encompassed in this large gain and was only present in one copy. All junction sequences were identical to the mother (supplemental table 5). FISH studies confirmed the gain of DSCR on the short arm of the derivative 21 (supplemental figure 1).

CMA and WGS (15.33X) in F1-fe2 found the deletion of the “R” fragment (chr21:41508869-41769255) identical to the deletion observed in F1-M in a chromosome 21 with a normal appearance on karyotype

(Figure 1a and b). F1-fe3 karyotype showed the same derivative 21 as F1-M with the same CMA profile (Figure 1a and b).

The derivative chromosome 21 was segmented in twenty-six fragments (A to Z) in order to propose a reconstruction according to WGS, CMA and FISH data in F1-M, F1-fe1, F1-fe2 and F1-fe3 (Figure 1c).

According to the reconstruction of the maternal and fetal derivatives, we could hypothesize what happened during meiosis to explain the transmission of these rearrangements.

The transmission from der21-F1-M to der21-F1-fe1 may have implied a maternal crossing over between the H fragment from the normal chromosome 21 and H fragment on the short arm of the derivative chromosome 21. It resulted in the formation of a der21-F1-fe1 consisting of the long arm and the pericentromeric region of der21-F1-M (F fragment) and a part of the long arm of normal chromosome 21 ranging from H fragment to 21qter (Figure 1c).

In the case of F1-fe2, the maternal crossing-over took place at a locus between the centromere and the Q fragment. F1-fe3 has exactly the same rearrangement as its mother's, the crossing-over took place after the deletion, therefore between the S fragment and 21qter and did not affect the structure of the derivative.

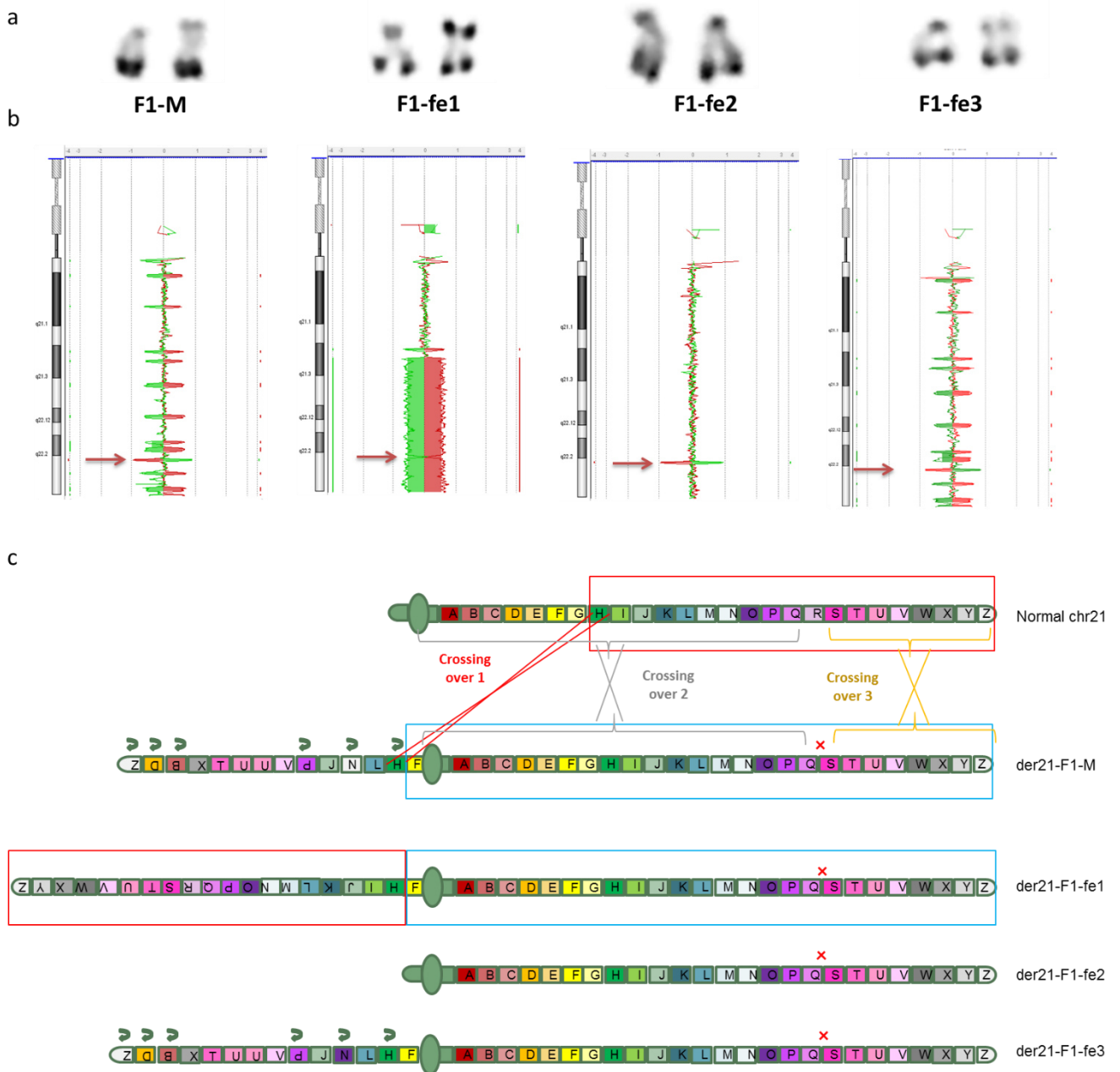


Figure 1. Derivatives 21 identified in family 1 and meiotic recombination hypothesis.

a. Partial karyogram showing normal chromosome 21 and derivative 21 in F1-M, F1-fe1, F1-fe2 and F1-fe3.

b. CMA results. A red arrow indicates the deletion (or return to normal for fetus 1). Red curve: Patient in Cy5 – Control 1 in Cy3. Green curve: Control 2 in Cy5 – Patient in Cy3

c. Representation of the derivative according to FISH and WGS results. The chromosome 21 was segmented into 26 fragments (scheme not at scale) according to the identified breakpoints. The red cross indicates the deletion of R fragment. Green arrows represent inverted fragments. The crossing-over events during maternal meiosis leading to the fetal rearrangement are indicated in red (F1-Fe1), grey (F1-fe2) and yellow (F1-fe3) recombination during maternal meiosis resulting in rearranged F1-fe1, F1-fe2 and F1-fe3's chromosome 21 respectively. Crosses are representation of "crossing over".

Family 2

In order to characterize the rearrangement of the *F9* locus identified in F2-C, CMA was first performed and found a one-copy gain of 514kb in Xq26.3q27.1 (chrX:139492966-140007131). This gain was interrupted by a deletion of the exon 6 of the *F9* gene (NM_000133.4) in chrX:139551007-139557201 (figure 2a and supplemental table 6). It is to note that a de novo deletion of 6.2Mb was also identified on chromosome 12 in chr12:74537664-80738940 that could explained the developmental and language delay of the proband (supplemental figure 2). WGS study (45X) revealed in Xq26.3q27.1 locus three gains of one copy of chrX:139491996-139550775 (fragment B), chrX:139557243-140017... (D) and chrX:140174831-140215... (F) as well as the loss of chrX:139550775-139557243 (fragment C) containing *F9* exon 6 (figure 2b and supplemental table 6).

Two breakpoints could not be characterized at the base-pair level because they were located in two segmental duplications in chrX:140004210-140020416 and in chrX:140213663-140229404 (supplemental table 7).

CMA was also carried out in the mother and found a gain of two copies of the same 514kb region in Xq26.3q27.1 in chrX:139492966-140007131, interrupted by the presence of only two copies of the exon 6 of the *F9* gene, in chrX:139551007-139557201 (figure 2a and supplemental table 6).

The WGS (42.6X) found the same three fragments (B, D, F) present in four copies, intersected by a region in two-copy number (fragment C) corresponding to *F9* exon 6 (figure 2b and supplemental table 6). Junction sequences were identical in son and mother (supplemental table 7). The presence of micro-homology at junctions and multiple gains were suggestive of a replication-based mechanism.

The derivative X chromosome was segmented in seven fragments in order to propose a reconstruction according to WGS data in mother and son. (Figure 2c). F2-M presented the gain of two blocks B-D-F (lacking *F9* exon 6) and a normal copy of *F9*. However, the exact insertion of the blocks B-D-F could not be completely resolved through short-reads data. So there were two alternative constructions for maternal derivative X. We worked with construction 1 (for construction 2 see supplemental figure 3).

During meiosis, a crossing over between D fragments generated derX-F2-C, containing only one "B-D-F" block and lacking the C fragment (exon 6 of the *F9* gene).

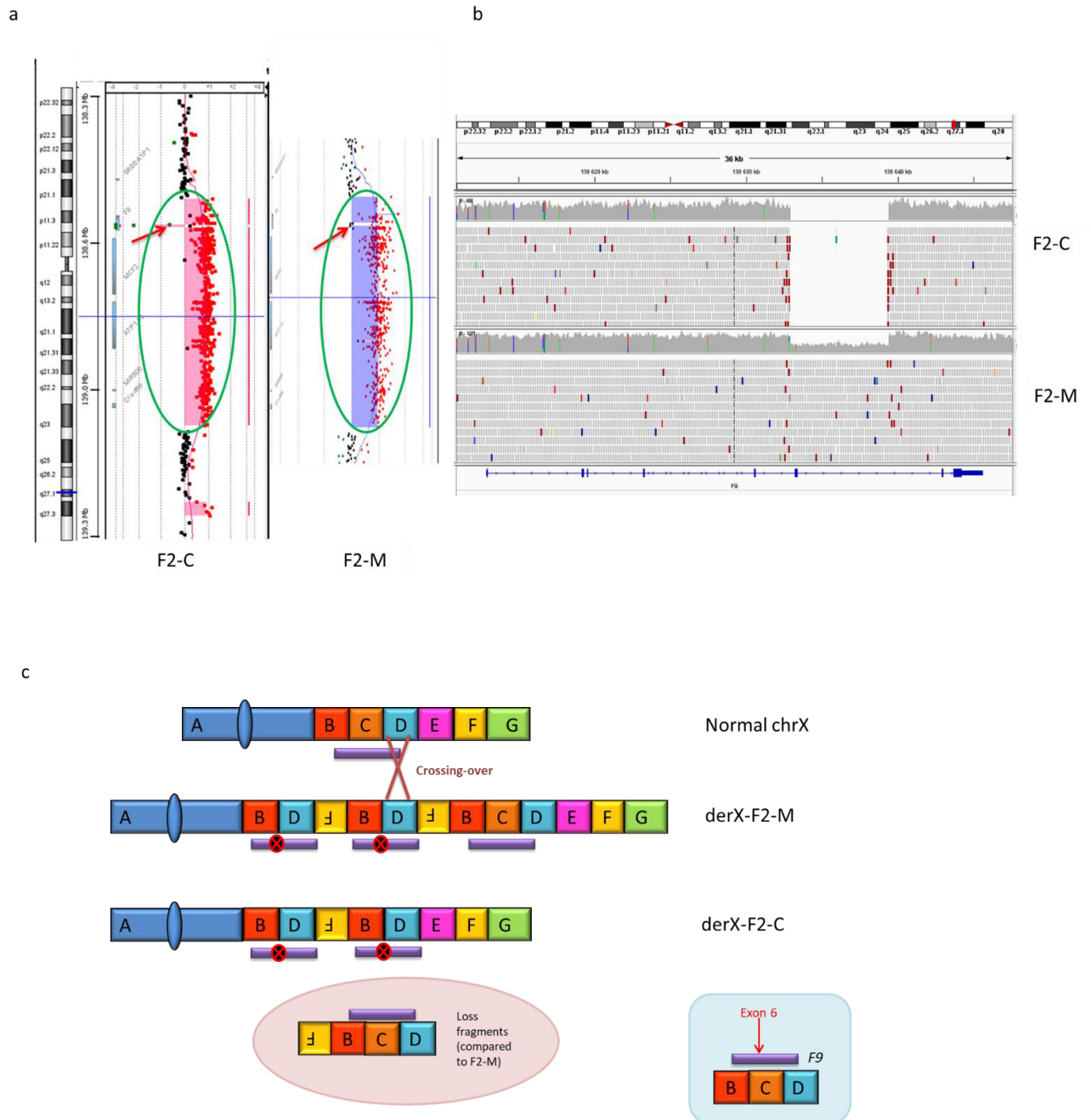


Figure 2. Family 2 rearrangement.

a. CMA profiles performed in F2-C (left) and F2-M (right). The duplication of 514kb (green circle) in Xq26.3q27.1 is identical in F2-C and F2-M. At exon 6 of the *F9* gene (red arrow), F2-C has a deletion and F2-M a normal copy number.

b. Visualization of *F9* on IGV software with WGS data of F2-C and F2-M. The red reads on each side of the deletion present in F2-C and F2-M represent pairs of reads that are too close. Exon 6 of *F9* (NM_000133.4) is deleted in F2-C.

c. Schematic representation of derivate chromosomes X for F2-M and F2-C thanks to WGS data. The purple line represents the *F9* which is included in fragments B, C and D and the red cross in black circle indicates the deletion of exon 6. Red cross represents the crossing-over leading to derX-F2-C. The fragments size is not at scale.

Family 3

The proband showed a rearranged chromosome 10 on its short arm at karyotype (Figure 3a). CMA found three one-copy gains on chromosome 10, of 11.1Mb (chr10:5796821-16902584), 212kb (chr10:18150911-18363263) and 740kb (chr10:19361585-20102260) as well as a one-copy loss of 10.5Mb in chr10:43858258-54371980 (Supplemental table 8 and supplemental figure 4). The 10.5Mb deletion and the 11.1Mb gain were considered as probably pathogenic. This result was confirmed by FISH showing the 11.1Mb deletion and the presence of two signals for probe corresponding to the 11.1Mb gain, one on the long arm and one on the short arm, indicating a complex rearrangement (supplemental figure 4). The WGS (37.8X) revealed 10 gains ranging from 13.1kb to 10.1Mb, one 10.5Mb deletion and a breakpoint in in chr10:18523298-18523302. Six of the gains corresponded to the 11.1Mb and 740kb gains already identified but insufficiently resolved by CMA (Supplemental table 8).

Parental study showed a rearranged chromosome 10 in the father (Figure 3a). CMA only showed a 471kb loss in chr10:18890552-19361644, considered as probably benign (Supplemental table 8 and supplemental figure 4). Interestingly, one of the boundary of this deletion coincided with the boundary of one of the duplications in F3-C. FISH showed that the LSI 10p14 signal was present on the long arm of derivative 10 whereas the RP11-541M12 signal (10q11.22) was present on the short arm, suggesting a double intrachromosomal insertion (Figure 3b).

The WGS study (42.6X) revealed 37 breakpoints including 15 deletions ranging from 1.4kb to 58.4kb besides the larger 566.4kb deletion already found by CMA (Supplemental table 8). Thirty-one of 37 breakpoints were located on the short arm of chromosome 10, mainly clustered into 10Mb in 10p13p12.31 for 27 of them.

All junctions could be defined at the base-pair level in F3-F and F3-C. All junctions of F3-C were identical to F3-F (Supplemental table 9).

Among the 22 junctions in F3-F, 9 showed microhomology (from 2 to 5bp), and one showed a 240pb insertion corresponding to a SVA-F type retrotransposon present at multiple positions in the genome (including chr10:21430989-21431228). This could be suggestive of a large replicative mechanism. However, the other junctions could also have resulted from NHEJ repair.

The chromosome 10 was divided in 38 fragments (A to AL) in order to reconstruct the derivative 10 in F3-F according to the WGS data (Figure 3c). The model was consistent with FISH results showing the displacement of chr10:5784216-15930452 (fragment E) locus from 10p15p13, to the 10q11q21 band and the displacement of chr10:43814192-54397546 (fragment AH) locus from 10q11q21 to 10p15p13.

During paternal meiosis, crossing-over between AF fragments of the normal and the derivative chromosomes 10 yielded the rearranged chromosome 10 observed in the son and composed of the short arm of the normal chromosome 10 and the long arm of der10-F3-F. It resulted in the gain of 10 fragments (E, K, M, P, R, U, V, Z, AB, AD) both present on the short arm and on the long arm of the der10-F3-C and the loss of fragments (AH) that were present on the paternal derivative short arm.

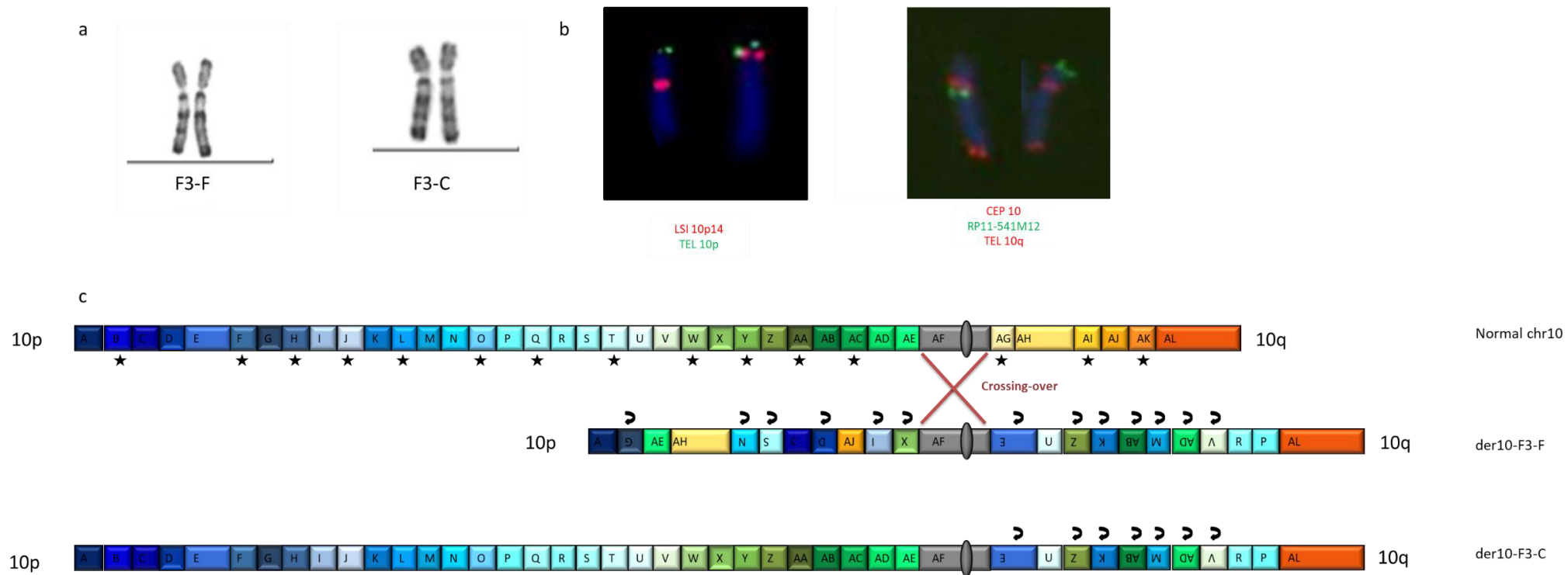


Figure 3. Family 3 rearrangement

a. Partial karyogram (R-band) showing chromosomes 10 in father (F3-F) and son (F3-C).

b. FISH images for F3-F. left: The 10p15p13 region was explored using the LSI 10p14 probe (red) and 10pter (green) as control probe; right :the 10q11q21 region was explored using the RP11-541M12 probe (green) with CEP10 (red) and 10qter (red) as control probes. . The LSI 10p14 is located on the long arm of derivative 10 whereas the RP11-541M12 signal is located on the short arm of the derivative.

c. Schematic representation of derivative chromosomes for F3-F and F3-C thanks to WGS data. The arrows indicate the inverted fragments, and the stars indicated the deleted fragments in der10-F3-F. Red cross represented the crossing-over leading to der10-F3-C. The fragments size is not at scale

Family 4

CMA was performed in F4-C due to intellectual disability. It found four one-copy gains on the long arm of chromosome 1 ranging from 230kb to 15.9Mb: chr1:169542256-169772213 (1q24.2); chr1:176409669-178390838 (1q25.2); chr1:178551617-194401653 (1q25.2q31.3); and chr1:195658048-199165384 (1q31.3q32.1) (Figure 4b and supplemental table 10). A one-copy loss was also found on chromosome 2 long arm, spanning 610kb (chr2:206906076-207515805; 2q33.3). Subsequent karyotype showed a chromosome 2 with rearranged long arms (Figure 4a). FISH confirmed the 2q33.3 deletion (only one signal for RP11-10O12 probe). It also showed one additional signal for probes RP11-126C20 (1q24.2.), RP3-423N22 (1q25.2), RP11-239J11 (1q31.2) and RP11-75C23 (1q31.3) on the long arm of a chromosome 2 (supplemental figure 5). This result was suggestive of an imbalanced insertion of the 1q24.2q31.3 region in chromosome 2 long arm.

For F4-fe, karyotype showed an interstitial deletion on the long arm of a chromosome 1 (Figure 4a). CMA found two deletions of 18.2Mb (chr1:176209674-194401653) and 3.4Mb (chr1:195718529-199149645) on the long arm of chromosome 1 considered as pathogenic (Figure 4b and supplemental table 10). FISH confirmed the presence of a derivative chromosome 1 carrying two deletions on the long arm at 1q25.2q31.3 and in 1q31.3q32.1 (supplemental figure 6).

F4-M karyotype showed an apparently balanced insertion of a part of the long arm of chromosome 1 into the long arm of chromosome 2 (Figure 4a). FISH supported this finding (supplemental figure 7). CMA revealed three deletions ranging from 47kb to 610kb: chr1:176328453-176374994 (1q25.2); chr1:178402889-178548610 (1q25.2); and chr2:206906076-207515805 (2q33.3), and a 168kb one-copy gain in chr2:229851943-230020147 (2q36.3) (Figure 4b and supplemental table 10). None of these CNVs was considered as pathogenic.

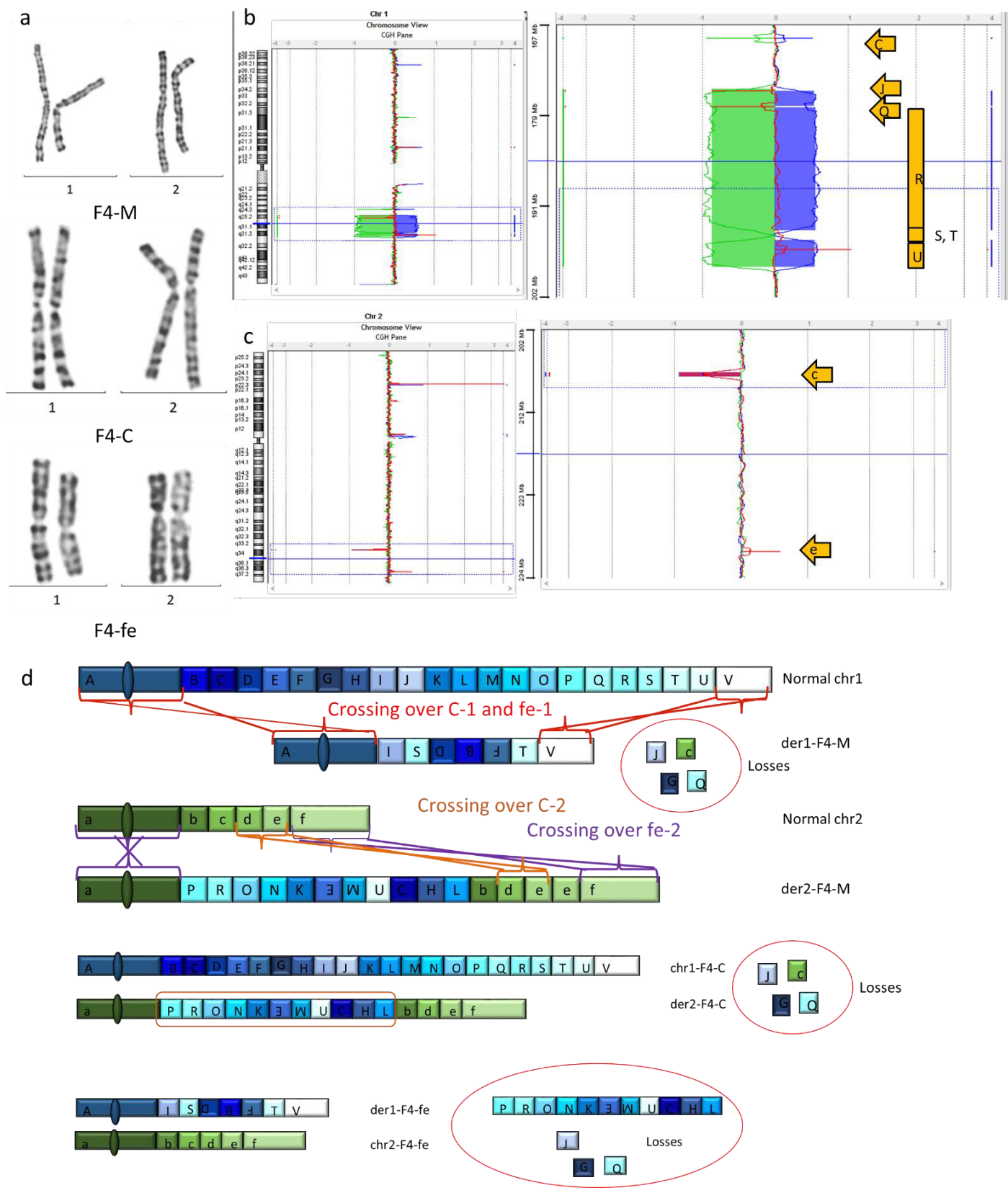


Figure 4. Family 4 rearrangement

a. Partial karyogram (G-band) Showing chromosomes 1 and 2 in F4-M, F4-C and F4-fe

b and c. Superposition of CMA results for F4-M (red), F4-C (bleu) and F4-fe (green) for chromosome 1 (b) and chromosome 2 (c). Letters in yellow arrows and rectangles represent fragments determined thanks to WGS. For

F4-M, we could see the three chromosome 1 losses (C, J and Q) and one loss (c) and one gain (e) for chromosome 2. For F4-C, we could see four chromosome 1 gains (C; between J and Q, J and Q excluded; R and U) and one chromosome 2 deletion (c). For F4-fe, we could see three chromosome 1 deletions (C; from J to R, and U) and no CNV on chromosome 2.

d. Schematic representation of der1-F4-M, der2-F4-M, der1-F4-C, der2-F4-C, der1-F4-fe, der2-F4-fe thanks to WGS data. Fragment size are not at scale. Crosses represented crossing-over (C for F4-C, fe for F4fe, 1 for chromosome 1, 2 for chromosome 2). Red circle represented losses fragments and orange rectangle represented duplicated fragments.

In order to better characterize these rearrangements, a WGS was performed (14.80X) in F4-M. The WGS revealed a complex rearrangement with 23 breakpoints including four deletions ranging from 6.1kb to 644kb, three on chromosome 1 and one on chromosome 2. It also found the tandem duplication in 2q36.3 but was not able to specify if this duplication belong to the derivative chromosome 2 (Supplemental table 10).

All junctions were characterized at the base-pair level (Supplemental table 11). Three of them showed microhomology (3 to 5 bp) and four showed templated insertions suggesting a replicative mechanism. Using WGS, F4-M rearrangement was reconstructed with chromosome 1 divided in 22 fragments (A to V) and chromosome 2 divided in 6 fragments (a to f) (Figure 4d and supplemental figure 8).

The rearrangements observed in the descents are secondary to chromosome missegregation during maternal meiosis. Crossing-over probably took place at each derivative extremities and did not affect their structure, except for the duplication of the “e” fragment on chromosome 2 (Figure 4d). The rearrangement was then transmitted in an unbalanced and in a mirror manner, the derivative 2 for F4-C, resulting in multiple gains of 1q24.2q31.3 region of approximately 22 Mb, and the derivative 1 for F4-f, resulting in the loss of 14 fragments of chromosome 1 of approximately 22 Mb.

Family 5

Karyotype showed rearranged chromosome 1 with abnormal extremity of short arm for F5-C1, F5-C2 and F5-F (Figure 5a).

CMA performed in F5-C1 revealed three deletions from 88kb to 4.3Mb and six gains from 35 to 759kb distributed all along the short arm of chromosome 1 (Figure 5b and supplemental table 12). WGS data (19.92X) showed a complex rearrangement composed of 22 breakpoints and including three losses from 97kb to 4.1Mb and eight gains from 15 to 72kb (supplemental table 12). All junctions were resolved at the base-pair level (supplemental table 13).

F5-C2 presented according to CMA two deletions of 88kb and 737kb and five gains on the short arm of chromosome 1 (Figure 5b and supplemental table 12). All the CNVs boundaries were identical to those identified in F5-C1.

F5-F was first explored by CMA. It showed multiple CNVs on the short arm of chromosome 1 between 1p36.32 and 1p21.2: four deletions ranging from 88kb to 737kb and eight gains ranging from 46kb to 716kb (Figure 5b and supplemental table 12). Two of the deletions were in common with F5-C1 and F5-C2; six gains were shared with F5-C1 and five gains were shared with F5-C2.

To further elucidate the rearrangement, WGS (14.99X) was performed in F5-F. It identified a complex rearrangement of the short arms of chromosome 1 composed of 40 breakpoints including six losses ranging from 44kb to 752kb and 12 gains ranging from 7kb to 724kb (supplemental table 12). All junctions could be resolved at the base-pair level (supplemental table 13). However for three breakpoints, the accurate position could not be determined as they were located in segmental duplication. For these breakpoints, two alternative positions were possible. Only five junctions showed microhomology of 2 to 3bp. The molecular signatures were thus rather suggestive of a NHEJ repair. The junctions identified in F5-C1 were identical to those in common with F5-F.

Taken into account WGS data, there were several ways to reconstruct the derivative 1 in F5-F. To better understand the order of the different fragments on the derivative 1, FISH analyses were performed using RP11-168D14 (1p35.3p35.2), RP11-479E21 (1p32.1) probes and 1pter subtelomeric probes. It

showed the sequence 1pter- RP11-168D14- RP11-479E21 signal, enabling to better define the rearrangement (supplemental figure 9). Integrating all the chromosomal and genomic data allowed proposing the most probable reconstruction of the rearrangements in F5-F and his two sons with the chromosome 1 divided in 41 fragments (A to AO) (Figure 5c).

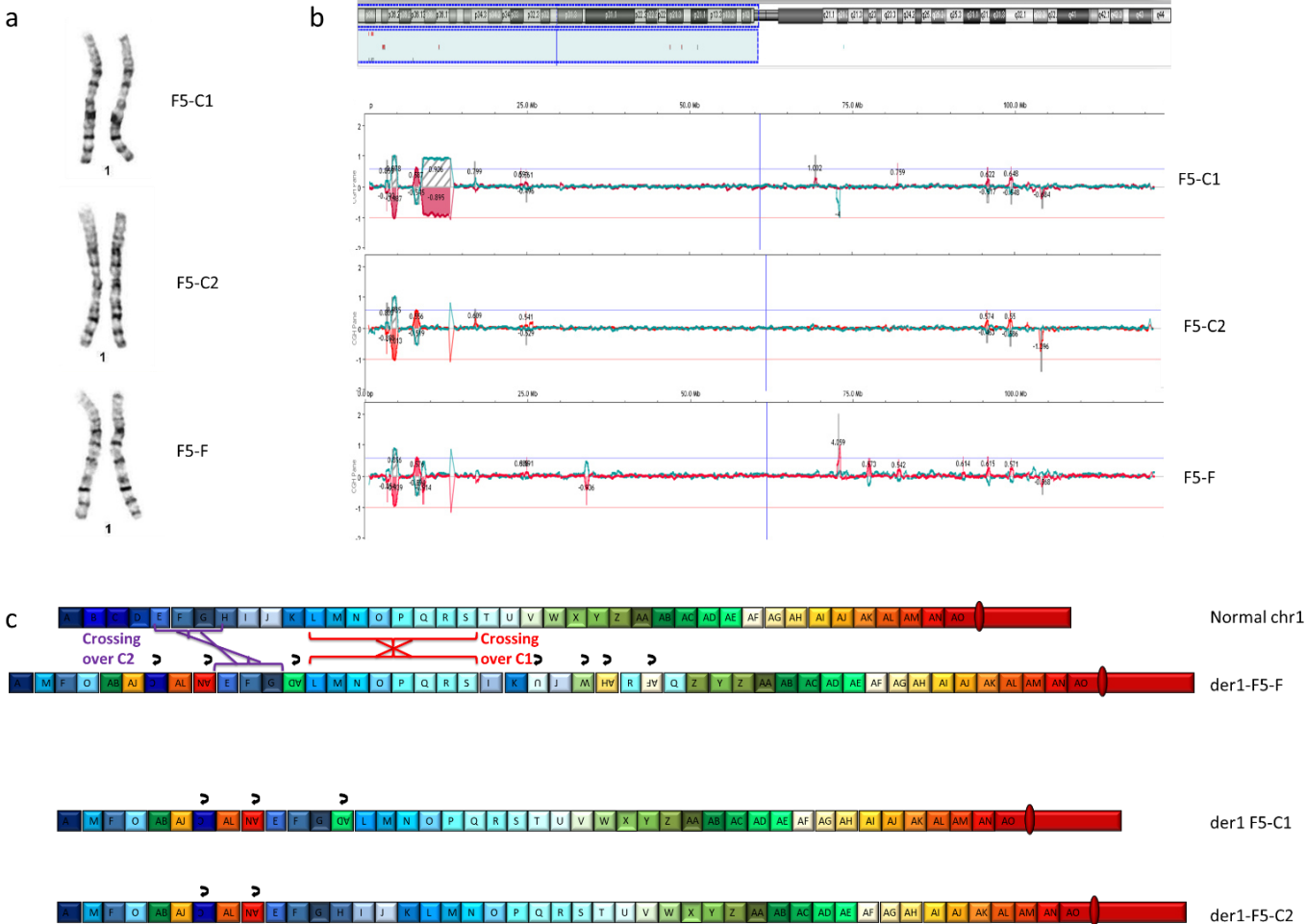


Figure 5. Family 5 rearrangement.

- a. Partial karyogram (G-band) showing chromosome 1 for F5-F, F5-C1 and F5-C2 with abnormal chromosome on right.
- b. CMA profile of short arm of chromosome 1 for F5-C1 (up), F5-C2 (middle) and F5-F (down). Red curve: Patient in Cy5 - Control 1 in Cy3; green curve: Control 2 in Cy5 - Patient in Cy3.
- c. Schematic representation of der1 thanks to WGS data for F5-C1 and F5C, and thanks to CMA data and familial rearrangement for F5-C2. Arrows represented inverted fragments. Fragment size are not at scale.

During paternal meiosis, a crossing over occurred in fragments L to S (chr1:13290274-23734482) and resulted to der1-F5-C1 composed by proximal part of the der1-F5-F (11 fragments) and normal chromosome 1 from chr1:13290274 position. A second crossing-over occurred between EFG fragments and lead to der1-F5-C2.

Family 6

F6-C karyotype revealed an apparently balanced translocation between the short arms of a chromosome 1 and the long arms of a chromosome 3 (Figure 6a). This result was confirmed by FISH using chromosome painting probes (Figure 6b). CMA found four CNVs: a gain of one copy of 269kb in the PAR1 region (Xp22.33 / Yp11.32) at chrX:537515-813813 including the *SHOX* gene, that could explain the tall stature of the patient, a loss of one copy of 1.3Mb in 1p31.1 at chr1:73960120-75313006, a gain of one copy of 3.4Mb in 1q24.2.q24.3 at chr1:169007884-172463708 and a gain of one copy of 1Mb in 3q13.11 at chr1:104033454-105110730 (supplemental table 14 and supplemental figure 10). These three latter CNVs were classified as VUS (variants of unknown significance). FISH showed a complex rearrangement with an additional signal for RP11-126C20 and RP11-277C14 probes (1q24.2q24.3) on the short arm of derivative 1 and an additional signal for RP11-91B3 probe (3q13.11) on the short arm of derivative 1 (supplemental figure 11).

Parental study identified in F6-F a complex rearrangement between chromosome 1 and 3 (Figure 6a). This was further supported by FISH studies that showed an insertion of a fragment of the long arm of chromosome 3 into the short arm of chromosome 1 and a translocation of the 3qter part onto the long arm of a chromosome 1 (Figure 6b).

F6-F rearrangement was then characterized by WGS (13.95X) (Supplemental table 14). It showed nine breakpoints, five in chromosome 1 and four in chromosome 3. This rearrangement was balanced and no CNV was detected.

All junctions were characterized at the base-pair level (supplemental table 15). Two of them showed microhomology (2 and 3 bp) and five of them insertions (from 1 to 90pb), that could suggest a replication-based mechanism.

Using all these data, rearrangement of the F6-F and F6-C were reconstructed with chromosomes 1 divided into 6 fragments (A to F) and chromosome 3 divided into 5 fragments (a to e) (Figure 6c).

During paternal meiosis, a crossing-over between D fragments of normal chromosome 1 and der1-F6-F occurred to give the der1-F6-C and a second crossing over occurred between d fragments of normal chromosome 3 and der1-F6-F to generate der3-F6-C.

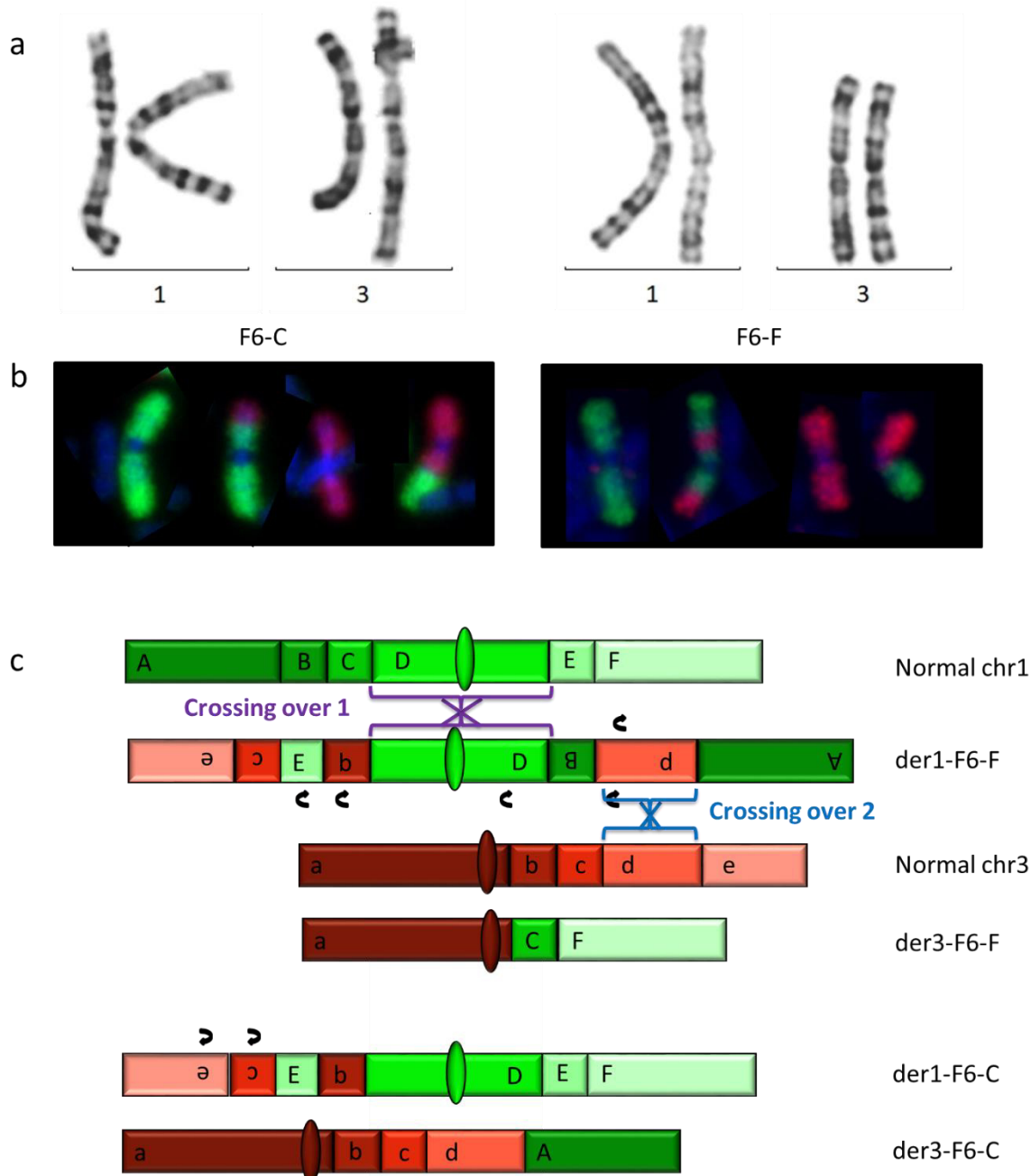


Figure 6. Family 6 rearrangement.

a. Karyotype images for F6-C (on left) and F6-F (on right), on G-banding

b. WCP of chromosome 1 (on green) and 3 (on red) of F6-C (on right) and F6-F (on left)

c. Schematic representation of WGS results. Fragments are not at scale. Arrows represented inverted fragments. The der1-F6-F is represented inversely. Purple cross represented crossing-over leading to der1-F6-C and blue cross represented crossing-over leading to der3-F6-C.

Discussion

The aim of this study was to characterize complex chromosomal rearrangements in six families in order to understand their transmission through meiosis.

The rearrangements that were identified in parents presented multiple breakpoints affecting a single or two chromosomes and were clustered on regions of 1 to about 30 Mb. So they could be defined as chromoanagenesis. Molecular signatures at the breakpoint gave clues about their mechanisms. Four of them could be defined as chromoanasythesis (F1, F2, F4 and F6) whereas two could not be unambiguously classified (F3 and F5). It was already mentioned that the distinction between chromothripsis and chromoanagenesis was not always that simple, with the presence of replicative mechanisms and NHEJ signatures in the same rearrangements (Chatron et al., 2020).

The whole genome paired-end sequencing allowed us to determine the sequences of the junction fragments and precisely characterize the breakpoints. Indeed, “short-read” WGS appeared to be an efficient method to characterize chromoanagenesis (Chatron et al., 2020; Kloosterman et al., 2011). However, this approach, in which only 100 to 150bp of short 350 bp fragments are sequenced, may be limited by the identification of breakpoints in highly repeated loci, such as short arms of acrocentrics, constitutive heterochromatin and segmental duplications (Redin et al., 2017; Schluth-Bolard et al., 2019). The reads in these sequences fail to be aligned correctly with the reference genome, either because of an incomplete reference or multiple loci localization. In the present study, the position of a few breakpoints could not be completely resolved, one breakpoint in the alpha-satellite sequences of chromosome 21 in F1 and 3 breakpoints in segmental duplications in F5. Moreover, it may be sometimes difficult to determine unambiguously the order in which the gained fragments have been inserted, especially when several fragments are clustered together. This was the case for derX-F2-M and der2-F4-M in which two hypotheses were equally possible from the WGS data (supplemental figures 3 and 8). The combination with karyotype and FISH results may help to resolve these cases. Additional solutions are now available to overcome these problems thanks to 3rd generation

technologies, making it possible to study large fragments by long read sequencing (Single Molecule Real Time Sequencing developed by Pacific Bioscience or Oxford Nanopore Technologies) and by optical mapping (BioNano Genomics) (Neveling et al., 2020). Unfortunately, there was no more sample available for our patients to perform such approaches.

In this study, the parents, three mothers and three fathers, transmitted their rearrangement to their descent either balanced or unbalanced, through meiotic recombination generating the formation of a new derivative in the children or through chromosome missegregation. Twelve families with stable or unbalanced transmission of chromoanagenesis have already been described to date (Bertelsen et al., 2016; de Pagter et al., 2015; Gu et al., 2013; Gudipati et al., 2019; Kloosterman et al., 2012; Kurtas et al., 2019; Sabatini et al., 2018; Weckselblatt et al., 2015), including three cases of chromoanasythesis (Gudipati et al., 2019; Kloosterman et al., 2012; Sabatini et al., 2018) and nine cases of chromothripsis (Bertelsen et al., 2016; de Pagter et al., 2015; Gu et al., 2013; Kurtas et al., 2019; Weckselblatt et al., 2015) (Table 1). For eight of these families, the parent was asymptomatic (Bertelsen et al., 2016; de Pagter et al., 2015; Gu et al., 2013; Kurtas et al., 2019; Sabatini et al., 2018) and the rearrangement was discovered by chance during a parental study. The discovery of these chromoanagenesis in asymptomatic people leads us to believe that the prevalence of this type of rearrangement in the general population may be underestimated.

In four families, (Bertelsen et al., 2016; Gudipati et al., 2019; Kloosterman et al., 2012; Sabatini et al., 2018), the parents transmitted their rearrangement unchanged to their offspring.

In other cases the transmission was unbalanced and lead to new rearrangements in the children with the presence of losses or gains responsible for their phenotype (de Pagter et al., 2015; Gu et al., 2013; Kurtas et al., 2019; Weckselblatt et al., 2015). This was also the case for the six families we described, with meiotic recombination issues for five families and one missegregation of derivatives.

Only three cases of paternal transmission of chromoanagenesis were described in the literature (Bertelsen et al., 2016; Kurtas et al., 2019; Sabatini et al., 2018). This was already known for complex

inherited rearrangements (Pellestor et al., 2011). Indeed, men carrier of complex rearrangement are mostly infertile or subfertile due to the deficiency or cessation of spermatogenesis (Madan, 2012). The meiosis I prophase checkpoint is supposed to be less permissive in men than in women, and spermatogenesis is then strongly disturbed by the presence of complex changes (Kurahashi et al., 2009; Sciurano et al., 2012).

In our study, derivative reconstitution using WGS data helped to elucidate where the crossing-over may have taken place to generate the transmitted derivatives. Interestingly, in most cases, it supposed segments of asynapsis between normal and derivative chromosome. To reduce the length of asynapsis, complex loops between homologues may be required (supplemental figures 12 to 17).

For F1-M, the hypothesis would be the formation of a folding loop leading to the pairing between 3 strands, chromosome 21 long arm, der21-F1-M long arm and der 21-F1-M short arm, resulting in a 1.259 Mb asynapsis fragment (Fragments A-B-D). For F2-M, we hypothesized that pairing of the two maternal X chromosomes implicated short asynapsis involving fragments of 6kb (C), 559kb (D-F-B) and 565kb (F-B-C-D).

In F3-F, we could suppose that the centromeric region (AF fragment) and the end of the long arm of chromosome 10 (AL fragment) are paired as well as the 10p15p13 region (E fragment) and 10q11q21 region (AH fragment) thanks to loops formation. This configuration would leave seven asynaptic regions of 5.785Mb (A-B-C-D), 4.422Mb (F to AE), 7kb (AG), 79kb (AI-AJ-AK), 1.873Mb (U-Z-K-AB-M-AD-V-R-P), 4.297Mb (X-I-AJ-D-C-S-N) and 3.271Mb (A-G-AE).

During F4-M meiosis, chromosome pairing would have resulted in a total asynapsis of 30 Mb composed of five regions of 260kb (I-S), 931kb (F), 646kb, 21.745 Mb (P-R-O-N-K-E-M-U-C-H-L) and 190kb (e). We can also hypothesized that der2-F4-M may also be able to pair with normal chromosome 1 reducing the length of the asynapsis.

In F5-F, the global length of asynapsis was evaluated to 24.5Mb including regions of 1.398Mb (B-C-D), 4.521kb (H-I-J-K), 13.612Mb (T-U-V-W-X), 1.875kb (M-F-O-AB-AJ-C-AL-AN), 45kb AD and 18.074Mb (I-

K-U-J-W-AH-R-AF-Q-Z). In this latter segment, a loop formation could enable the pairing of the fragment W, reducing the asynapsis of 9.7Mb for both chromosomes.

Finally, for F6-F, the pairing may have occurred through a tetravalent completed by loops formation to allow pairing of extremities A and e. That resulted in small asynapsis of 1.013Mb (b), 133kb (c), 42kb (B), 1.328Mb (C) and 3.481Mb (E).

In all these cases, meiosis was not stopped despite the presence of asynaptic regions, in particular in male meiosis in which the association of the asynaptic segments with the X-Y body was described to disrupt and stop spermatogenesis (Sciurano et al., 2012). There are unfortunately very few data about the maximal length of asynapsis tolerated in human meiosis. One would expect that short asynapsis would not disturb meiosis. In the cases with larger asynapsis (> 5Mb), we cannot exclude the formation of more complex loops allowing local pairing in order to minimize the asynapsis length.

In this study, we showed that healthy individuals, male and female, may be carrier of chromoanagenesis and could transmit these complex rearrangements to their descent in a balanced or unbalanced manner. The unbalanced transmission could involve chromosome missegregation or meiotic recombination and the resulting gamete composition is unpredictable. Prenatal diagnosis should then be proposed in these families. CMA seems to be more appropriate since it offers a rapid pangenomic detection of CNVs that could affect fetal phenotype. For couples with infertility or repeated miscarriages, diagnosis by WGS could be interesting. Indeed, the presence of very complex rearrangements can lead to a high proportion of miscarriages by recombination or by segregation of derivatives (Madan, 2012). These couples are also at risk of having a child with an imbalance.

Table 1. Review of reported inherited chromoanagenesis

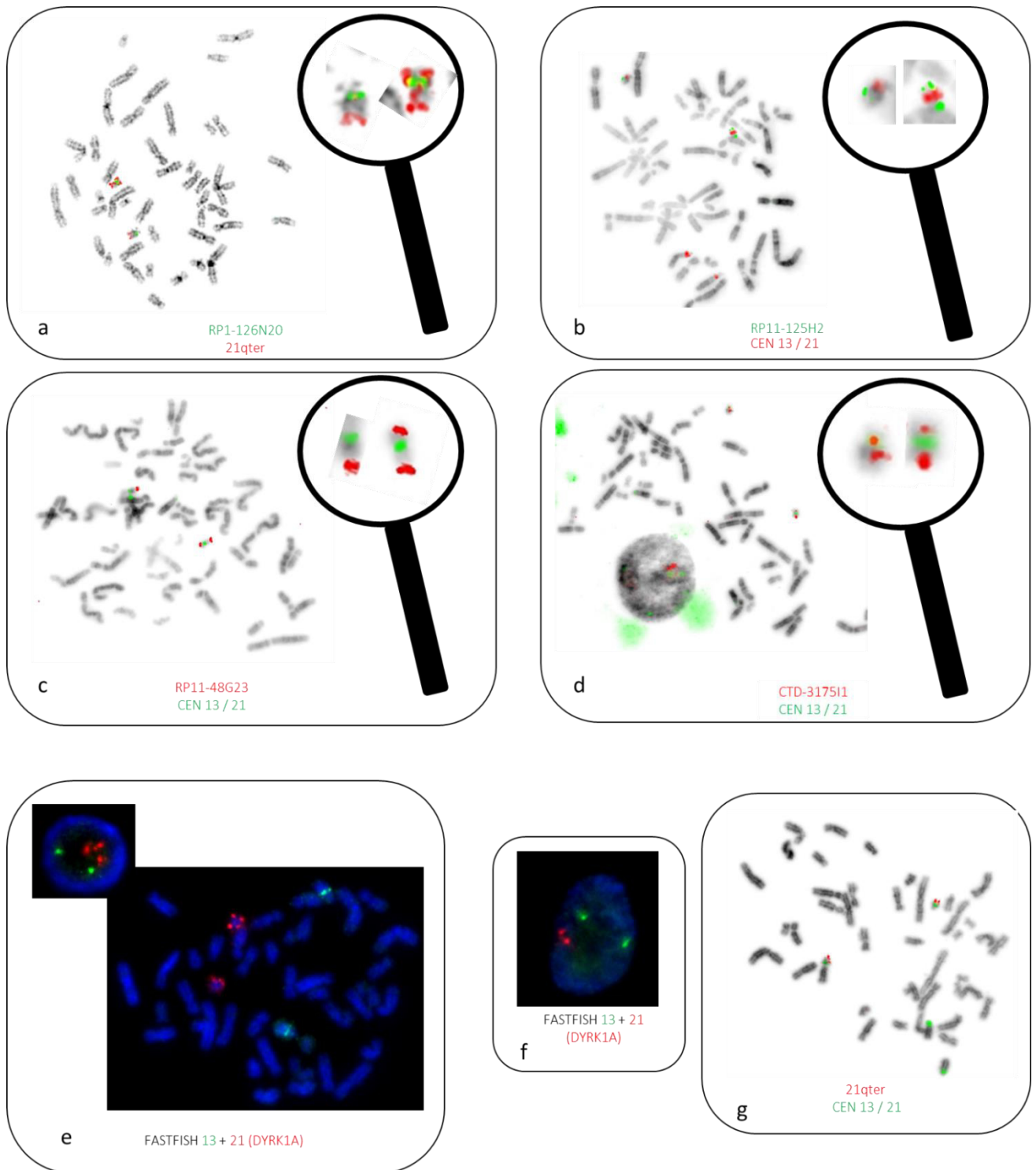
Publication	Number of cases	Chromosome(s)	Chromoanagenesis type	Transmission	Carrier parent	Parental phenotype	Family transmission
Kloosterman et al., 2012	2	X	chromoanasythesis	stable	mother	NS	NA
Gu et al., 2013	5	5, 11, 21	chromothripsis	unbalanced	mother	asymptomatic	variable
Bertelsen et al., 2015	11 3 generations	3, 5	chromothripsis	stable	1 father 3 mother	asymptomatic	stable
de Pagter et al., 2015	2	9, 10, 14, 16	chromothripsis	unbalanced	mother	asymptomatic	NA
de Pagter et al., 2015	2	6, 7, 9, 10, 12	chromothripsis	unbalanced	mother	Developmental delay	NA
de Pagter et al., 2015	2	1, 3, 5	chromothripsis	unbalanced	mother	asymptomatic	NA
Weckselblatt et al., 2015	2	1, 4, 7, 15	chromothripsis	unbalanced	mother	NS	NA
Kurtas et al., 2019	2	3, 8	chromothripsis	unbalanced	mother	asymptomatic	NA
Kurtas et al., 2019	2	6, 14	chromothripsis	unbalanced	father	asymptomatic	NA
Kurtas et al., 2019	2	6, 7, 15	chromothripsis	unbalanced	mother	asymptomatic school difficulties	NA
Sabatini et al., 2018	3 3 generations	21	chromoanasythesis	stable	father mother	asymptomatic	stable
Gudipati et al., 2019	2	1	chromoanasythesis	stable	mother	Intellectual disability	NA

NS : non specified, NA : not applicable.

- Bertelsen, B., Nazaryan-Petersen, L., Sun, W., Mehrjouy, M.M., Xie, G., Chen, W., Hjermind, L.E., Taschner, P.E.M., Tümer, Z., 2016. A germline chromothripsis event stably segregating in 11 individuals through three generations. *Genet. Med.* 18, 494–500. <https://doi.org/10.1038/gim.2015.112>
- Chatron, N., Giannuzzi, G., Rollat-Farnier, P.-A., Diguët, F., Porcu, E., Yamine, T., Uguen, K., Bellil, Z.-L., Zillhardt, J.L., Sorlin, A., Ader, F., Afenjar, A., Andrieux, J., Bardel, C., Calpena, E., Chantot-Bastarud, S., Callier, P., Chelloug, N., Chopin, E., Cordier, M.-P., Dubourg, C., Faivre, L., Girard, F., Heide, S., Herenger, Y., Jaillard, S., Keren, B., Knight, S.J.L., Lespinasse, J., Lohmann, L., Marle, N., Maroofian, R., Masurel-Paulet, A., Mathieu-Dramard, M., Metay, C., Pagnamenta, A.T., Portnoi, M.-F., Prieur, F., Rio, M., Siffroi, J.-P., Valence, S., Taylor, J.C., Wilkie, A.O.M., Edery, P., Reymond, A., Sanlaville, D., Schluth-Bolard, C., 2020. The enrichment of breakpoints in late-replicating chromatin provides novel insights into chromoanagenesis mechanisms. <https://doi.org/10.1101/2020.07.17.206771>
- Chen, K., Wallis, J.W., McLellan, M.D., Larson, D.E., Kalicki, J.M., Pohl, C.S., McGrath, S.D., Wendl, M.C., Zhang, Q., Locke, D.P., Shi, X., Fulton, R.S., Ley, T.J., Wilson, R.K., Ding, L., Mardis, E.R., 2009. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat. Methods* 6, 677–681. <https://doi.org/10.1038/nmeth.1363>
- de Pagter, M.S., van Roosmalen, M.J., Baas, A.F., Renkens, I., Duran, K.J., van Binsbergen, E., Tavakoli-Yaraki, M., Hochstenbach, R., van der Veken, L.T., Cuppen, E., Kloosterman, W.P., 2015. Chromothripsis in healthy individuals affects multiple protein-coding genes and can result in severe congenital abnormalities in offspring. *Am. J. Hum. Genet.* 96, 651–656. <https://doi.org/10.1016/j.ajhg.2015.02.005>
- Gu, H., Jiang, J., Li, J., Zhang, Y., Dong, X., Huang, Y., Son, X., Lu, X., Chen, Z., 2013. A familial Cri-du-Chat/5p deletion syndrome resulted from rare maternal complex chromosomal rearrangements (CCRs) and/or possible chromosome 5p chromothripsis. *PloS One* 8, e76985. <https://doi.org/10.1371/journal.pone.0076985>
- Gudipati, M.A., Waters, E., Greene, C., Goel, N., Hoppman, N.L., Pitel, B.A., Webley, M.R., Zou, Y., 2019. Stable transmission of complex chromosomal rearrangements involving chromosome 1q derived from constitutional chromoanagenesis. *Mol. Cytogenet.* 12, 43. <https://doi.org/10.1186/s13039-019-0455-z>
- Holland, A.J., Cleveland, D.W., 2012. Chromoanagenesis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements. *Nat. Med.* 18, 1630–1638. <https://doi.org/10.1038/nm.2988>
- Jourdy, Y., Bardel, C., Fretigny, M., Diguët, F., Rollat-Farnier, P.-A., Mathieu, M.-L., Labalme, A., Sanlaville, D., Edery, P., Vinciguerra, C., Schluth-Bolard, C., 2021. Complete characterisation of two new large Xq28 duplications involving F8 using whole genome sequencing in patients without haemophilia A. *Haemoph. Off. J. World Fed. Hemoph.* <https://doi.org/10.1111/hae.14402>
- Kloosterman, W.P., Guryev, V., van Roosmalen, M., Duran, K.J., de Bruijn, E., Bakker, S.C.M., Letteboer, T., van Nesselrooij, B., Hochstenbach, R., Poot, M., Cuppen, E., 2011. Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. *Hum. Mol. Genet.* 20, 1916–1924. <https://doi.org/10.1093/hmg/ddr073>
- Kloosterman, W.P., Tavakoli-Yaraki, M., van Roosmalen, M.J., van Binsbergen, E., Renkens, I., Duran, K., Ballarati, L., Vergult, S., Giardino, D., Hansson, K., Ruivenkamp, C.A.L., Jager, M., van Haeringen, A., Ippel, E.F., Haaf, T., Passarge, E., Hochstenbach, R., Menten, B., Larizza, L., Guryev, V., Poot, M., Cuppen, E., 2012. Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms. *Cell Rep.* 1, 648–655. <https://doi.org/10.1016/j.celrep.2012.05.009>
- Kurahashi, H., Bolor, H., Kato, T., Kogo, H., Tsutsumi, M., Inagaki, H., Ohye, T., 2009. Recent advance in our understanding of the molecular nature of chromosomal abnormalities. *J. Hum. Genet.* 54, 253–260. <https://doi.org/10.1038/jhg.2009.35>

- Kurtas, N.E., Xumerle, L., Giussani, U., Pansa, A., Cardarelli, L., Bertini, V., Valetto, A., Liehr, T., Clara Bonaglia, M., Errichiello, E., Delledonne, M., Zuffardi, O., 2019. Insertional translocation involving an additional nonchromothriptic chromosome in constitutional chromothripsis: Rule or exception? *Mol. Genet. Genomic Med.* 7, e00496. <https://doi.org/10.1002/mgg3.496>
- Liu, P., Erez, A., Nagamani, S.C.S., Dhar, S.U., Kołodziejska, K.E., Dharmadhikari, A.V., Cooper, M.L., Wiszniewska, J., Zhang, F., Withers, M.A., Bacino, C.A., Campos-Acevedo, L.D., Delgado, M.R., Freedenberg, D., Garnica, A., Grebe, T.A., Hernández-Almaguer, D., Immken, L., Lalani, S.R., McLean, S.D., Northrup, H., Scaglia, F., Strathearn, L., Trapane, P., Kang, S.-H.L., Patel, A., Cheung, S.W., Hastings, P.J., Stankiewicz, P., Lupski, J.R., Bi, W., 2011. Chromosome Catastrophes Involve Replication Mechanisms Generating Complex Genomic Rearrangements. *Cell* 146, 889–903. <https://doi.org/10.1016/j.cell.2011.07.042>
- Madan, K., 2012. Balanced complex chromosome rearrangements: Reproductive aspects. A review. *Am. J. Med. Genet. A.* 158A, 947–963. <https://doi.org/10.1002/ajmg.a.35220>
- Neveling, K., Mantere, T., Vermeulen, S., Oorsprong, M., Beek, R. van, Kater-Baats, E., Pauper, M., Zande, G. van der, Smeets, D., Weghuis, D.O., Stevens-Kroef, M.J., Hoischen, A., 2020. Next generation cytogenetics: comprehensive assessment of 48 leukemia genomes by genome imaging. *bioRxiv* 2020.02.06.935742. <https://doi.org/10.1101/2020.02.06.935742>
- Pellestor, F., Anahory, T., Lefort, G., Puechberty, J., Liehr, T., Hédon, B., Sarda, P., 2011. Complex chromosomal rearrangements: origin and meiotic behavior. *Hum. Reprod. Update* 17, 476–494. <https://doi.org/10.1093/humupd/dmr010>
- Redin, C., Brand, H., Collins, R.L., Kammin, T., Mitchell, E., Hodge, J.C., Hanscom, C., Pillalamarri, V., Seabra, C.M., Abbott, M.-A., Abdul-Rahman, O.A., Aberg, E., Adley, R., Alcaraz-Estrada, S.L., Alkuraya, F.S., An, Y., Anderson, M.-A., Antolik, C., Anyane-Yeboah, K., Atkin, J.F., Bartell, T., Bernstein, J.A., Beyer, E., Blumenthal, I., Bongers, E.M.H.F., Brilstra, E.H., Brown, C.W., Brüggewirth, H.T., Callewaert, B., Chiang, C., Corning, K., Cox, H., Cuppen, E., Currall, B.B., Cushing, T., David, D., Deardorff, M.A., Dheedene, A., D’Hooghe, M., Vries, B.B.A. de, Earl, D.L., Ferguson, H.L., Fisher, H., FitzPatrick, D.R., Gerrol, P., Giachino, D., Glessner, J.T., Gliem, T., Grady, M., Graham, B.H., Griffis, C., Gripp, K.W., Gropman, A.L., Hanson-Kahn, A., Harris, D.J., Hayden, M.A., Hill, R., Hochstenbach, R., Hoffman, J.D., Hopkin, R.J., Hubshman, M.W., Innes, A.M., Irons, M., Irving, M., Jacobsen, J.C., Janssens, S., Jewett, T., Johnson, J.P., Jongmans, M.C., Kahler, S.G., Koolen, D.A., Korzelius, J., Kroisel, P.M., Lacassie, Y., Lawless, W., Lemyre, E., Leppig, K., Levin, A.V., Li, Haibo, Li, Hong, Liao, E.C., Lim, C., Lose, E.J., Lucente, D., Macera, M.J., Manavalan, P., Mandrile, G., Marcelis, C.L., Margolin, L., Mason, T., Masser-Frye, D., McClellan, M.W., Mendoza, C.J.Z., Menten, B., Middelkamp, S., Mikami, L.R., Moe, E., Mohammed, S., Mononen, T., Mortenson, M.E., Moya, G., Nieuwint, A.W., Ordule, Z., Parkash, S., Pauker, S.P., Pereira, S., Perrin, D., Phelan, K., Aguilar, R.E.P., Poddighe, P.J., Pregno, G., Raskin, S., Reis, L., Rhead, W., Rita, D., Renkens, I., Roelens, F., Ruliera, J., Rump, P., Schilit, S.L.P., Shaheen, R., Sparkes, R., Spiegel, E., Stevens, B., Stone, M.R., Tagoe, J., Thakuria, J.V., Bon, B.W. van, Kamp, J. van de, Burgt, I. van D., Essen, T. van, Ravenswaaij-Arts, C.M. van, Roosmalen, M.J. van, Vergult, S., Volker-Touw, C.M.L., Warburton, D.P., Waterman, M.J., Wiley, S., Wilson, A., Vega, M. de la C.A.Y., Zori, R.T., Levy, B., Brunner, H.G., Leeuw, N. de, Kloosterman, W.P., Thorland, E.C., Morton, C.C., Gusella, J.F., Talkowski, M.E., 2017. The Genomic Landscape of Balanced Cytogenetic Abnormalities Associated with Human Congenital Anomalies. *Nat. Genet.* 49, 36. <https://doi.org/10.1038/ng.3720>
- Riggs, E.R., Andersen, E.F., Cherry, A.M., Kantarci, S., Kearney, H., Patel, A., Raca, G., Ritter, D.I., South, S.T., Thorland, E.C., Pineda-Alvarez, D., Aradhya, S., Martin, C.L., 2020. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet. Med.* 22, 245–257. <https://doi.org/10.1038/s41436-019-0686-8>
- Sabatini, P.J.B., Ejaz, R., Stavropoulos, D.J., Mendoza-Londono, R., Joseph-George, A.M., 2018. Stable transmission of an unbalanced chromosome 21 derived from chromoanasythesis in a

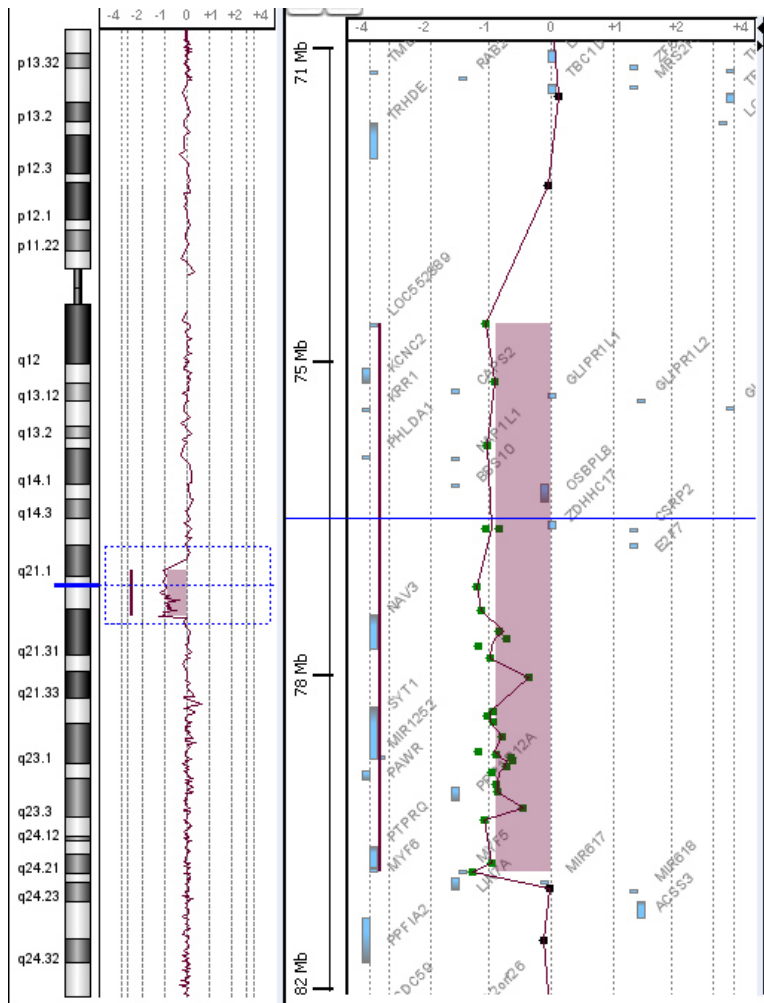
- patient with a SYNGAP1 likely pathogenic variant. *Mol. Cytogenet.* 11, 50.
<https://doi.org/10.1186/s13039-018-0394-0>
- Schluth-Bolard, C., Diguët, F., Chatron, N., Rollat-Farnier, P.-A., Bardel, C., Afenjar, A., Amblard, F., Amiel, J., Blesson, S., Callier, P., Capri, Y., Collignon, P., Cordier, M.-P., Coubes, C., Demeer, B., Chaussonot, A., Demurger, F., Devillard, F., Doco-Fenzy, M., Dupont, C., Dupont, J.-M., Dupuis-Girod, S., Faivre, L., Gilbert-Dussardier, B., Guerrot, A.-M., Houlier, M., Isidor, B., Jaillard, S., Joly-Hélas, G., Kremer, V., Lacombe, D., Caignec, C.L., Lebbar, A., Lebrun, M., Lesca, G., Lespinasse, J., Levy, J., Malan, V., Mathieu-Dramard, M., Masson, J., Masurel-Paulet, A., Mignot, C., Missirian, C., Morice-Picard, F., Moutton, S., Nadeau, G., Pebrel-Richard, C., Odent, S., Paquis-Flucklinger, V., Pasquier, L., Philip, N., Plutino, M., Pons, L., Portnoï, M.-F., Prieur, F., Puechberty, J., Putoux, A., Rio, M., Rooryck-Thambo, C., Rossi, M., Sarret, C., Satre, V., Siffroi, J.-P., Till, M., Touraine, R., Toutain, A., Toutain, J., Valence, S., Verloes, A., Whalen, S., Edery, P., Tabet, A.-C., Sanlaville, D., 2019. Whole genome paired-end sequencing elucidates functional and phenotypic consequences of balanced chromosomal rearrangement in patients with developmental disorders. *J. Med. Genet.* 56, 526–535.
<https://doi.org/10.1136/jmedgenet-2018-105778>
- Sciurano, R.B., Rahn, M.I., Rey-Valzacchi, G., Coco, R., Solari, A.J., 2012. The role of asynapsis in human spermatocyte failure. *Int. J. Androl.* 35, 541–549. <https://doi.org/10.1111/j.1365-2605.2011.01221.x>
- Stephens, P.J., Greenman, C.D., Fu, B., Yang, F., Bignell, G.R., Mudie, L.J., Pleasance, E.D., Lau, K.W., Beare, D., Stebbings, L.A., McLaren, S., Lin, M.-L., McBride, D.J., Varela, I., Nik-Zainal, S., Leroy, C., Jia, M., Menzies, A., Butler, A.P., Teague, J.W., Quail, M.A., Burton, J., Swerdlow, H., Carter, N.P., Morsberger, L.A., Iacobuzio-Donahue, C., Follows, G.A., Green, A.R., Flanagan, A.M., Stratton, M.R., Futreal, P.A., Campbell, P.J., 2011. Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. *Cell* 144, 27–40.
<https://doi.org/10.1016/j.cell.2010.11.055>
- Weckselblatt, B., Hermetz, K.E., Rudd, M.K., 2015. Unbalanced translocations arise from diverse mutational mechanisms including chromothripsis. *Genome Res.* 25, 937–947.
<https://doi.org/10.1101/gr.191247.115>
- Zhu, M., Need, A.C., Han, Y., Ge, D., Maia, J.M., Zhu, Q., Heinzen, E.L., Cirulli, E.T., Pelak, K., He, M., Ruzzo, E.K., Gumbs, C., Singh, A., Feng, S., Shianna, K.V., Goldstein, D.B., 2012. Using ERDS to infer copy-number variants in high-coverage genomes. *Am. J. Hum. Genet.* 91, 408–421.
<https://doi.org/10.1016/j.ajhg.2012.07.004>



Supplemental figure 1.

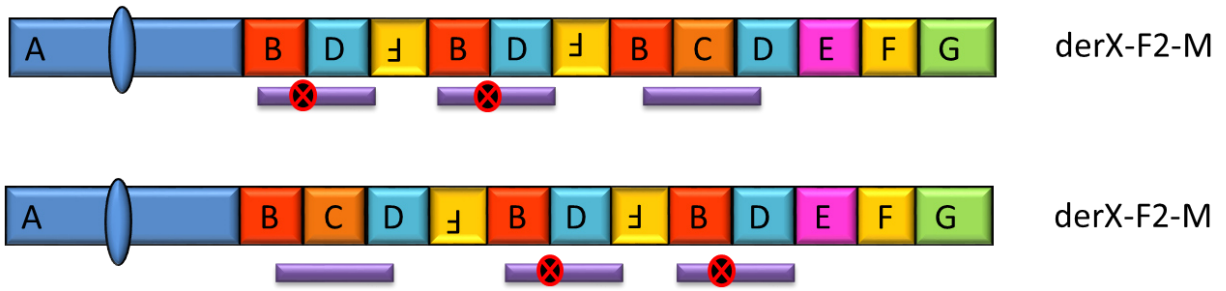
FISH images in F1-M (a, b, c, d), F1-fe1 (e), F1-fe-2 (f, g).

The der21-F1-M has two hybridization signals for the 21qter (a), RP11-125H2 (b), RP11-48G23 (c) and CTD-3175I1 (d) probes, one on the short arm of chromosome 21 and one on the long arm. The der21-F1-fe1 has two hybridization signals for the DSCR probe (DYRK1A), at each end (e). For fetus 2, a single hybridization signal for the 21qter probe is present on each chromosome 21 (f, g).

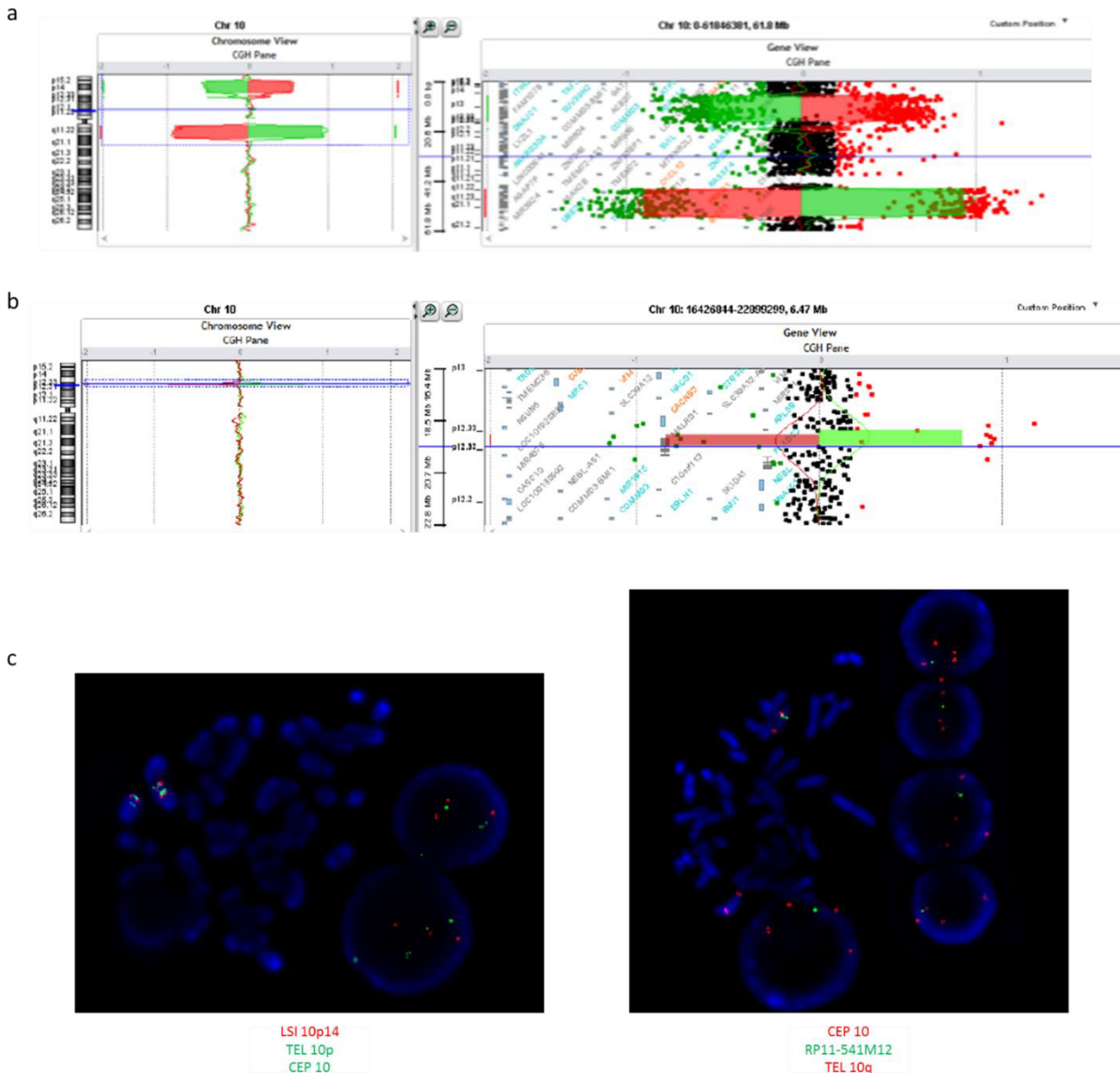


Supplemental figure 2

CMA profile performed in F2-C showing chr12:74537664-80738940 deletion



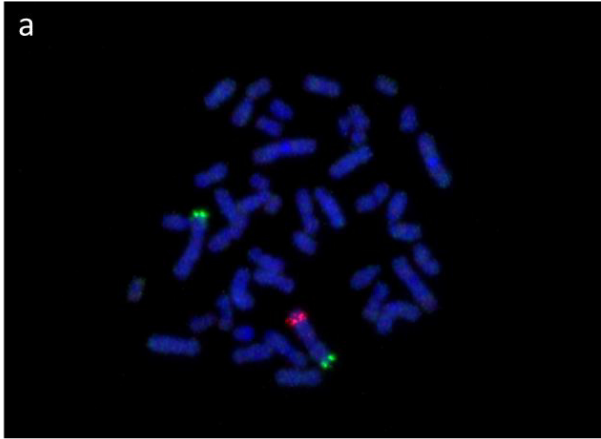
Supplemental figure 3: two possible reconstructions in F2-M according to WGS data



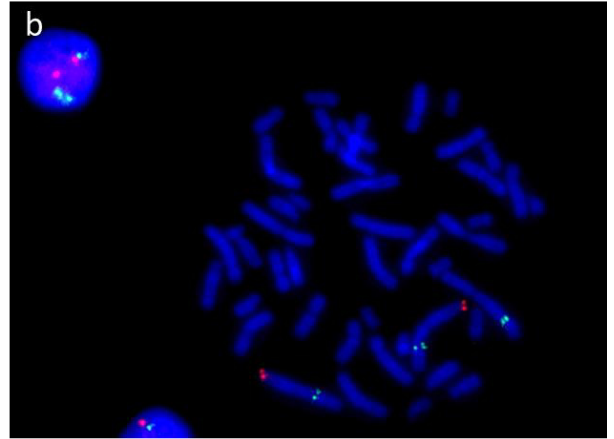
Supplemental figure 4. CMA profiles and FISH images for F3-C and F3-F.

a and b. CMA profiles for F3-C (a) and F3-F (b). Red curve: Patient in Cy5 - Control 1 in Cy3; green curve: Control 2 in Cy5 - Patient in Cy3

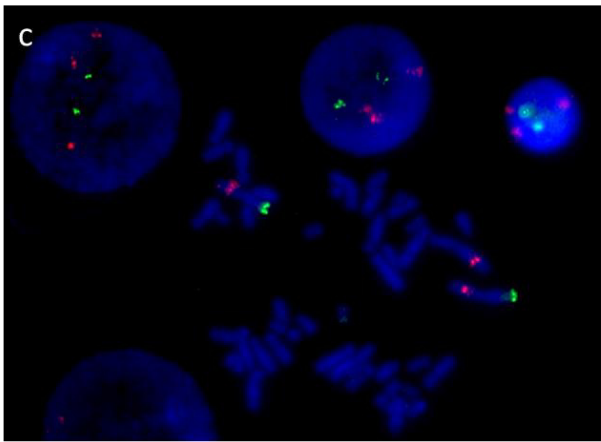
c. FISH images for F3-C. left: The 10p15p13 region was explored using the LSI 10p14 probe (red) and 10pter (green) and a centromeric probe (CEP 10) as control probe; right: the 10q11q21 region was explored using the RP11-541M12 probe (green) with CEP10 (red) and 10qter (red) as control probes. There are two signals for LSI 10p14 probe and only one signal for RP11-541M12 probe.



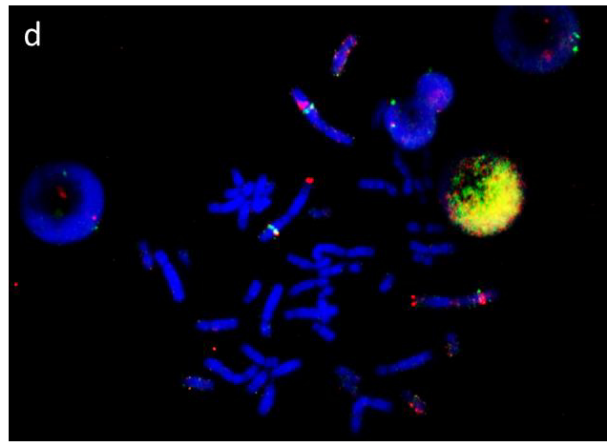
RP11-10012
2pter



RP11-126C20
1pter



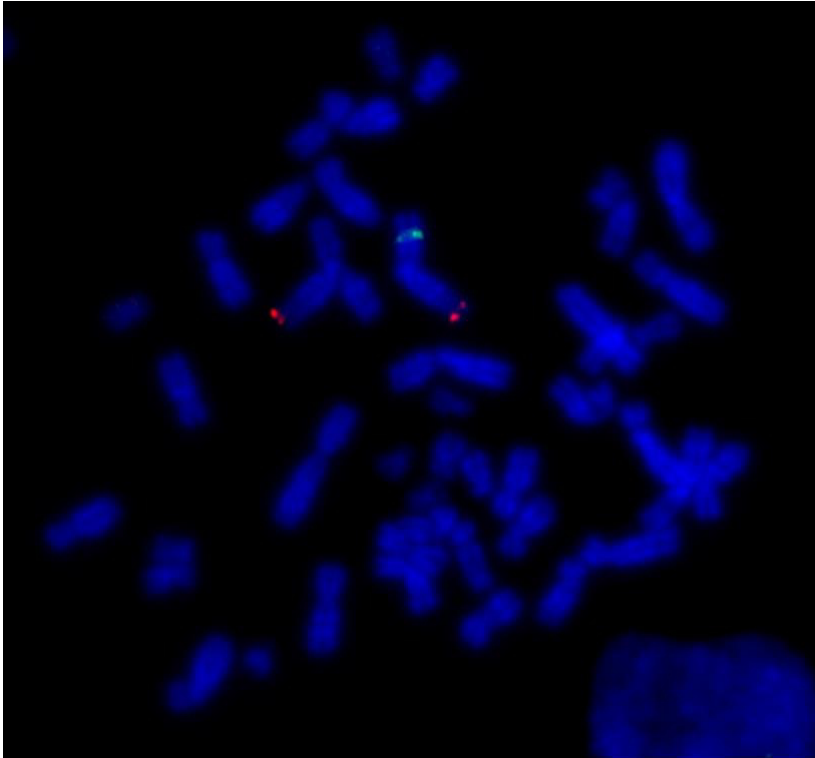
RP3-423N22
1pter



RP11-239J11
RP11-75C23
1pter

Supplemental figure 5

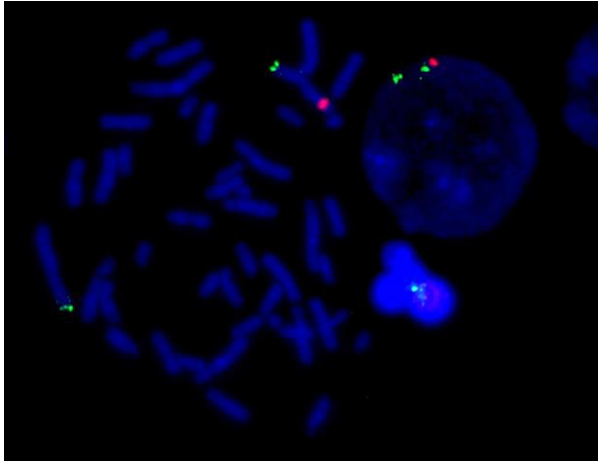
FISH images for F4-C. Only one signal was present for RP11-10012 confirming the 2q33.3 deletion (a). It also showed one additional signal for probes RP11-126C20 (b), RP3-423N22 (c), RP11-239J11 (d) and RP11-75C23 (d) on the long arm of a chromosome 2.



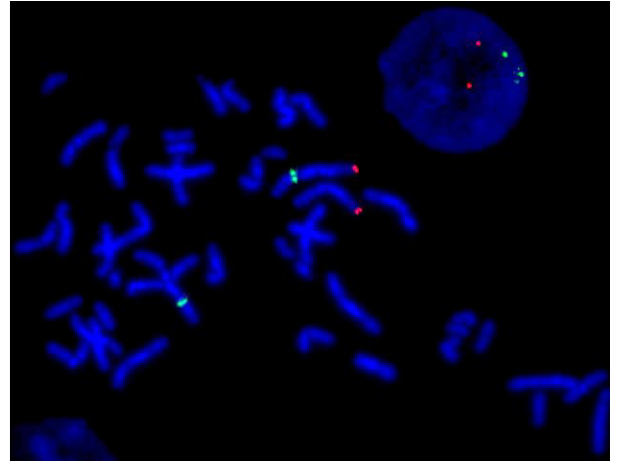
RP11-239J11
1pter

Supplemental figure 6

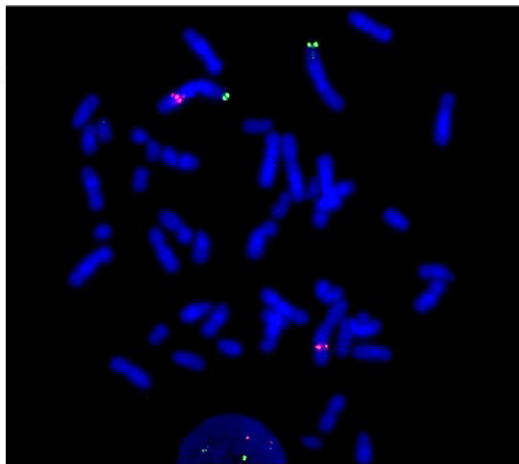
FISH image for F4-fe. It showed only one signal for probes RP11-239J11.



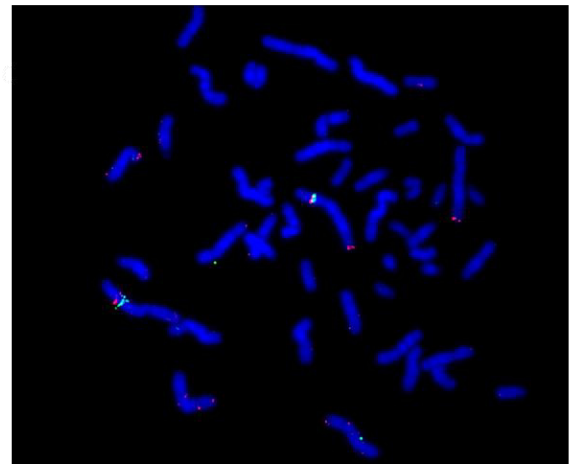
RP11-10012
2pter



RP11-126C20
1pter



RP3-423N22
1pter



RP11-239J11
RP11-75C23
1pter

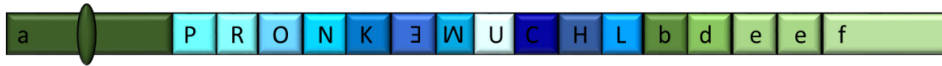
Supplemental figure 7

FISH images for F4-M.

FISH showed an apparently balanced insertion of part of the long arms of a chromosome 1 (RP11-126C20, RP3-423N22, RP11-239J11 and RP11-75C23) in the chromosome 2 long arms. This rearrangement is associated with a deletion of the 2q33.3 region (RP11-10012).



Normal chr2



der2-F4-M



Normal chr2



der2-F4-M

Supplemental figure 8

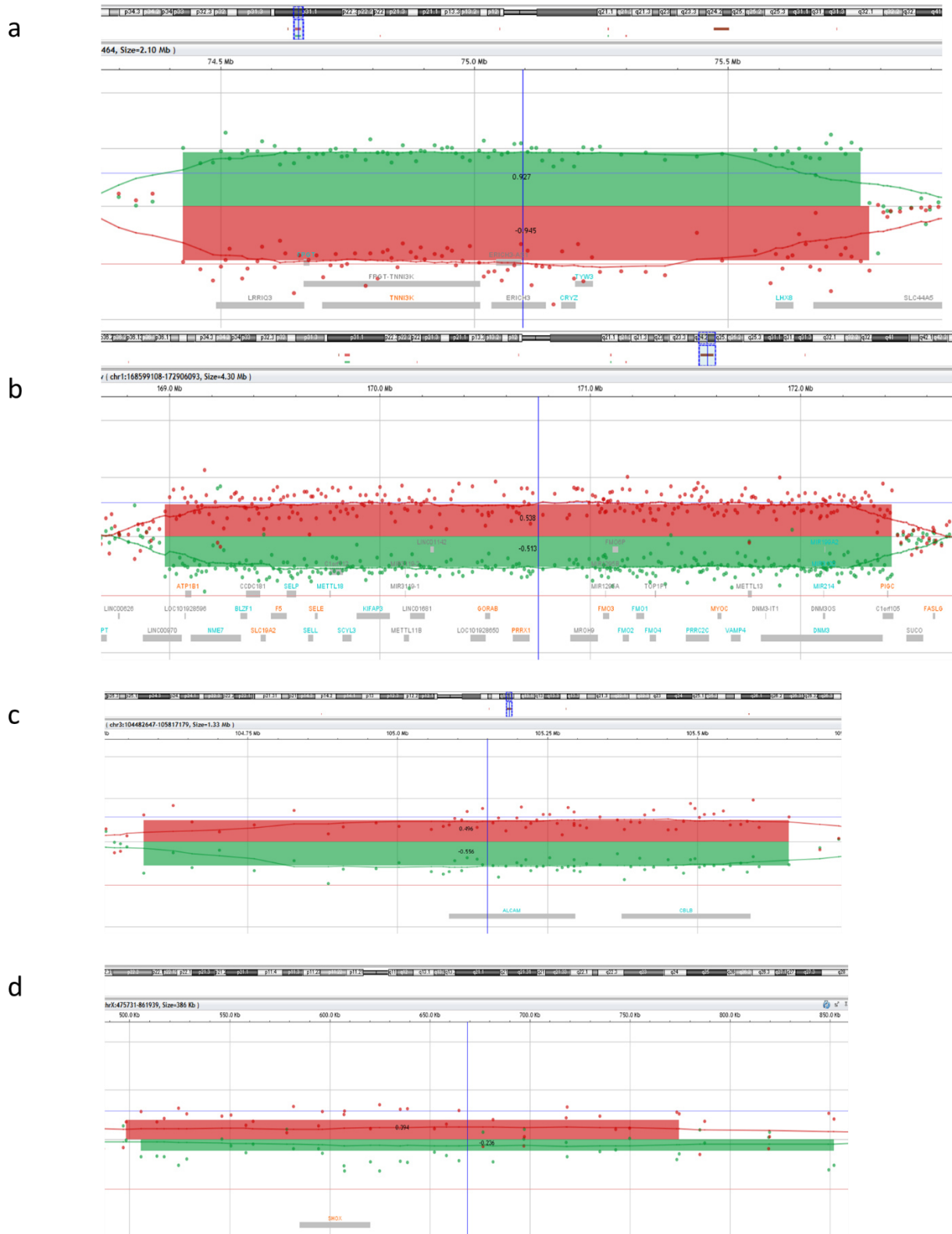
Two possible reconstructions in F4-M according to WGS data. The gain of “e” fragment could be in normal chromosome 2 or in der-2-F4-M.



RP11-168D14
RP11-479E21
1qter

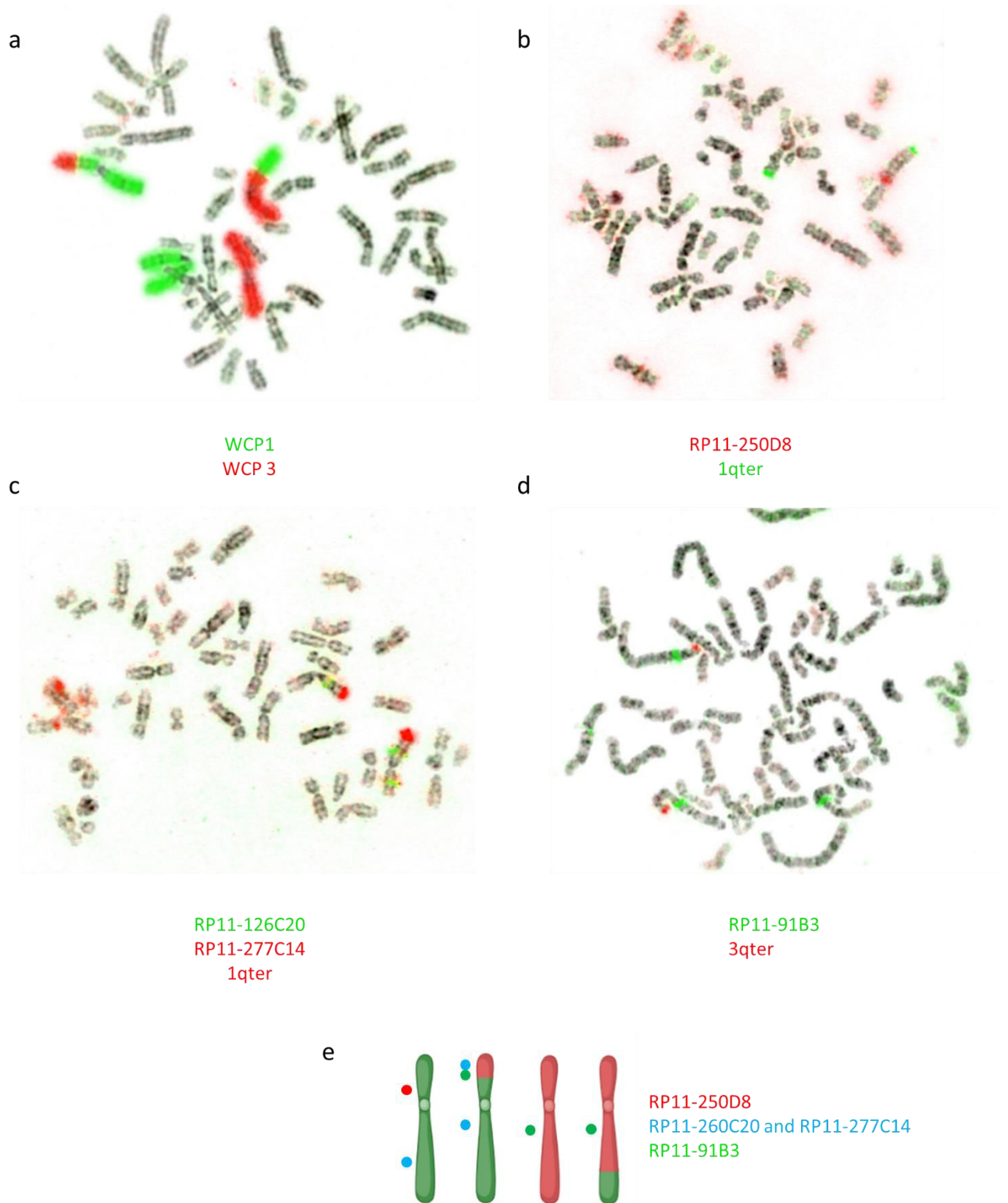
Supplemental figure 9. FISH performed in F5-F.

It showed the sequence 1pter- RP11-168D14- RP11-479E21 signal.



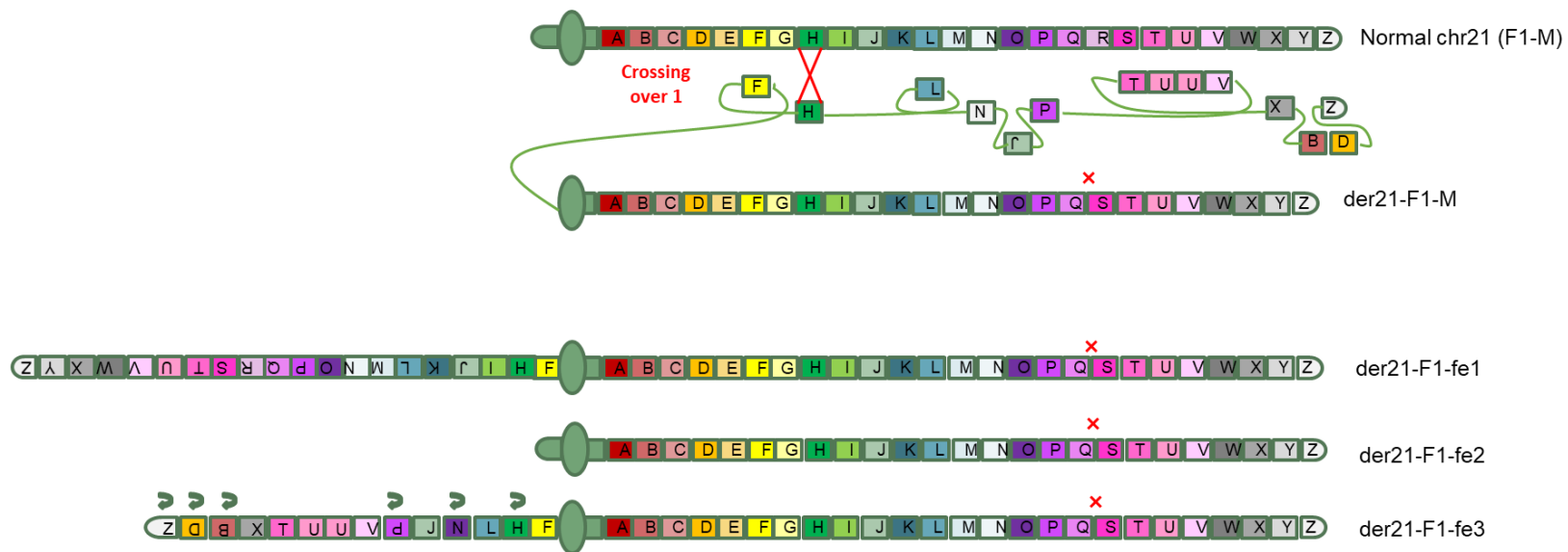
Supplemental figure 10. CMA profile of F6-C.

a-d: Red curve: Patient in Cy5 - Control 1 in Cy3; green curve: Control 2 in Cy5 - Patient in Cy3. It showed a loss of one copy of 1.3Mb in 1p31.1 at chr1:73960120-75313006 (a), a gain of one copy of 3.4Mb in 1q24.2.q24.3 at chr1:169007884-172463708 (b) and a gain of one copy of 1Mb in 3q13.11 at chr1:104033454-105110730 (c). d.Gain of one copy of 269kb in the PAR1 region (Xp22.33 / Yp11.32) at chrX:537515-813813 including the *SHOX* gene.



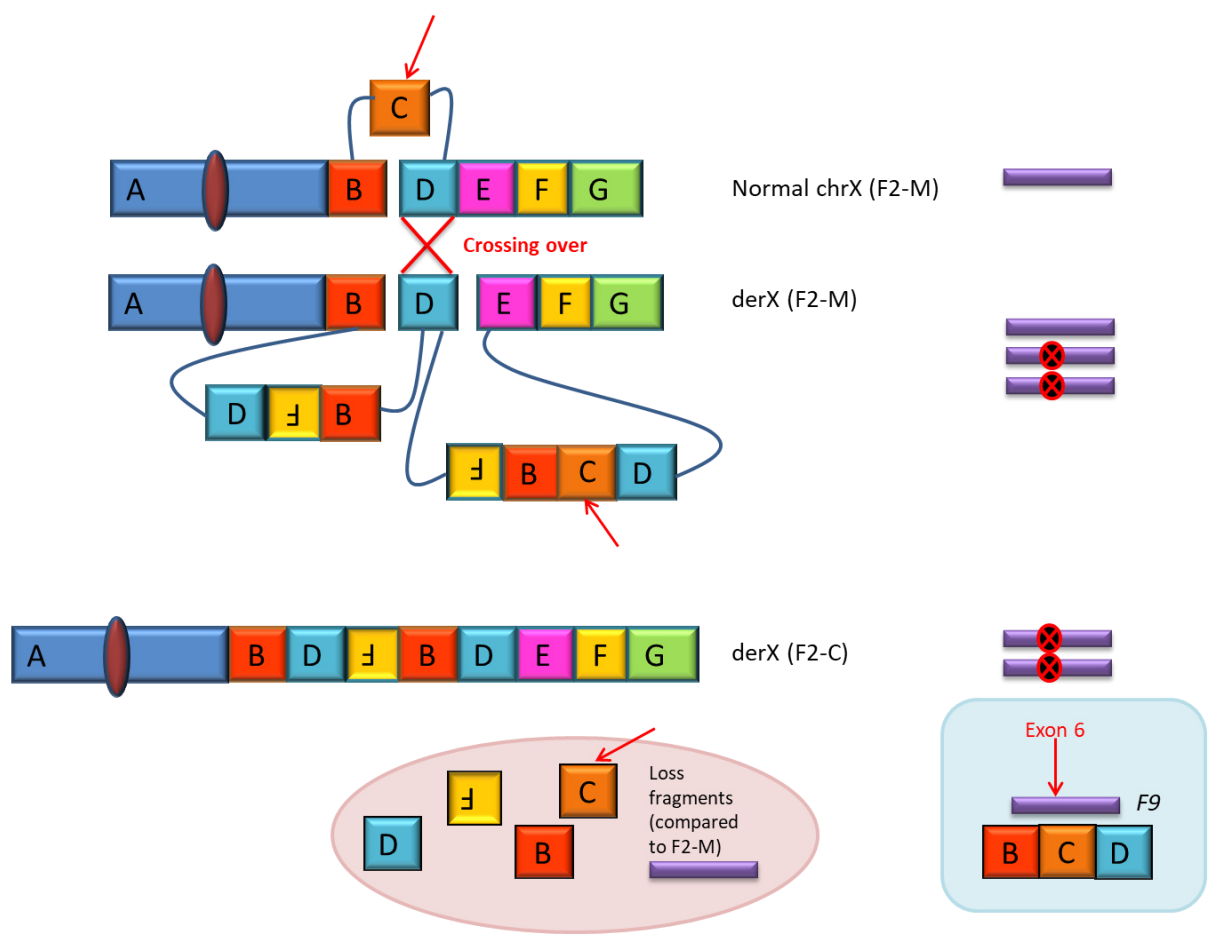
Supplemental figure 11. FISH performed in F6-C.

a. Whole chromosome painting of chromosome 1 (green) and 3 (red) resulting in apparently balance translocation between chromosomes 1 and 3. b. One signal for RP11-250D8 showing loss of the 1p31.1 region. c and d showed an additional signal for RP11-126C20 (c) and RP11-277C14 (c) probes (1q24.2q24.3) on the short arm of derivative 1 and an additional signal for RP11-91B3 (d) probe (3q13.11) on the short arm of derivative 1. Schematic representation of FISH results using probes RP11-250D8 (in red), RP11-260C20 and RP11-277C14 (in blue) and RP11-91B3 (in green).

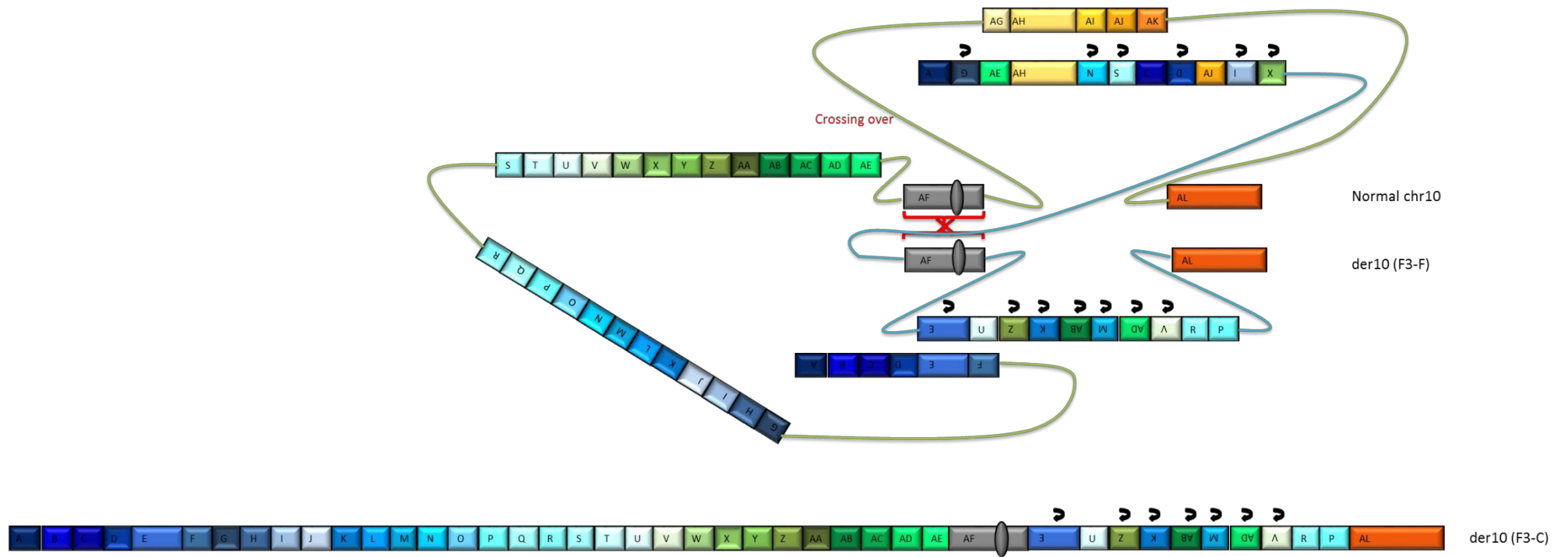


Supplemental figure 12. Representation of recombination during F1-M meiosis.

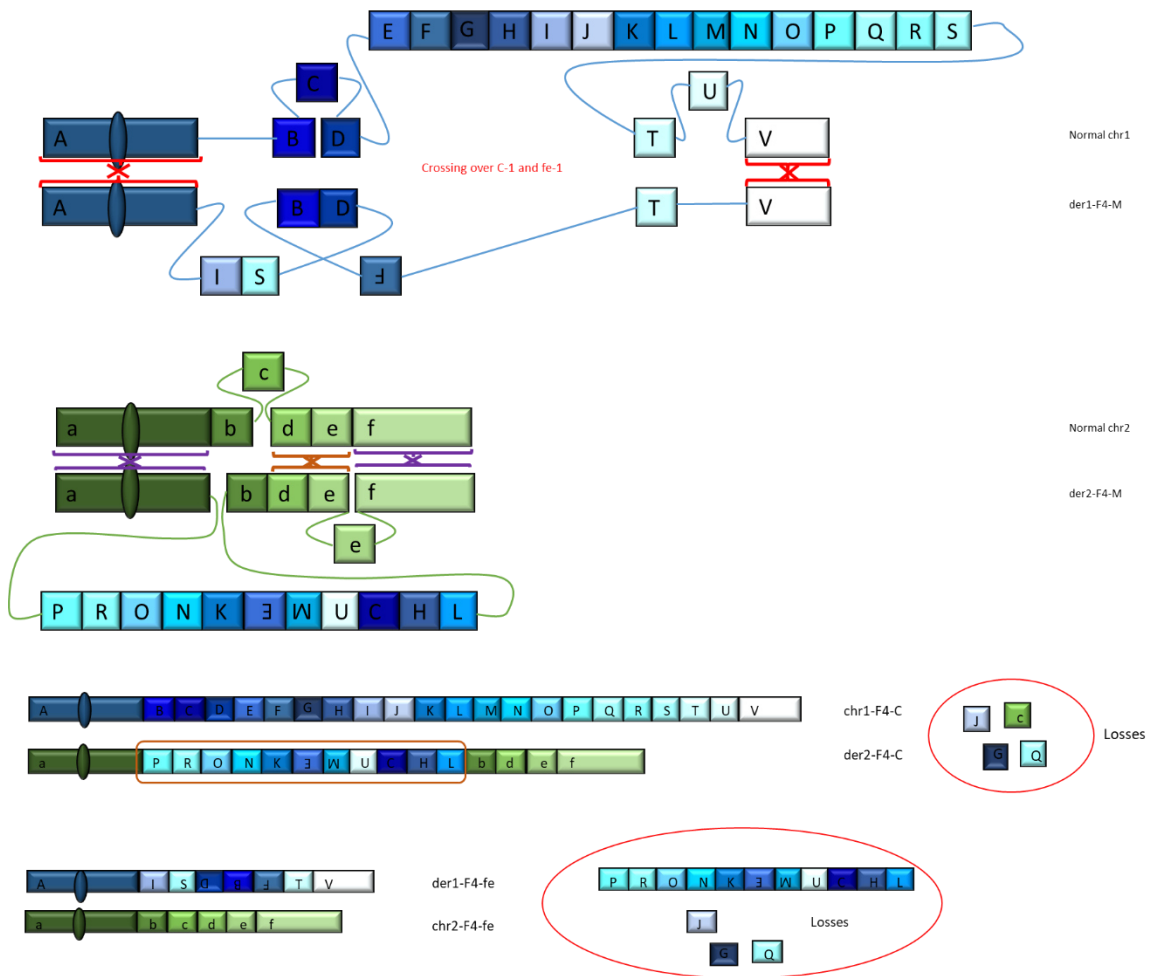
Crosses represented deleted fragment. Arrows represented inverted fragments.



Supplemental figure 13. Representation of recombination during meiosis resulting in rearranged F2-C chromosome X: derX-F2-C-. Red cross represent the crossing over resulting in derX-F2-C.

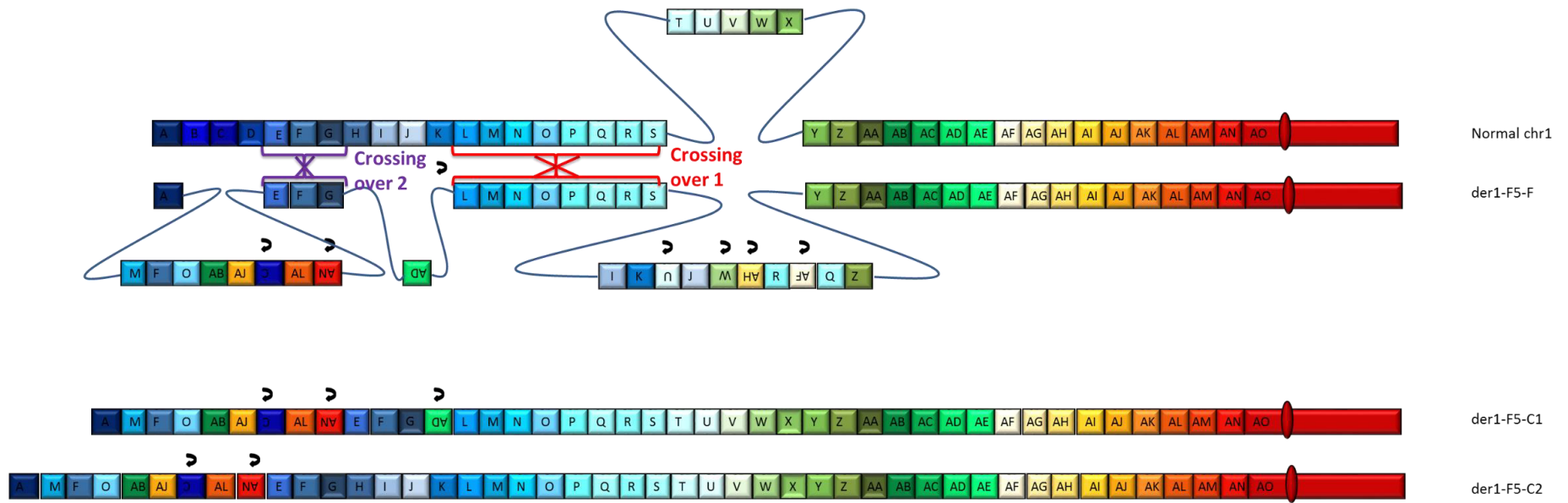


Supplemental figure 14. Representation of recombination during meiosis resulting in rearranged der10-F3-C. Red cross represent the crossing over resulting in der10-F3-C formation. Arrows represent inverted fragments.

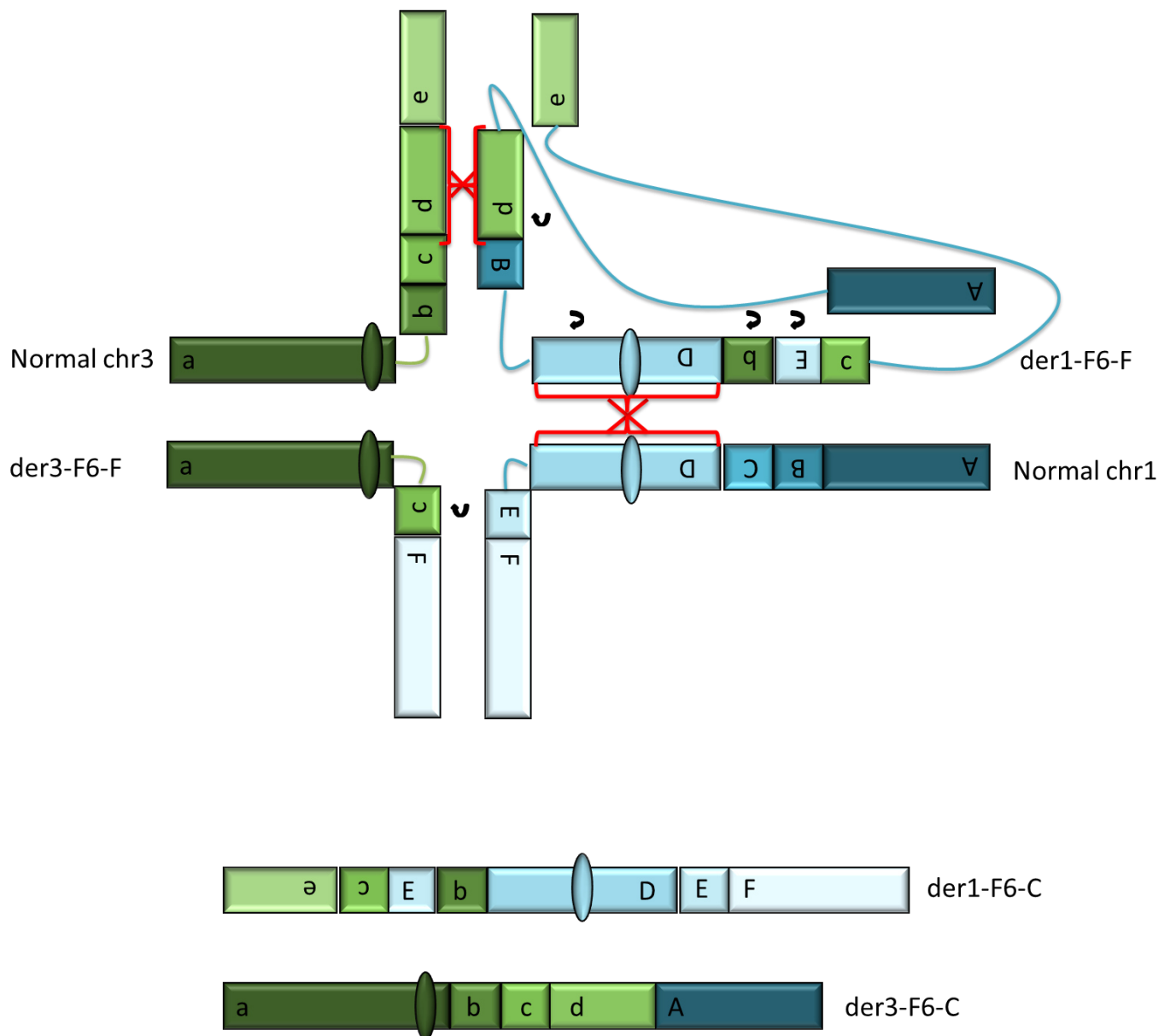


Supplemental figure 15. Representation of recombination during meiosis resulting in rearranged F4-C and F4-fe chromosomes 1 and 2.

Crossing over C-1 and fe-1 are crossing over between chr1 and der1-F4-M resulting in chromosomes 1 (chr1 and der1) for F4-C and F4-fe respectively. Crossing over C-2 and fe-2 are crossing over between chr2 and der2-F4-M resulting in chromosomes 2 (chr2 and der2) for F4-C and F4-fe respectively.



Supplemental figure 16. Representation of recombination during meiosis resulting in rearranged F5-C1 and F5-C2 der1. Crossing over 1 resulting in der1-F5-C1 and crossing over 2 resulting in der1-F5-C2



Supplemental figure 17: Representation of recombination during F6-F meiosis resulting in der1-F6-C and der3-F6-C.

Supplemental table 1. PCR primers

Patient	Forward primer	Reverse primer	Junction to confirm
F2-C	ACTGTGTTGTGCTTCCAGGT	TCAGGGTGCTGAGAGTGA CT	Fragment D - Fragment F
F2-C	GCAGCAGAAGTCCCATT TACC	CTGTCTCAGTGCAAAGCCCT	Fragment B - Fragment D
F2-C	ACAGCCTCCCCATAAGCAAC	AGGCCTCACGAGAATCTGGA	Fragment F - Fragment B
F3-F	TCCCTTAAGCAGCAAACCCA	ATGCCCAGGTGCTGTCATTT	Fragment S - Fragment C
F5-C, F5-F	ATCACTGCACTCTTTCCTCTG	GCCAGCCTGACATTCTTTTGTG	Fragment M - Fragment S chr1:98853025
F5-C, F5-F	AATTACCTGCCCTCTTCCC	TGCTTATTGTCTGCCTGTCTG	Fragment C - Fragment AL
F5-C, F5-F	CACAATCTCAGCTCACCACAAC	GCACATTCACAACAGTAACCCC	Fragment G – Fragment AD

Supplemental table 2. qPCR primers

Patient	Locus	Forward primer	Reverse primer
F1-M, F1-fe1, F1-fe2, F1-fe3	Fragment C (deletion)	GATGGCTCGTGCCTGGAAT	GCCTTCACCACCCTGTTTTG
F2-M, F2-C	Fragment F (gain)	TGCCACTTCCTGCTTCCATT	GGCAACAGCGTAAACAAGGG
F2-M, F2-C	Fragment D (Gain)	GTCTCTGGTCACTTATTGCCCA	TGGGTGTTGCATGTGGATGA
F2-M, F2-C	Fragment E (normal)	GAGTCGGGGCAAAGTCAGAA	GCGGAAATGCTAAGCCTGTG
F3-F	Deletion 1	CAATGTGTGTGCCTCTGTGC	TGGATGATTTCTGGTGCCGT
F3-F	Deletion 2	GGCAGTCACAATCCATCCC	AAGTCTCCTGTTGCAGCACA
F3-F	Deletion 3	AGCAGCCTGAGTCACCTAGA	ATGCACATGTTCTCTGGCGA
F3-F	Deletion 4	TCTCATCCGTTCACTTCGCC	GCTGGATTTGGACTTGGGCT
F3-F	Deletion 5	ACATGACGTTCTGGCTTCCA	GGATGTTGCTGGGTGACTGA
F3-F	Deletion 6	CATTCCAGCACGCACACAAA	AACACCAACTCTCAGGCTCC
F3-F	Deletion 7	AGTGGAGATGGGGTTTTGCC	AAACGCATGTAGAAGGCAGG
F3-F	Deletion 8	TGACATGCAAGGGGAAAGGG	ATCGATGGCTGTGGGACTTG
F3-F	Deletion 9	TGCAATGGAGAAAAGAGAACATGG	ACTGCATTCCATCCGCTTCG
F3-F	Deletion 11	TCCCCTTTTCCATCCTCCCT	ATCGTCCATGTGCAAGGGTG
F3-F	Deletion 12	CCCTGCTAATTCCCAACCGT	CCCTGCTAATTCCCAACCGT
F3-F	Deletion 13	CCCCGTGCCCTACTTGAAAT	TCAACCTCAGTGCCTTGCC
F3-F	Deletion 14	TGTGACCTCCCCTTCTAGCT	ATCCTGAGACTGGGCACACT
F3-F	Deletion 15	ACTCTTCCGGTGCTGTTTCCAG	GGACCAGACCCCACCATAGA

Supplemental table 3. FISH probes

Patient	Probes	Coordinates (hg19)	Locus	Provider
F1-M	RP11-513H12	chr21:15439855-15643116	21q21.1 (Fragment B)	Rainbow Fish
F1-M	RP11-950G19	chr21:27299710-27476582	21q21.3 (Fragment G)	Local probe labeling
F1-M	CTD-3175I1	chr21:30831613-30988738	21q22.11 (Fragment J)	Local probe labeling
F1-M	RP11-125H2	chr21:35411448-35563427	21q22.12 (Fragment L)	BlueFish
F1-M	RP11-48G23	chr21:45594131-45777878	21q22.3 (Fragment X)	Maison
F1-M	21qter (D21S1575)		21q22.3 (Fragment Z)	Cytocell
F1-M, F1-fe1, F1-fe2, F1-fe3	DSCR (<i>DYRK1A</i>) FAST FISH KIT (13+21)		DSCR	Cytocell
F3-F, F3-C	RP11-541M12	chr10:48352184-48559099	10q11q21 (Fragment AH)	Maison
F3-F, F3-C	LSI 10p14		10p15p13 (Fragment E)	Kreatech
F4-C, F4-M	RP11-100I2	chr2:207086913-207253605	2q33.3 (Fragment c)	Rainbow
F4-C, F4-M	RP11-126C20	chr1:169556113-169722361	1q24.2 (Fragment C)	Rainbow
F4-C, F4-fe, F4-M	RP11-239J11	chr1:193121619-193242482	1q31.2 (Fragment R)	Local probe labeling
F4-C, F4-fe, F4-M	RP11-75C23	chr1:197378486-197541678	1q31.3 (Fragment U)	Local probe labeling
F4-C, F4-M	RP3-423N22	Chr1:176901282-177036288	1q25.2 (Fragment L)	Local probe labeling
F5-F	RP11-168D14	chr1:29831258-29994457	1p35.2 (Fragment S)	Local probe labeling
F5-F	RP11-479E21	chr1:59868024-60058178	1p32.1 (Fragment Y)	BlueFish
F6-C, F6-F	RP11-126C20	chr1:169556113-169722361	1q24.2q24.3 (Fragment E)	Rainbow
F6-C, F6-F	RP11-277C14	chr1:171894064-172049904	1q24.2q24.3 (Fragment E)	Local probe labeling
F6-C, F6-F	RP11-250D8	chr1:74236637-74264832	1p31.1 (Fragment C)	Local probe labeling
F6-C, F6-F	RP11-91B3	chr3:105585432-105764957	3q13.11 (Fragment b)	Local probe labeling

Supplemental table 4. CNV identified by CMA and WGS in family 1

Patient	Coordinates (hg38)	Height (kb)	Rearrangement type	CMA coordinates (hg38)	Fragment
F1-M	chr21:15336063-15766884	430	Gain	chr21:15341716-15766035	B
	chr21:19787002-20073048	286	Gain	chr21:19797689-20042400	D
	chr21:26176397-26498731	322	Gain	chr21:26182546-26497696	F
	chr21:27479451-27844411	364	Gain	chr21:27485771-27814179	H
	chr21:30801284-31356845	555	Gain	chr21:30802834-31350990	J
	chr21:35294252-35763015	468	Gain	chr21:35307593-35762913	L
	chr21:39037793-39358479	320	Gain	chr21:39040470-39345911	N
	chr21:39895599-40502742	607	Gain	chr21:39911705-40505019	P
	chr21:41508869-41769255	260	Loss	chr21:41509750-41763347	R
	chr21:43376266-43402926	26	Gain	chr21:43378865-43695028	T
	chr21:43402926-43418143	15	Gain (4 copies)		U
	chr21:43418143-43700465	282	Gain		V
	chr21:45508296-46124868	616	Gain	chr21:45512557-46122813	X
	chr21:46469920-qter	1659	Gain	chr21:46473088-46670405	Z
F1-fe1	chr21:26176397-26498731	322	Gain	chr21:26182546-26497696	F
	chr21:27479451-qter	20650	Gain	chr21:27485771-46670405	H → Z
	chr21:43402820-43418143	15	Gain (4 copies)		U
F1-fe2	chr21:41508869-41769255	260	Loss	chr21:41509750-41763347	R
F1-fe3			Gain	chr21:15341716-15766035	B
			Gain	chr21:19797689-20042400	D
			Gain	chr21:26182546-26497696	F
			Gain	chr21:27485771-27814179	H

			Gain	chr21:30802834-31350990	J
			Gain	chr21:35307593-35762913	L
			Gain	chr21:39040470-39345911	N
			Gain	chr21:39911705-40505019	P
			Loss	chr21:41509750-41763347	R
			Gain	chr21:43378865-43695028	T
			Gain (4 copies)		U
			Gain		V
			Gain	chr21:45512557-46122813	X

Supplemental table 5: junction sequences identified in Family 1

Patient	junction	fragment 1 implicated	fragment 2 implicated	fragment 1 position	strand	fragment 2 position	strand	sequence
F1-M	1	A	F	?		27,870,932	-	GATTTTATGTGAAGGTATTAAGTTTTCTCCATAGGCTGCAAAGCGTGAGAAATGTCCAC/TTATTGTTATGTTCCAGTTTTCATGTTAGCATGGCTCATA
	2	F	H	27548715-27548720	-	28,851,770-28,851,775	+	TCCAAAGTTCTCTAAGGAAGAGAAGTACGACGTAAGACCCACATG/GCTCC/CAGCAGCTCTGTTCCGTGGTTTGCTAGGCACAACCCATGCTTCAGCACTT
	3	H	L	29216730	+	37,135,313	-	TTCCAAAACCTTGACATTTCCAAAGGTTCTCGGGACTGGCTCTCGACAAGCTCCTGGGAG/CTATTAATGTAGAGTGCTTACTAGCAAAAATGACTTGCAA
	4	L	N	36666550	-	40,409,718	+	TTTCTTCCCAGCGTATGTACTCCATGCTCCCTAAAGCATTT/AAAAGAAAGG/CAGAAGAGCTTGCTGGA AAACGCTGATAAGCCATGGTAAGAAACAA
	5	N	J	40730405	+	32,729,160	-	ATCCCTTCATAGCTTCGATATTTTCATGCTCCTCTAAGAAATGTAGATTCATCTGGGCGCAG/CTAATTG GATAATTACACAGCTCTGCTCTGCTCTCCCAAAGACCATATACATGGAATAGATATATCAGGGAT ATTTGTCTATT/AGCCCCATTAGCATGCCTGGCTAATTTTTGTGTTTTAGTAGAGATGGTGTTCCTACTGTGT TGGCCAGGCTGGACTGA
	6	J	P	32173602-32173603	-	41,267,524-41,267,525	+	TCACTTTACCAATTATATCCTGAGGCACAGAGAGGAATTTGCTCAAGC/C/TCTATTAGCAAATGTTCACTG TATAATAACAATCTTATGAACATAGTTTAC
	7	P	V	41874669	+	45,120,234	-	CCAGATAATCCACAATAATTTTCTTATGTCAAGGACAGCTG/TGTATAGACAGCTGGAACGGTGGCCAGG AGAATCCCAAAGTCTGGGATTACAGGTG/TGATGCTTCTTGAGTCTGCTGTGGCCCCACGGTGCCGG CTC
	8	U	U	44822700-44822709	-	44,838,014-44,838,023	-	TTTATGGGTCAGGAAACTGAGGCACACATGGCCGGAGGACAGAGG/TGGGTGGGA/CCCTCAGTCTCT GGGCAGGCTGGGGTCTCAGTCTCTGGTGGCTTGGGGCGTC
	9	T	X	44796146-44796151	-	47,544,777-47,544,782	-	ACTCTCCAGCACCTGGAGGCCTTTGGCATGGGGGACGAGTTGGG/GAGGC/TGAGACTCTGGGGCCAAG GCTGGGTGGCTGAGTCTCTCCAGGCCACTGC
	10	X	B	46927927	-	16,708,382	+	TATATATTGTTCTGTATCCTACAACCTTGTGTAATTCATTTATTAGCTCT/GTGGATGGAGAGGTGGTGA GTGGATAGGTTGGTAGATGGGGAGTTGGA
	11	B	D	17139203	+	21,159,316	+	GTTTGATTATTAACATATATTTGGAGAGAATTTCTTATT/CTTATTTGTTAATGTCCAAATG/TCTAAAG CTACTCTTTTTGCAGGAGTGCAGTTAGTT
	12	D	Z	21445358-21445361	+	47,889,833-47,889,836	+	AGTTGATGCAGTTAGATACAATTTGTACACAACTAGAGCAGCAGAAAAAAGTGTCTCTATTTGCA/GA J/ACCTGTTCCATCTTTTGAAGAAAAAGATAGAATTCAGAGCTG
	13	Q	S	42880796	+	43,189,415	+	ATGTTTGCAATTGTGGGGCTCAGAATTCACCAGAATTTAATAGGAC/TCGATAATGTGTATAAACAGATA AACTCGAGCCGTTCCAGGTGTGGGCGCAGTGGCTCAGGCCTGTAATCTTAGCACTTTGGGATACTGAGA CAGGTGGCCGTTTACTCACATGGCCACACTTAGCTGCCAGAGCTGGCT/CTGGGTGAAAGCAGGGCGC AGCCACGGAGGCACAGTA
F1-fe1	1	A	F	?		27870932	-	TTATGTGAAGGTATTAAGTTTTCTCCATAGGCTGCAAAGCGTGAGAAATGTCCAC/TTATTGTTATGTT CAGTTTTCATGTTAGCATGGCTCATAAGGC
	2	F	H	27548715-27548720	-	28851770-28851775	+	AGTTCTCTAAGGAAGAGAAGTACGACGTAAGACCCACATG/GCTCC/CAGCAGCTCTGTTCCGTGGTTTTGCT AGGCACAACCCATGCTTCAGCACTTATGCA

	3	Q	S	42880796	+	43189415	+	ATGTTTGCAATTGTGGGGGCTCAGAATTCACCAGAATTTAATAGGAC/ <u>TCGATAATGTGTATAAACAGATA</u> <u>AACTCGAGCCGTTCCAGGTGTGGGCGCAGTGGCTCAGGCCTGTAATCTTAGCACTTTGGGATACTGAGA</u> <u>CAGGTGGCCGTTTACTCACATGGCCACACTTAGCTGCCAGAGCTGGCT</u> /CTGGGTGGAAAGCAGGGCGC AGCCACGGAGGCACAGTAGG
F1-fe3	1	Q	S	42880796	+	43189415	+	ATGTTTGCAATTGTGGGGGCTCAGAATTCACCAGAATTTAATAGGAC/ <u>TCGATAATGTGTATAAACAGATA</u> <u>AACTCGAGCCGTTCCAGGTGTGGGCGCAGTGGCTCAGGCCTGTAATCTTAGCACTTTGGGATACTGAGA</u> <u>CAGGTGGCCGTTTACTCACATGGCCACACTTAGCTGCCAGAGCTGGCT</u> /CTGGGTGGAAAGCAGGGCGC AGCCACGGAGGCACAGTA

In green: inserted fragments of unknown origin. The "/" represents the transition from fragment 1 to fragment 2. The sequences underlined between two "/" are sequences of microhomology which may belong to either fragment. For junction 1 in F1-M and F1-fe1: the sequence in grey is a sequence of unknown position. For junction 5 in F1-M: the red sequence corresponds to the genomic position chr1: 168337343-168337362, and the orange sequence to chr4: 30251858-30251877. For junction 7 in F1-M: the sequence in dark blue corresponds to the genomic position chr1: 171686840-171686871. For junction 13 in F1-M, 3 in F1-fe1 and 1 in F1-fe2: the sequence in light blue corresponds to the genomic position chr18: 12340461-12340481, the violet chr21: 32729152-32729204 and the yellow chr21: 43550704-43550726. The genomic positions are in hg19 version of the genome. Coordinates are in hg19.

Supplemental table 6. CNV identified by CMA and WGS in family 2

Patient	Coordinates (hg38)	Height (kb)	Rearrangement type	CMA coordinates (hg38)	Fragment
F2-C	chrX:139491996-139550775	58	Gain (2 copies)	chrX:139492966-140007131	B
	chrX:139557243-140017...	459	Gain (2 copies)		D
	chrX:140174831-140215...	40	Gain (2 copies)		F
	chrX:139550775-139557243	6	Deletion	chrX:139551007-139557201	C
F2-M	chrX:139491996-139550775	58	Gain (4 copies)	chrX:139492966-140007131	B
	chrX:139557243-140017...	459	Gain (4 copies)		D
	chrX:140174831-140215...	40	Gain (4 copies)		F

Supplemental table 7 junction sequences identified in Family 2

Patient	junction	fragment 1 implicated	fragment 2 implicated	fragment 1 position	strand	fragment 2 position	strand	sequence
F2-C	1	B	D	138632932-138632934	+	138639402-138639404	+	CAGGGATGGCAATTGGGAGTAAGACTTTTTAGTAAAGAACTAAAC ACAAAGTCATTAGACT/ <u>CT</u> /TGTCCTTGACCGAGCAGACATCTCATAC CCCAAATAGCTAATATTTTGATAGCTATGATCCTGAACGGCCAAACAT TCCAAAACCA
	2	D	F	139100...		139298...		
	3	F	B	139256989-139256992	-	138574155-138574158	+	TTGTGCTTCAAATGTCCCAACAAGAATGTAAAAAGCCAACAC/ <u>ATA</u> / CACAGAAGGAAGACCAAAAAGGGATAAAGGTTGAGATATGATATG GTCATGGAAATTGGCTGGGAGTAGGGTGCTAGAAGAGGCTGGAGA ATGTTTTACAGGTG
F2-M	1	B	D	138632932-138632934	+	138639402-138639404	+	AAGACTTCCATTCAGGGATTGCAATTGGGAGTAAGACTTTTTAGTAA AGAACTAAACACAAAGTCATTAGACT/ <u>CT</u> /TGTCCTTGACCGAGCAG ACATCTCATACCCCAAATAGCTAATATTTTGATAGCTATGATCCTGAA CGGCCAAAC
	2	D	F	139100...		139298...		
	3	F	B	139256989-139256992	-	138574155-138574158	+	TGGACTTTTCCAAAATTGAAAACTTTTGTGCTTCAAATGTCCCAACA AGAATGTAAAAAGCCAACAC/ <u>ATA</u> /CACAGAAGGAAGACCAAAAAG GGATAAAGGTTGAGATATGATATGGTCATGGAAATTGGCTGGGAGT AGGGTGCTAGAA

The "/" represents the transition from fragment 1 to fragment 2. The sequences in red underlined between two "/" are areas of microhomology which may belong to either fragment. Coordinates are in hg19.

Supplemental table 8. CNV identified by CMA and WGS in family 3

Patient	Coordinates (hg38)	Height (kb)	Rearrangement type	CMA coordinates (hg38)	Fragment
F3-F	chr10:3037021-3041562	4,5	Deletion		B
	chr10:15930452-15947809	17,4	Deletion	*	F
	chr10:15950919-15978316	27,4	Deletion		H
	chr10:16006038-16030820	24,8	Deletion		J
	chr10:16803317-16823031	19,7	Deletion		L
	chr10:18133008-18137719	6,7	Deletion		O
	chr10:18363990-18365371	1,4	Deletion		Q
	chr10:18451843-18510213	58,4	Deletion	*	T
	chr10:18556171-18559659	3,5	Deletion		W
	chr10:18799523-19365912	566,4	Deletion	chr10:18890552-19361644	Y
	chr10:19791615-19797086	5,4	Deletion		AA
	chr10:19914147-19954892	40,7	Deletion	*	AC
	chr10:43814192-43820799	6,6	Deletion		AG
	chr10:54318592-54342758	24,2	Deletion		AI
chr10:54350099-54397546	47,5	Deletion	*	AK	
F3-C	chr10:5784216-15930452	10146	Duplication	chr10:5796821-16902584	E
	chr10:16030821-16803317	772,5	Duplication		K
	chr10:16823031-16915828	92,7	Duplication		M
	chr10:18137719-18363988	226,3	Duplication	chr10:18150911-18363263	P
	chr10:18365371-18388110	22,7	Duplication		R
	chr10:18510213-18523298	13,1	Duplication		U
	chr10:18523302-18556075	32,7	Duplication		V
	chr10:19365912-19791615	425,7	Duplication	chr10:19361585-20102260	Z
	chr10:19797086-19914147	117,1	Duplication		AB
	chr10:19954892-20125435	170,5	Duplication		AD
	chr10:43814192-54397546	10583	Deletion	chr10:43858258-54371980	AG → AK

*: CNV found in CMA data after reanalysis (CNV not initially identified due to number of oligonucleotides presents in CNV lower to three)

F3-F	1	A	G	3079211-3079213	+	15992916-15992918	-	ATTCGATGCATAAGATACAGAGGCACACAGTCTCTAGCCTTAAATGTTTCTGTTGAATTTGATAGTTCGTA GGG/ TA /ACTTCTATTCCTGGCTTTGTCTACTAATTAATAATGAACTAATCCCAGTGATAATTTTTCAATACCAT TTCC
	2	G	AE	15989808	-	20414367	+	AAGATACCATCTCACACCAGTCAGAATGGCCATTTTTAAAAAATCAAAAAATAACACAT/ AA /GAAGGGCTG GATAGAGAGATTGTATACCATCACCTTGGCGGACCAGACTACTAAAACATCTGAGAAAATGTAGGCACGAT CACAACCTC
	3	AE	AH	20644913	+	44316247	+	ATTGCCATGTATTCTCTCGGGAAAAAATGTGCCTGGCATAGGAGAACACTCAGCCGCATGAATACTTAATA AA/ AA /AGCCTGGATAGTTTTTGTCTCCAAGCTCTCCCTGAAGCTTTCATAACTATCTAATGAATGCATCCAG GAAATG
	4	AH	N	56078352	+	18421937	-	TACATTGTTACTAACTAAAACCCAAAGTTTCTTTAGATTTCCCTAGTTTTACTTAATGTCAGTTTTCTGTTCA GGATCCCAATCAGGATAC/GATGGAGGGTAAGGAGAGGGAGAGCATTAGGACAAACTTAATGTATGTG GGGCTT
	5	N	S	16957820-16957821	-	18740771-18740772	-	CTTTTGAAGACTTTAACCTTCAGAATCTTCTGGCTGTGAAAAAGACTTCGTGGAGATCTGGGACAATCATAC CTCT/ G /CATTCTGTTCAAAGATGTGTGCCTTCAAAGTTAAATGCAATGAAAACGCAGGTTTACTAGCTTGT CA
	6	S	C	18677043	-	3083754	+	CTTTTGTTCATGAAAATATCTGCTCTATTTTGCATTTGAAGGTGTAATGGACTTTAATGATAGCAAACTTC CATCAAGCAGGGC/ CATTTTGCTTGAAGGCAGCATGCTCGTTAAGAGTCATCACCCTCCCTAATCTCAAGT AATC/????/TTTGACCAATCACCCGCCAAAACCCCTTTTACAAAACCAAATAATAAAAAAAAAAAAA AAATAAATTAATAAAAAAAAAAAAAAAAAAAAAA /GAAAAATGGTACAACGAGTTTGAAAAACAGTTTGAC CATTCTTTAAAAATTAATA
	7	C	D	5670457-5670459	+	5826179-5826181	-	GTGTCAGATCTCCTTGATTCTACTAACTCCAGGCATCTTCTTCTCTGTCTATCCAGTCCACCATAGACT CATTCTGGAAAGCGGGGAGGG/ AT /AGAATTGAATAAGTAGAAGAAAGAAATTTCCAGAGTTCGAAGACAA GGCTTTTG
	8	D	AJ	5670458-5670460	-	56109857-56109859	-	TCCCAGACCCAGGGAGAGCAGTGTTCTCCGGACCTACAGCGTAATAGTTTCTTTCTGCTCCCTCTCTGC C/ AT /GATTATAACCTTTATATAATATCATTATAGAAAAAGGTTTTCTATTCTACGGCTAATAACAATTGT TTCT
	9	AJ	I	56102518-56102520	-	16048035-16048037	-	TCGAAATAATCTCATTGACTCCATGTCTCACATCCAGGGCACATTAATGCAAAGGGTAAGCTTCCAAGGCCT TGG/ GC /GAAGTGGTTCATTGATAATTATTCATTAGTTGAAAATTAATTTAGTTGATTACCTTAAGTACAA AAG
	10	I	X	16020315-16020316	-	19088451-19088452	-	AAGGAGTTGCAAAGGGTTTCTGCAAAGAGGAATAGAGAAATGGGCATTTGCAATGAGGCATATAGAAGT ATGGGAT/ G /TCTTGGAAGCTGCATTTCTAGGTATTTTATTCTTTTGTGGCTATTGCAATGAGACTGTGTTT TGGATTTG
	11	X	AF	18848588-18848589	-	20644915-20644916	+	TCCATTGTCCATTCATTTCTTCTTATTGCATTCCATATTTCAAATGCATTACATTTACATTCCATTGCCG/ I / ACCAAGAAATGCCTAAACCTTCCATATACCAGAATAAAGCTTATGCTCAGAAAGTCTTGTTTAGCATGTT
	12	AF	E	44309637-44309640	+	15972448-15972451	-	ACAGCACTTGCCATCATAGGAATAGAATGTATGTATCACACCCACATGCACACACACTGCCTAGGAAGGCA ATTGCTG/ CAT /CATGAGGAAGAAATGAGACAAGGTCCAAAAGGTAGACTATTCAACAAGACAATTGGTCA AGTCCCTGT
	13	E	U	5826182-5826179	-	18799142-18799145	+	ATAAATAATCACAACCTTCTGAAAATCAAGGGACACACTTAGAGACATGCAAAATGCACTGGAAAGTCTCAGC A/ ATA /CAGGTAAATTTCAAACCTGACTCAGGCTATGTGTGGAAAGCCAAACACCTGGGCAACGTTGCAG CCAAAGAT

14	U	Z	18812222-18812227	+	20080539-20080544	-	CATGAGCCTGAACACTAAAAAAATTCTGCTAACAAAGGCACTGTGTTTTCCATGCCTGAGTCAACCCACTGCCCTTTCCTAA/ <u>ATGCC</u> /TATCCAAATAATCCAGAATCATCCATCTCAAGGTGCTTAAATTAATGTCTACAAAGTCCCTT
15	Z	K	19654841-19654842	-	16845315-16845316	-	TCACACAATTGATAGGAAGCTCATGACAAACAAATCTAATGTATGGTAATAGACATCAAATACTGGTTGCCTC/ <u>T</u> /AGAGCTTGTACATATTTTTGTTAGCTATATATCTAAATATTTTCATTTGTTTGGTGTTAATGTAATGGTGTGTTT
16	K	AB	16072819-16072820	-	20203075-20203076	-	TTCTGTTTGGCAAATTCAGTGCCAGGTATGAGCTGCAGCAGTACCTGGCAAGCCCAGTGTCCCTCTCCACACAA/ <u>T</u> /CCTTCCTTCTTTCTTCTCCCTACCTCCCTACCTCCCTCCCTTCTTCCCTCCCTTCTTCTCCTTCCCT
17	AB	M	20086015	-	16957827	-	GGCACAATCATAGCTCACTGCAACTCTGACCTCCTGGACTCAAGCAATTCCCTCCTCAGTCTCCC/ <u>TC</u> /TACCTCTGGTTAGTAAAAATATCTGCTCTCTTTTGGGTGTTTATTTTACATGTACTTATTCAGCCTCTGAAGTTT
18	M	AD	16865030	-	20414364	-	TGTACTCTGTGTGCCTGCAGGTTTAAACACCCCGTGGAGGCTGCCAAGGCTTATGGCTTTCATTCTCCAAAGTGGCAGCCCAAGCTG/CCAAGACTACAAACAGGTTGGATTTTGTTCAGCTTCAATCAGTTCTACACCACTTCAAATAC
19	AD	V	20243821-20243823	-	18845098-18845100	-	ATTTGCTTTAGGAAAAGGGTCCTGAAAGAAGTGACAGATAGATACGACAGGG/ <u>AT</u> /TCCATTCCATTCTCCATTCCATTCCATTCTCCATTCTCCATTCTCCATTCAATTCTCGATTC
20	V	R	18812231-18812235	-	18654300-18654304	+	TCCAAATTCAAACTTACTCAGGTTTAAAGTACAAGTAGTAATCCAGGAAATTCATCCATTCCC/ <u>CATC</u> /CATATTTGACAGTTATTTGGTGACGGCCAATTATTTACAAATAAAAGCGATTTGGACTTTACATTAGTTTGC
21	R	P	18677038-18677039	+	18426648-18426649	+	ATAGGAGTGAGCCACTGCACCTGGCCCAAGCATTCTCATCCTTACCAAATGTATTCTTACCAAAGTCAAGCCAT/ <u>T</u> /ACAGGTGGTATTCGGTTACATGAGTAAGTTCTTTAGTGGAGATTTGTGAGAACCTGGTACCCATCAAC
22	P	AL	18652918-18652919	+	56157306-56157307	+	CGAATAGTAGGGTCTTGAAGGCAGTTTAAATTTTTTTTTTTTTTTTTTTAGATGATGTTTTGCTCGTTGCC/ <u>T</u> /TAGTCAAGTTCTATTTTGGCAATGGGATTCATATTGGAACCTAATTTTATATTATAAAAAATGGATAAATAATCACAACA

In green: inserted fragments of unknown origin. The "/" represents the transition from fragment 1 to fragment 2. The sequences in red underlined between two "/" are areas of microhomology which may belong to either fragment. Coordinates are in hg19.

Supplemental table 10. Family 4 CNVs identified by WGS

Patient	Coordinates (hg38)	Height (kb)	Rearrangement type	CMA coordinates (hg38)	Fragment
F4-M	chr1:176200525-176206664	6,1	Deletion		G
	chr1:176304475-176405780	101,3	Deletion	chr1:176328453-176374994	J
	chr1:178402635-178550945	148,3	Deletion	chr1:178402889-178548610	Q
	chr2:206882041-207527634	645,6	Deletion	chr2:206906076-207515805	c
	chr2:229848424-230038609	190,2	Gain	chr2:229851943-230020147	e
F4-C			Gain	chr1:169542256-169772213	C
			Gain	chr1:176409669-178390838	L-M-N-O-P
			Gain	chr1:178551617-194401653	R
			Gain	chr1:195658048-199165384	U
			Deletion	chr2:206906076-207515805	c
F4-fe			Deletion	chr1:176209674-194401653	H-I-J-K-L-M-N-O-P-Q-R
			Deletion	chr1:195718529-199149645	U

Supplemental table 11: junction sequences identified in F4-M

junction	fragment 1 implicated	fragment 2 implicated	fragment 1 position	strand	fragment 2 position	strand	sequence
1	a	P	chr2:177091762	+	chr1:178402635	-	CCAGCTTGGTCTAAAACCTTTTAAATAACTCTGCAAGACCCAAAGAGAACCAAAGAA/GTCTATCAGGATGGATTGTTTTTCTAAGGATAAAATGAAGG
2	P	R	chr1:177893173	-	chr1:194431617	-	GGTGTGACATCAGCTCATTGCAACCTCTG/TGTTCAAATGTTAACCAGCTGTAGGGCAGGTTTCCTTAGCTATTCGTGCCTAAATTGTTCTTCTTTATAA
3	R	O	chr1:178550945	-	chr1:177892351	+	TGTACTTGCTCCGCTGGAGTGTCTACCCCGCCCCGGTGCTTCC/TGCATATTGAACCAATATGAGTTGTTCAATACAGAGCCTGAACAGCCAAAGC
4	O	N	chr1:177893115-116	+	chr1:177808890-891	+	TGGTGAACCCCATCTCTATTAATAAATAACAAAATTAGCCAGGCATGGTGAC/C/ACTTAAGGTTCTCCAACAGCTAAACCTGAGTCTCAGTGAGGAGACAA
5	N	K	chr1:177892314	+	chr1:176405780	+	TGTGACCTAAAGATTCAATGCTACTCTACTAATGTCATAAACTACTGT/TGTCAGACTACCA CCAATGTTCCCTTAAGGCCAAGGTCTCTAAGGCAG
6	K	E	chr1:176406041	+	chr1:175269090	-	ATGTGTTTAGAAATGTCACCTGGGAGCTAGGTGCTGGAACAAGGACCTCTCAACTCTG/GGCACAGACTCAACTGCTGTCCCACTGGCAACTGATGCT
7	E	M	chr1:175227097	-	chr1:177808887	-	TGGATCCAAAGGGAAAGCCCTATGTAAGCATGCATCAAAGCGGCCCTGCTTCTCTCTAG /GTGAG/GTATTCTGGAAAGAGGCCAAAGTATCTCTACCC
8	M	U	chr1:177808732	-	chr1:195596333	+	TGGGTGGCTGAGAAGAACCCTAGACAGATAAAG/GATT/TACACCATATCCACCCCTTCATT TTTATAGAATCTCAACTCTTTAAAGAACCTATTTTTAT
9	U	C	chr1:199167388	+	chr1:169786327	-	TTCGGGTATGGCTGACAGATTTTCAGTCATTACATATTA/GAAAAACAAGAGTTTTCTTCTA GTAGAATGTTAAATAATTAATAATA/TCCAGGATTTTTGCATCTCTGAGGTATTTTTCTTA GAAGTTATTGT
10	C	H	chr1:169541777	-	chr1:176206664	+	GCCTCACTTCTGACCATAATCAGAATTCCT/TGGAGAAGTGGTCAGGCAA/GCTGG/AGCTG CTGCCTCTACGCCGGCGCTGGCCCTGGCCGCGCAGCTGT
11	H	L	chr1:176216838	+	chr1:176406063	+	TAACAAGCATAACATAAGCTTAGAGGCTTAACAACATTCTTATATAGTTCATGACTCC/GG CTGAGCTGGTATCCTAGATGCAAGACAAAGTCTTCCCT
12	L	b	chr1:177808739	+	chr2:206883041	-	ACCTTCTCCACACAAATACATGCATCCGACCAGAAGGAATTCTTTATCT/TTA/TCACCATC ACACCAAACAGATAAATGCCACTCTTTTTGTGAAAA
13	b	d	chr2:177091812	-	chr2:207527634	+	CGCTGAGCTCATTATTTTTGGTTGCTTGTTAACTAACGTGGTTATATG/CTAGAATGGGAAT GAAAAAATGGTCATGGGAACAGAGAGAGGTCACAAATC
14	e	e	chr2:230038608-609	+	chr2:229848424-825	+	TTTTTTCCCTTTAGTAAGTACATTAATTAGCAGAGACTTAAAT/ATTCTATAAACAAGAAA ATGAAAAACAAATCACCTCTTGAAATAAAACATGGC
15	A	l	chr1:169541620-623	+	chr1:176216847-850	+	TTATATCTGTGGGGAAGGACTTGTGACTTCGGTCATACTCAA/GCA/ATTTTTACTGAGCTCA GATGGGAAATCCTTTAGTCTCCCTGGGACTCCCTTCTA
16	l	S	chr1:176304474	+	chr1:194431618	+	TTTTACCCCAATCCTATTTCTCTCAGATTTGTGTAATAAGC/T/GGGGCCAGGCATGGTAGC TTACGCCTGTAATCCAGCACTTTGGGAGACTGAGGCA
17	S	D	chr1:194604247	+	chr1:175227003	-	ATTCCGACATTGAAAACGACAATTAGA/AGTGAATGCTTGAGTATTCACTCAAGAGTA/TT GGAGAACAATGGAATTAATCTTTTGTAGCCTAACAAAT

18	D	B	chr1:1697863 86-388	-	chr1:169541761 -163	-	AAAAATCTTCAGAGGAAGACCTATTTTTAATGATTTGCTTTTCT/TC/AGGCAAGCTGCCTCC AGGTCTTTATCAGACAGTGCCCCAGAGGAACACTATC
19	B	F	chr1:1695416 45	-	chr1:176200525	-	TGATCAAATGCACTCTACTTCAGACCCAGTCACAGATCCT/ <u>GGAGAAGCTTGAACACAGCA</u> <u>AAGCAAACAGGATGCAC</u> /TGTGGTTGGGGTGGTACATTTTTTCCCAAATTCAGAATTGG
20	F	T	chr1:1752691 49	-	chr1:194604343	+	TTGAAGGATATTGTGTTACTCATGGAAGTCCAGGAGTGTGAGAATCATCTTAAATGCCACC CACCCAAGAAAA/ <u>CATCATGAATGTTTTATAAATTTATATATATTTTGAGTAAGTAATC/AA</u> <u>AATATACCTC/TTCCTTGGTAAATACATACTCTTTTA/AAAACATACAACCTTCAGTAT/GTTC</u> <u>AAAATTCCTCATATAAC/CCCCGTCTATTACTAGTTTGTTAAGCATTCT</u>
21	T	V	chr1:1955962 68	+	chr1:199167537	+	ACTTTTTATTGATAGCTATTATTTAGAAAAAAGTGTAGTATTGAAA/ <u>AG</u> /TACCATACTTAA GAAACAATAATGGTTCAGTGAATATATATGCATAGA

In green: inserted fragments of unknown origin. The "/" represents the transition from fragment 1 to fragment 2. The sequences in red underlined between two "/" are areas of microhomology which may belong to either fragments. The purple sequence corresponds to the genomic position chr1:169786325-169786360, the yellow sequence corresponds to the genomic position chr1:194460246-194604291, the orange sequence corresponds to the genomic position chrX :132504466-132504491, the blue sequence corresponds to the genomic position chr1:175269125-175269146. Coordinates are in hg38.

Supplemental table 12. Family 5 CNVs identified by WGS

Patient	Coordinates (hg38)	Length (kb)	Rearrangement type	CMA coordinates (hg38)	Fragment
F5-F	Chr1:3457262-3554385	97	Deletion	chr1:3460011-3548495	B
	chr1:4102907-4855050	752	Deletion	chr1:4112842-4850566	D
	chr1:7575555-8300115	724	Gain	chr1:7578163-8295050	F
	chr1:8769956-9012435	242	Deletion	chr1:8770143-9009346	H
	chr1:23734482-23792054	58	Gain	chr1:23736438-23782108	M
	chr1:24506596-24567981	61	Gain	chr1:24514420-24566620	O
	chr1:25848572-25855812	7	Gain	*	Q
	chr1:33483135-33694225	211	Deletion	chr1:33491851-33684918	T
	chr1:37295836-37348069	52	Deletion	*	V
	chr1:47050613-47094887	44	Deletion	*	X
	chr1:76892687-77170084	277	Gain	chr1:76894559-77181400	Z
	chr1:77170536-77185979	15	Gain		AB
	chr1:81497622-81542296	45	Gain	chr1:81497847-81682708	AD
	chr1:81568908-81683491	114	Gain		AF
	chr1:91474906-91567433	92	Gain	chr1:91486520-91549883	AH
	chr1:95212805-95438325	225	Gain	chr1:95221101-95408336	AJ
chr1:98853016-98982297	129	Gain	chr1:98875184-98978528	AL	
chr1:101466242-101578673	112	Gain	*	AN	
F5-C1	chr1:3457262-3554385	97	Deletion	chr1:3460011-3548495	B
	chr1:4102907-4855050	752	Deletion	chr1:4112842-4850566	D
	chr1:7575555-8300115	724	Gain	chr1:7578163-8295050	F
	chr1:8769956-13290514	4520	Deletion	chr1:8770143-13111056	H, I, J, K
	chr1:23734482-23792054	58	Gain	chr1:23736438-23782108	M
	chr1:24506596-24567981	61	Gain	chr1:24514420-24566620	O
	chr1:77170536-77185979	15	Gain	*	AB
	chr1:81497622-81542296	45	Gain	chr1:81497847-81532343	AD
	chr1:95212805-95438325	225	Gain	chr1:95221101-95408336	AJ
	chr1:98853016-98982297	129	Gain	chr1:98875184-98978528	AL
	chr1:101466242-101578673	112	Gain	*	AN

F5-C2			Deletion	chr1:3460011-3548495	B
			Deletion	chr1:4112842-4850566	D
			Gain	chr1:7578163-8295050	F
			Gain	chr1:23736438-23782108	M
			Gain	chr1:24514420-24566620	O
			Gain	chr1:95221101-95408336	AJ
			Gain	chr1:98875184-98978528	AL

*: CVN found in CMA data after reanalysis (CNV not initially identified due to number of oligonucleotides presents in CNV lower to three)

Supplemental table 13: junction sequences identified in Family 5

Family 5: F5-F

Patient	junction	fragment 1 implicated	fragment 2 implicated	fragment 1 position	strand	fragment 2 position	strand	sequence
F5-F	1	A	M	chr1:3457262	+	chr1:23734482	+	AAGTTTGACGGGCAGAATTACCGGAGTACAAGGGCATCATTGTAG CTTGAGTTT/CAGTATCAGATAGATTGTAGTCCTTAATTTTTATTTT ATTTTTTTT
	2	M	F	chr1:23792052(054)	+	chr1:7575497(499)	+	GCAGTCAGTGGCAGCAGCAGCAGGTAGCTCCGTTTAGGGGGGCG TGGAGCACCAGTT/ TG /TACAGCCCTGAGGCAGGAGAAAGGAAAT GAGATCTAAGCCT
	3	F	O	chr1:8300112(115)	+	chr1:24506596(599)	+	GGAGAGGTGAGTTTAGGGCAGATGTAGT/ GGT /CAAAACAGAA GGAATTGGTGACTCTGGGTAAAGGATACAGCATACCCGAGTTCTCT TGTGCTATTCTTA
	4	O	AB	chr1:24567981	+	chr1:77170536	+	GAGTGGGCTTGATTCTACCCATAGAGAGAATCGACCGAGGGTTCC GCCA/GATATATAAGCTCTCTGAGGGCAGGGGCTATATCTGCTCT GTTCAACTG
	5	AB	AJ	chr1:77185978(979)	+	chr1:95212805(806)	+	GGAAACTGAACATTTGACTATGTGTCATCAAGTCACCATGCGACCT GAACCGCC/ T /GACATTGGCTGTGGGTTTGCATAAATAGCTCTTA TTATTTTGAG
	6	AJ	C	chr1:95438402	+	chr1:4102907	-	GTGTTAATGGATGCTGGGAGTTCACACACAATGCCCTTGCTCTC/ GATAGCTCTTATTTTTGAGATTCGTCCCATCAATACCTAATTTAT TGAGAGTT
	7	C	AL	chr1:3554287(289)	-	chr1:98853027(025)	+	CCGCTGGCTTTCCAGTGTTCTGCCATCCCCAGCGACCCTCTCTC GGAGCCGCC/ TG /TTTTATTTTATTATTATTGTTGTTAATCTCTTAC TGTGCC
	8	AL	AN	chr1:98982294(297)	+	chr1:101578670(673)	-	GGTGTAGGTAATAATTTATAGGATGCCAATGGTTTGGACTGAGCT CCTGCAC/ TAG /TATATCAAATGTTTGATTGATTTTTGAAGCCATT AGGGTAAACT
	9	AN	E	chr1:101466242	-	chr1:4855050	+	ACTGGACCTAGTGGCTCTTTTGTAAACAATAAAATACAGCAATAGT G/TATCCATATTTATATGATACAGATGTAGAGATATGAACATGGAT ATGTTAGGT
	10	G	AD	chr1:8769956	+	chr1:81542296	-	AGCCTCCCAGTAGCTGGGATCCGGGCGCACACCACTGCACCTG GCTAAT/GGCTCTGGCATCCTTTACTTTCCTATTCTGCTGCTCT CGACAGAC

	11	AD	L	chr1:81497622	-	chr1:13290574	+	AAATTGAGGTAGAGAAAGGTTAAACAACCTTGCTTAAAGTCACACAGCTCATA/CCCTGTGGCCTCAACTGTTTTCTGATTATGCCTGTTGCTTTGAATG
	12	S	I	33483135	+	9012435	+	ATTTCTTTCATGGATTATGCCTTTGATTTGTATCTGAAAAGTCATTG/GCAGGAGGAAGGGGTTGAGACTTCAGCCTGGGGACCTCTGCTCTTCCCTT
	13	I	K	13011467	+	13027392	+	
	14	K	U	13290574	+	37295836	-	TGTGAAGCTTGTGTCCTGGCAGAGGGAGTGGTTGGCACAGGAGAGGACTCTGCCTTAGGGCTGAGTTATCC/ACTTGGGCAGCTAAAACCTCCACTTACT
	15	U	J	33694225	-	13011467	+	CTCCCCGTTGTCCCCTTCCCAGGCCCTGGCAACCACTAACG/GTAGGGTTTCTTACCTGAGTCCTCAGCAGGGTCTCACCTGGATGCCGTTGTGTAGG
	16	J	W	13027392(393)	+	47050751(752)	-	TTATTTATTCTTTTACTTCCATAAATGTGTATCTAATTCAATCGATT/A/TAAGGAATAAGGTGAATTTTTATTAAGTGAAAAAATCAATAACAATATAGG
	17	W	AH	37348069(072)	-	91567434(437)	-	CCTCCTGGGGAAGATGAAGGATGTGTCCTGGTGCAAATCATCACACAAGC/CCA/GTGCATTCCAGCCAAAAACAAAACAAAAGACCACATGGCAGTGA
	18	AH	R	91474906	-	25855438	+	CTCTAAAAGTAATGAGTGGGAATCGGGGTGAAAGAGAGAGA/AGGCACGAGCCACAGTCCTGTTAATTTTTTTAAACAAACATATATTATTTATAA
	19	R	AF	25855811(812)	+	81683490(491)	-	ACTCAAAGGACTGTATGCCTATTTGATAGCCACTCTAGGAAGAGGCACCTTGTAAGGGAATCTTGTAGT/A/TAACAATGATGATGGACTTTGGAGATTTG
	20	AF	Q	81568908	-	25848572	+	TGACAGTCTACCTAAAATAAGATCTGTATTATT/ACCTGATAGGAGTGTTTTGTTTTTTAACCTGGGAGCTTTGGGCCACACCAAAATAATCTGTG
	21	Q	Z	25855437(438)	+	76892687(688)	+	TATTGCCAGGCTGGTCCAAATTCCTGGCTTCCCAAAGTGCTGAGCTTCCCAAAGTGCTGAGATTAC/A/TTGATATAGAGGAAATTTGGGAAAGAAGAAA
	22	Z	Y	77170084	+	47094750	+	TCCCTACCTGGCATCTGCTTCTACTCTACCCTCTTTCAGTTCAT/GGG/TGTAATCCAGCATTTTGGGAGGCCAAGGCGGGCAGATCACTTGAGGTCAGGAGTTTGGAGACAAGCCTGGCCAACATAGTGAACC
F5-C	1	A	M	chr1:3457262	+	chr1:23734482	+	GGCAGAATTACCGGAGTACAAGGGCATATTGTAGCTTGAGTTT/CAGTATCAGATAGATTGTAGTCTTAATTTTTATTTATTTTTTTGAGACAGAGT
	2	M	F	chr1:23792052(054)	+	chr1:7575497(499)	+	AGTCAGTGGCAGCAGCAGCAGGTAGCTCCGTTTAGGGGGCGTGAGACACAGTT/TG/TACAGCCCTGAGGCAGGAGAAAAGGAAATGAGATCAAGCCTGC

3	F	O	chr1:8300112(115)	+	chr1:24506596(599)	+	CACACAATCCTCTGAGCTCTGCTTATCCCCTGGAGAGGTGAGTTA GGGCAGATGTAGTG/ <u>GGT</u> /CAAAACAGAAGGAATTGGTGACTCTG GGTAAAGGATA
4	O	AB	chr1:24567981	+	chr1:77170536	+	ACCCATAGAGAGAATCGACCGAGGGTTCGCGAGAT/ATATAAGC TCTCTGAGGGCAGGGCTATATCTGTCCTGTTCACTGTATCAC TCATACTACA
5	AB	AJ	chr1:77185978(979)	+	chr1:95212805(806)	+	TGTCATCAAGTCACCATGCGACCTGAACCGCC/ <u>T</u> /GACATTGGCTG TGGGTTTGCATAAATAGCTCTTATTATTTTGGATACATTCCATCA ATACTTAGT
6	AJ	C	chr1:95438402	+	chr1:4102907	-	CCACTTCTCCCCTTTAGCTGCTTGTGTTAATGGATGCTGGGAGT TCACACACAATGCCCTTGCTCTC/GATAGCTCTTATTATTTGAGAT TCGTCCCATCAATACCTAATTTATTGAGAGTTTTAGCATGAAGGG TTGTTGAATT
7	C	AL	chr1:3554287	-	chr1:98853027	+	TAGACCCCGCTGGCTTTCCAGTGTTCTGCCATCCCCAGCGACC CTCTCTCGGAGCCGCC/ <u>TG</u> /TTTTATTTATTATTATTGTTGTTAAT CTCTTACT
8	AL	AN	chr1:98982294(297)	+	chr1:101578670(673)	-	TTTACTGGCGGTGTTAGGTAAAATTTATAGGATGCCAATGGTTTG GACTGAGCTCCTGCAC/ <u>TAG</u> /TATATCAAATGTTTGATTGATTTT GAAGCCATTA
9	AN	E	chr1:101466242	-	chr1:4855050	+	CTTTTGTAACAAATAAATACAGCAATAGTG/TATCCATATTTATAT GATACAGATGTAGAGATATGAACATGGATATGTTAGGTTGGTGCA AAACTACTT
10	G	AD	chr1:8769956	+	chr1:81542296	-	CTCCCGAGTAGCTGGGATTCCGGGCGCACACCACTGCACCTGGCT AAT/GGCCTCTGGCATCCTTTTACTTTCTCATTCTGCTGCTCTCG ACAGACTTG
11	AD	L	chr1:81497622	-	chr1:13290574	+	GAGAAAGGTTAAACAACCTTGCTTAAAGTCACACAGCTCATA/CCCT GTGGCCTCAACTGTTTTCTGATTATGCCTGTTGTCTTTGAATGTCA ACAAAAGT

The "/" represents the transition from fragment 1 to fragment 2. The sequences in red underlined between two "/" are areas of microhomology which may belong to either fragment. Coordinates are in hg38.

Supplemental table 14: CNV identified by CMA in family 6

Patient	Length (kb)	Rearrangement type	CMA coordinates (hg38)	Fragment
F6-C	1,3 Mb	Deletion	chr1:73960120-75313006	C
	3,4 Mb	Gain	chr1:169007884-172463708	E
	1 Mb	Gain	chr1:104033454-105110730	B

Supplemental table 15: junction sequences identified in F6-F

junction	fragment 1 implicated	fragment 2 implicated	fragment 1 position	strand	fragment 2 position	strand	sequence
1	A	d	chr1:73928858	+	chr1:150114108	-	AATTAGATGCCAGGGTCAAACTGCAAAATATTAATAA TTCACAAAACAG/ AGA /GGGCTCAGCACCATTGGGAAT GCTCTGAGTCTGAGTGAACATGC
2	d	B	chr3:105976284	-	chr1:73928859	+	CTCAGTATTTGACTTAAATACTCCCGAC/ AGGAGTATTTT AAGTGGCATAAAAATAGCATAATGC ACTATTAAGATAATA GTGGCATT TT(TG) TATGGCAACCAATGATTGATGATCAA ATGCTGATCATCAAAT /ATAGTCTCCAGAATGAAGATGT GAAAATTATTGA
3	B	D	chr1:73971357	+	chr1:168986010	-	ACTACTAGTTTTTTAAAGGATACAACCTCAGGAACAGCCA AATGGAAGAGCTGTGTAGGG/ CA /TGAGGTCAGGAAAGT GAAAGTTCTCCCTAAATATTTAA
4	D	b	chr1:75299367	-	chr3:105842466	-	CAATAAGAGAAATACTTTGAGAACACCAACTGAAAGTCA TTTCCAAAAGATA/ C /TCTAAATTGTCAAAGGAGATAGCC AGATATCTCTTTTGGAGAGATAA
5	b	E	chr3:104829919	-	chr1:172466533	-	TGTGTGTATAGTGGATAATTTGCTGAGATAAATTACTAGT TT/ G /CAGGTCAAAGGAAACTCAGAATGAGTGTGTGGC TGGCACAGATAACTATGAAATCA
6	E	c	chr1:168986022	-	chr3:105842479	+	AGATGAAGAATCTTTCTCTAAGAGAAGTCCCT/ACACCG TTTCTCCCTTGCCCTTCTTACTATGGAGGCTGAAATTTCA AAGTTTCAATATACCGTAC
7	c	e	chr3:105975668	+	chr3:150114179	+	CCTTGATCTTGGACTTCCAGCCCCATAACTGTGAGAAA TAAATTTCTGTT/ TTGGTCTTCTC /CTTCTACTCCCTCAC AGTGTTCCTGGGATCTCCT
8	a	C	chr3:104829879	+	chr1:75299323	-	TTATTATTCTATAATTAATTAATAGTCTAAAGGATTCTG TGATAATAGTGGCAT/ AAAATGCACTATTTTGCCATAAA AATGC /CAAAAATTAAGACAGGAAAATAGGGAGGGAG AGAAAAATAAAGG
9	C	F	chr1:73971383	-	chr1:172467147	+	TGTGCTTGGAGGGTGGTGTGCCTGGAGAGGCATGGAA/ T TGATACTATAAA /CCTACTGGATTATAGAGCTTAGACATG TCCATACTGTGGTACTTCTGCT
10	insertion		chr3:126559168	-	chr1:16352402	-	TGGTGTCCACTCCTCCACAGAAACAGCTTTTCTGCAGCCT CAGTGTCTGTACCTATTTCTGGGCGTCTCTTCTG/GGCGTC TCGGGGTCCCAGCTGTG

11	insertion	chr1:16352221	-	chr3:126559162	-	TGGAGCGGGCTGGGCCAGGTGAGGCCTCCTCCTCCCC GGGCCGAGAGGG/AGAACCTGAACAGGGAGAGGTCTC TGCTATGAAGAGTGCAGATGTGGA
----	-----------	---------------	---	----------------	---	--

In green: inserted fragments of unknown origin. The "/" represents the transition from fragment 1 to fragment 2. The sequences in red underlined between two "/" are areas of microhomology which may belong to either fragment. The purple sequence corresponds to the genomic position chr3:104829866-104829914. Coordinates are in hg38.

DISCUSSION

Nous avons vu en introduction que la chromatine n'est pas organisée de manière aléatoire dans le noyau interphasique. Elle est structurée de manière hiérarchique en territoires chromosomiques, compartiments chromosomiques, domaines associés topologiquement et boucles de chromatine. L'expression génique et les interactions chromatiniennes sont liées à cette architecture. Lors de la méiose, la chromatine atteint son niveau de compaction maximal et forme des chromosomes qui s'apparient entre homologues. Les chromosomes vont subir des cassures doubles brins, et vont se réparer par recombinaison homologue, en prenant le chromosome homologue comme matrice de réparation. Cela va former des jonctions de Holliday, qui en se résolvant peuvent résulter en des *crossing-over*. Cela aboutit alors à l'échange de matériel entre les deux chromosomes homologues.

Nous nous sommes alors intéressés à l'impact des remaniements chromosomiques sur le noyau, tant au moment de l'interphase qu'au moment de la méiose.

Nous avons vu en introduction que l'expression des gènes est régulée au niveau du génome lui-même, par la présence de séquences promotrices.

Nous avons étudié dans une première partie l'impact de remaniements chromosomiques en fonction de leur position par rapport au site d'initiation de la transcription dans un gène avec des nombreuses isoformes, le gène *TCF4*. Les remaniements chromosomiques impactent l'expression des gènes, en altérant les interactions entre les séquences régulatrices et le gène. Cette altération est, dans le cas du gène *TCF4*, différente en fonction de la position des remaniements chromosomiques. Nous avons pu observer un gradient de sévérité phénotypique corrélé au nombre de sites d'initiation de la transcription touchés.

L'expression des gènes est également régulée au niveau épigénétique par la modification des histones, et le positionnement nucléaire, régulant les interactions chromosomiques et la proximité de certaines régions génomiques avec des corps nucléaires.

Pour explorer cela nous avons entrepris, dans une deuxième partie, d'explorer les conséquences des remaniements chromosomiques sur l'architecture du génome. Nous avons présenté les résultats de l'étude en 3D-FISH de huit patients présentant une déficience intellectuelle et/ou une malformation congénitale. Cette partie nous a permis de conclure que les remaniements chromosomiques n'impactent pas drastiquement la position nucléaire des loci impliqués. Des légères relocalisations sont néanmoins observées chez la moitié des patients. La relocalisation périphérique du dérivé X de la patiente LD/0108 par rapport au chromosome 1 normal, peut participer au phénotype de la patiente

par diminution de l'expression des gènes situés sur le fragment terminal du bras court du chromosome 1.

Afin de compléter cette analyse des conséquences sur l'organisation nucléaire, nous allons réaliser une étude d'expression des gènes par RNA-seq, une étude de la modification des marques d'histones par Chip-seq et une analyse des interactions chromatinienne globales du génome par Hi-C afin d'évaluer les modifications des interactions à courte ou longue distance du point de cassure. L'ensemble de ces données nous permettra d'avoir une vision globale de l'impact des remaniements chromosomiques sur le noyau interphasique.

Lors de la méiose, les remaniements chromosomiques perturbent l'appariement des chromosomes homologues. Dans le cas de remaniements simples, les chromosomes réussissent à s'apparier par la formation de tétravalent dans le cas des translocations réciproques, la formation de trivalent pour les translocations robertsoniennes et la formation de boucle en cas d'inversion. Les risques associés sont alors la formation de gamètes déséquilibrés par mauvaise ségrégation des chromosomes ou par la formation d'un *crossing-over* dans la boucle d'inversion. Nous avons alors exploré, dans la troisième partie de ce travail, l'impact des remaniements complexes sur la méiose, par la caractérisation de chromoanagenesis chez six familles avec une transmission déséquilibrée à la descendance. La caractérisation des remaniements chromosomiques par caryotype, FISH, ACPA et séquençage haut débit du génome entier (WGS) nous a permis de reconstituer les chromosomes dérivés des parents et des enfants. Cela nous a permis d'élucider la région dans laquelle a eu lieu le crossing-over qui a mené à la formation du chromosome dérivé remanié de l'enfant. Nous avons pu émettre des hypothèses d'appariement des chromosomes parentaux au moment de la méiose, dans lesquelles des segments d'asynapsis sont présents entre le chromosome normal et le chromosome dérivé. Dans certains cas, la formation de figures complexes, avec des boucles, permet de réduire la longueur des fragments d'asynapsis. La méiose de ces parents, et en particulier celle des trois pères n'a pas été arrêtée malgré la présence de fragments d'asynapsis.

Ce travail nous montre l'importance d'explorer le génome sur le plan fonctionnel. En effet, le WGS permet de mettre en évidence un grand nombre de variations, mais l'ensemble des informations nécessaires à l'expression des gènes n'est pas contenu uniquement dans la séquence génomique. Il existe d'autres niveaux d'information comme la structure de la chromatine et l'architecture tridimensionnelle. La modification des marques d'histones et la méthylation de l'ADN, font partie des mécanismes épigénétiques qui aident à compacter la chromatine et structurer le génome en domaines chromatinienne.

Il est alors important d'explorer le génome sur le plan fonctionnel et d'explorer les interactions longues distances afin de comprendre les corrélations génotype-phénotype. L'équipe de Mundlos a démontré l'importance de l'architecture du génome dans l'expression des gènes en montrant que l'interruption des TADs peut perturber les régulations longues distances et entraîner l'apparition de phénotypes [146].

Cependant, les études fonctionnelles peuvent s'avérer compliquées. En effet bien que le génome soit identique dans chacune des cellules de l'organisme, les gènes ont un niveau d'expression qui peut être différent au cours du temps, par exemple propre à un stade du développement, ou dans l'espace avec des gènes exprimés différemment en fonction du type cellulaire. L'expression de certains gènes peut également être modifiée en fonction de certains stimuli, comme le stress. Les analyses d'expression de gènes, comme la RT-qPCR ou le RNA-seq, permettent donc de quantifier des produits d'expression dans un tissu donné, à un moment donné du développement. Dans le cas de pathologies du neurodéveloppement, le tissu le plus approprié pour réaliser les études d'expression serait le tissu nerveux. Or celui-ci n'est pas accessible. Il pourrait être intéressant de créer un modèle cellulaire pour explorer les remaniements. Nous pourrions utiliser des cellules IPs qui pourraient être reprogrammées en tissu nerveux. Grâce à la technologie CRISPR/Cas9, les remaniements chromosomiques pourraient être recréés dans ces cellules. Cela permettrait d'étudier ces remaniements uniques, en s'affranchissant des variations interindividuelles. Les différences observées ne seraient expliquées que par le remaniement.

BIBLIOGRAPHIE

- 1 Olins AL, Olins DE. Spheroid chromatin units (v bodies). *Science*. 1974 Jan;183(4122):330–2.
- 2 Olins DE, Olins AL. Chromatin history: our view from the bridge. *Nat Rev Mol Cell Biol*. 2003 Oct;4(10):809–14.
- 3 Brown CE, Lechner T, Howe L, Workman JL. The many HATs of transcription coactivators. *Trends Biochem Sci*. 2000 Jan;25(1):15–9.
- 4 Li B, Carey M, Workman JL. The role of chromatin during transcription. *Cell*. 2007 Feb;128(4):707–19.
- 5 Izzo A, Schneider R. Chatting histone modifications in mammals. *Brief Funct Genomics*. 2010 Dec;9(5–6):429–43.
- 6 Sawan C, Herceg Z. Histone Modifications and Cancer. *Advances in Genetics*. Elsevier; 2010; pp 57–85.
- 7 Nowak SJ, Corces VG. Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet*. 2004 Apr;20(4):214–20.
- 8 Passarge E. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am J Hum Genet*. 1979 Mar;31(2):106–15.
- 9 Bhat J, Kabelitz D. An introduction to immunology and epigenetics. *Epigenetics of the Immune System*. Elsevier; 2020; pp 1–23.
- 10 Kempfer R, Pombo A. Methods for mapping 3D chromosome architecture. *Nat Rev Genet*. 2020 Apr;21(4):207–26.
- 11 Umlauf D. Le génome intime... et en trois dimensions. *Med Sci (Paris)*. 2015 Mar;31(3):304–11.
- 12 Croft JA, Bridger JM, Boyle S, Perry P, Teague P, Bickmore WA. Differences in the localization and morphology of chromosomes in the human nucleus. *J Cell Biol*. 1999 Jun;145(6):1119–31.
- 13 Cremer M, Grasser F, Lanctôt C, Müller S, Neusser M, Zinner R, et al. Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. *Methods Mol Biol*. 2008;463:205–39.
- 14 Xie SQ, Lavitas L-M, Pombo A. CryoFISH: fluorescence in situ hybridization on ultrathin cryosections. *Methods Mol Biol*. 2010;659:219–30.
- 15 Chen B, Hu J, Almeida R, Liu H, Balakrishnan S, Covill-Cooke C, et al. Expanding the CRISPR imaging toolset with *Staphylococcus aureus* Cas9 for simultaneous imaging of multiple genomic loci. *Nucleic Acids Res*. 2016 May;44(8):e75.
- 16 Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science*. 2002 Feb;295(5558):1306–11.

- 17 Zhao Z, Tavoosidana G, Sjölander M, Göndör A, Mariano P, Wang S, et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet.* 2006 Nov;38(11):1341–7.
- 18 Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nat Genet.* 2006 Nov;38(11):1348–54.
- 19 Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, et al. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* 2006 Oct;16(10):1299–309.
- 20 Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 2009 Oct;326(5950):289–93.
- 21 Stack SM, Brown DB, Dewey WC. Visualization of interphase chromosomes. *Journal of Cell Science.* 1977 Aug;26(1):281–99.
- 22 Zorn C, Cremer C, Cremer T, Zimmer J. Unscheduled DNA synthesis after partial UV irradiation of the cell nucleus: Distribution in interphase and metaphase. *Experimental Cell Research.* 1979 Nov;124(1):111–9.
- 23 Cremer T, Peterson SP, Cremer C, Berns MW. Laser microirradiation of Chinese hamster cells at wavelength 365 nm: effects of psoralen and caffeine. *Radiat Res.* 1981 Mar;85(3):529–43.
- 24 Meaburn KJ, Misteli T. Chromosome territories. *Nature.* 2007 Jan;445(7126):379–81.
- 25 Lichter P, Cremer T, Borden J, Manuelidis L, Ward DC. Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. *Hum Genet.* 1988 Nov;80(3):224–34.
- 26 Manuelidis L. Individual interphase chromosome domains revealed by in situ hybridization. *Hum Genet.* 1985;71(4):288–93.
- 27 Schardin M, Cremer T, Hager HD, Lang M. Specific staining of human chromosomes in Chinese hamster x man hybrid cell lines demonstrates interphase chromosome territories. *Hum Genet.* 1985;71(4):281–7.
- 28 Zink D, Cremer T. Cell nucleus: chromosome dynamics in nuclei of living cells. *Curr Biol.* 1998 Apr;8(9):R321–324.
- 29 Visser AE, Aten JA. Chromosomes as well as chromosomal subdomains constitute distinct units in interphase nuclei. *J Cell Sci.* 1999 Oct;112 (Pt 19):3353–60.
- 30 Zink D, Cremer T, Saffrich R, Fischer R, Trendelenburg MF, Ansorge W, et al. Structure and dynamics of human interphase chromosome territories in vivo. *Hum Genet.* 1998 Feb;102(2):241–51.
- 31 Habermann FA, Cremer M, Walter J, Kreth G, von Hase J, Bauer K, et al. Arrangements of macro- and microchromosomes in chicken cells. *Chromosome Res.* 2001 Oct;9(7):569–84.

- 32 Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, et al. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol.* 2005 May;3(5):e157.
- 33 Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum Mol Genet.* 2001 Feb;10(3):211–9.
- 34 Tanabe H, Müller S, Neusser M, Hase J von, Calcagno E, Cremer M, et al. Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *PNAS.* 2002 Apr;99(7):4424–9.
- 35 Cremer M, Küpper K, Wagler B, Wizelman L, Hase J v., Weiland Y, et al. Inheritance of gene density–related higher order chromatin arrangements in normal and tumor cell nuclei. *Journal of Cell Biology.* 2003 Sep;162(5):809–20.
- 36 Parada LA, Sotiriou S, Misteli T. Spatial genome organization. *Experimental Cell Research.* 2004 May;296(1):64–70.
- 37 Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet.* 2001 Apr;2(4):292–301.
- 38 Cremer T, Kreth G, Koester H, Fink RH, Heintzmann R, Cremer M, et al. Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. *Crit Rev Eukaryot Gene Expr.* 2000;10(2):179–212.
- 39 Visser AE, Jaunin F, Fakan S, Aten JA. High resolution analysis of interphase chromosome domains. *J Cell Sci.* 2000 Jul;113 (Pt 14):2585–93.
- 40 Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature.* 2015 Feb;518(7539):331–6.
- 41 Marchal C, Sima J, Gilbert DM. Control of DNA replication timing in the 3D genome. *Nat Rev Mol Cell Biol.* 2019 Dec;20(12):721–37.
- 42 Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell.* 2014 Dec;159(7):1665–80.
- 43 Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 2012 Apr;485(7398):376–80.
- 44 Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, Jewett AI, et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A.* 2015 Nov;112(47):E6456–65.
- 45 Ganji M, Shaltiel IA, Bisht S, Kim E, Kalichava A, Haering CH, et al. Real-time imaging of DNA loop extrusion by condensin. *Science.* 2018 Apr;360(6384):102–5.
- 46 Wutz G, Várnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* 2017 Dec;36(24):3573–99.

- 47 Dolgin E. DNA's secret weapon against knots and tangles. *Nature*. 2017 Apr;544(7650):284–6.
- 48 Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, et al. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell*. 2015 May;161(5):1012–25.
- 49 Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell*. 2015 Aug;162(4):900–10.
- 50 Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell*. 2017 Oct;171(3):557-572.e24.
- 51 Zufferey M, Tavernari D, Oricchio E, Ciriello G. Comparison of computational methods for the identification of topologically associating domains. *Genome Biology*. 2018 Dec;19(1):217.
- 52 Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell*. 2002 Dec;10(6):1453–65.
- 53 Palstra R-J, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W. The β -globin nuclear compartment in development and erythroid differentiation. *Nat Genet*. 2003 Oct;35(2):190–4.
- 54 Lanzuolo C, Roure V, Dekker J, Bantignies F, Orlando V. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol*. 2007 Oct;9(10):1167–74.
- 55 Bantignies F, Roure V, Comet I, Leblanc B, Schuettengruber B, Bonnet J, et al. Polycomb-Dependent Regulatory Contacts between Distant Hox Loci in *Drosophila*. *Cell*. 2011 Jan;144(2):214–26.
- 56 De Magistris P, Antonin W. The Dynamic Nature of the Nuclear Envelope. *Current Biology*. 2018 Apr;28(8):R487–97.
- 57 Zhao Y, Abreu E, Kim J, Stadler G, Eskiocak U, Terns MP, et al. Processive and distributive extension of human telomeres by telomerase under homeostatic and nonequilibrium conditions. *Mol Cell*. 2011 May;42(3):297–307.
- 58 Sutherland H, Bickmore WA. Transcription factories: gene expression in unions? *Nat Rev Genet*. 2009 Jul;10(7):457–66.
- 59 Spector DL, Lamond AI. Nuclear Speckles. *Cold Spring Harb Perspect Biol*. 2011 Feb;3(2):a000646.
- 60 Chen Y, Zhang Y, Wang Y, Zhang L, Brinkman EK, Adam SA, et al. Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *Journal of Cell Biology*. 2018 Aug;217(11):4025–48.
- 61 Chen Y, Belmont AS. Genome Organization around Nuclear Speckles. *Curr Opin Genet Dev*. 2019 Apr;55:91–9.
- 62 Lallemand-Breitenbach V, de Thé H. PML nuclear bodies. *Cold Spring Harb Perspect Biol*. 2010 May;2(5):a000661.

- 63 Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*. 2008 Jun;453(7197):948–51.
- 64 Meuleman W, Peric-Hupkes D, Kind J, Beaudry J-B, Pagie L, Kellis M, et al. Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res*. 2013 Feb;23(2):270–80.
- 65 Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science*. 2002 Apr;296(5565):158–62.
- 66 Bickmore WA. The spatial organization of the human genome. *Annu Rev Genomics Hum Genet*. 2013;14:67–84.
- 67 Cremer T, Cremer M. Chromosome territories. *Cold Spring Harb Perspect Biol*. 2010 Mar;2(3):a003889.
- 68 Cremer M, von Hase J, Volm T, Brero A, Kreth G, Walter J, et al. Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res*. 2001;9(7):541–67.
- 69 Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum Mol Genet*. 2001 Feb;10(3):211–9.
- 70 Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, et al. Three-Dimensional Maps of All Chromosomes in Human Male Fibroblast Nuclei and Prometaphase Rosettes. *PLOS Biology*. 2005 Apr;3(5):e157.
- 71 Pliss A, Fritz AJ, Stojkovic B, Ding H, Mukherjee L, Bhattacharya S, et al. Non-Random Patterns in the Distribution of NOR-Bearing Chromosome Territories in Human Fibroblasts: A Network Model of Interactions. *J Cell Physiol*. 2015 Feb;230(2):427–39.
- 72 Kumaran RI, Spector DL. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J Cell Biol*. 2008 Jan;180(1):51–65.
- 73 Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P, et al. Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells. *PLOS Genetics*. 2008 Mar;4(3):e1000039.
- 74 Reddy KL, Zullo JM, Bertolino E, Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature*. 2008 Mar;452(7184):243–7.
- 75 Pellestor F, Anahory T, Lefort G, Puechberty J, Liehr T, Hédon B, et al. Complex chromosomal rearrangements: origin and meiotic behavior. *Hum Reprod Update*. 2011 Jul;17(4):476–94.
- 76 Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, et al. Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. *Cell*. 2011 Jan;144(1):27–40.
- 77 Nielsen J, Wohlert M. Chromosome abnormalities found among 34,910 newborn children: results from a 13-year incidence study in Arhus, Denmark. *Hum Genet*. 1991 May;87(1):81–3.

- 78 Warburton D. De novo balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. *Am J Hum Genet.* 1991 Nov;49(5):995–1013.
- 79 Madan K, Nieuwint AW, van Bever Y. Recombination in a balanced complex translocation of a mother leading to a balanced reciprocal translocation in the child. Review of 60 cases of balanced complex translocations. *Hum Genet.* 1997 Jun;99(6):806–15.
- 80 Halgren C, Nielsen NM, Nazaryan-Petersen L, Silaharoglu A, Collins RL, Lowther C, et al. Risks and Recommendations in Prenatally Detected De Novo Balanced Chromosomal Rearrangements from Assessment of Long-Term Outcomes. *Am J Hum Genet.* 2018 07;102(6):1090–103.
- 81 Schluth-Bolard C, Delobel B, Sanlaville D, Boute O, Cuisset J-M, Sukno S, et al. Cryptic genomic imbalances in de novo and inherited apparently balanced chromosomal rearrangements: Array CGH study of 47 unrelated cases. *European Journal of Medical Genetics.* 2009 Sep;52(5):291–6.
- 82 Schluth-Bolard C, Labalme A, Cordier M-P, Till M, Nadeau G, Tevissen H, et al. Breakpoint mapping by next generation sequencing reveals causative gene disruption in patients carrying apparently balanced chromosome rearrangements with intellectual deficiency and/or congenital malformations. *J Med Genet.* 2013 Mar;50(3):144–50.
- 83 Frühmesser A, Blake J, Haberlandt E, Baying B, Raeder B, Runz H, et al. Disruption of EXOC6B in a patient with developmental delay, epilepsy, and a de novo balanced t(2;8) translocation. *Eur J Hum Genet.* 2013 Oct;21(10):1177–80.
- 84 Kalscheuer VM, Tao J, Donnelly A, Hollway G, Schwinger E, Kübart S, et al. Disruption of the Serine/Threonine Kinase 9 Gene Causes Severe X-Linked Infantile Spasms and Mental Retardation. *Am J Hum Genet.* 2003 Jun;72(6):1401–11.
- 85 Klein-Tasman BP, van der Fluit F, Mervis CB. Autism Spectrum Symptomatology in Children with Williams Syndrome Who Have Phrase Speech or Fluent Language. *J Autism Dev Disord.* 2018 Sep;48(9):3037–50.
- 86 Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature.* 2009 Apr;458(7239):719–24.
- 87 Liu P, Erez A, Nagamani SCS, Dhar SU, Kołodziejska KE, Dharmadhikari AV, et al. Chromosome Catastrophes Involve Replication Mechanisms Generating Complex Genomic Rearrangements. *Cell.* 2011 Sep;146(6):889–903.
- 88 Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, et al. Punctuated Evolution of Prostate Cancer Genomes. *Cell.* 2013 Apr;153(3):666–77.
- 89 Kloosterman WP, Guryev V, van Roosmalen M, Duran KJ, de Bruijn E, Bakker SCM, et al. Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. *Hum Mol Genet.* 2011 May;20(10):1916–24.
- 90 Pellestor F. Chromoanagenesis: cataclysms behind complex chromosomal rearrangements. *Mol Cytogenet.* 2019 Dec;12(1):6.
- 91 Liu P, Carvalho CMB, Hastings PJ, Lupski JR. Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev.* 2012 Jun;22(3):211–20.

- 92 Pellestor F. Chromothripsis: how does such a catastrophic event impact human reproduction? *Hum Reprod.* 2014 Mar;29(3):388–93.
- 93 Tubio JMC, Estivill X. Cancer: When catastrophe strikes a cell. *Nature.* 2011 Feb;470(7335):476–7.
- 94 Ernst A, Jones DTW, Maass KK, Rode A, Deeg KI, Jebaraj BMC, et al. Telomere dysfunction and chromothripsis. *Int J Cancer.* 2016 Jun;138(12):2905–14.
- 95 Crasta K, Ganem NJ, Dagher R, Lantermann AB, Ivanova EV, Pan Y, et al. DNA breaks and chromosome pulverization from errors in mitosis. *Nature.* 2012 Jan;482(7383):53–8.
- 96 Chatron N, Giannuzzi G, Rollat-Farnier P-A, Diguët F, Porcu E, Yammine T, et al. The enrichment of breakpoints in late-replicating chromatin provides novel insights into chromoanagenesis mechanisms. *bioRxiv.* 2020 Jul;2020.07.17.206771.
- 97 Ly P, Teitz LS, Kim DH, Shoshani O, Skaletsky H, Fachinetti D, et al. Selective Y centromere inactivation triggers chromosome shattering in micronuclei and repair by non-homologous end joining. *Nature Cell Biology.* 2017 Jan;19(1):68–75.
- 98 Hatch EM, Fischer AH, Deerinck TJ, Hetzer MW. Catastrophic Nuclear Envelope Collapse in Cancer Cell Micronuclei. *Cell.* 2013 Jul;154(1):47–60.
- 99 Pantelias A, Karachristou I, Georgakilas AG, Terzoudi GI. Interphase Cytogenetic Analysis of Micronucleated and Multinucleated Cells Supports the Premature Chromosome Condensation Hypothesis as the Mechanistic Origin of Chromothripsis. *Cancers (Basel).* 2019 Aug;11(8). DOI: 10.3390/cancers11081123
- 100 Goedecke W, Eijpe M, Offenberg HH, Aalderen M van, Heyting C. Mre11 and Ku70 interact in somatic cells, but are differentially expressed in early meiosis. *Nature Genetics.* 1999 Oct;23(2):194–8.
- 101 Currall BB, Chiang C, Talkowski ME, Morton CC. Mechanisms for Structural Variation in the Human Genome. *Curr Genet Med Rep.* 2013 Jun;1(2):81–90.
- 102 Kim S, Peterson SE, Jasin M, Keeney S. Mechanisms of germ line genome instability. *Semin Cell Dev Biol.* 2016;54:177–87.
- 103 Babcock M, Pavlicek A, Spiteri E, Kashork CD, Ioshikhes I, Shaffer LG, et al. Shuffling of Genes Within Low-Copy Repeats on 22q11 (LCR22) by Alu-Mediated Recombination Events During Evolution. *Genome Res.* 2003 Dec;13(12):2519–32.
- 104 Sasaki M, Lange J, Keeney S. Genome destabilization by homologous recombination in the germline. *Nat Rev Mol Cell Biol.* 2010 Mar;11(3):182–95.
- 105 Goldman AS, Lichten M. Restriction of ectopic recombination by interhomolog interactions during *Saccharomyces cerevisiae* meiosis. *Proc Natl Acad Sci USA.* 2000 Aug;97(17):9537–42.
- 106 Bhargava R, Onyango DO, Stark JM. Regulation of Single-Strand Annealing and its Role in Genome Maintenance. *Trends in Genetics.* 2016 Sep;32(9):566–75.
- 107 Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol.* 2017;18(8):495–506.

- 108 Ghezraoui H, Piganeau M, Renouf B, Renaud J-B, Sallmyr A, Ruis B, et al. Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. *Mol Cell*. 2014 Sep;55(6):829–42.
- 109 Lieber MR. Mechanisms of Human Lymphoid Chromosomal Translocations. *Nat Rev Cancer*. 2016 May;16(6):387–98.
- 110 McVey M, Lee SE. MMEJ repair of double-strand breaks (director’s cut): deleted sequences and alternative endings. *Trends Genet*. 2008 Nov;24(11):529–38.
- 111 Yu X, Gabriel A. Ku-dependent and Ku-independent end-joining pathways lead to chromosomal rearrangements during double-strand break repair in *Saccharomyces cerevisiae*. *Genetics*. 2003 Mar;163(3):843–56.
- 112 Ottaviani D, LeCain M, Sheer D. The role of microhomology in genomic structural variation. *Trends Genet*. 2014 Mar;30(3):85–94.
- 113 Hunter N. Meiotic recombination: The essence of heredity. *Cold Spring Harbor Perspectives in Biology*. 2015 Dec;7(12):a016618.
- 114 Handel MA, Schimenti JC. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. *Nature Reviews Genetics*. 2010 Feb;11(2):124–36.
- 115 MacGregor IA, Adams IR, Gilbert N. Large-scale chromatin organisation in interphase, mitosis and meiosis. *Biochem J*. 2019 Aug;476(15):2141–56.
- 116 Brown MS, Bishop DK. DNA Strand Exchange and RecA Homologs in Meiosis. *Cold Spring Harb Perspect Biol*. 2015 Jan;7(1). DOI: 10.1101/cshperspect.a016659
- 117 Jasin M, Rothstein R. Repair of Strand Breaks by Homologous Recombination. *Cold Spring Harb Perspect Biol*. 2013 Nov;5(11). DOI: 10.1101/cshperspect.a012740
- 118 Miura T, Yamana Y, Usui T, Ogawa HI, Yamamoto M-T, Kusano K. Homologous Recombination via Synthesis-Dependent Strand Annealing in Yeast Requires the Irc20 and Srs2 DNA Helicases. *Genetics*. 2012 May;191(1):65–78.
- 119 Zhang B, Tang Z, Li L, Lu L-Y. NBS1 is required for SPO11-linked DNA double-strand break repair in male meiosis. *Cell Death & Differentiation*. juill 2020;27(7):2176–90.
- 120 Hillers KJ, Villeneuve AM. Chromosome-Wide Control of Meiotic Crossing over in *C. elegans*. *Current Biology*. 2003 Sep;13(18):1641–7.
- 121 Tock AJ, Henderson IR. Hotspots for Initiation of Meiotic Recombination. *Front Genet*. 2018 Nov;9. DOI: 10.3389/fgene.2018.00521
- 122 Borde V, Robine N, Lin W, Bonfils S, Géli V, Nicolas A. Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *EMBO J*. 2009 Jan;28(2):99–111.
- 123 Myers S, Freeman C, Auton A, Donnelly P, McVean G. A common sequence motif associated with recombination hot spots and genome instability in humans. *Nature Genetics*. 2008 Sep;40(9):1124–9.

- 124 Lane S, Kauppi L. Meiotic spindle assembly checkpoint and aneuploidy in males versus females. *Cell Mol Life Sci*. 2019 Mar;76(6):1135–50.
- 125 Batista DAS, Pai GS, Stetten G. Molecular analysis of a complex chromosomal rearrangement and a review of familial cases. *American Journal of Medical Genetics*. 1994;53(3):255–63.
- 126 Saadallah N, Hulten M. A complex three breakpoint translocation involving chromosomes 2, 4, and 9 identified by meiotic investigations of a human male ascertained for subfertility. *Hum Genet*. 1985 Dec;71(4):312–20.
- 127 Kloosterman WP, Tavakoli-Yaraki M, van Roosmalen MJ, van Binsbergen E, Renkens I, Duran K, et al. Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms. *Cell Rep*. 2012 Jun;1(6):648–55.
- 128 Gu H, Jiang J, Li J, Zhang Y, Dong X, Huang Y, et al. A Familial Cri-du-Chat/5p Deletion Syndrome Resulted from Rare Maternal Complex Chromosomal Rearrangements (CCRs) and/or Possible Chromosome 5p Chromothripsis. *PLoS One*. 2013 Oct;8(10). DOI: 10.1371/journal.pone.0076985
- 129 Bertelsen B, Nazaryan-Petersen L, Sun W, Mehrjouy MM, Xie G, Chen W, et al. A germline chromothripsis event stably segregating in 11 individuals through three generations. *Genet Med*. 2016 May;18(5):494–500.
- 130 de Pagter MS, van Roosmalen MJ, Baas AF, Renkens I, Duran KJ, van Binsbergen E, et al. Chromothripsis in healthy individuals affects multiple protein-coding genes and can result in severe congenital abnormalities in offspring. *Am J Hum Genet*. 2015 Apr;96(4):651–6.
- 131 Weckselblatt B, Hermetz KE, Rudd MK. Unbalanced translocations arise from diverse mutational mechanisms including chromothripsis. *Genome Res*. 2015 Jul;25(7):937–47.
- 132 Kurtas NE, Xumerle L, Giussani U, Pansa A, Cardarelli L, Bertini V, et al. Insertional translocation involving an additional nonchromothriptic chromosome in constitutional chromothripsis: Rule or exception? *Mol Genet Genomic Med*. 2019;7(2):e00496.
- 133 Sabatini PJB, Ejaz R, Stavropoulos DJ, Mendoza-Londono R, Joseph-George AM. Stable transmission of an unbalanced chromosome 21 derived from chromoanasythesis in a patient with a SYNGAP1 likely pathogenic variant. *Mol Cytogenet*. 2018 Dec;11(1):50.
- 134 Gudipati MA, Waters E, Greene C, Goel N, Hoppman NL, Pitel BA, et al. Stable transmission of complex chromosomal rearrangements involving chromosome 1q derived from constitutional chromoanagenesis. *Mol Cytogenet*. 2019 Dec;12(1):43.
- 135 Spielmann M, Lupiáñez DG, Mundlos S. Structural variation in the 3D genome. *Nat Rev Genet*. 2018 Jul;19(7):453–67.
- 136 Schluth-Bolard C, Diguët F, Chatron N, Rollat-Farnier P-A, Bardel C, Afenjar A, et al. Whole genome paired-end sequencing elucidates functional and phenotypic consequences of balanced chromosomal rearrangement in patients with developmental disorders. *Journal of Medical Genetics*. 2019 Aug;56(8):526–35.
- 137 Ottaviani A, Schluth-Bolard C, Rival-Gervier S, Boussouar A, Rondier D, Foerster AM, et al. Identification of a perinuclear positioning element in human subtelomeres that requires A-type lamins and CTCF. *EMBO J*. 2009 Aug;28(16):2428–36.

- 138 Kunwar F, Tandel K, Bakshi SR. Recurrent chromosomal translocations: Is proximity a rule? *INDIAN J EXP BIOL.* 2017;6.
- 139 Brianna Caddle L, Grant JL, Szatkiewicz J, van Hase J, Shirley B-J, Bewersdorf J, et al. Chromosome neighborhood composition determines translocation outcomes after exposure to high-dose radiation in primary cells. *Chromosome Res.* 2007 Nov;15(8):1061.
- 140 Medina-Martínez O, Ramírez-Solis R. In vivo mutagenesis of the Hoxb8 hexapeptide domain leads to dominant homeotic transformations that mimic the loss-of-function mutations in genes of the Hoxb cluster. *Dev Biol.* 2003 Dec;264(1):77–90.
- 141 Popova BC, Tada T, Takagi N, Brockdorff N, Nesterova TB. Attenuated spread of X-inactivation in an X;autosome translocation. *Proc Natl Acad Sci U S A.* 2006 May;103(20):7706–11.
- 142 Russell LB. Mammalian X-Chromosome Action: Inactivation Limited in Spread and in Region of Origin. *Science.* 1963 May;140(3570):976–8.
- 143 Teller K, Illner D, Thamm S, Casas-Delucchi CS, Versteeg R, Indemans M, et al. A top-down analysis of X_a- and X_i-territories reveals differences of higher order structure at ≥ 20 Mb genomic length scales. *Nucleus.* 2011 Sep;2(5):465–77.
- 144 Bello S, Rodríguez-Moreno A. [An updated review of 1p36 deletion (monosomy) syndrome]. *Rev Chil Pediatr.* 2016 Oct;87(5):411–21.
- 145 Lessel D, Gehbauer C, Bramswig NC, Schluth-Bolard C, Venkataramanappa S, van Gassen KLI, et al. BCL11B mutations in patients affected by a neurodevelopmental disorder with reduced type 2 innate lymphoid cells. *Brain.* 2018 Aug;141(8):2299–311.

