



HAL
open science

Analyse quantitative et systématique de la dynamique spatio-temporelle de l'activité de la voie de signalisation ERK au cours de l'embryogenèse des ascidies

Kilian Biasuz

► To cite this version:

Kilian Biasuz. Analyse quantitative et systématique de la dynamique spatio-temporelle de l'activité de la voie de signalisation ERK au cours de l'embryogenèse des ascidies. Biologie cellulaire. Université de Montpellier, 2022. Français. NNT : 2022UMONT030 . tel-04074747

HAL Id: tel-04074747

<https://theses.hal.science/tel-04074747v1>

Submitted on 19 Apr 2023

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.

**THÈSE POUR OBTENIR LE GRADE DE DOCTEUR
DE L'UNIVERSITÉ DE MONTPELLIER**

En Biologie Santé

École doctorale CBS2 n°168 Sciences Chimiques et Biologiques pour la Santé

Unité Mixte de Recherche 5237 Centre de Recherche en Biologie cellulaire de Montpellier (CRBM)

**Analyse quantitative et systématique de la dynamique
spatio-temporelle de l'activité de la voie de
signalisation ERK au cours de l'embryogenèse des
ascidies**

Présentée par Kilian BIASUZ

Le 25 octobre 2022

Sous la direction de Patrick LEMAIRE

Équipe de recherche : Contrôle transcriptionnel de la morphogenèse des chordés

Devant le jury composé de

Pr. François FAGOTTO, CRBM Montpellier	Président du jury
Dr. Clare HUDSON, LBDV Villefranche-sur-Mer	Rapporteuse
Dr. Romain LEVAYER, Institut Pasteur Paris	Rapporteur
Dr. Charles KERVRANN, INRIA Rennes	Examineur
Dr. Matteo RAUZI, IBV Nice	Examineur
Dr. Emmanuel FAURE, LIRMM Montpellier	Membre invité
Dr. Grégoire MALANDAIN, INRIA Sophia-Antipolis Nice	Membre invité
Dr. Patrick LEMAIRE, CRBM Montpellier	Directeur de thèse



**UNIVERSITÉ
DE MONTPELLIER**

À mes deux grands parents,

REMERCIEMENTS

Cinq années de thèse, beaucoup de chemin parcouru, tout cela n'aurait évidemment pas pu être possible sans les nombreuses personnes que j'ai eu l'opportunité de côtoyer durant ces années et que j'aimerais remercier.

Tout d'abord mes premiers mots sont pour toi Patrick. Je te remercie énormément de m'avoir accueilli dans l'équipe et de m'avoir guidé depuis le premier jour jusqu'au dernier (qui n'est d'ailleurs pas encore pour tout de suite). Tu es un exemple pour moi aussi bien scientifiquement qu'humainement et j'ai vraiment eu de la chance d'avoir pu passer tout ce temps à tes côtés et que ce soit toi qui m'aies formé pour devenir chercheur. Tu m'as accordé ta confiance et j'espère en retour avoir pu apporter ma contribution aux travaux et à la vie de l'équipe.

Et quelle équipe! Un grand merci à vous tous d'avoir été là pour moi tout au long de cette aventure humaine qu'est la thèse, toujours de bonne humeur et bienveillants, encore plus particulièrement au cours de ces (quelques) derniers mois de rédaction. Je n'oublierai jamais tout votre soutien sur cette dernière ligne droite ! Alors merci Christelle pour tout, c'est grâce à toi si j'ai "dépilé, dépilé" ! Merci Benjamin, vraiment ! Je t'ai emmené dans mes galères et tu as toujours été là pour moi ! Merci Jean-Philippe, pour tous ces moments passés à partager notre passion commune pour le Paris Saint Germain (j'espère que ta vue te permettra de lire ces quelques lignes) ! Merci M^{me} Moring, la boss ! Merci Lisa et là, bien évidemment, je suis obligé de penser à tes cookies ! Merci Grégoire ! Merci pour tous tes conseils et ton soutien pendant la rédaction de mon manuscrit, toujours à portée de zoom ! MorphoMerci Emmanuel ! Pour tous ces moments passés au labo et en dehors ! Merci Jacques! Merci pour ta présence, pour tous ces moments avec toi à Garrigues, pour ces sorties en vélo et rando et j'espère qu'on pourra encore se retrouver sur la ligne de départ d'un triathlon ! Merci Julien ! Pour toute ton aide et tes encouragements depuis le tout début de ma thèse jusqu'à aujourd'hui ! Merci Panajot ! Le meilleur stagiaire que j'ai eu à encadrer ! Merci Delphine, pour ta bonne humeur ! Merci également à tous les autres anciens membres de l'équipe : Ulla, Matthieu, Justine, Matija et à tous les stagiaires pour tous les bons moments partagés !

Je souhaiterais ensuite remercier l'ensemble du CRBM, en particulier Anne Debant et Serge Roche qui m'ont accordé un financement pour ma 4^{ème} année de thèse lorsqu'ils étaient à la direction du laboratoire; les membres du conseil du laboratoire avec qui j'ai siégé et tous les membres du mouvement Green-CRBM; "Monsieur" Thierry Lorca et son équipe; Dimitri Xirodimas et son équipe pour son aide pour la protéomique ainsi que Serge Urbach et son équipe de la Plateforme de Protéomique de Montpellier pour l'acquisition et l'analyse des données; Philippe et Marc pour l'animalerie ascidies; la plate-forme d'imagerie MRI et plus particulièrement Sylvain pour tous ces moments passés à nous émerveiller de la capacité du MuViSPIM à créer de nouveaux problèmes; Mickaël et Ghislain pour m'avoir toujours accueilli dans la bonne humeur; Vinh, Monique et Laetitia pour l'administration; tout le service informatique, promis je ferai un ticket GLPI; Deni et Antoni pour la laverie; tout le service Atelier et plus particulièrement Jean Casanova, ça a toujours été un plaisir de venir te voir dans ton atelier. Enfin merci à toute l'équipe du restaurant du campus pour votre accueil, votre cuisine, votre générosité et pour toutes ces omelettes! Merci aux membres du service de sécurité.

Je voudrais aussi remercier Clare Hudson et Romain Levayer pour avoir accepté de lire et corriger ma thèse et pour m'avoir soutenu pendant la rédaction de mon manuscrit, ainsi que François Fagotto pour avoir présidé mon jury et Charles Kervrann et Matteo Rauzi pour avoir participé au jury de ma thèse.

Un grand merci également à Claire Chazaud, Serge Pelet et Stéphane Bodin pour avoir participé à mon comité de suivi de thèse, pour votre bienveillance, pour nos discussions et tous vos précieux conseils.

Merci au Labex EpiGenMed pour avoir financé mon contrat doctoral, à l'Université de Montpellier pour avoir financé une prolongation de 5 mois de ce dernier en raison de la crise COVID-19 et à l'école doctorale CBS2.

Je remercie le département Biologie Mécanismes du Vivant (BMV) de la Faculté des Sciences de l'Université de Montpellier pour m'avoir donné l'opportunité de donner des enseignements de TD et TP au cours de mes deux premières années de thèse.

Et enfin, parce que le doctorat ne s'arrête vraiment pas une fois les portes du labo refermées, un grand merci à ma famille pour m'avoir toujours soutenu et pour avoir toujours cru en moi. Si j'en suis là c'est en grande partie grâce à vous et pour vous. Alors merci à ma mère, à mon père, à mon oncle, mes frères et soeurs (Juliette, Agathe, Louise, Robin, Elena), merci JC, merci Sonia, merci Geneviève et Daniel, merci mes demi-frères et soeurs (Nathan, Noa, Angela), merci Thibault, merci Mael, merci Ilona, merci Olivier et Cécilia, merci Stéphanie. Merci à mon meilleur ami Jason, tout est dit, et à ma deuxième famille : mamie, Severine, M^r John, Guillaume, parrain et sa famille. Merci à mes amis docteurs de la Bioshield Company Rafath et Maxence, même si je ne vous ai pas vu autant que j'aurais aimé pendant la thèse, c'est grâce à tous ces bons moments passés ensemble que j'ai tenu pendant le master et j'espère vous retrouver pour ma soutenance et pour la suite. Finalement, je voudrais terminer en remerciant Julia. Ca n'a pas toujours été simple d'avoir nos thèses respectives à gérer en parallèle, surtout à la fin, mais merci d'être toujours là, dans les bons comme les moins bons moments, merci de m'avoir soutenu, écouté et encouragé quand il le fallait.

RÉSUMÉ

Titre : Analyse quantitative et systématique de la voie de signalisation ERK au cours de l'embryogenèse des ascidies.

Le succès de l'embryogenèse nécessite la formation des types cellulaires corrects dans des positions appropriées. Afin de prendre les bonnes décisions, les cellules communiquent en utilisant un répertoire conservé de facteurs de signalisation et de voies de transduction. La voie de signalisation ERK est l'une de ces importantes voies utilisée à de nombreuses reprises au cours du développement chez les animaux. Cependant, nous ne comprenons toujours pas comment la signalisation ERK participe à des fonctions aussi diverses. Il a été proposé que les caractéristiques de la dynamique spatio-temporelle codent en elles-mêmes la diversité des effets observés et une variété de cinétiques d'activité de ERK est maintenant observée dans de nombreux types de cellules ou en réponse à différents stimuli. Ainsi, bien que l'étude de la dynamique de l'activité ERK au cours du développement soit essentielle pour approfondir notre compréhension du fonctionnement de cette voie et de son impact sur les comportements cellulaires, les études quantitatives de l'activité ERK à la résolution cellulaire dans des embryons vivants commencent seulement à émerger et notre compréhension reste limitée à quelques éléments non reliés entre eux.

L'étude de la voie ERK chez les vertébrés est complexifiée par la duplication de leurs génomes et la présence de multiples paralogues. Ainsi, pendant mon doctorat, j'ai étudié la voie ERK chez les ascidies, de petits invertébrés marins membres des tuniciers, groupe frère des vertébrés. Ces animaux possèdent des gènes uniques des composants de la voie ERK et l'utilisent pour contrôler de nombreux événements d'inductions au cours de leur embryogenèse. Les embryons de l'ascidie *Phallusia mammillata* ont deux autres propriétés favorables : ils sont petits et entièrement transparents, donc faciles à imager, et ils se développent rapidement avec des lignages cellulaires invariants, de sorte que chaque cellule peut être nommée et trouvée à la même position, ce qui fournit un cadre rigoureux pour des analyses à la résolution cellulaire. De plus, avant le début de mon doctorat, notre équipe a développé une capacité unique à identifier/segmenter/suivre les cellules au cours du développement de *Phallusia*.

Mon projet de thèse a donc combiné imagerie quantitative *in vivo*, informatique, modélisation théorique et perturbations expérimentales afin d'étudier de façon quantitative et systématique la dynamique spatio-temporelle de la voie de signalisation ERK au cours de l'embryogenèse de *Phallusia*. Je me suis tout d'abord concentré sur le développement des méthodologies expérimentales permettant la quantification manuelle *in vivo* de l'activité ERK au stade 64-cellules en utilisant le biosenseur ERK-KTR. Comme preuve de principe, nous avons utilisé cet outil dans le cadre de la révision d'un article visant à tester l'hypothèse selon laquelle les surfaces de contact entre cellules émettrices et réceptrices fournissent les informations quantitatives qui déterminent le résultat des inductions chez les ascidies. J'ai ensuite participé à l'élaboration d'un cadre théorique et des méthodes algorithmiques permettant d'identifier et de nommer automatiquement chaque cellule des embryons reconstruits de *Phallusia*. En connectant les embryons dans un référentiel commun, ce travail établit les embryons stéréotypés d'ascidies comme un cadre numérique de choix pour construire et intégrer des atlas 4D fonctionnels à la résolution cellulaire. Enfin, je me suis concentré sur le développement et à la validation de la méthodologie permettant de mesurer automatiquement l'activité de ERK pour chaque cellule d'embryons vivants de *Phallusia* toutes les 2 minutes. L'ensemble de ce travail permettra d'identifier de nouvelles inductions et d'étudier la variabilité de la signalisation ERK entre embryons en lien notamment avec la variabilité des contacts cellulaires afin de tester la généralité du modèle d'induction surface-de-contact dépendant chez les ascidies.

RÉSUMÉ GRAND PUBLIC

Titre : Analyse quantitative et systématique de la voie de signalisation ERK au cours de l'embryogenèse des ascidies.

Le succès de l'embryogenèse requiert la formation des bons types cellulaires aux bons endroits. Pour cela, les cellules communiquent entre elles via des signaux chimiques qui sont intégrés par des voies de signalisation. La voie ERK est l'une de ces importantes voies utilisée à de nombreuses reprises au cours du développement. Cependant, nous ne comprenons toujours pas comment elle participe à des fonctions aussi diverses. Il a été proposé que la dynamique de l'activation de la voie encode la diversité des effets observés mais notre compréhension reste limitée. Mon projet de thèse, qui profite de notre capacité unique à reconstruire en 3D le développement des embryons d'ascidies, des invertébrés marins du groupe frère des vertébrés, combine microscopie avancée, perturbations expérimentales et modélisation théorique afin d'étudier quantitativement et systématiquement la dynamique de la voie ERK dans chaque cellule d'embryons vivants.

ABSTRACT

Title : Quantitative and systematic analysis of the ERK signalling pathway during ascidian embryogenesis.

Successful embryogenesis requires the formation of the correct cell types in appropriate positions. In order to make the right decisions, cells communicate using a conserved repertoire of signalling factors and transduction pathways. The ERK signalling pathway is one of these important pathways that is used repeatedly during development in animals. However, we still do not understand how ERK signalling is involved in such diverse functions. It has been proposed that the characteristics of the spatio-temporal dynamics themselves encode the diversity of effects observed and a variety of kinetics of ERK activity is now observed in many cell types or in response to different stimuli. Thus, although the study of the dynamics of ERK activity during development is essential to further our understanding of the function of this pathway and its impact on cellular behaviour, quantitative studies of ERK activity at cellular resolution in living embryos are only beginning to emerge and our understanding remains limited to a few unrelated elements.

The study of the ERK pathway in vertebrates is complicated by the duplication of their genomes and the presence of multiple paralogs. Thus, during my PhD, I studied the ERK pathway in ascidians, small marine invertebrates that are members of the tunicates, the sister group of vertebrates. These animals possess unique genes for each component of the ERK pathway and use this pathway to control many induction events during their embryogenesis. Embryos of the ascidian *Phallusia mammillata* have two other favourable properties: they are small and fully transparent, thus easy to image, and they develop rapidly with invariant cell lineages, so that each cell can be named and found at the same position, providing a rigorous framework for analyses at cellular resolution. In addition, prior to the start of my PhD, our team developed a unique ability to identify/segment/track cells during *Phallusia* development.

My thesis project therefore combined quantitative *in vivo* imaging, computer science, theoretical modelling and experimental perturbation to quantitatively and systematically study the spatio-temporal dynamics of the ERK signalling pathway during *Phallusia* embryogenesis. I first focused on the development of experimental methodologies allowing the manual *in vivo* quantification of ERK activity at the 64-cell stage using the ERK-KTR biosensor. As a proof of principle, we used this tool as part of a paper review to test the hypothesis that contact surfaces between transmitter and receiver cells provide the quantitative information that determines the outcome of inductions in ascidians. I then participated in the development of a theoretical framework and algorithmic methods to automatically identify and name each cell in reconstructed *Phallusia* embryos. By connecting the embryos into a common repository, this work establishes stereotyped ascidian embryos as a numerical framework of choice for building and integrating functional 4D atlases with cellular resolution. Finally, I focused on the development and validation of the methodology to automatically measure ERK activity for each cell of living *Phallusia* embryos every 2 minutes. This work will allow us to identify new ERK-dependent inductions and to study the variability of ERK signalling between embryos in relation to the variability of cell contacts in order to test the generality of the surface-contact dependent induction model in ascidians.

PUBLIC SUMMARY

Title : Quantitative and systematic analysis of the ERK signalling pathway during ascidian embryogenesis.

Successful embryogenesis requires the formation of the right cell types in the right places. To do this, cells communicate with each other via chemical signals that are integrated by signalling pathways. The ERK pathway is one of these important pathways that is used repeatedly during development. However, we still do not understand how it is involved in such diverse functions. It has been proposed that the dynamics of pathway activation encode the diversity of effects observed but our understanding remains limited. My thesis project, which takes advantage of our unique ability to reconstruct in 3D the development of ascidian embryos, marine invertebrates of the vertebrate sister group, combines advanced microscopy, experimental perturbations and theoretical modelling to quantitatively and systematically study the dynamics of the ERK pathway in every cell of living embryos.

TABLE DES MATIÈRES

INTRODUCTION	1
LA COMMUNICATION CELLULAIRE ORCHESTRE LA COORDINATION ET LA SPÉCIFICATION DES CELLULES AU COURS DU DÉVELOPPEMENT EMBRYONNAIRE	2
Introduction à la communication cellulaire	2
L'expérience de Spemann et Mangold, 1924, révèle l'importance du transfert d'informations entre cellules au cours du développement embryonnaire	2
Découverte des principales voies de communication cellulaire	3
Mode de fonctionnement des voies de signalisation impliquées dans la spécification des destins cellulaires	4
Communication cellulaire par émission de ligands diffusibles : le concept de gradient de morphogène et d'information de position	6
Origines du concept de "gradient de morphogène"	6
Principes de la formation d'un gradient morphogène	8
L'interprétation du gradient de concentration de morphogène par les réseaux de régulation génétiques permet l'établissement de frontières entre tissus	9
Communication cellulaire par contact direct entre cellules émettrices et réceptrices : comment spécifier le destin d'une cellule au sein d'un champ de cellules équivalentes	11
Mécanismes émergents de communication cellulaire et directions futures : vers une compréhension opérationnelle et quantitative des voies de signalisation	13
LA VOIE DE COMMUNICATION ERK	17
Composition moléculaire de la voie ERK	17
La dynamique de l'activité de la voie ERK joue un rôle important dans le contrôle des réponses cellulaires induites	18
Les biosenseurs génétiquement codés permettent de faire la lumière sur la voie ERK	20
Biosenseurs basés sur le FRET	21
Biosenseurs basés sur un seul fluorophore	21
Propriétés de transmission des signaux dans le module RAS/ERK	22
Au cours du développement précoce de la souris, la dynamique de l'activité de ERK après la mitose détermine la spécification du destin entre les cellules de l'épiblaste et de l'endoderme primitif	24
Dans l'embryon précoce de drosophile, la spécification du destin des cellules latérales dépend de l'activité intégrée de ERK dans le temps plutôt que d'autres caractéristiques de la dynamique.....	26
L'utilisation du ERK-KTR a révélé des oscillations de l'activité ERK modulée par la concentration de ligand au cours du patterning de la vulve chez <i>C. elegans</i>	28
Des ondes d'activité de ERK dans la formation de la placode chez la drosophile sont associées à une contractilité de l'actomyosine plus élevée dans les cellules du front d'onde ERK qui est requise pour l'invagination	29
Le nombre d'études de biologie du développement utilisant ces biosenseurs pour suivre la dynamique spatio-temporelle de l'activité ERK est encore limité	30
LES ASCIDIES : UN BON MODÈLE POUR L'ÉTUDE DE LA DYNAMIQUE SPATIO-TEMPORELLE DE LA VOIE ERK AU COURS DU DÉVELOPPEMENT EMBRYONNAIRE	32
Introduction aux ascidies	32

Les ascidies fournissent un cadre unique pour l'imagerie quantitative in toto sur des embryons vivants avec une résolution subcellulaire	36
Les embryons d'ascidie : un cadre rigoureux pour l'étude de l'embryogenèse au niveau cellulaire	36
Les ascidies possèdent des propriétés permettant l'imagerie in toto de leurs embryons vivants	39
Une capacité unique de segmenter et suivre chaque cellule au cours du développement précoce des ascidies.....	41
La voie de signalisation ERK est simplifiée chez les ascidies	42
Les ascidies font un usage intensif de la voie FGF/RAS/ERK au cours de leur embryogenèse précoce	42
Activité de la signalisation ERK surface-de-contact dépendante au cours de l'induction neurale chez les ascidies	43
OBJECTIFS DU TRAVAIL DE THÈSE	46
RÉSULTATS	48
CHAPITRE 1 : L'ANALYSE QUANTITATIVE IN VIVO DE L'ACTIVITÉ ERK AU STADE 64-CELLULES SUPPORTE LE MODE DE COMMUNICATION CELLULAIRE SURFACE-DE-CONTACT DÉPENDANT DANS LES EMBRYONS D'ASCIDIES	49
Introduction	49
Manuscrit de l'article : Guignard, Léo et al. "Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis." Science (New York, N.Y.) Vol. 369,6500 (2020): EAAR5663.	53
Extraits du supplément d'information : "Validation biologique des prédictions du modèle à partir de mesures de l'activité ERK in vivo pour chaque cellule au stade 64-cellules."	63
Discussion	71
CHAPITRE 2 : L'IDENTIFICATION ET LE NOMMAGE SYSTÉMATIQUE ET AUTOMATIQUE DES CELLULES HOMOLOGUES DANS LE EMBRYONS STÉRÉOTYPÉS D'ASCIDIES ÉTABLIT LES ASCIDIES COMME UN CADRE NUMÉRIQUE POUR LA CONSTRUCTION ET L'INTÉGRATION D'ATLAS 4D FONCTIONNELS MULTI-ÉCHELLES À LA RÉOLUTION CELLULAIRE	73
Introduction	73
Manuscrit de l'article : Biasuz, Kilian et al. "Systematic automated cell naming of segmented ascidian embryos establishes ascidians as a digital framework for constructing, analysing and integrating multi-scale functional 4D atlases at cellular resolution."	76
Discussion	109
CHAPITRE 3 : ANALYSE QUANTITATIVE ET SYSTÉMATIQUE DE LA DYNAMIQUE SPATIO-TEMPORELLE DE L'ACTIVITÉ DE LA SIGNALISATION ERK AU COURS DU DÉVELOPPEMENT PRÉCOCE DE PHALLUSIA MAMMILLATA	110
Introduction	110
Manuscrit de l'article : Biasuz, Kilian et al. "Systematic and quantitative monitoring of the spatio-temporal dynamic of the ERK signalling at single cell resolution during early ascidian embryogenesis."	112
Discussion	161

DISCUSSION ET PERSPECTIVES	164
RÉSUMÉ	165
DIFFICULTÉS RENCONTRÉES	169
DIRECTIONS FUTURES	172
Vers une exploration multi-échelle de la variabilité de l'embryogénèse des ascidies	172
Vers une compréhension mécanistique du mécanisme original de signalisation surface-de-contact dépendent dans les embryons d'ascidies	174
PUBLICATIONS ANNEXES	177
RÉFÉRENCES	190

INTRODUCTION

LA COMMUNICATION CELLULAIRE ORCHESTRE LA COORDINATION ET LA SPÉCIFICATION DES CELLULES AU COURS DU DÉVELOPPEMENT EMBRYONNAIRE

La formation d'un organisme complexe à partir d'une cellule unique, l'ovocyte fécondé (ou zygote), requiert la formation des différents types de cellules qui composent ses tissus en nombre précis et leur organisation précise dans l'espace. La constitution des différents tissus implique de la division cellulaire et de la différenciation afin que des cellules assurent des fonctions différentes. Par exemple, les cellules musculaires se contractent, les cellules épidermiques protègent du milieu extérieur, les cellules pancréatiques sécrètent les enzymes digestives. Ces différents types de cellules se forment progressivement au cours du développement embryonnaire. Parmi les différents processus qui interviennent pour orchestrer le développement embryonnaire chez les animaux, la communication cellulaire joue un rôle majeur.

Introduction à la communication cellulaire

Notre compréhension de ces mécanismes de communication cellulaire a évolué au fil de nombreuses décennies de recherches.

L'expérience de Spemann et Mangold, 1924, révèle l'importance du transfert d'informations entre cellules au cours du développement embryonnaire

Au début du XX^{ème} siècle, des expériences classiques de greffe de tissus dans des embryons ont révélé le rôle crucial de la communication entre cellules dans la prise de décision concernant leur destin et d'autres processus. L'expérience de Spemann et Mangold, 1924, est l'une de ces expériences les plus connues (Spemann and Mangold, 1924). Elle a conduit, plus que toute autre, au modèle actuel selon lequel le développement est le résultat d'un transfert d'informations entre cellules. Cette expérience a révélé que certains territoires de l'embryon étaient capables d'organiser les territoires voisins (Figure 1). Plus spécifiquement, si la lèvre dorsale du blastopore (le site où commence la gastrulation) d'un embryon d'amphibien (Triton) est transplantée dans une région indifférenciée d'un autre embryon, elle est capable de recruter des cellules hôtes et de provoquer la formation d'un embryon secondaire (Figure 1). Les cellules de l'implant sont donc capables d'affecter le destin des cellules hôtes, un processus appelé "induction" (Spemann and Mangold, 1924). C'est également avec ce travail qu'est apparu le concept d'"organisateur" : "une région de l'embryon d'où émanent des informations de détermination d'une certaine quantité dans certaines directions" (Spemann and Mangold, 1924). Ce concept propose que la détermination des destins cellulaires ne commence pas simultanément dans toutes les parties de l'embryon, mais progresse de cellules en cellules à partir d'un endroit (l'organisateur) dans une direction définie.

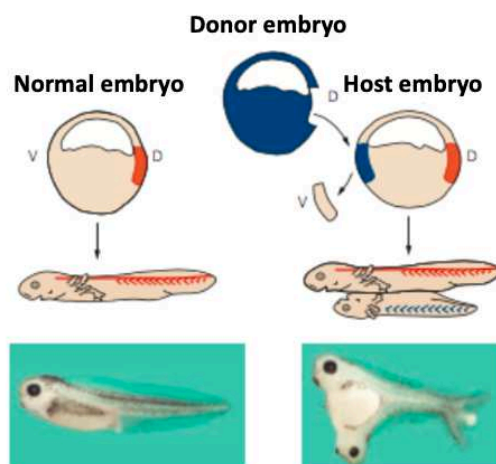


Figure 1 : L'expérience de Spemann-Mangold 1924. Deux embryons de triton sont choisis au début de la gastrulation. Au préalable, l'un d'eux a été marqué peu après la fécondation (bleu). L'embryon marqué sera le donneur de lèvre dorsale. L'embryon non marqué sera le receveur. Sur l'embryon receveur, on découpe une fenêtre réceptrice ventrale diamétralement opposée à la lèvre dorsale du blastopore. Sur l'embryon donneur marqué on prélève la lèvre dorsale du blastopore et on l'implante dans la fenêtre réceptrice de l'embryon receveur. La cicatrisation est complète une heure après l'implantation. La greffe de la lèvre dorsale du blastopore provoque la formation d'un axe dorsal secondaire et la formation d'un embryon secondaire. D, dorsal ; V, ventral. Adapté de (Lemaire and Kodjabachian, 1996; De Robertis, 2009).

Découverte des principales voies de communication cellulaire

Ainsi, au cours du développement embryonnaire, un groupe particulier de cellules peut influencer le destin d'un groupe adjacent de cellules. Des décennies de recherches ont ensuite permis de découvrir un ensemble de voies de communication cellulaires centrales qui sont responsables de ces inductions, notamment les systèmes de signalisation Notch, Transforming Growth Factor beta (TGF- β), Wnt, Hedgehog (HH) et Récepteurs Tyrosine Kinase (RTK), qui jouent des rôles essentiels dans une gamme étonnamment large de processus de développement (Figure 2) (voir pour revue (Perrimon et al., 2012)). Les composants moléculaires principaux de ces voies ont été identifiés, leurs interactions caractérisées et leurs rôles dans de nombreux contextes développementaux ont été élucidés. En raison de la prépondérance et de la diversité de leurs rôles, ainsi que de leur conservation générale entre les espèces, ces voies de signalisation intercellulaire font désormais partie des systèmes les mieux étudiés en biologie (Antebi et al., 2017). Ces travaux fournissent un cadre fondamental pour comprendre la signalisation au cours du développement embryonnaire. Ces voies de signalisation se répartissent en 10 classes majeures (Tableau 1), un nombre étonnamment faible compte tenu de la diversité des événements de spécification du destin qui se produisent pendant l'embryogenèse (Perrimon et al., 2012; Biasuz et al., 2018; Patel and Shvartsman, 2018).

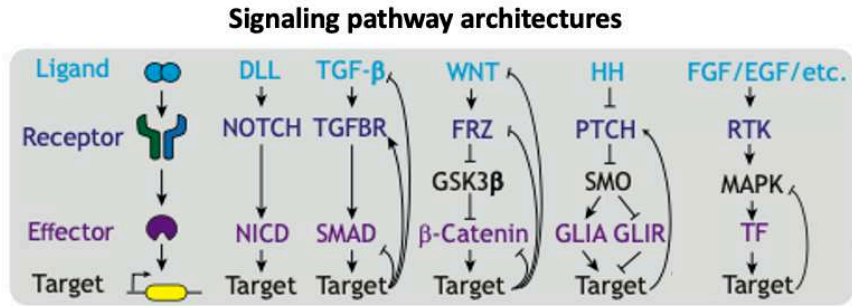


Figure 2 : Schémas des principales voies de signalisation du développement. Les voies de signalisation utilisées pour spécifier les destins cellulaires utilisent diverses architectures pour contrôler la communication entre cellules selon un plan commun : les interactions ligand-récepteur activent des effecteurs intracellulaires, qui régulent ensuite l'expression des gènes cibles. L'activité de signalisation intracellulaire induit également de nombreuses boucles de rétroaction qui modulent davantage la capacité de traitement du signal de la voie. TF : Transcription factor, facteur de transcription. Adapté de (Li and Elowitz, 2019).

Noms des voies de signalisation
Notch
Facteur de croissance transformant β (Tgf- β)
Wnt
Hedgehog (Hh)
Récepteurs tyrosine kinases (RTK)
JAK-STAT
Hippo
Jun Kinase (JNK)
NF- $\kappa\beta$
Récepteur de l'acide rétinoïque (RAR)

Tableau 1 : Principales voies de signalisation qui orchestrent le développement. Adapté de (Perrimon et al., 2012).

Mode de fonctionnement des voies de signalisation impliquées dans la spécification des destins cellulaires

Ces voies agissent à travers une architecture commune dont la conception leur permet de moduler l'expression des gènes afin de spécifier des destins cellulaires particuliers en réponse à des signaux émis par les cellules "inductrices" (Figure 2). Cette capacité à moduler l'expression génétique s'effectue par le contrôle de l'activité de facteurs de transcriptions : des protéines capables de se lier à l'ADN sur des sites spécifiques et de réguler l'expression de gènes cibles (voir pour revue (Spitz and

Furlong, 2012)). Certains de ces facteurs de transcription sont communs à de nombreux types de cellules, tandis que d'autres ne sont présents que dans des types de cellules spécifiques. C'est le cas par exemple pour GATAa, un facteur de transcription membre de la famille de GATA et médiateur de la signalisation RTK(FGFR)/ERK dans les embryons d'ascidies, qui est activé spécifiquement dans les blastomères ectodermiques et cible l'induction de destins neuraux en réponse à la signalisation RTK(FGFR)/ERK uniquement aux cellules ectodermiques alors que toutes les cellules de l'embryon reçoivent la signalisation (Bertrand et al., 2003).

Les signaux utilisés sont essentiellement chimiques (protéines, stéroïdes, ...). Ils sont émis par les cellules inductrices soit sécrétés librement dans le milieu extracellulaire, où ils diffusent en fonction des propriétés du milieu, soit ancrés à la membrane des cellules qui les émettent, nécessitant que la cellule émettrice du signal soit en contact direct avec ses cellules cibles (Figure 2) (voir pour revue ((Perrimon et al., 2012)). D'un autre côté, des cellules "réceptrices" sont compétentes à répondre lorsqu'elles expriment un récepteur et un appareil intracellulaire de transduction du signal pour la voie de communication considérée (Waddington, 1940). À l'exception de l'acide rétinoïque, qui traverse la membrane plasmique et active l'expression de gènes cibles en se liant directement à des récepteurs intracellulaires qui agissent comme facteurs de transcription dans le noyau, les ligands des autres voies se lient à des récepteurs transmembranaires sur les cellules cibles (Figure 2) (voir pour revue ((Perrimon et al., 2012)). Ces récepteurs transmembranaires, une fois activés, déclenchent une cascade d'événements de transduction du signal qui propagent le message depuis la membrane vers un effecteur terminal nucléaire qui contrôle l'activité de facteurs de transcriptions (Figure 2) (voir pour revue ((Perrimon et al., 2012)). De plus, de nombreuses voies activent des boucles de rétroaction qui modulent, amplifient ou interrompent le signal entrant (Figure 2) (Freeman, 2000). Les cellules réceptrices, à leur tour, peuvent devenir inductrices et changer le destin de leurs voisines en produisant de nouveaux signaux, générant ainsi des événements inductifs séquentiels qui augmentent la diversité des destins cellulaires dans les tissus.

Les conséquences de l'activation de chacune de ces voies peuvent être diverses en fonction de l'état interne de chaque cellule qui dépend de son passé (types de récepteurs et de facteurs de transcription exprimés, régions de l'ADN accessibles) et qui influence l'interprétation du signal (Sagner and Briscoe, 2017). Ainsi, les cellules réceptrices peuvent être amenées à se différencier, à se diviser, à changer leur morphologie, à migrer, à adhérer les unes aux autres ou, dans certains cas, à subir une apoptose (Perrimon et al., 2012). Les changements générés par les signaux au cours du développement sont généralement irréversibles et nécessitent donc une grande précision dans leur transmission et leur interprétation. Comme les perturbations de ces processus de communication peuvent affecter le développement embryologique normal, les mécanismes de signalisation qui opèrent au cours du développement doivent être fiables et reproductibles.

Bien que des voies de signalisation particulières aient des composants moléculaires et des mécanismes de régulation distincts, des principes communs impliqués dans la formation des différents types de cellules et dans l'organisation spatiale précise des tissus (formations de motifs, de tissus différents, de frontières entre les tissus, ...) au sein de l'embryon qui se développe ont été identifiés (Perrimon et al., 2012; Li and Elowitz, 2019). Parmi eux, on trouve deux principes fondamentaux et historiques (Wolpert, 1994). Un premier, basé sur les ligands diffusibles sécrétés, stipule que des cellules adjacentes équivalentes adoptent des destins distincts en fonction de petites

différences dans la concentration locale du ligand : c'est le concept de gradient de morphogène et d'information de position (Čapek and Müller, 2019). Un second mécanisme d'inhibition latérale, basé sur la signalisation Notch dépendante du contact, joue un rôle essentiel dans la détermination de cellules individualisées au sein d'un groupe de cellules équipotentes (Schweisguth, 2000). Ces deux mécanismes sont détaillés dans les prochaines sections.

Communication cellulaire par émission de ligands diffusibles : le concept de gradient de morphogène et d'information de position

Au cœur des mécanismes de spécification des destins cellulaires par la communication cellulaire se trouve le concept de morphogène, qui propose que la diffusion des molécules de signalisation à partir d'une source coordonne la spécification des destins et la formation des tissus en créant un gradient de concentration. Ainsi, un champ de cellules naïves traversé par le gradient est partitionné en domaines spatiaux de types de cellules entre des seuils de concentration (Wolpert, 1969; Sagner and Briscoe, 2017) (Figure 3).

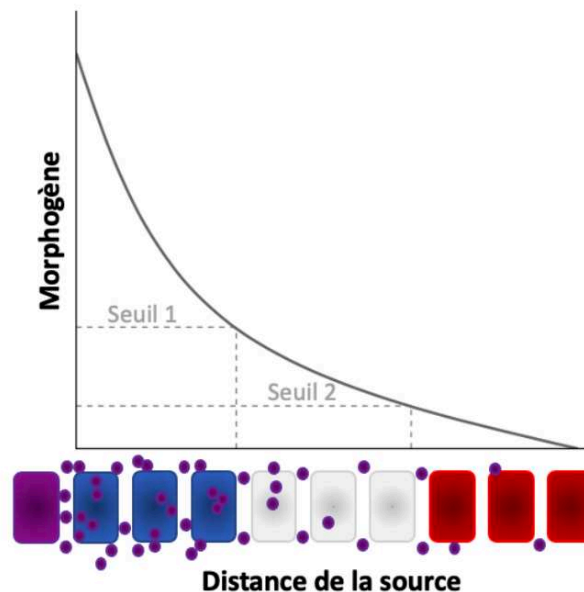


Figure 3 : Le modèle du drapeau français. Un morphogène est sécrété à partir d'une cellule source (violette) et sa diffusion crée un gradient de concentration au sein du tissu. Les cellules exposées à des concentrations de morphogène au-dessus du seuil 1 adoptent une réponse distincte (bleue) des cellules exposées à des niveaux intermédiaires de concentrations de morphogène (entre les seuils 1 et 2, réponse blanche), tandis que les cellules exposées à des niveaux de morphogènes en dessous du seuil 2 adoptent une troisième réponse (rouge). Ainsi, un gradient de concentration d'une seule substance biochimique qui diffuse peut donner lieu à de multiples statuts cellulaires et attribuer des valeurs de position aux cellules.

Origines du concept de "gradient de morphogène"

Le concept de "gradient" est reconnu depuis longtemps pour rendre compte de la formation de tissus dans les embryons en développement et en régénération et a retenu l'attention des biologistes expérimentaux et théoriques pendant des décennies, en partie en raison de son élégance et de sa simplicité. Dès le début du XIX^{ème} siècle, Thomas Hunt Morgan constate que la régénération se produit à des rythmes différents lorsque des verres de terre ou des hydres sont coupés à différents

niveaux le long de l'axe du corps. Afin d'expliquer ces résultats, il suggère qu'il existe un gradient de "substances" le long de l'animal et que ce gradient détermine à la fois la polarité du système et le rythme de la régénération (Morgan, 1901). Le terme "morphogène" est quant à lui introduit plus tard par Alan Turing, qui l'utilise pour désigner "toute substance biochimique qui diffuse entre les cellules et génère des réponses spécifiques à des concentrations particulières" (Turing, 1952). C'est ensuite Lewis Wolpert qui en 1969 marque un pas déterminant en fournissant un cadre théorique crucial pour l'étude des mécanismes de formation des patterns au cours du développement en présentant un modèle dans lequel les gradients peuvent générer des patterns de destinée cellulaire à travers le cadre conceptuel de "l'information positionnelle" (Wolpert, 1969). Dans ce cadre de l'information positionnelle, chaque cellule d'un tissu en développement ou en régénération acquiert une valeur positionnelle unique en mesurant la concentration du "morphogène" qui présente un gradient spatial à travers le tissu. La cellule se différencie alors en fonction de cette valeur positionnelle. L'une des conditions essentielles pour que ce modèle de gradient fonctionne repose sur l'hypothèse selon laquelle les cellules d'un champ présentent plusieurs seuils de réponse à des concentrations spécifiques de morphogènes. Ce modèle utilise le drapeau tricolore français comme représentation visuelle pour expliquer comment le gradient de morphogène est lu à des seuils de concentration fixes et donne lieu à trois types de tissus différents : bleu, blanc ou rouge (Figure 3). Cependant, les mécanismes qui sous-tendent la formation de gradients restaient inconnus. Deux ans plus tard, Francis Crick a suggéré un modèle simple de gradient unidimensionnel avec un morphogène diffusant à partir d'une source localisée d'un côté du champ de cellule et un puits de l'autre côté (Crick, 1970)). La source produit le morphogène et le maintient à un niveau stable, tandis qu'à l'autre extrémité le puits dégrade les molécules de morphogène qui l'atteignent.

Ces concepts ont eu une profonde influence dans la communauté des biologistes du développement. Ils ont ouvert la voie à des études plus détaillées sur les gradients morphogènes. La révolution de la génétique moléculaire à la fin des années 1980 et au début des années 1990 a commencé à révéler l'identité de plusieurs molécules qui présentaient les caractéristiques de morphogènes (Gerhart, 1999). Aujourd'hui, on considère généralement que la majorité des signaux sécrétés pour les principales voies de signalisation (Tableau 1) agissent comme morphogènes (Sagner and Briscoe, 2017). C'est le cas par exemple des Wnts (Figure 2; Tableau 1), des protéines de signalisation sécrétées, conservées au cours de l'évolution, qui, dans divers contextes de développement, se propagent à partir de leur site de synthèse pour former un gradient et activer l'expression de gènes cibles à distance (Nusse, 2012; Perrimon et al., 2012).

Ce modèle a longtemps prévalu pour expliquer le rôle de la signalisation Wnt dans la formation des disques imaginaux des ailes de la drosophile : des poches épithéliales qui passent de 50 à ~50 000 cellules au cours des stades larvaires chez la drosophile et qui donnent naissance aux ailes entières dans la drosophile adulte (Zecca et al., 1996; Neumann and Cohen, 1997; Swarup and Verheyen, 2012). Wingless, la principale protéine Wnt chez la drosophile, est initialement produite dans tous les précurseurs de l'aile à des stades précoces et se limite ensuite à une bande étroite de cellules à la limite dorso-ventrale (D-V) de ce disque imaginal. Depuis cette zone Wingless diffuse pour former un gradient et il était généralement accepté que le signal Wingless active l'expression de gènes cibles en fonction de sa concentration (Couso et al., 1994; Zecca et al., 1996; Neumann and Cohen, 1997; Nolo et al., 2000; Jafar-Nejad et al., 2006). Pour tester spécifiquement le rôle de la diffusion de Wingless, Alexandre et al. [2014] ont remplacé le gène endogène Wingless, par un gène qui exprime une forme

non-diffusible de Wingless qui reste attaché à la membrane des cellules qui l'expriment. Étonnamment, les adultes obtenus sont viables avec des ailes normales (Alexandre et al., 2014). Afin d'expliquer comment la signalisation juxtacrine de Wingless suffit pour la formation des ailes normales, les auteurs ont ensuite montré que l'expression précoce de Wingless dans tout le primordium de l'aile ainsi que l'effet persistant de la signalisation assurent une activation continue des gènes cibles même en l'absence de diffusion du Wingless depuis les cellules de la limite D-V. Cette étude révèle que dans certains contextes développementaux, contrairement au mécanisme de gradient qui peut être généralement admis, le développement peut opérer sans gradient instructif, ce qui souligne l'importance de tester directement l'hypothèse du gradient de morphogène.

L'identification et la caractérisation des voies de signalisation qui agissent par l'intermédiaire de signaux sécrétés diffusibles a été suivie d'études approfondies portant principalement sur deux questions : comment se forme un gradient ? Et comment les cellules répondent-elles aux signaux graduels pour former des groupes de cellules avec des destins distincts et des frontières précises ?

Principes de la formation d'un gradient morphogène

Tout d'abord, des descriptions théoriques ont aidé à comprendre comment la forme du gradient est contrôlée par les taux de production, de diffusion et de dégradation des morphogènes (voir pour revue (Wartlick et al., 2009)). Des mécanismes cellulaires et des interactions moléculaires tels que le trafic cellulaire du morphogène dans des vésicules et la liaison des molécules morphogènes à d'autres composants extracellulaires ont également été identifiés pour leur rôle dans le contrôle du transport des morphogènes (Yan and Lin, 2009). Cependant, un défi majeur a longtemps été le développement d'outils expérimentaux qui permettent de tester l'apport de ces différents mécanismes dans le contrôle de la formation d'un gradient de morphogène dans un contexte *in vivo* (Kicheva et al., 2012).

Dpp est un morphogène naturel de la famille TGF β produit par une bande de cellules du disque imaginal de l'aile de drosophile et qui s'étend pour former un gradient le long de l'axe antéro-postérieur (Affolter and Basler, 2007). Une étude récente à partir d'un système synthétique de morphogène dans le disque imaginal, a permis de tester les conditions minimales requises pour la formation de ce gradient de morphogène (Stapornwongkul et al., 2020). Dans ce contexte expérimental, les auteurs ont tout d'abord montré qu'une protéine inerte (GFP, Green Fluorescent Protein, protéine fluorescente verte) peut former un gradient par diffusion en présence de molécules extracellulaires à la surface des cellules du disque imaginal capables d'interagir avec elle (nanocorps anti-GFP). Ensuite, la modification des récepteurs Dpp pour les rendre réceptifs à la GFP extracellulaire, leur a permis de montrer qu'en présence de ces récepteurs, le signal GFP peut agir comme morphogène et remplacer le signal Dpp pour organiser les axes antéro-postérieurs *in vivo*. Enfin, l'expression simultanée des récepteurs Dpp réceptifs à la GFP avec des récepteurs non transmembranaires, mobiles, eux aussi réceptifs à la GFP a encore amélioré la formation de l'aile, pour atteindre une qualité proche du type sauvage. Ces résultats, combinés à des arguments théoriques, suggèrent que la portée et l'activité du gradient de morphogène Dpp dans le disque imaginal de l'aile de drosophile impliquent une combinaison de diffusion libre et de diffusion assistée par les récepteurs non signalants mobiles qui étendent le gradient. A l'inverse, ces résultats suggèrent que le trafic vésiculaire du ligand Dpp n'est pas requis dans ce processus. Le

développement et l'application de tels systèmes synthétiques de morphogène dans différents contextes développementaux *in vivo* permettra de mieux comprendre comment les signaux sécrétés diffusent pour former un gradient de concentration qui fournit une information positionnelle adaptée et reproductible.

L'interprétation du gradient de concentration de morphogène par les réseaux de régulation génétiques permet l'établissement de frontières entre tissus

Des études moléculaires et génétiques rigoureuses sur l'action des morphogènes, ont permis de mieux comprendre comment les cellules interprètent les différences de concentration du gradient de morphogène pour se différencier en un type cellulaire spécifique, révélant des circuits intracellulaires complexes, notamment des réseaux de régulation génétique (voir pour revue (Sagner and Briscoe, 2017)).

Ces circuits ont été particulièrement bien documentés dans le contexte de la formation du tube neural le long de l'axe dorso-ventral (D-V) chez les embryons de vertébrés par le gradient de morphogène Sonic Hedgehog (Shh) (Figure 4) (Dessaud et al., 2007; Balaskas et al., 2012; Zagorski et al., 2017; Li et al., 2018).

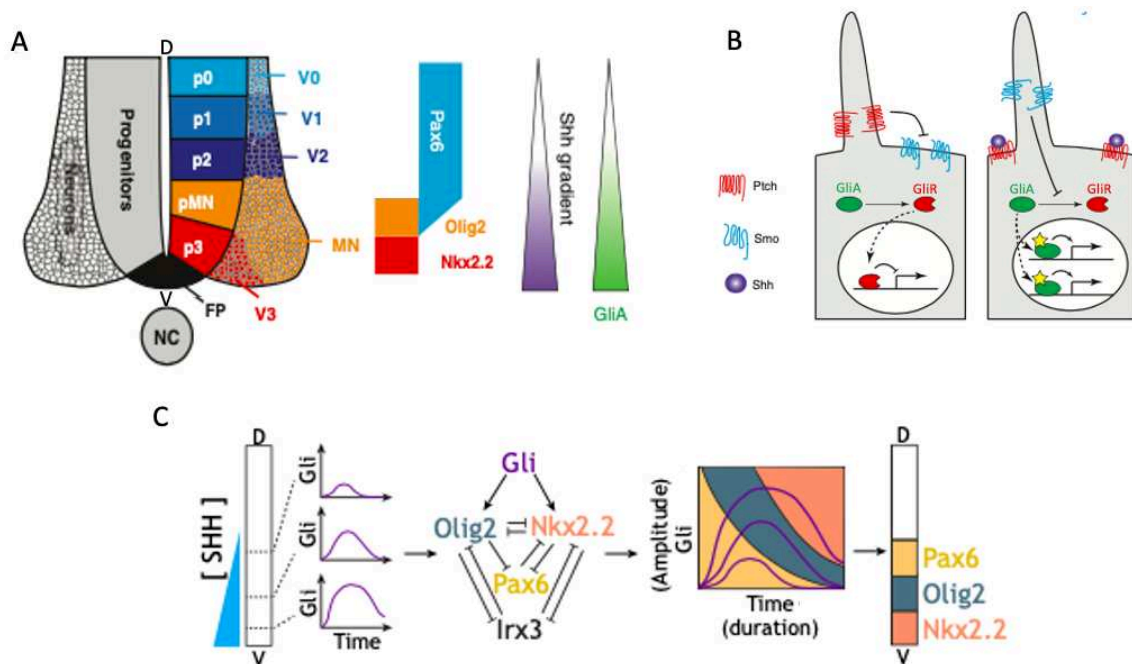


Figure 4 : La formation des domaines spatiaux des progéniteurs neuraux en réponse au gradient Shh est contrôlée par l'interprétation de la dynamique de l'activité de la voie Shh par un réseau de régulation génétique.

A) Le tube neural ventral est subdivisé en différents domaines de progéniteurs neuraux (p0-p3 et pMN) qui forment des classes distinctes d'interneurones (V0-V3) et de motoneurones (MN). La subdivision des domaines progéniteurs les plus ventraux (p1, p2, pMN, p3) est contrôlée par l'expression différentielle de Pax6, Olig2 et Nkx2.2. Ce schéma est établi en réponse à l'activité graduelle de Shh sécrété par les cellules de la notochorde (NC) et de la plaque de plancher (Floor Plate FP), qui génère un gradient de Gli activé (GliA). B) Modulation de l'activité de Gli par la signalisation Shh. C) Le signal Shh sécrété par le côté ventral du tube neural forme un gradient de concentration. Les cellules situées à différentes positions encodent les différentes concentrations de SHH en profils d'activité Gli avec une amplitude et une durée distinctes. L'activité Gli est ensuite décodée et raffinée par un circuit de régulation génétique situé en aval. Une carte 2D basée à la fois sur l'amplitude et la durée de l'activité Gli détermine la spécification du destin cellulaire, et donc les

domaines spatiaux des différents types de cellules progénitrices. D, dorsal ; V, ventral. Adapté de (Balaskas et al., 2012; Sagner and Briscoe, 2017; Li and Elowitz, 2019).

Le signal Shh est sécrété par les cellules de la notochorde ventrale et de la plaque de plancher (Floor Plate (FP)). Il diffuse pour former un gradient le long de l'axe D-V du tube neural (Figure 4 A). Ce gradient est responsable de la subdivision de la moitié ventrale du tube neural en cinq domaines de progéniteurs neuraxiaux : p3, pMN, p2, p1 et p0. Chacun d'entre eux va générer différents types de neurones : p0, p1, p2 et p3 génèrent respectivement des interneurons V0, V1, V2 et V3 (Ericson et al., 1997; Ribes and Briscoe, 2009) tandis que pMN génère des motoneurons (MN). L'identité de chacun de ces types de neurones est dictée par différentes combinaisons de facteurs de transcription exprimés dans les progéniteurs neuraxiaux. Les progéniteurs p3 expriment Nkx2.2 tandis que les progéniteurs pMN expriment Olig2 et un faible niveau de Pax6. Enfin, les progéniteurs p2 et p1 expriment un niveau élevé de Pax6 mais pas Olig2 (Figure 4 A). L'ordre d'apparition de ces facteurs de transcription au cours de la formation du tube neural est corrélé à la fois avec la concentration et la durée d'exposition au signal Shh (Roelink et al., 1995; Dessaud et al., 2007). Le signal Shh module l'activité de son effecteur intracellulaire, Gli, un facteur de transcription (Figure 2; Figure 4 B) (Ruiz i Altaba, 1998; Matise and Joyner, 1999). En l'absence de ligand, les récepteurs Patched (Ptch) bloquent l'entrée de Smoothed (Smo) dans le cilium primaire (Figure 4 B). Cela conduit à la dégradation protéasomique de Gli complet (GliA) en sa version répressive (GliR) (Park et al., 2000). Lors de l'exposition aux ligands Shh, Shh se lie à Ptch et inhibe son activité, de sorte que Smo peut entrer dans le cilium (Figure 4 B) (Ingham and McMahon, 2001). L'activité de Smo bloque la dégradation de Gli, ce qui entraîne la stabilisation et l'activation de GliA, qui active les gènes cibles dans le noyau.

Alors que le rôle du gradient Shh dans la subdivision de la moitié ventrale du tube neural en cinq domaines de progéniteurs neuraxiaux et les acteurs moléculaires impliqués sont bien connus, la façon dont les frontières précises de ces domaines sont établies alors que des cellules proches, de part et d'autre d'une frontière, reçoivent des concentrations de ligand très similaires, reste mal comprise. De récents travaux ont permis de mettre en évidence un schéma complexe qui relie la lecture de la concentration du signal Shh par les cellules du tube neural ventral à la spécification des domaines des progéniteurs neuraxiaux ventraux (Dessaud et al., 2007; Balaskas et al., 2012). L'analyse quantitative de la dynamique de la signalisation Shh, rapportée par l'activité de Gli a tout d'abord permis de révéler que les cellules situées à différentes positions du gradient encodent les différentes concentrations de Shh en profils d'activité Gli avec une amplitude et une durée distinctes (Figure 4 C) (Dessaud et al., 2007). Cette activité Gli est ensuite décodée par un réseau de répressions croisées entre les facteurs de transcription qui contrôlent le destin des cellules progénitrices, à savoir : Pax6 réprime Nkx2.2, Olig2 inhibe l'expression de Pax6 et Nkx2.2 réprime Pax6 et Olig2 (Figure 4 C) (Balaskas et al., 2012). La structure du réseau qui relie ces facteurs de transcription à l'amplitude et la durée de l'activité Gli détermine la prise de décision cellulaire, et donc les domaines spatiaux des différents types de cellules progénitrices. Par exemple, des niveaux plus élevés d'activité Gli sont nécessaires pour induire Nkx2.2 plutôt que Olig2, car la répression de Nkx2.2 par Pax6 et Olig2 doit être surmontée. Enfin, en plus d'expliquer le mécanisme qui entraîne la formation de schémas spatiaux différentiels d'expression génétique, ces travaux ont révélé que le réseau de régulation génétique confère de la robustesse à la fluctuation du signal Shh (Balaskas et al., 2012). Cependant, à ce jour, il n'est pas clair si ce type de réseau de régulation génétique, contrôlé par l'amplitude et la

durée de l'activité de la signalisation, est utilisé pour la formation de domaines spatiaux dans d'autres tissus. Des travaux chez le poisson zèbre suggèrent, par exemple, que les régions antérieures et postérieures du tube neural en formation sont sensibles à des caractéristiques distinctes de l'activité de la signalisation Shh (Xiong et al., 2018). En effet, alors que le niveau transitoire maximal et le niveau moyen d'activité de la signalisation Shh sont de bons prédicteurs du choix des destins des progéniteurs de la région antérieure, ces caractéristiques n'expliquent pas les choix des destins des progéniteurs de la région postérieure. Ces études soulèvent donc une nouvelle série d'interrogations à savoir combien de types de circuits permettant de décoder différentes concentrations de signal existent, comment ils varient selon les différents contextes tissulaires et les espèces et comment ils sont impactés par des facteurs tels que la géométrie des tissus et la vitesse de développement.

L'interprétation du morphogène par les réseaux de régulation génétiques comporte plusieurs propriétés intéressantes pour la formation des tissus. Ces réseaux permettent la conversion d'un signal graduel en domaines discrets d'expression génétique. De plus, des interactions répressives croisées entre les gènes induits dans des domaines adjacents peuvent affiner la formation des frontières entre les différents domaines et protéger le système contre les fluctuations. Enfin, les réseaux de régulation génétique complexes permettent aux cellules d'interpréter différentes propriétés du morphogène telles que les niveaux, la durée et la dynamique de la signalisation (voir pour revue (Sagner and Briscoe, 2017)). Cependant, les modèles d'études actuels sont principalement restreint à des réseaux de régulation génétique relativement simples et il est difficile de généraliser les conclusions de ces travaux.

Communication cellulaire par contact direct entre cellules émettrices et réceptrices : comment spécifier le destin d'une cellule au sein d'un champ de cellules équivalentes

Comme nous l'avons vu, l'utilisation de ligands diffusibles qui forment des gradients permet la régionalisation de grands domaines. Cependant, dans de nombreux contextes développementaux, il faut pouvoir également spécifier précisément le destin de cellules individuelles isolées. C'est le cas, par exemple, dans le neuroépithélium embryonnaire de la drosophile où des cellules ectodermiques équivalentes se différencient en une mosaïque poivre et sel de cellules neuroblastes ou épithéliales sous l'action de la signalisation Notch (Figure 5) (voir pour revue ((Perrimon et al., 2012))).

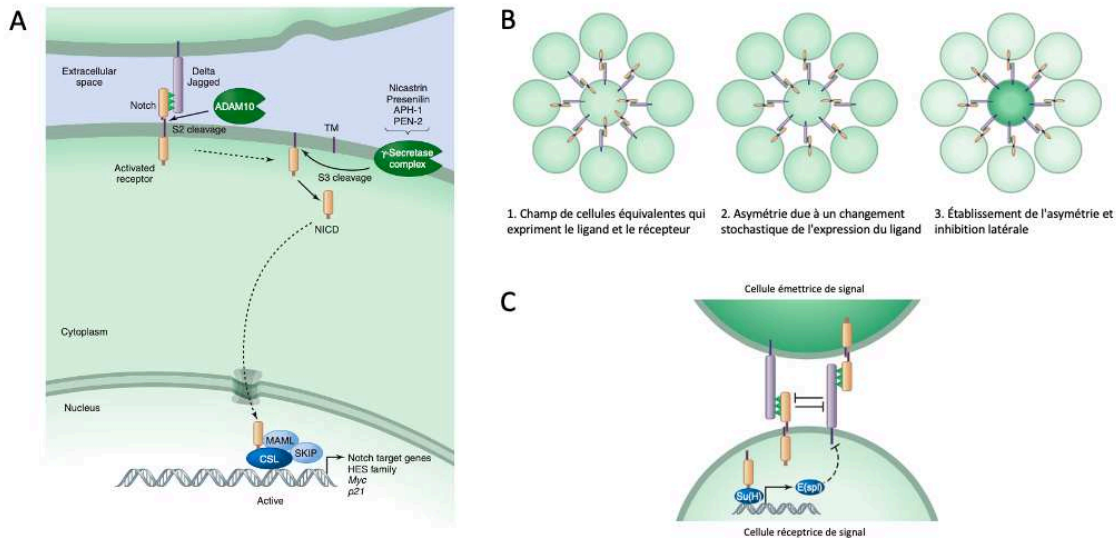


Figure 5. A) Schéma simplifié de la voie de signalisation Notch. B) L'inhibition latérale. Le processus est progressif et peut être séparé en trois étapes : (1) Initialement, toutes les cellules du champ expriment à la fois Delta et Notch et sont équivalentes. (2) Des changements stochastiques dans l'expression des gènes modifient l'équilibre des molécules du ligand et du récepteur, de sorte que la cellule du milieu exprime davantage Delta. (3) L'asymétrie est établie lorsque l'expression de Delta dans la cellule du milieu est stabilisée par des boucles de rétroaction. C) Au cours de la neurogenèse de la drosophile, l'asymétrie entre les cellules voisines est établie par une boucle de rétroaction négative qui inhibe Delta dans les cellules réceptrices du signal par l'intermédiaire de répresseurs du complexe E(spl) (action indirecte, en pointillées). La cellule qui active Delta dans l'amas proneural devient un neuroblaste, tandis que ses voisines inhibées latéralement adoptent le destin de cellule épidermique. Adapté de (Kopan, 2012; Perrimon et al., 2012).

Contrairement aux systèmes qui utilisent des ligands diffusibles pour transmettre les signaux, le récepteur Notch et ses ligands, qui appartiennent à la famille Delta/Serrate/Lag2 (DSL), sont tous deux des protéines transmembranaires liées aux cellules (Figure 5 A) (voir pour revue (Henrique and Schweisguth, 2019)). La communication se fait donc par contact direct entre cellules adjacentes (signalisation juxtacrine), ce qui lui confère un ensemble de propriétés différentes dont les mécanismes sont notamment utilisés pour la spécification de destin de cellules individuelles isolées

L'interaction des récepteurs Notch avec les ligands DSL présentés par les cellules voisines déclenche deux clivages protéolytiques au sein du récepteur. Le premier est extra-membranaire, exécuté par des métalloprotéases de la famille ADAM et permet le second clivage, qui est intra-membranaire, exécuté par des sécrétases et qui libère le domaine intracellulaire de Notch (NICD) (Figure 5 A) (Kopan, 2012). Le NICD est ensuite transporté vers le noyau où il agit comme un coactivateur transcriptionnel et active l'expression de gènes cibles. Parmi les gènes cibles, on trouve le répresseur E(spl) qui régule, indirectement, à la baisse l'expression des ligands DSL (Kopan, 2012).

Ainsi, alors qu'initialement, toutes les cellules neuroépithéliales expriment de faibles niveaux du ligand Delta et du récepteur Notch, certaines cellules commencent à exprimer des niveaux plus élevés de Delta, probablement à la suite de variations stochastiques dans les processus de transcription ou traduction (Figure 5 B) (voir pour revue (Perrimon et al., 2012)). L'activation de la signalisation Notch qui en résulte dans les cellules adjacentes entraîne une régulation négative de la production du ligand Delta dans ces cellules (via l'activation du répresseur E(spl)) qui, à son tour, réduit la signalisation Notch dans les cellules voisines (Figure 5 B C). Cette diminution de la

signalisation Notch restreint la répression de l'expression du ligand Delta qui se concentre alors à la membrane et renforce encore plus l'accroissement de la signalisation Notch dans les cellules adjacentes (Figure 5 B C) (Liao and Oates, 2017). Le système se résout donc rapidement en cellules émettrices de signal exprimant Delta et en cellules réceptrices de signal qui activent la signalisation Notch qui se différencient en cellules neuronales et épithéliales, respectivement (Figure 5 B).

En dépit de l'apparente simplicité de la voie DSL/Notch, des boucles de rétroaction complexes amplifient les différences entre cellules adjacentes et confèrent une directionnalité dans la signalisation Notch, depuis les cellules émettrices vers les cellules réceptrices, permettant la formation de destins qui sont anti-corrélés et spatialement organisés dans un champ de cellules voisines (Collier et al., 1996; Sprinzak et al., 2011). Un tel système au cours duquel une cellule qui adopte un destin particulier empêche ses voisines immédiates de faire de même (Figure 5 B) est appelé inhibition latérale et constitue l'un des rôles les mieux caractérisés de la signalisation Notch. Il est largement utilisé au cours du développement pour la formation de motifs poivre et sel (Sprinzak et al., 2011). Cependant, par exemple, dans la voie des mammifères, quatre récepteurs Notch interagissent avec un ensemble de ligands Notch, dont delta like 1 (Dll1) et delta like 4 (Dll4) et ce mécanisme ne fournit aucun moyen évident pour expliquer comment une cellule réceptrice détermine quel ligand est responsable de son activation (Nandagopal et al., 2018).

Il est intéressant de noter que des études sur la spécification des organes sensoriels dans le thorax de la drosophile indiquent que, dans certains cas, les ligands DSL peuvent activer la signalisation Notch au-delà des cellules directement adjacentes, car les cellules Delta qui envoient le signal étendent des filopodes qui peuvent atteindre des cellules situées à quelques diamètres de cellules (De Jossineau et al., 2004).

Enfin, la signalisation Notch intervient également dans le contrôle des décisions entre des cellules non équivalentes pour la formation de frontières précises (Sprinzak et al., 2011). Dans ces cas, les cellules sont initialement distinctes les unes des autres, soit parce qu'elles expriment de manière asymétrique des régulateurs de la voie Notch, soit parce que le ligand et le récepteur sont distribués de manière différentielle dans les cellules adjacentes (voir pour revue (Bray, 2006)). Par exemple, au cours de la spécification des veines de l'aile chez la drosophile, l'expression de Delta dans les régions des veines induit la signalisation Notch dans les cellules inter-veineuses pour inhiber le destin des veines.

Mécanismes émergents de communication cellulaire et directions futures : vers une compréhension opérationnelle et quantitative des voies de signalisation

Nous avons vu que la communication cellulaire est fondamentale pour le développement embryonnaire et le maintien de l'homéostasie des tissus adultes et qu'un nombre étonnamment restreint de voies de signalisation agissent de manière répétée dans des contextes cellulaires très différents pour produire des résultats divers. Au niveau moléculaire, leurs ligands, récepteurs, effecteurs intracellulaires, facteurs de transcription et modulateurs ont été identifiés et nombre de leurs interactions ont été caractérisées. Nous possédons aujourd'hui une liste presque complète des principales molécules impliquées dans ces systèmes et des nombreux processus cellulaires et tissulaires qu'ils contrôlent.

L'un des grands mystères du développement animal est la manière dont un petit nombre de signaux intracellulaires peuvent être réutilisés à différents endroits et à différents moments pour coordonner un large éventail de décisions relatives au destin cellulaire. Classiquement, la concentration absolue et la durée d'exposition au morphogène, ainsi que le passé de la cellule, sont depuis longtemps considérés comme les principaux déterminants des événements de communication cellulaire, suffisants pour induire différents destins cellulaires (Gurdon et al., 1998; Sagner and Briscoe, 2017). Cependant, un nombre croissant de travaux impliquant des approches expérimentales à l'échelle de la cellule unique avec des outils tels que des rapporteurs fluorescents de l'activité des voies de signalisation, l'imagerie quantitative ou la microfluidique combinés avec de la modélisation suggère désormais que le répertoire des modalités d'interprétation des signaux est beaucoup plus large qu'on le pensait initialement (Biasuz et al., 2018).

Par exemple, des combinaisons différentes de ligands et de récepteurs pour la même voie peuvent conduire à des résultats distincts. Dans la crête neurale du poussin, les ligands Delta-like 1 (Dll1) et Delta-like 4 (Dll4) déclenchent des dynamiques distinctes d'activité de la voie Notch, pulsatile ou soutenue, par le biais du même récepteur Notch1, induisant ainsi différents destins cellulaires (Figure 6 A) (Nandagopal et al., 2018).

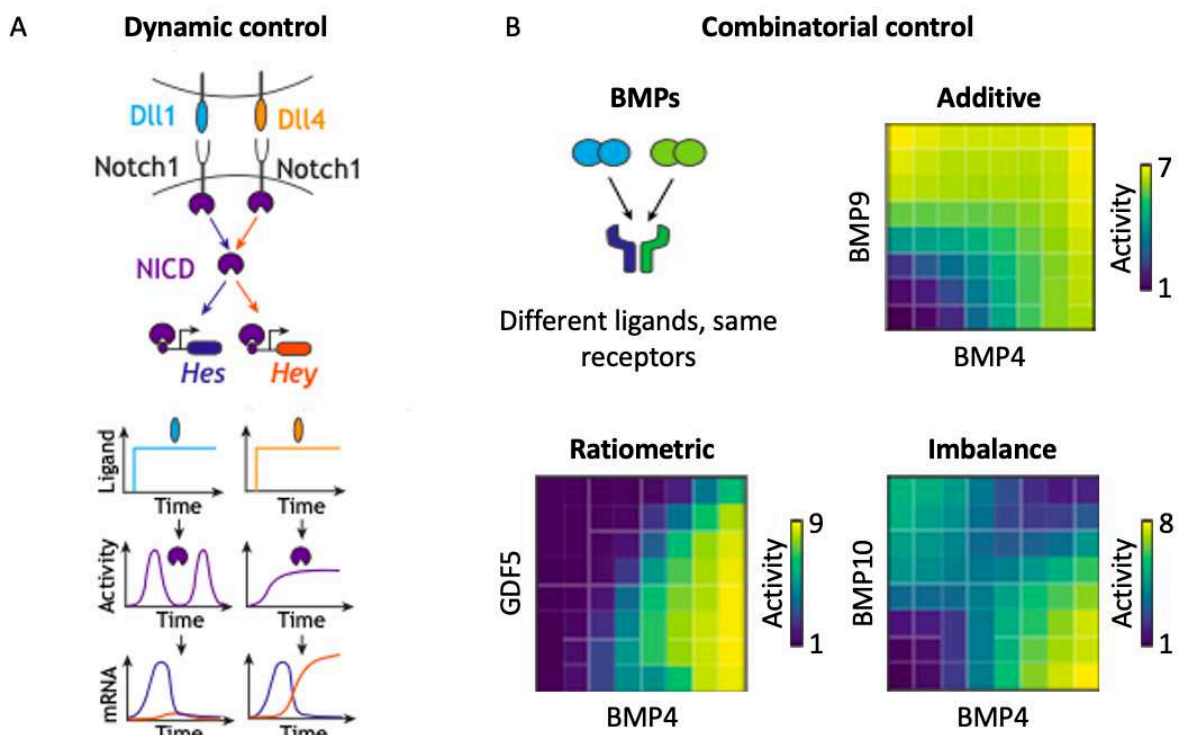


Figure 6 : Reconnaissance de l'identité d'un ligand par le contrôle dynamique ou combinatoire de l'activité de signalisation.

A) Les ligands de la voie Notch Dll1 et Dll4 se lient tous deux au récepteur Notch1 mais activent l'effecteur en aval NICD avec des dynamiques différentes. Dll1 induit des réponses pulsatiles, qui activent préférentiellement la cible transcriptionnelle Hes1, tandis que Dll4 induit des réponses soutenues, qui sont nécessaires pour activer Hey1. B) Plusieurs ligands de BMP peuvent se lier aux mêmes récepteurs. En étudiant quantitativement la réponse de la voie BMP à la présentation simultanée de plusieurs ligands, Antebi et al. ont montré que les cellules peuvent répondre aux concentrations relatives (plutôt qu'absolues) de deux ligands de manière complexe (Antebi et al., 2017). Par exemple, la voie peut répondre au rapport des concentrations de deux ligands (relation ratiométrique) ou s'activer lorsque les concentrations des ligands sont proches, de façon additive, ou éloignés d'un rapport de concentration spécifique (détection de déséquilibre). Moléculairement, ces calculs peuvent émerger d'interactions compétitives entre récepteurs et ligands qui forment des complexes de signalisation protéique distincts, lesquels phosphorylent ensuite les effecteurs en aval à des vitesses différentes. Adapté de (Li and Elowitz, 2019).

La situation devient plus complexe lorsque les cellules sont exposées à plusieurs ligands et récepteurs pour la même voie. C'est le cas par exemple pour la voie BMP pour laquelle dans de nombreux processus développementaux, de multiples ligands BMP sont présents et les cellules individuelles expriment généralement de multiples variantes de récepteurs (Danesh et al., 2009; Salazar et al., 2016). En étudiant quantitativement la réponse de la voie BMP à l'exposition simultanée de plusieurs ligands BMP, Antebi et al. [2017], ont montré que les cellules répondent de manière complexe aux concentrations relatives de deux ligands (Figure 6 B) (Antebi et al., 2017). Par ailleurs, le changement des récepteurs exprimés peut modifier la réponse observée pour un ensemble donné de ligands. Ces résultats suggèrent donc que pour des voies de communication avec de multiples ligands et récepteurs, l'activité de la signalisation est encodée dans un espace combinatoire à haute dimension, les cellules pouvant moduler leur réponse par le biais de leur profil d'expression des récepteurs pour détecter les informations codées dans les combinaisons de ligands.

Enfin, les choses ne sont pas plus simples à l'intérieur de la cellule. Dans les voies Wnt et TGF- β , par exemple, la concentration du ligand contrôle la variation du niveau des effecteurs intracellulaires de ces voies plutôt que leur niveau absolu pour déterminer les réponses cellulaires, un régime plus robuste à la variabilité cellulaire des niveaux d'expression des composants de la voie de signalisation (Goentoro and Kirschner, 2009; Frick et al., 2017).

Dans l'ensemble, ces travaux révèlent deux nouveaux modèles de communication cellulaire : dans le cas d'un signal unique, les réponses cellulaires peuvent être fonction de caractéristiques dynamiques de l'activité de la voie telles que l'amplitude, la durée ou la fréquence, un modèle appelé "contrôle dynamique" (Purvis and Lahav, 2013), ou bien, les réponses cellulaires peuvent être spécifiées par le chevauchement spatial de combinaisons particulières de ligands par des cellules exprimant des combinaisons particulières de récepteurs, un modèle appelé "contrôle combinatoire" (Rahimi et al., 2016; Johnson and Toettcher, 2019).

Cependant, de nombreuses questions simples, liées à la compréhension quantitative des mécanismes bien étudiés au niveau moléculaire et cellulaire restent sans réponse. C'est le cas par exemple au niveau de l'initiation de la réponse à un signal : "Combien de complexes ligand-récepteur sont nécessaires pour déclencher un signal ?"; "Comment les complexes ligand-récepteur sont-ils répartis dans l'espace dans un champ de cellules qui répondent ?"; Comment les concentrations absolues de ligands et de récepteurs, ainsi que la cinétique de leurs interactions, ont un impact sur les aspects quantitatifs et qualitatifs de l'activité de la signalisation ?"; "Comment la concentration du

ligand et le temps d'exposition sont-ils intégrés par les cellules ?”; “Comment le contexte cellulaire influence-t-il l'interprétation des signaux ?”. Les mécanismes par lesquels l'activité des voies de signalisation modifie les activités des facteurs de transcription en aval et la machinerie de transcription pour contrôler l'expression des gènes sont également relativement inexplorés. Les réponses à ces questions permettront de mieux comprendre comment chaque voie peut générer des résultats aussi diversifiés, ce qui fait la robustesse des systèmes de signalisation, et comment des perturbations de la transduction des signaux peuvent provoquer des anomalies du développement.

Les études quantitatives sur la transduction du signal au niveau de la cellule unique n'en sont qu'à leurs débuts et permettent d'adresser certaines des questions opérationnelles les plus fondamentales - mais encore obscures - concernant ces voies (Antebi et al., 2017). Les questions opérationnelles portent moins sur la description d'interactions moléculaires spécifiques que sur la façon dont la voie dans son ensemble perçoit, traite et représente les signaux extracellulaires dans la cellule, qui ne peut être déduite de la seule connaissance des interactions moléculaires. Sur le plan expérimental, le déchiffrement de ces codes nécessite de combiner un contrôle quantitatif des signaux d'entrée dans l'espace et dans le temps avec le suivi simultané de l'activité des voies dans divers contextes cellulaires. Les réponses à ces questions sont essentielles tant pour la compréhension fondamentale que pour les applications émergentes qui cherchent à utiliser ces voies pour diriger les cellules vers des destins spécifiques de manière prévisible. Les mesures quantitatives sont également nécessaires pour modéliser de manière plus réaliste le répertoire croissant des modalités de signalisation entre les cellules (Biasuz et al., 2018).

LA VOIE DE COMMUNICATION ERK

Parmi les voies de signalisation, la voie ERK (Extracellular signal Regulated Kinase), qui conduit à l'activation de la kinase ERK dans le noyau, est une voie majeure qui participe au contrôle de nombreux processus cellulaires tels que la différenciation, la transcription, la progression du cycle cellulaire, la prolifération, la migration, la survie, l'adhésion et le métabolisme et qui est mutée dans la plupart des cancers (Sanchez-Vega et al., 2018).

La voie de signalisation ERK est utilisée au cours de l'embryogenèse chez les vertébrés et invertébrés (Lavoie et al., 2020) et est largement étudiée en biologie du développement (Patel and Shvartsman, 2018). Elle est, par exemple, importante dans des processus précoces ou tardifs de l'embryogenèse des mammifères tels divers que la placentation, la gastrulation, la spécification des neurones, la vasculogenèse et le développement des membres (Deng et al., 1994; Min et al., 1998; Hatano et al., 2003; Newbern et al., 2008; Srinivasan et al., 2009). Dans les études de biologie du développement de modèles animaux, la perte ou le gain de fonction des composants de la voie ERK entraînent de graves défauts de développement. Des mutations germinales dans les gènes de la voie ERK sont associées à des troubles du développement chez l'homme appelés RASopathies (Rauen, 2013). Les individus affectés présentent des malformations congénitales, notamment des malformations craniofaciales et des anomalies cardiaques, ainsi qu'un retard cognitif.

Composition moléculaire de la voie ERK

La cascade de signalisation depuis l'activation du récepteur transmembranaire jusqu'à la phosphorylation des cibles nucléaires de ERK est bien décrite (Figure 7) (Kolch, 2000).

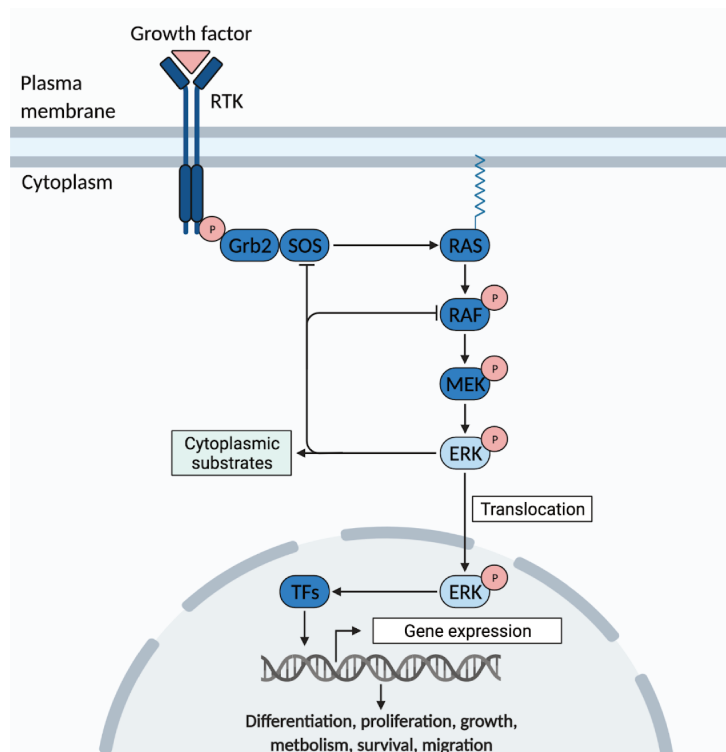


Figure 7 : Schéma simplifié de la signalisation ERK. Les récepteurs à tyrosine kinase activent d'autres voies (notamment PLC γ ,PI3K) qui ne sont pas représentées ici.

L'activation de récepteurs à tyrosine kinase (RTK) par des ligands extracellulaires est fréquemment employée pour déclencher l'activation de ERK. Lors de l'activation des récepteurs, les facteurs d'échange de guanine cytoplasmiques (GEF, typiquement SOS) sont recrutés vers la membrane plasmique où ils activent une petite GTPase RAS. RAS est ancrée dans la membrane et présente deux états de régulation : un état inactif lié au GDP et un état actif lié au GTP. L'état de régulation de RAS est contrôlé par les activités opposées de facteurs d'échange de nucléotides de guanine (GEF), dont SOS, qui catalyse l'échange du GDP pour le GTP, et les protéines d'activation de la GTPase (GAP), qui augmentent le taux d'hydrolyse du GTP en GDP. L'activation de RAS déclenche la cascade de sérine thréonine kinase RAF/MEK/ERK. Tout d'abord RAS-GTP ancré à la membrane plasmique y recrute RAF. Concentré à la membrane, RAF est phosphorylé et il dimérise, ce qui conduit à la formation de son site catalytique complet et actif (Lavoie and Therrien, 2015). RAF à son tour phosphoryle et active MEK qui ensuite phosphoryle et active ERK.

Alors que les familles des kinases RAF et MEK ont une spécificité de substrat étroite, ERK catalyse la phosphorylation de centaines de substrats cytoplasmiques et nucléaires dont des facteurs de transcription (Morrison, 2012). La stimulation de la voie ERK induit une translocation rapide de ERK vers le noyau (Watson et al., 2018).

Enfin, de multiples boucles de rétroaction entre ces composants ont été décrites, et des phosphatases répriment également de multiples composants (Shin et al., 2009).

La dynamique de l'activité de la voie ERK joue un rôle important dans le contrôle des réponses cellulaires induites

Comme de nombreuses autres voies de signalisation, la configuration convergente de la voie ERK pose un premier problème conceptuel : malgré la régulation étroite et distincte des ligands qui déclenchent leurs récepteurs RTK respectifs, tous les signaux passent par ERK (Figure 8 A) (Shilo and Barkai, 2017). Comment alors des réponses distinctes à ces différents signaux activant la voie ERK sont-elles codées ?

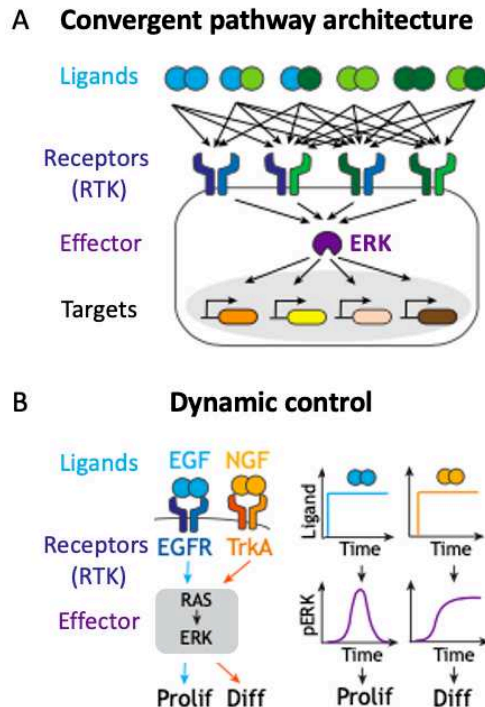


Figure 8 : La dynamique de la voie ERK discrimine l'identité du couple ligand-récepteur. A) Comme de nombreuses voies de signalisation, la voie ERK présente une structure convergente : de nombreux ligands interagissent avec de nombreux récepteurs RTK de manière multiple et les informations convergent au niveau de ERK. B) Malgré cette convergence, l'identité du couple ligand-récepteur peut être discriminée pour induire des programmes d'expression génique différentiels. Par exemple, le traitement des cellules PC12 par EGF/NGF induit une activité de ERK transitoire/soutenue qui est corrélée à la prolifération/différenciation (Marshall, 1995). Adapté de (Li and Elowitz, 2019).

Des études fondamentales sur une lignée cellulaire de rat (PC12) ont permis de montrer que les caractéristiques dynamiques de la signalisation ERK jouent un rôle important dans la reconnaissance et la transmission de ces signaux. En effet, alors que le traitement de ces cellules PC12 par le ligand EGF provoque leur prolifération, le traitement par le ligand NGF induit lui une différenciation neuronale (Greene and Tischler, 1976; Huff et al., 1981), bien que les deux voies passent par ERK (Figure 8 B). Des expériences d'immuno-marquages sur la forme phosphorylée et activée de ERK dans des populations synchronisées de ces cellules ont montré que la prolifération cellulaire induite par le ligand EGF repose sur une activation ERK transitoire, tandis que la différenciation neuronale induits par le ligand NGF implique une activation de ERK soutenue (Marshall, 1995). Dans leur ensemble, ces travaux ont conduit à formuler l'hypothèse selon laquelle la dynamique de l'activité de ERK pourrait représenter les identités des couples ligand-récepteur pour induire les phénotypes cellulaires distincts observés (Marshall, 1995). De nombreuses études ultérieures dans le même système expérimental ont révélé comment le codage et le décodage de la durée de la phosphorylation de ERK découlent à la fois de l'activité des complexes ligand-récepteur et des mécanismes de rétroaction positive et négative qui agissent à plusieurs niveaux de la voie (Traverse et al., 1992; Marshall, 1995; Kao et al., 2001; Murphy et al., 2002, 2004; Sasagawa et al., 2005; Santos et al., 2007; Whitmarsh, 2007; Uhlitz et al., 2017; Blum et al., 2019).

La façon dont cette architecture façonne la réponse dynamique de ERK en réponse aux signaux d'entrée reste cependant à ce jour mal comprise (Patel and Shvartsman, 2018). Nos connaissances

actuelles de la voie ERK sont pour la plupart statiques, sans compréhension intégrée de la dynamique spatio-temporelle. Les études dans ce domaine sont également motivées par l'existence des RASopathies : des troubles du développement chez l'homme d'origine génétique associés à une dérégulation de l'activité de la voie ERK (Rauen, 2013). En effet, bien que l'on pense que ces anomalies reflètent des changements quantitatifs dans les dynamiques spatio-temporelles des signaux ERK développementaux, nous comprenons mal les limites entre lesquelles la voie ERK doit fonctionner pour assurer un développement normal (Goyal et al., 2017). Quelles sont, par exemple, les amplitudes et durées maximales et minimales de l'activité ERK qui conduisent encore à un développement normal ? Dans quelle mesure l'activité de ERK doit-elle être limitée dans l'espace ? La quantification de ces paramètres est nécessaire pour sonder la robustesse intrinsèque des signaux ERK développementaux et mieux comprendre l'origine de ces syndromes (Patel and Shvartsman, 2018).

Les biosenseurs génétiquement codés permettent de faire la lumière sur la voie ERK

Classiquement, l'activation spatio-temporelle de ERK au cours de l'embryogenèse a été étudiée dans des cellules fixées à l'aide d'immuno-marquages avec des anticorps dirigés contre des formes phosphorylées de ERK (Shilo and Barkai, 2017). L'avancée récente des biosenseurs fluorescents génétiquement codés, a permis de visualiser les systèmes de transduction du signal *in vivo* au niveau de la cellule unique. Ces études ont ouvert la voie à l'étude quantitative de la dynamique spatiale et temporelle de ERK (Biasuz et al., 2018; Nakamura et al., 2021). Les protéines fluorescentes et les biosenseurs ont permis la visualisation de diverses molécules biologiques (protéines, lipides, nucléotides, ions, etc.), de réactions biochimiques (association, dissociation, réactions enzymatiques, etc.) et de propriétés biophysiques (diffusivité, température, force, etc.) ayant lieu à l'intérieur d'une cellule avec de hautes résolutions temporelles et spatiales (Zhang et al., 2002; Miyawaki, 2003). Les biosenseurs actuels pour ERK sont principalement de deux types : les biosenseurs basés sur le transfert d'énergie par résonance de Förster (FRET) entre deux molécules fluorescentes (Figure 9 A) (Komatsu et al., 2011) et les biosenseurs basés sur un seul fluorophore, dont la localisation sub-cellulaire dépend de l'activité de ERK (Figure 9 B C) (Regot et al., 2014).

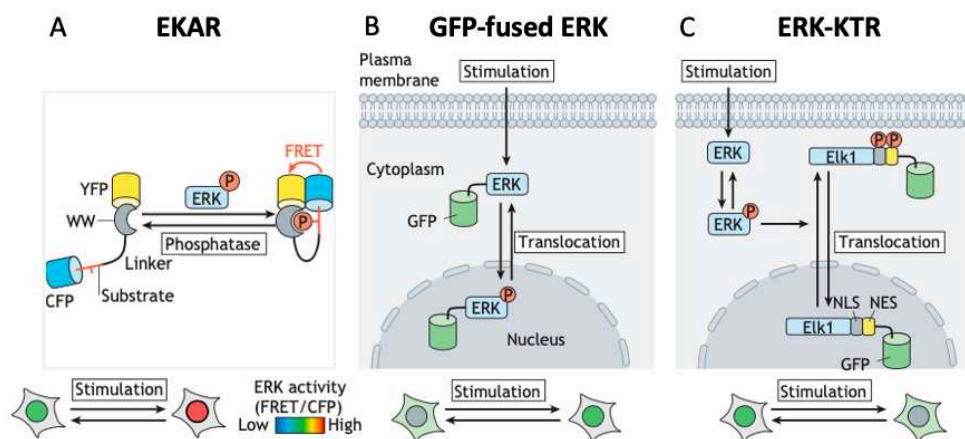


Figure 9 : Biosenseurs génétiquement codés pour ERK. Schémas illustrant les constructions et les changements de fluorescence en réponse à l'activation de ERK pour A) EKAR. EKAR est composé : d'une paire de protéines fluorescentes : bleue (Cyan Fluorescent Protein, CFP) et jaune (Yellow Fluorescent Protein, YFP) capables de FRET (le spectre d'émission de la CFP chevauche le spectre d'excitation de la YFP); d'un peptide phosphorylable contenant un motif

d'accueil pour ERK, d'un domaine WW capable de se lier au peptide phosphorylé par ERK et d'un peptide flexible constitué d'une répétition de glycines. L'activation de ERK conduit à la phosphorylation du peptide phosphorylable et à sa liaison au domaine WW et le changement conformationnel qui en résulte déclenche le FRET entre le fluorophore donneur (CFP) et accepteur (YFP). Différentes versions successives d'EKAR, dont EKAREV, ont été produites visant à augmenter le gain du signal FRET (Harvey et al., 2008; Fritz et al., 2013; Vandame et al., 2014). B) ERK fusionné à une GFP et C) ERK-KTR. Adapté de (Nakamura et al., 2021).

Biosenseurs basés sur le FRET

Le FRET est un processus de transfert d'énergie dans lequel l'énergie d'excitation d'un fluorophore donneur est transférée à un fluorophore accepteur lorsque ces deux molécules sont suffisamment proches (de l'ordre du nm) (Förster, 1948). A partir des propriétés des fluorophores, des biosenseurs ERK basés sur le FRET ont été conçus pour mesurer l'activité ERK à partir des variations des signaux FRET.

Parmi eux, le biosenseurs EKAR (ERK Kinase Activity Reporter, rapporteur de l'activité kinase de ERK) a eu un impact particulièrement important (Figure 9 A) (Harvey et al., 2008). Ils possèdent cependant certaines limitations. Ils ne sont pas totalement spécifiques de ERK. En effet, le peptide phosphorylable des biosenseurs EKAR (Figure 9 A), dérivé à l'origine de Cdc25, est phosphorylé de manière non spécifique par la kinase cycline-dépendante 1 (CDK1), ce qui entraîne une augmentation artificielle du signal FRET immédiatement avant le début de la phase M et une diminution du signal FRET à la fin de la phase M, même dans des conditions d'inhibition de MEK (Aoki et al., 2013). De plus, les approches FRET sont également associées à plusieurs limitations inhérentes au processus physique du FRET ainsi qu'aux technologies utilisées pour mesurer le FRET. La plus grande de ces limitations des mesures FRET est sans doute le faible rapport signal/bruit associé à ce type d'imagerie par rapport à l'imagerie d'un seul marqueur fluorescent, principalement à cause la perte d'énergie associée au processus FRET et au fait que deux molécules fluorescentes contribuent au signal FRET mesuré (Leavesley and Rich, 2016). Leur cinétique de désactivation est également lente et ils ne reflètent pas précisément la désactivation de ERK, très probablement parce que la conformation fermée du capteur FRET est très stable et que les phosphatases ne peuvent accéder à ses sites phosphorylés (Komatsu et al., 2011). Enfin, ils nécessitent deux protéines fluorescentes et limitent ainsi le nombre de composants qui peuvent être observées simultanément dans une cellule donnée alors que les réseaux de signalisation sont connus pour être hautement intégrés. Cependant, il est intéressant de noter que l'utilisation des biosenseurs de type FRET permet de suivre l'activité de ERK avec une résolution spatiale au sein d'une cellule.

Biosenseurs basés sur un seul fluorophore

Contrairement aux biosenseurs conçus pour mesurer l'activité de ERK à partir des signaux FRET entre deux fluorophores, les biosenseurs basés sur un fluorophore unique permettent de mesurer l'activité de ERK par des changements dans la localisation subcellulaire du rapporteur fluorescent (Nakamura et al., 2021).

L'activation de ERK dans des cellules vivantes a classiquement été suivie par l'accumulation nucléaire de ERK fusionné à une protéine fluorescente (Figure 9 B) (Nakamura et al., 2021). Ce biosenseur est encore largement utilisé, mais comme l'ajout d'ERK induit une rétroaction négative excessive (Shin et

al., 2009), il est fortement recommandé d'utiliser un ERK muté pour son activité kinase (Nakamura et al., 2021). Enfin, l'enrichissement de ERK dans le noyau n'implique pas nécessairement un niveau élevé d'activité ERK nucléaire (Wilson et al., 2017).

Un nouveau système de Kinase Translocation Reporters (KTR, rapporteurs de l'activité kinase par translocation) a été développé pour suivre l'activité en temps réel de kinases, dont ERK, qui jouent un rôle fondamental dans de nombreuses voies de signalisation (Figure 9 C) (Regot et al., 2014). Chaque rapporteur est composé d'un substrat de kinase, qui assure sa spécificité, fusionné à un signal de localisation nucléaire bipartite (bNLS), un signal d'exportation nucléaire (NES) et une protéine fluorescente (Figure 9 C). L'introduction de la charge négative par la phosphorylation du substrat module l'import/export nucléocytoplasmique, régulant ainsi la distribution subcellulaire du rapporteur qui est mesurée comme un proxy de l'activité kinase. Cette méthode récapitule de manière précise à la fois les dynamiques d'activation et de désactivation pour chaque kinase mais ne permet cependant pas de visualiser la localisation subcellulaire des kinases activées (Regot et al., 2014) (Figure 9 C). Enfin, les biosenseurs à base de fluorophore unique, tels que les KTR, en libérant une longueur d'onde d'acquisition par rapport aux biosenseurs de type FRET, permettent l'imagerie multiplexée de l'activité kinase au niveau de la cellule unique (Maryu et al., 2016; Jacques et al., 2021) et la combinaison avec les systèmes optogénétiques qui permettent de contrôler finement l'activité de protéines de signalisation dans des cellules vivantes grâce à la lumière (Goglia et al., 2020; Johnson et al., 2020; Gagliardi et al., 2021).

Propriétés de transmission des signaux dans le module RAS/ERK

Comme nous l'avons vu, la cascade RAS/MEK/ERK est un module de signalisation partagé qui est activé par de nombreux signaux extracellulaires et qui induit des phénotypes cellulaires bien distincts tels que la prolifération, la différenciation (Figure 8). Ces différentes réponses semblent être dynamiquement encodées : la prolifération cellulaire étant associée à une activation de ERK transitoire tandis que la différenciation implique une activation de ERK soutenue (Figure 8 B) (Marshall, 1995). Cependant la façon dont ce module RAS/ERK façonne une réponse dynamique reste mal comprise (Patel and Shvartsman, 2018).

L'utilisation d'un biosenseur fluorescent pour l'activité de ERK (ERK2-tagué) combiné à un système optogénétique permettant l'activation quantitative spatio-temporelle précise de RAS par la lumière a permis de sonder les propriétés intrinsèques de transmission du signal dans ce module RAS/ERK (Figure 10 A) (Toettcher et al., 2013). L'activation sélective de RAS par la lumière repose sur le système optogénétique Phy-PIF et implique le recrutement membranaire dépendant de la lumière bleue du domaine catalytique de SOS fusionné à PIF via PhyB lié à la membrane (Figure 10 A). Cet outil est appelé opto-SOS.

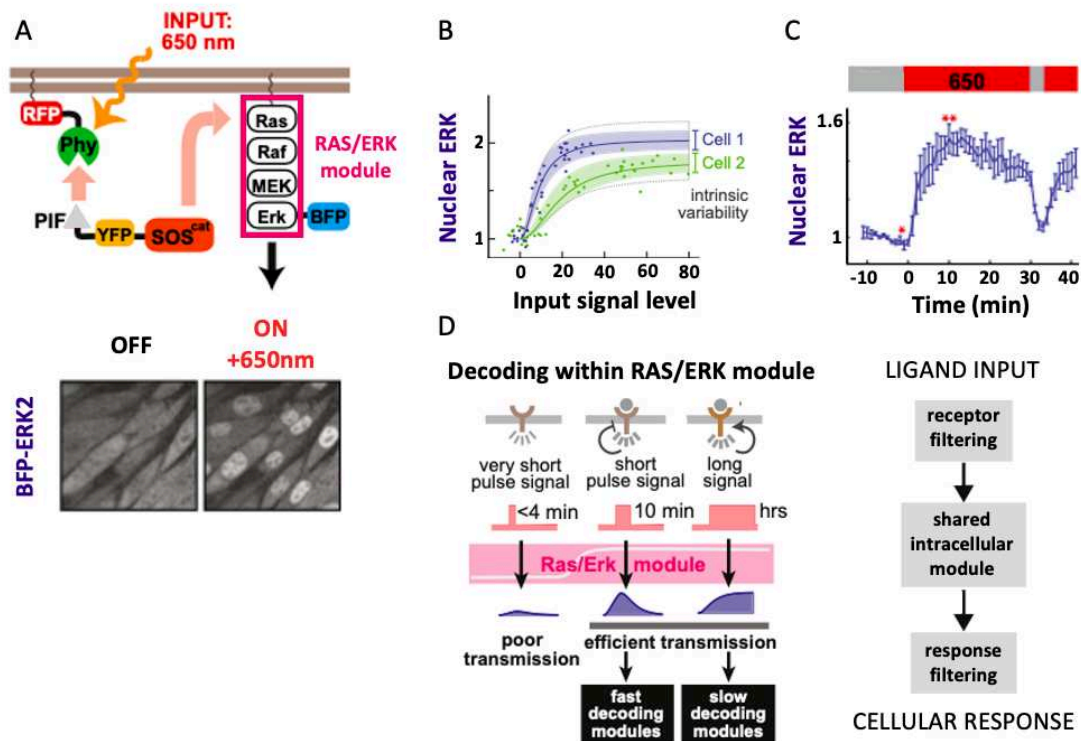


Figure 10 : Le module Ras/Erk transmet fidèlement des signaux dynamiques de 4 minutes à 2 heures. A) La lumière entraîne l'hétérodimérisation de Phy, localisé dans la membrane, avec PIF fusionné à SOScat, ce qui conduit à l'activation de Ras et à la translocation nucléaire de BFP-ERK2. B) Doses-réponses de l'activité de ERK en fonction des niveaux de signaux en entrée pour deux cellules représentatives. L'intervalle de confiance de la courbe dose-réponse pour chaque cellule est représentée. C) Variations de l'intensité nucléaire de ERK en réponse à des signaux lumineux au cours du temps (n = 3 cellules). D) Modèle de transmission de signaux par le module RAS/ERK. Les stimuli d'entrée très courts sont filtrés et ignorés par le module RAS/ERK, tandis que les entrées de quelques minutes à quelques heures sont fidèlement et efficacement transmises par la cascade. Les schémas distincts de l'activité ERK sont ensuite probablement décodés par les modules de filtrage dynamique en aval (boîtes noires). Adapté de (Toettcher et al., 2013).

Dans des cultures cellulaires de fibroblastes de souris (NIH-3T3), ce système expérimental a tout d'abord permis de montrer que, bien qu'il existe une variabilité intercellulaire de la réponse, le module RAS/ERK transmet avec précision des informations quantitatives sur le niveau du signal d'entrée pour chaque cellule (Figure 10 B) (Toettcher et al., 2013). En mesurant la corrélation entre les modulations de l'activation de RAS et l'activation de ERK, le délai de transmission a également pu être extrait : il faut 3 min pour transmettre le signal depuis opto-SOS jusqu'à la translocation nucléaire de ERK. L'activation de ERK diminue en quelques minutes après l'arrêt du signal d'entrée (Figure 10 C). En mesurant la transmission du signal du module RAS/ERK sur toute une gamme d'échelle de temps physiologiquement pertinente (allant de 1 minute à plusieurs heures), les auteurs ont également montré que, alors que les signaux d'entrées de quelques minutes ou moins sont fortement atténués, les signaux d'entrées avec un temps d'activation allant de 4 minutes à plusieurs heures sont eux fidèlement transmis avec une amplitude égale (Toettcher et al., 2013). Une telle capacité de filtrage dans le module peut être intéressante pour empêcher des réponses à des événements stochastiques parasites autour de la cellule. A l'inverse, les signaux d'entrées qui persistent au-delà de quelques minutes, et qui correspondent probablement à de véritables événements d'activation des récepteurs, sont efficacement transmis par le module RAS/ERK.

Ces résultats indiquent qu'il peut exister une organisation hiérarchique globale de la manière dont les voies de signalisations cellulaires traitent dynamiquement les signaux (Figure 10 D). Les composants membranaires en amont, tels que les récepteurs, peuvent coder des informations de manière dynamique en convertissant des signaux extracellulaires distincts en signaux intracellulaires avec des fréquences intrinsèques spécifiques (par exemples, par le biais d'un contrôle par rétroaction de la durée de signalisation des récepteurs ou par le biais de l'activation différentielle de modules de signalisation parallèles selon les couples ligands-récepteurs). Ces informations sont ensuite transmises efficacement, et potentiellement amplifiées, par un module intracellulaire partagé, comme le module RAS/ERK. Enfin, l'information transmise fidèlement par un tel module peut être interprétée par des modules de décodage dynamiques situés en aval ou en parallèle du module central. Par exemple, dans les cellules NIH-3T3, des programmes de signalisation intracellulaires qui répondent de manière différentielle à la dynamique du module RAS/ERK ont été identifiés (Toettcher et al., 2013). En revanche, peu de choses sont connues sur les mécanismes à l'interface signalisation/expression génétique qui permettent de décoder la dynamique du module RAS/ERK pour induire des schémas précis et reproductibles d'expression génétique (Johnson and Toettcher, 2019).

Au cours du développement précoce de la souris, la dynamique de l'activité de ERK après la mitose détermine la spécification du destin entre les cellules de l'épiblaste et de l'endoderme primitif

L'étude de l'activité de la voie ERK dans des cellules isolées néglige les rôles de la géométrie, de la polarité des tissus, de la sécrétion de ligands et de l'expression de récepteurs spatialement organisés dans le comportement de signalisation (Butler and Wallingford, 2017). Il est donc essentiel d'étudier le rôle de l'activité de la voie ERK dans des contextes développementaux multicellulaires et organisés spatio-temporellement.

La signalisation ERK est importante dans de nombreux aspects du développement des mammifères (Lavoie et al., 2020). Cependant, l'imagerie en direct de la dynamique d'activation de ERK à partir des biosenseurs n'en est encore qu'à ses débuts.

De récents travaux ont permis d'établir une lignée de souris exprimant le ERK-KTR et d'observer la dynamique de ERK dans le temps pendant la différenciation des blastocystes de souris (Pokrass et al., 2020).

Le blastocyste de souris préimplantatoire subit deux événements de différenciation consécutifs et produit trois types de lignage cellulaire (Figure 11). Le premier événement de différenciation est régi par la signalisation Hippo qui sépare le trophoctoderme (TE) et la masse cellulaire interne (MCI) (Nishioka et al., 2009). Ensuite, les cellules de l'épiblaste (Epi) et de l'endoderme primitif (PrE) apparaissent au sein de la masse cellulaire interne (MCI). Cet événement est marqué par l'expression mutuellement exclusive de NANOG dans les cellules de l'épiblaste (Epi) et de GATA6 dans les cellules de l'endoderme primitif (PrE) (Figure 11) (Chazaud et al., 2006; Guo et al., 2010). La signalisation FGFR/ERK joue un rôle central dans la différenciation des cellules de l'épiblaste (Epi) et de l'endoderme primitif (PrE) : l'inhibition de la signalisation FGFR/ERK par des inhibiteurs chimiques ou des manipulations génétiques empêche toutes les cellules de la masse cellulaire interne (MCI) d'adopter le destin PrE, et toutes les cellules de la masse cellulaire interne (MCI) deviennent Epi. A

l'inverse, suite au traitement avec une forte dose de FGF toutes les cellules de la masse cellulaire interne (MCI) adoptent le destin PrE (Chazaud et al., 2006; Yamanaka et al., 2010).

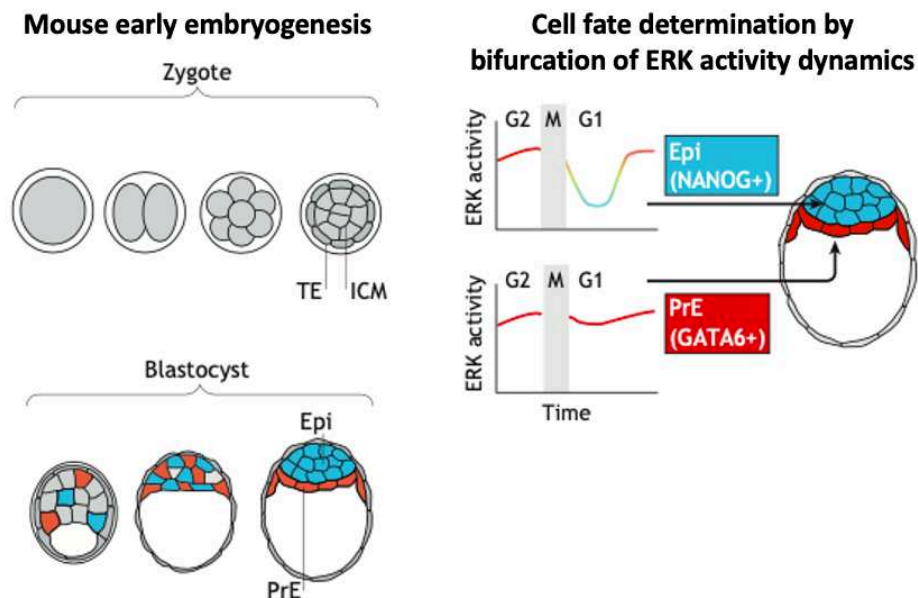


Figure 11 : La dynamique de l'activité ERK pendant la phase G1 détermine la spécification du destin cellulaire entre les cellules de l'épiblaste (Epi) ou de l'endoderme primitif (PrE). Gauche : schéma de

l'embryogenèse précoce de la souris. Les cellules rouges et bleues correspondent aux cellules spécifiées à l'endoderme primitif (PrE) et à l'épiblaste (Epi), respectivement. TE = trophoblaste. MCI = masse cellulaire interne. Droite : à la fin de la mitose, certaines cellules montrent une inactivation transitoire de ERK après la division cellulaire et se différencient en cellules de l'épiblaste (Epi, expriment NANOG), tandis que d'autres conservent une activité de ERK élevée et se différencient en cellules de l'endoderme primitif (PrE, expriment GATA6). Adapté de (Nakamura et al., 2021).

L'imagerie time-lapse de l'activité de ERK, toutes les 15 minutes pendant 6 heures, grâce au ERK-KTR suivie d'immunomarquage pour les marqueurs des destins cellulaires NANOG (pour le destin Epi) et GATA6 (pour le destin PrE), a permis de relier les schémas temporels de l'activité de ERK aux destins cellulaires terminaux. En effet, l'activité de ERK dans les blastomères de la masse cellulaire interne (MCI) bifurque à la fin de la mitose : certaines cellules montrent une inactivation transitoire de ERK après la division cellulaire, tandis que d'autres conservent une activité de ERK élevée (Figure 11) (Pokrass et al., 2020). Cette bifurcation de l'activité de ERK à la sortie de la mitose régit l'engagement de la spécification vers les destins Epi ou PrE : l'expression de NANOG et GATA6 observée par immunofluorescence après imagerie time-lapse est directement corrélée avec l'activité ERK à la fin de la sortie de la mitose (Figure 11). L'activité de ERK à la fin de l'imagerie time-lapse ne montre cependant aucune corrélation avec le niveau d'expression des marqueurs du destin cellulaire. Ainsi, il semble que ce soit la dynamique temporelle de l'activité ERK mais pas l'activité au point final de la mesure qui joue un rôle central dans la détermination des destin cellulaires au sein de la masse cellulaire interne (MCI).

Ces résultats révèlent donc le rôle de la dynamique de l'activité de ERK au cours de cet événement du développement embryonnaire précoce chez la souris et soulignent l'importance de l'imagerie de l'activité de ERK grâce aux biosenseurs dans des embryons qui se développent.

Dans l'embryon précoce de drosophile, la spécification du destin des cellules latérales dépend de l'activité intégrée de ERK dans le temps plutôt que d'autres caractéristiques de la dynamique

Afin de mieux comprendre comment la dynamique de l'activité ERK est intégrée dans les cellules de l'embryon pour moduler des événements de prise des décisions, l'utilisation du système optogénétique qui permet le contrôle de l'activation de RAS (opto-SOS) a récemment été adaptée chez la drosophile (Johnson et al., 2017; Johnson and Toettcher, 2019).

Dans l'embryon précoce de drosophile, l'activité de la voie ERK est nécessaire pour la différenciation des cellules contractiles de l'endoderme intestinal au pôle postérieur et pour former des neuroblastes intermédiaires, une sous-population de cellules progénitrices neuronales, sur la surface latérale de l'embryon (Johnson and Toettcher, 2019).

Grâce à la stimulation optogénétique de RAS, les auteurs ont tout d'abord pu montrer que la modulation de l'activité de ERK est suffisante pour changer le destin des cellules latérales en cellules progénitrices neuronales ou endodermiques : une activation transitoire de la voie ERK (30 minutes) conduit à la formation du destin neuronal tandis qu'une activité soutenue (> 1 heure) conduit à la formation de cellules de l'endoderme intestinal (Figure 12 A) (Johnson and Toettcher, 2019).

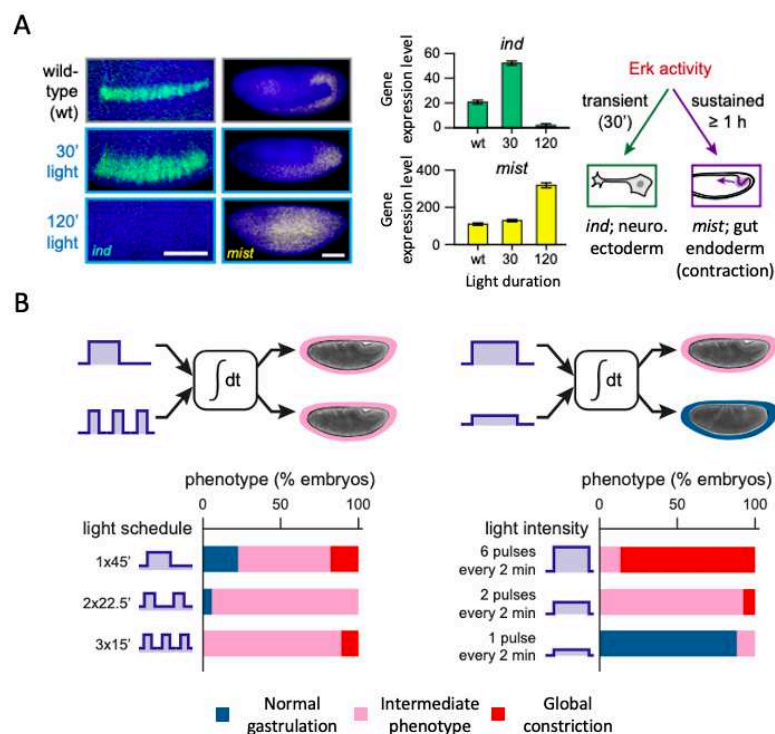


Figure 12 : L'activité de ERK intégrée détermine la prise de décision cellulaire dans l'embryon précoce de drosophile. A) Une courte impulsion d'activation de ERK (30 minutes) augmente la population de neuroblastes intermédiaires, tandis qu'une longue impulsion (120 minutes) fait basculer les mêmes cellules vers le stade endodermique intestinal. Gauche : images représentatives de RNA FISH pour *ind* (marqueur de la population de neuroblastes intermédiaires) et *mist* (marqueur des cellules contractiles de l'endoderme intestinal) dans 3 conditions : embryons "sauvages" (WT), embryons optoSOS illuminés pendant 30 minutes ou 120 minutes. Barres d'échelle, 100 μ m. Milieu : quantification des niveaux d'expression pour chaque marqueur dans les 3 conditions. La stimulation avec une impulsion lumineuse transitoire de 30 minutes a augmenté l'expression de *ind* mais n'a pas induit l'expression de *mist* dans

les cellules latérales. En revanche, 120 minutes de lumière ont complètement aboli l'expression de *ind*, et induit au contraire l'expression de *mist* dans les mêmes cellules latérales. Droite : modèle conceptuel de la façon dont les destins neurogènes et de l'endoderme intestinal induits par ERK sont spécifiés. L'activité transitoire de ERK donne lieu à l'expression de *ind* et au destin neurogène, tandis que l'activité soutenue conduit à l'expression de *mist* et à la spécification du destin endodermique intestinal. B) La spécification de ces destins est déterminée par la charge cumulée de la signalisation ERK. Les cellules contractiles de l'endoderme intestinal participent aux mouvements morphogénétiques pendant la gastrulation. La mesure du degré d'invagination au début de la gastrulation permet donc de suivre la spécification de cette population de cellules : plus le nombre de cellules ayant adopté le destin endodermique intestinal est grand, plus l'invagination sera prononcée, jusqu'à une contraction globale de l'embryon. Cet événement de prise de décision peut être déclenché par deux classes de décodeurs dynamiques de l'activité ERK : un détecteur de persistance ou un détecteur de charge cumulative. Ces deux décodeurs peuvent être distingués par leurs réponses à certains schémas lumineux. Un détecteur de persistance peut être déclenché par un signal de lumière unique, mais pas lorsque ce signal est divisé en plusieurs épisodes lumineux courts. A l'inverse, un capteur de charge cumulative se déclenche en réponse aux deux entrées à dose égale. Aussi, un capteur de charge cumulative peut être déclenché par une illumination de haute intensité, mais pas par une entrée de plus faible intensité et de même durée alors qu'un détecteur de persistance répondrait de façon similaire aux deux entrées de durée égale. Les phénotypes de gastrulation pour différents programmes d'éclairage ont été mesurés : les embryons optoSOS ont été exposés à une seule impulsion de 45 minutes de lumière ou à la même dose de 45 minutes divisée en 2 ou 3 impulsions égales et délivrée sur une période de 90 minutes (gauche) ou les embryons optoSOS ont été exposés à différentes amplitudes lumineuses sur la même durée. Chaque barre représente le pourcentage d'embryons présentant différents phénotypes de gastrulation selon la légende indiquée. Sur la période étudiée du développement embryonnaire de la drosophile, les durées des cycles cellulaires sont comprises entre 18-75 min (Ferree et al., 2016; Vergassola et al., 2018). Adapté de (Johnson and Toettcher, 2019).

Au moins deux mécanismes ont été proposés pour expliquer comment la dynamique de la voie ERK est décodée par la cellule pour spécifier son destin. Le modèle classique est celui d'un détecteur de persistance, où les gènes en aval détectent la durée d'une impulsion unique de l'activité de ERK (Murphy et al., 2002, 2004; Nakakuki et al., 2010). Un autre modèle est celui de la détection de la charge cumulative, où des destins distincts sont déclenchés lorsque le signal total est intégré (Hannanta-anan and Chow, 2016; Gillies et al., 2017).

En délivrant différents programmes d'illumination sur les cellules latérales, les auteurs ont pu tester lequel de ces mécanismes est utilisé pour la spécification des destins neuronaux ou endodermiques. Ce travail a révélé que c'est la dose totale d'activité de ERK et non la durée ou l'amplitude qui dicte cette prise de décision. Ainsi, alors que la même dose totale de stimulation de l'activité ERK peut être délivrée en une seule fois ou divisée dans le temps; des signaux de temps identiques mais de différentes amplitudes produisent des phénotypes différents (Figure 12 B) (Johnson and Toettcher, 2019).

Il manque cependant encore grandement une compréhension mécanistique de la manière dont l'activité intégrée de ERK est décodée pour déclencher des schémas précis et reproductibles d'expression génétique. Le modèle le plus simple nécessite deux comportements distincts parmi les gènes en aval de ERK (Johnson and Toettcher, 2019). Certains pourraient agir comme des accumulateurs, c'est-à-dire dont les niveaux augmentent de manière linéaire à l'activité de ERK au cours du temps, fournissant une mémoire de la quantité totale de signalisation ERK qui a été transmise. D'autres agiraient comme détecteur de seuil, en aval de l'accumulateur, comparant le signal accumulé à un seuil fixe et ne déclenchant une réponse que lorsque le seuil est franchi.

L'utilisation du ERK-KTR à révélé des oscillations de l'activité ERK modulée par la concentration de ligand au cours du patterning de la vulve chez *C. elegans*

L'application récente du ERK-KTR pour mesurer l'activité de ERK au cours de la spécification des destins cellulaires des cellules précurseurs de la vulve (CPV) chez *C. elegans* a révélé des niveaux de complexités supplémentaires dans la dynamique de ERK (de la Cova et al., 2017).

L'étude de la dynamique temporelle de l'activation de ERK à la résolution cellulaire au cours du développement embryonnaire est difficile. L'organisme modèle invertébré, *C. elegans*, présente un avantage pour aborder cette question en raison d'embryons qui se développent avec un lignage cellulaire invariable ce qui permet d'examiner la même cellule chez différents individus. La formation des cellules précurseurs de la vulve (CPV) chez *C. elegans* via la signalisation ERK est l'un des modèles développementaux les mieux compris concernant la signalisation ERK et les études intensives de ce système ont révélé de nombreux mécanismes fondamentaux de la signalisation ERK (Figure 13) (Patel and Shvartsman, 2018).

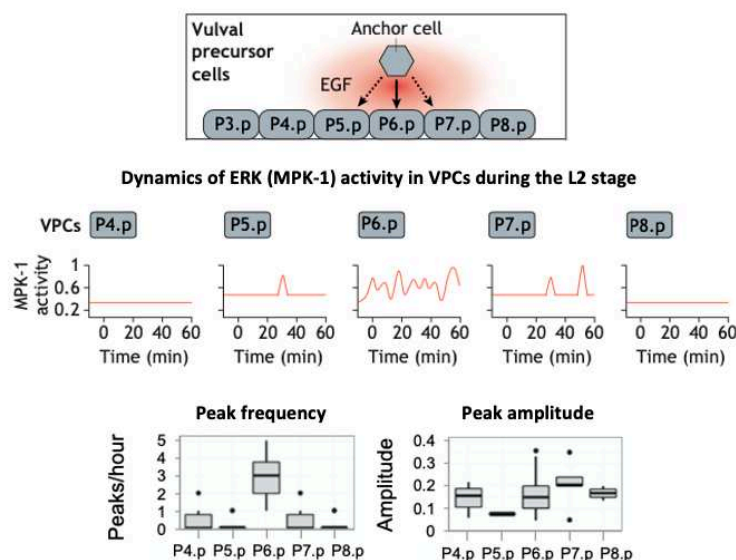


Figure 13 : Dynamique de l'activité ERK (MPK-1) dans les cellules précurseurs de la vulves (CPV) chez *C. elegans* au stade L2 visualisée avec le ERK-KTR. Adapté de (de la Cova et al., 2017; Nakamura et al., 2021).

Les CPV sont constituées de six cellules épithéliales ventrales alignées, qui sont nommées P3.p à P8.p et sont indifférenciées et équipotentes au stade L2 (second stade larvaire du vers) (Figure 13). La cellule d'ancrage (CA) dans la gonade sécrète de l'EGF (LIN-3) vers les CPV voisines (Figure 13) (Hill and Sternberg, 1992). Le récepteur EGFR (LET-23) active le module Ras/ERK. Cette induction paracrine repose en partie sur un mécanisme de "gradient de morphogène" : EGF active la voie ERK de telle sorte que trois destins cellulaires distincts, primaire (1°), secondaire (2°) ou tertiaire (3°) sont induits de manière dépendante de la distance à la cellule d'ancrage (CA) (Figure 13) (Sternberg and Horvitz, 1984; Katz et al., 1995). Le gradient d'activation de ERK atteint son maximum au niveau de P6.p, la CPV située le plus près de la cellule d'ancrage (CA) et qui adopte toujours le destin primaire. Les cellules P5.p et P7.p adjacentes à P6.p adoptent le destin secondaire. Enfin, les cellules P4.p et P8.p, plus périphériques deviennent des cellules tertiaires (Sternberg and Horvitz, 1986).

Les mesures *in vivo* du ERK-KTR ont tout d'abord montré un gradient spatial du niveau d'activité de ERK dans les CPV depuis P6.p (niveau le plus élevé) vers P4.p et P8.p (niveau le plus bas), ce qui soutient le modèle de gradient de concentration de l'EGF (LIN-3). La dynamique de ERK a ensuite été examinée avec le ERK-KTR tout au long du processus de différenciation des CVP du stade L2 à L3, toutes les 10 minutes pendant 13,5 heures, révélant une dynamique temporelle de ERK pulsatile (Figure 13) (de la Cova et al., 2017). Le signal graduel EGF (LIN-3) semble être converti en oscillations de l'activité de ERK modulées en fréquence, mais d'amplitudes comparables : P6.p présente les fréquences les plus élevées (3 impulsions par heure en moyenne, contre 0,5 impulsion par heure en moyenne dans P7.p par exemple) (Figure 13) (de la Cova et al., 2017).

On ne sait toujours pas comment la dynamique pulsatile de l'activation de ERK contribue à la détermination du destin cellulaire des CPV. La relation entre la dynamique ERK et la détermination du destin cellulaire dépend fortement du type de cellule et du contexte, et il serait donc nécessaire de valider plus directement la relation causale par une approche telle que l'optogénétique (Aoki et al., 2013; Toettcher et al., 2013; Johnson and Toettcher, 2019).

Au delà des dynamiques de l'activité ERK classiques : transitoire ou soutenue, ces résultats soulignent l'importance de la dynamique pulsatile de ERK qui commence maintenant à être observée dans différents contextes *in vitro* (Albeck et al., 2013; Aoki et al., 2013) et *in vivo* (Hiratsuka et al., 2015; Kumagai et al., 2015) et pour laquelle des modules génétiques capables de décoder la fréquence des oscillations ont été identifiés (Wilson et al., 2017).

Des ondes d'activité de ERK dans la formation de la placode chez la drosophile sont associées à une contractilité de l'actomyosine plus élevée dans les cellules du front d'onde ERK qui est requise pour l'invagination

Enfin, l'utilisation des biosenseurs pour ERK a permis d'étudier des types de dynamiques de l'activité de ERK qui se produisent au niveau des tissus et comment ces dynamiques ERK génèrent des réponses multi-cellulaires à l'origine de changements de formes dans des contextes de développement embryonnaire. C'est le cas par exemple au cours de la morphogenèse chez la drosophile.

Dans l'embryon de drosophile, à partir du stade 10 (~5h après fécondation), les placodes trachéales apparaissent et commencent à s'invaginer, des processus médiés par l'activité de la voie ERK (Figure 14) (Samakovlis et al., 1996; Hayashi and Ogura, 2020).

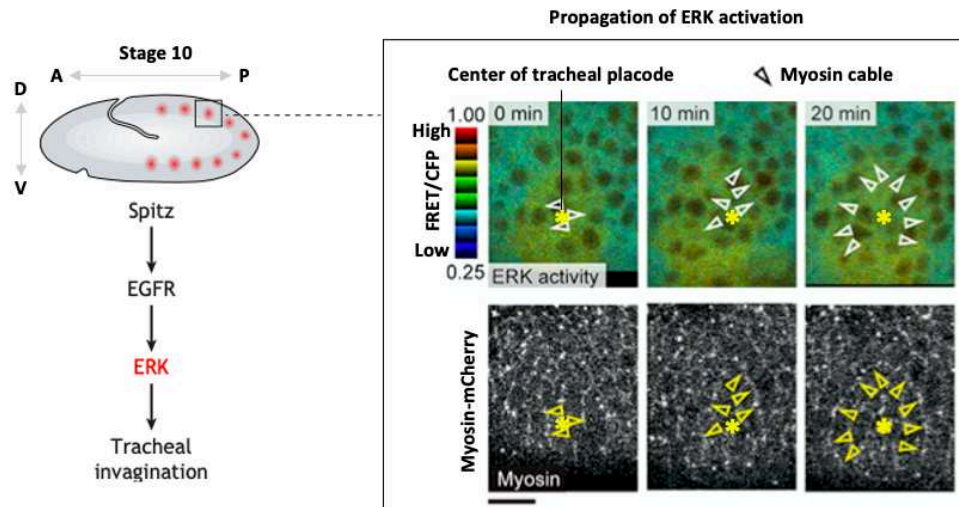


Figure 14 : Vague d'activité de ERK au cours de la formation de la placode chez la *Drosophile*. Au stade 10, dans le processus initial de formation de la trachée, la signalisation Spitz-EGFR-ERK est activée dans la placode trachéale. La dynamique de l'activité ERK est visualisée dans la placode à l'aide d'un biosenseur ERK FRET, EKAREV. Le motif circulaire des câbles de myosine se forme à l'interface de cellules ayant des activités ERK différentes et s'étend radialement en fonction de la propagation d'une onde d'activation de ERK. Adapté de (Ogura et al., 2018; Nakamura et al., 2021).

Le suivi de la dynamique de l'activité ERK au cours de la formation des placodes, à l'aide du biosenseur ERK FRET, EKAREV, à permis d'identifier une vague d'activation de ERK qui se propage de manière concentrique à partir du centre de la placode (Figure 14) (Ogura et al., 2018). Il a ensuite été démontré que cette vague est associée à une contractilité de l'actomyosine plus élevée dans les cellules à front d'onde ERK qui est requise pour l'invagination (Figure 14) (Ogura et al., 2018).

Le contrôle de la morphogenèse tissulaire par les ondes ERK a également été observé dans la migration collective des cellules de mammifères en culture, révélant l'existence d'un mécanisme moléculaire par lequel la contractilité de l'actomyosine est augmentée par la voie ERK (Hino et al., 2020).

Le nombre d'études de biologie du développement utilisant ces biosenseurs pour suivre la dynamique spatio-temporelle de l'activité ERK est encore limité

Le développement récent de rapporteurs génétiquement codés de l'activité de ERK a ouvert la voie à la mesure quantitative en temps réel avec une résolution cellulaire de l'activité de cette voie. Cependant, le nombre d'études de biologie du développement utilisant ces biosenseurs dans des contextes multi-cellulaires *in vivo* est encore limité et notre compréhension quantitative de la dynamique spatio-temporelle de la voie ERK est limitée à quelques éléments de compréhension non reliés entre eux (Biasuz et al., 2018).

Ainsi, par exemple, on ne comprend toujours pas très bien comment le contexte cellulaire affecte le décodage de la dynamique de l'activité ERK. En effet, alors que les cellules précurseurs de la vulve (CPV) chez *C. elegans* semblent être capables de lire des oscillations de ERK modulées en fréquence, dans l'embryon précoce de drosophile, les réponses cellulaires dépendent de l'activité de ERK intégrée dans le temps, plutôt que d'autres caractéristiques de la dynamique. Aussi, notre

compréhension sur la façon dont les différentes dynamiques de l'activité de ERK sont décodées par les mécanismes à l'interface signalisation/expression génétique pour induire des schémas spécifiques et reproductibles d'expression génétique reste très limitée. Pour résoudre ces questions, il sera important de combiner des mesures quantitatives en direct de l'activité des composants individuels avec une manipulation spatio-temporelle précise de leur activité et, enfin, avec une modélisation mathématique dynamique incluant des détails mécanistiques. En effet, de nombreux comportements de l'activité ERK ne peuvent pas être déduits de la seule connaissance des interactions moléculaires. C'est le cas, par exemple, avec l'activation constitutive de MEK qui présente des effets opposés : augmentation ou réduction des niveaux de l'activité ERK, dans différentes régions des embryons de drosophile et de poisson-zèbre (Goyal et al., 2017). C'est le cas également avec l'activation paradoxale de ERK en réponse à l'inhibition de RAF dans des cellules humaines de mélanome (Halaban et al., 2010; Fey et al., 2016).

Dans l'ensemble, ces observations soulignent l'importance de visualiser l'activité de la signalisation ERK en temps réel dans une variété de contextes cellulaires *in vitro* et *in vivo*. Plusieurs raisons peuvent expliquer le faible nombre d'études actuelles sur la dynamique de ERK dans des embryons qui se développent. Le développement de biosenseurs pour ERK est relativement récent et l'application de méthodes d'imagerie *in vivo* et d'analyse quantitative d'images sont encore difficiles. L'imagerie fluorescence d'événements biologiques sur de longues périodes de l'embryogenèse cumulée à un taux d'échantillonnage élevé, nécessaire pour capturer la dynamique spatio-temporelle de l'activation de ERK, reste également un défi notamment en raison de la toxicité induite par la lumière qui reste un obstacle fondamental à l'imagerie microscopique des embryons vivants (Débarre et al., 2014).

LES ASCIDIES : UN BON MODÈLE POUR L'ÉTUDE DE LA DYNAMIQUE SPATIO-TEMPORELLE DE LA VOIE ERK AU COURS DU DÉVELOPPEMENT EMBRYONNAIRE

Comme décrit précédemment, l'étude de la dynamique temporelle de l'activation de ERK à la résolution cellulaire au cours du développement embryonnaire est difficile et le nombre d'études utilisant des biosenseurs pour suivre l'activité de la voie ERK dans des contextes multi-cellulaires *in vivo* est encore limité. Dans ce contexte, les ascidies apparaissent comme un excellent modèle d'étude de la dynamique spatio-temporelle de la voie ERK sur une longue période de leur développement embryonnaire, avec une résolution cellulaire et une fréquence d'imagerie élevée.

Introduction aux ascidies

En regardant une ascidie adulte, il est difficile d'imaginer que nous sommes de proches cousins de ces animaux : fixés à une roche, ces invertébrés marins qui vivent dans le monde entier, généralement en eaux peu profondes, passent leur vie adulte à filtrer l'eau de mer, protégées des prédateurs par une épaisse tunique (Figure 15) (Lemaire, 2011).

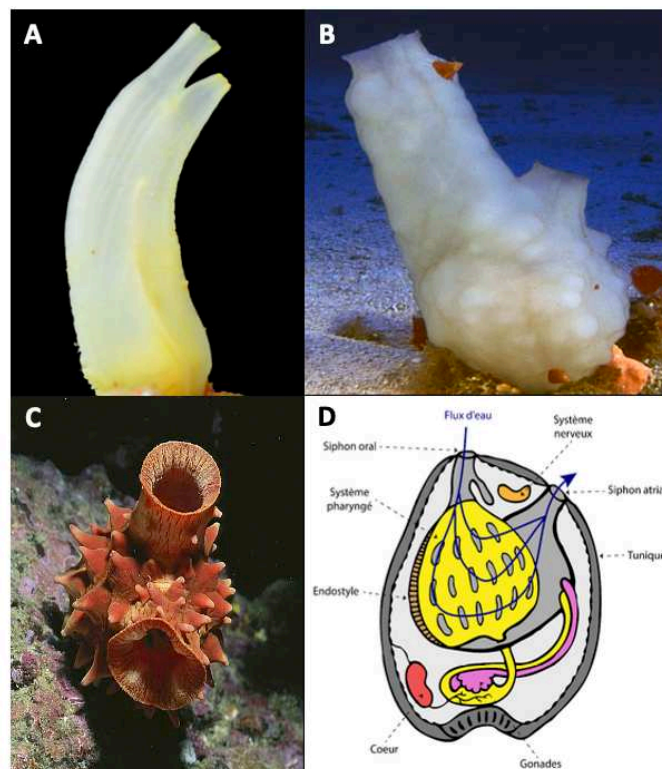


Figure 15 : Les ascidies solitaires adultes. A) *Ciona robusta*. B) *Phallusia mammillata*. C) *Halocynthia roretzi*. D) Schéma illustrant la morpho-anatomie d'une ascidie adulte.

Ces animaux ont pourtant conservé depuis plus de 400 millions d'années une forme larvaire ancestrale de têtard de chordés avec une notochorde axiale et un tube neural dorsal (Figure 16). Ces têtards se fixent et subissent une métamorphose pour adopter une forme adulte benthique solitaire ou coloniale (Figure 17).

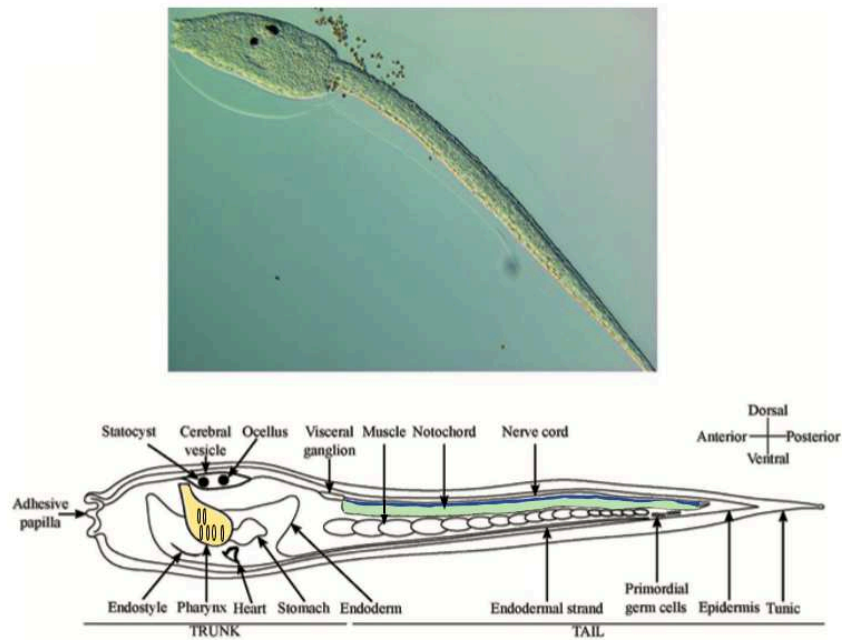


Figure 16 : La larve d'ascidie présente un plan d'organisation typique des chordés. En haut : photo d'une larve de *Ciona intestinalis*. En bas, schéma d'une larve d'ascidie montrant les principales synapomorphies des chordés : en vert, la notochorde, en bleu, le tube neural dorsal et en jaune, les fentes pharyngiennes. Adapté de (Boutet and Schierwater, 2021).

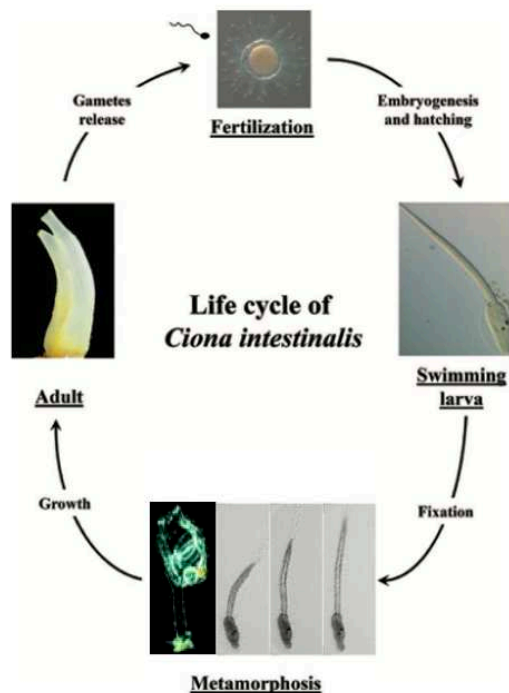


Figure 17 : Cycle de vie de l'ascidie solitaire *Ciona intestinalis*. Après la libération des gamètes, l'embryogenèse a lieu dans l'eau de mer et donne naissance à une larve nageuse en quelques heures. Après une période de nage libre (quatre à huit heures dans le cas de *C. intestinalis*), la larve adhère à un substrat et commence sa métamorphose,

la régression de la queue étant l'événement le plus spectaculaire de ce processus. Après la métamorphose, le juvénile produit un adulte sexuellement mature en un à deux mois selon les conditions d'alimentation. Adapté de (Boutet and Schierwater, 2021).

Alors que la position phylogénétique des tuniciers, auxquels appartiennent les ascidies, comme groupe frère des vertébrés est maintenant bien établie, les relations à l'intérieur des tuniciers sont plus débattues. Plusieurs phylogénies émergent fréquemment dans la littérature, bien qu'un consensus apparaisse actuellement (Figure 18) (Delsuc et al., 2006, 2018). On considère généralement que les tuniciers sont composés de cinq grands embranchements : Appendicularia, Stolidobranchia, Thaliacea, Phlebobranchia et Aplousobranchia (Figure 18). Les Appendicularia se caractérisent par un mode de vie pélagique et une forme adulte en forme de têtard, illustrée par l'espèce la plus connue, *Oikopleura dioika*, et sont souvent positionnés comme le groupe frère de tous les autres tuniciers (Figure 18) (Delsuc et al., 2018). Selon la phylogénie consensuelle actuelle, Stolidobranchia a été le premier groupe à diverger des autres (Delsuc et al., 2018). Ensuite, Thaliacea, Phlebobranchia et Aplousobranchia sont considérés comme monophylétiques (Figure 18). Les Thaliacea ont divergé en premier, et les Phlebobranchia sont regroupés avec les Aplousobranchia (Figure 18). Ainsi, les ascidies, dont les formes adultes sont fixées au substrat et que l'on retrouve solitaires et coloniales chez les Stolidobranchia et les Phlebobranchia et uniquement coloniales chez les Aplousobranchia forment un groupe paraphylétique. Les Thaliacea, y compris les salpes, sont uniquement pélagiques et forment des colonies planctoniques constituées de multiples individus (Delsuc et al., 2018).

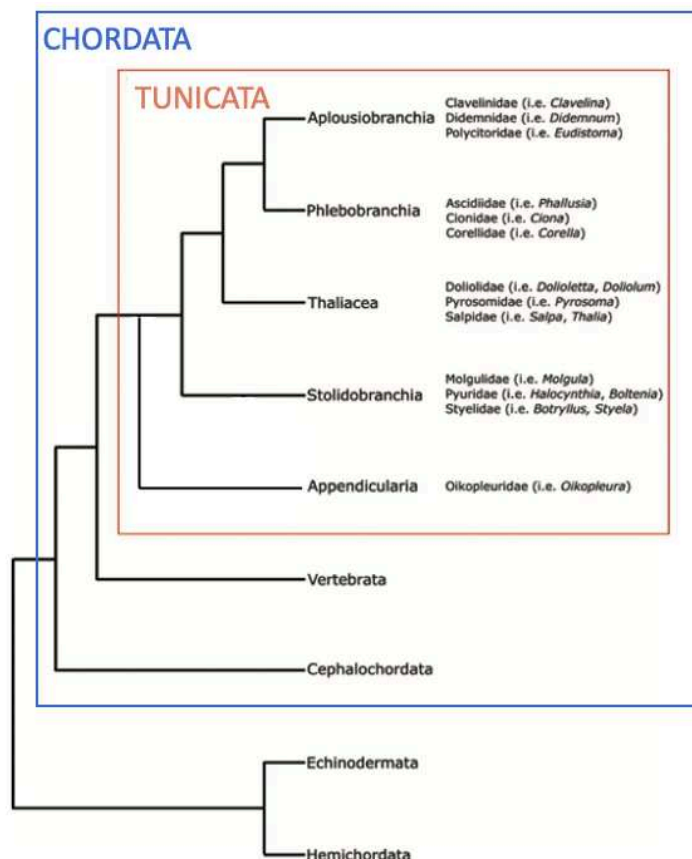


Figure 18 : Arbre phylogénétique détaillé des tuniciers au sein des deuterostomes. Adapté de (Boutet and Schierwater, 2021). Les ascidies, qui regroupent Aplousobranchia, Phlebobranchia et Stolidobranchia, forment un groupe paraphylétique.

Historiquement, les études sur les ascidies ont eu un grand impact conceptuel sur notre compréhension des règles qui guident un programme de développement. Laurent Chabry a utilisé des ascidies pour montrer pour la première fois qu'il était possible de prédire les conséquences d'une interférence avec le développement embryonnaire. Ses expériences sur les ablations de progéniteurs musculaires ont conduit à la théorie du développement mosaïque : les cellules se développent en fonction des informations maternelles qu'elles reçoivent au cours des divisions précoces (Chabry, 1887). Puis, en 1905, Edward Conklin montre que les embryons d'ascidies se développent de façon stéréotypée avec un lignage cellulaire invariant qu'il a pu décrire jusqu'à la gastrulation (Figure 19) (Conklin, 1905). Ces avancées majeures ont conduit à la proposition que le programme de développement des ascidies est piloté par l'héritage précis de déterminants maternels localisés, rendu possible par la position invariante des plans de clivage cellulaire. Ce point de vue initial a cependant évolué et on sait aujourd'hui que la plupart des destins cellulaires sont obtenus par des processus d'inductions cellulaires médiés par la communication cellulaire. Une exception à cette règle semble être la lignée musculaire primaire, étudiée par Chabry et Conklin, qui se forme suite à l'héritage d'un facteur de transcription maternel localisé, macho-1 (Nishida and Sawada, 2001).

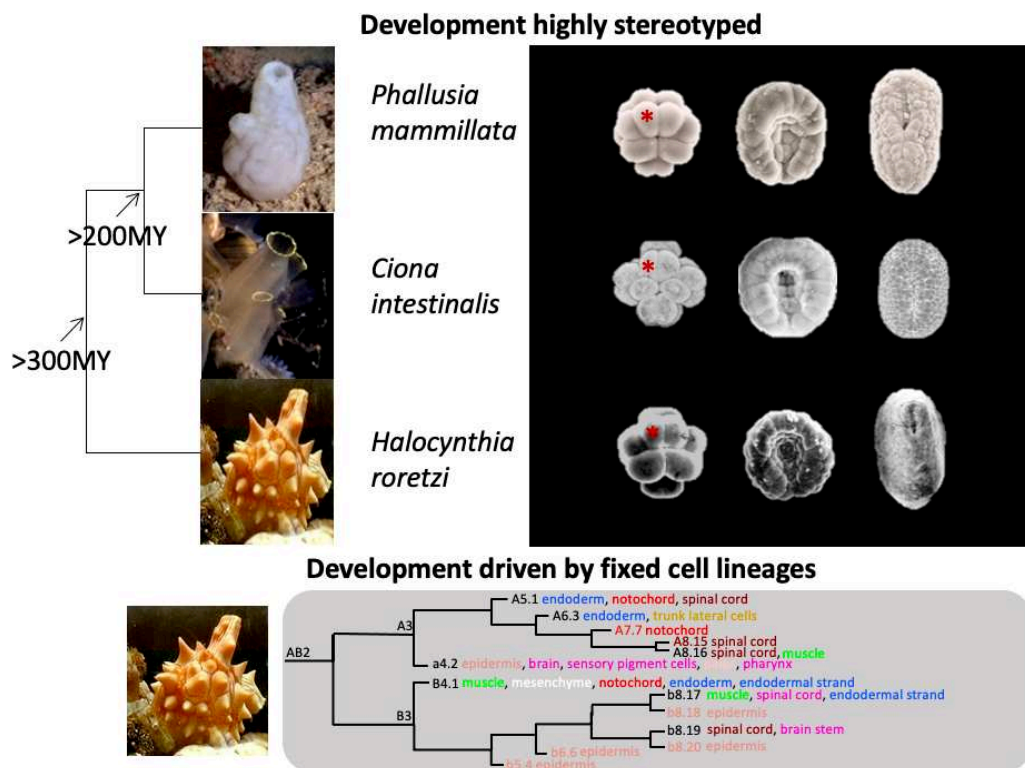


Figure 19 : Développement hautement stéréotypé chez les ascidies avec un lignage cellulaire fixe.

Haut : Comparaison de stade développementaux identiques (stade 16-cellule (St. 5a), early gastrula (St. 10), late neurula (St. 16)) chez 3 espèces d'ascidies (*Phallusia mammillata*, *Ciona intestinalis* et *Halocynthia roretzi*) montrant la reproductibilité de la géométrie du développement précoce, à tel point que chaque cellule peut être retrouvée à la même position (astérisque rouge par exemple). Les temps de divergence entre les différentes espèces en millions d'années (MY,

Million Years) sont indiqués au niveau des ancêtres communs. Bas : Exemple d'une partie du lignage cellulaire fixe chez *Halocynthia roretzi* (Nishida, 1987). Les noms et les destins des cellules sont indiqués et color-codés.

Après l'effervescence de la fin du XIX^{ème} et du début du XX^{ème} siècle, les ascidies se sont retirées pendant une longue période dans leur obscurité marine, d'où elles ont été tirées par la biologie moléculaire (Kuratani et al., 2006). Les premières études moléculaires ont été réalisées principalement sur des embryons de l'ascidie japonaise *Halocynthia roretzi*. Le nombre de laboratoires intéressés par les ascidies ayant augmenté, une espèce plus cosmopolite a été recherchée et *Ciona intestinalis* a été choisie. Plus de 50 laboratoires dans le monde étudient le développement des ascidies aujourd'hui. Les principales avancées opérées durant cette période comprennent l'analyse complète des lignées cellulaires chez différentes espèces (Nishida and Satoh, 1983, 1985; Nishida, 1987; Stolfi et al., 2014; Stach and Anselmi, 2015), le séquençage de plusieurs génomes d'ascidies (Dehal et al., 2002; Small et al., 2007; Stolfi et al., 2014; Brozovic et al., 2018), le développement de méthodes de transgénèse par électroporation robustes et simples (Lemaire, 2011; Sardet et al., 2011), et l'élucidation de réseaux de régulation génique précoces assez détaillés (Imai et al., 2006). Les données génomiques disponibles ont notamment révélé un contraste étonnant dans le taux d'évolution des gènes nucléaires codant pour les protéines entre les tuniciers et les vertébrés (Delsuc et al., 2008; Tsagkogeorga et al., 2010). L'évolution accélérée des gènes des tuniciers est également couplée à d'importants réarrangements structurels observés dans leurs génomes (Holland and Gibson-Brown, 2003; Seo et al., 2004; Denoeud et al., 2010). Les raisons derrière le taux rapide d'évolution génomique chez les tuniciers ne sont pas claires (Delsuc et al., 2018) et contrastent avec le niveau inhabituel de conservation des morphologies embryonnaires entre toutes les espèces d'ascidies étudiées jusqu'à présent (Figure 19).

Six espèces d'ascidies, *Halocynthia roretzi*, *Ciona intestinalis*, *Ciona robusta*, *Ciona savignyi*, *Phallusia mammillata* et *Botryllus schlosseri* représentent les principaux modèles d'ascidies pour les études d'embryologie (Figure 15; Figure 18). *Ciona intestinalis* et *Ciona robusta* ont longtemps été groupées sous le nom de *Ciona intestinalis* jusqu'à ce que des études moléculaires, morphologiques et sur les taux d'hybridations dans la nature montrent qu'en réalité *Ciona intestinalis* constitue un complexe d'espèces cryptiques, composé d'espèces distinctes dont *Ciona robusta* (anciennement connu sous le nom de *C. intestinalis* type A) et *Ciona intestinalis* (anciennement connu sous le nom de *C. intestinalis* type B) (Brunetti et al., 2015; Bouchemousse et al., 2016; Malfant et al., 2018).

Les ascidies fournissent un cadre unique pour l'imagerie quantitative *in toto* sur des embryons vivants avec une résolution subcellulaire

Les embryons d'ascidies présentent des caractéristiques qui en font un modèle unique pour l'étude des événements du développement embryonnaire par l'imagerie, sur de longues périodes avec une haute résolution spatio-temporelle.

Les embryons d'ascidie : un cadre rigoureux pour l'étude de l'embryogenèse au niveau cellulaire

Tout comme le nématode *C. elegans*, les ascidies présentent des caractéristiques fournissant un cadre rigoureux pour l'étude du programme développemental embryonnaire avec un niveau de résolution cellulaire.

En effet, leur embryogenèse se déroule avec un petit nombre de cellules qui présentent des cycles de divisions cellulaires et des orientations quasi invariables. Il est intéressant de noter également que contrairement aux nématodes (Goldstein, 2001), la géométrie et les lignées cellulaires des embryons d'ascidies est même restée essentiellement inchangée depuis l'émergence du groupe, il y a environ 400 millions d'années (Figure 19) (Lemaire, 2011; Delsuc et al., 2018). Grâce à cette invariance, les stades de développement correspondent à un nombre précis de cellules jusqu'au début de la gastrulation qui commence au stade 112-cellules. Les stades de développement ultérieurs sont ensuite définis par les événements majeurs du développement : l'internalisation du mésendoderme (gastrulation), la fermeture du tube neural (neurulation), l'extension de la queue (stade du bourgeon caudal) et le début du comportement de nage (larve éclos) (Figure 20) (Hotta et al., 2007).

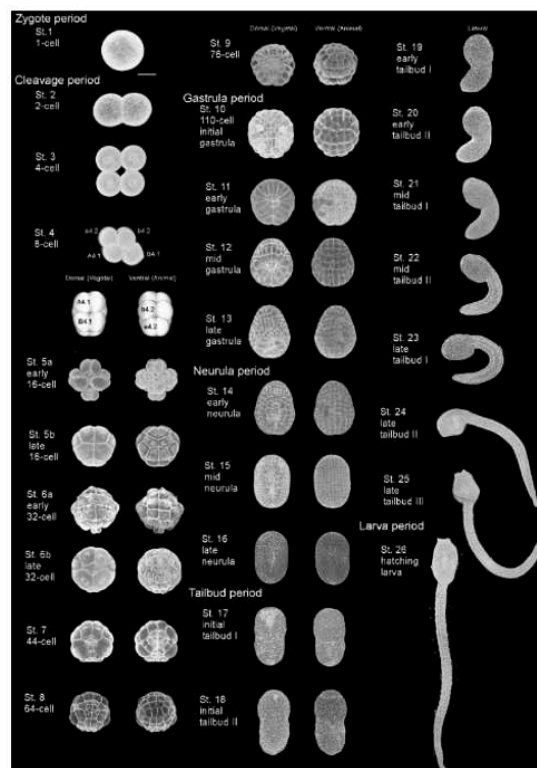


Figure 20 : Table du développement embryonnaire de *Ciona intestinalis*. D'après (Hotta et al., 2007).

Le développement des ascidies montre une telle reproductibilité au niveau cellulaire que chaque cellule a des homologues dans tous les embryons d'une espèce. Il est donc possible de définir des classes de cellules et de leur attribuer des noms canoniques uniques. Ainsi, les cellules embryonnaires des ascidies sont traditionnellement nommées selon un schéma proposé par Conklin (Conklin, 1905). Chaque cellule embryonnaire reçoit un nom de la forme Fp.q (par exemple A7.6), où F est une lettre qui indique la lignée fondatrice au stade 8 cellules, dont la cellule est issue ("A" correspond à végétal antérieur, "a" à animal antérieur, "B" à végétal postérieur et "b" à animal postérieur), p est le rang du cycle cellulaire mitotique (par ex. p = 3 signifie que la cellule est dans son troisième cycle cellulaire mitotique depuis la fécondation) et q est un nombre entier reflétant la position relative de la cellule par rapport au pôle végétal de l'embryon ou à la ligne médiane, par rapport à sa sœur (Figure 21). Pour distinguer les cellules symétriques bilatérales de gauche et de droite (en vue végétale), les cellules de droite reçoivent un suffixe "*" (par exemple A7.6*), tandis

que les cellules de gauche correspondantes ne reçoivent aucun suffixe ou, plus récemment, un suffixe "_" (par exemple A7.6_) (Figure 21). Lors de la division cellulaire, la syntaxe de nommage de Conklin spécifie que la cellule $Fp.q$ se divise en deux filles, nommées $F(p + 1).(2q)$ et $F(p + 1).(2q - 1)$. Les deux filles sont ensuite nommées individuellement en fonction de leur position relative dans l'embryon : $F(p + 1).(2q - 1)$ est donnée à la fille dont la surface apicale est géodésiquement plus proche du pôle végétal de l'embryon ou de la ligne médiane, si la distance géodésique au pôle végétal est très similaire entre les deux cellules filles. Cette utilisation de la distance géodésique limite l'applicabilité de ces règles de décisions aux stades où toutes les cellules ont un contact avec le milieu extérieur. A partir du stade 76 cellules, l'axe de certaines divisions cellulaires est orienté perpendiculairement à la surface apicale de l'embryon, générant une cellule interne dont la distance géodésique au pôle végétal n'est pas définie au-delà du début de la gastrulation. Des règles *ad hoc* inspirées de Conklin ont été utilisées pour nommer les cellules de la plaque neurale (Nicol and Meinertzhagen, 1988; Cole and Meinertzhagen, 2004).

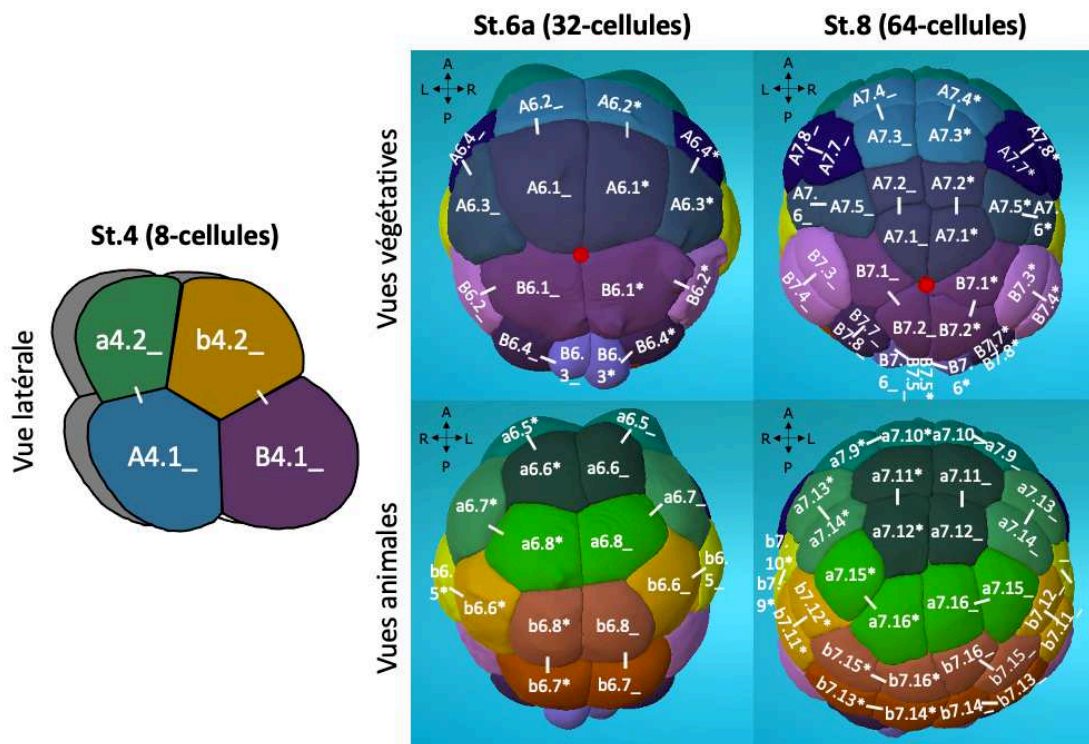


Figure 21 : Nommage cellulaire des embryons d'ascidies précoces selon la nomenclature de Conklin. A gauche : schéma latéral d'un embryon au stade 8 cellules, montrant la position des 4 lignées cellulaires fondatrices. Milieu et droite : vues végétatives (en haut) et animales (en bas) d'un embryon de *Phallusia mammillata* au stade 32 cellules (milieu) et 64 cellules (droite). Chaque paire de cellules bilatérales au stade 32 cellules s'est vue attribuer une couleur unique, qui a été propagée au stade 64 cellules. Les noms (en blanc) sont conformes à ceux de Conklin. Les lignes blanches relient les cellules sœurs. Les axes antéro-postérieur (AP) et gauche-droite (LR) sont représentés. Les disques rouges matérialisent le pôle végétal de l'embryon, utilisé comme référence pour la mesure de la distance géodésique dans le nommage selon Conklin. Milieu, gauche : clichés de l'embryon Astec-Pm-Stages au stade indiqué (Guignard et al., 2020) visualisés à l'aide de MorphoNet (Leggio et al., 2019).

Enfin, au-delà de l'extrême reproductibilité morphologique des embryons d'ascidies, les lignages cellulaires sont également invariants. Au démarrage de mon travail de thèse, ces lignages étaient connus pour toutes les cellules jusqu'au début de la gastrulation (Lemaire, 2009), et bien

documentés pour les stades ultérieurs notamment pour les lignées neuronales dans lesquelles des lignées cellulaires complètes sont connues pour certains types de cellules neuronales (Figure 19) (Nishida and Satoh, 1983; Nishida, 1987; Nicol and Meinertzhagen, 1988; Hudson and Lemaire, 2001; Cole and Meinertzhagen, 2004; Hudson and Yasuo, 2005; Stolfi and Levine, 2011).

De plus, au début de la gastrulation, au stade 112-cellules, le destin de presque toutes les cellules est restreint à un unique tissu larvaire et le moment et les mécanismes de la spécification du destin de chaque cellule au cours du développement précoce sont dans la plupart des cas connus avec précision (voir pour revue ((Lemaire, 2009)).

Les ascidies possèdent des propriétés permettant l'imagerie *in toto* de leurs embryons vivants

Les progrès de la biologie du développement ont souvent été motivés par les avancées des technologies d'imagerie, depuis la première utilisation de microscopes composés de plusieurs lentilles livrant des images de haute qualité par les embryologistes du milieu et de la fin du XIX^{ème} siècle, en passant par le développement des méthodes confocales et à deux photons dans les années 1980 et 1990 (White et al., 1987; Denk et al., 1990; Amos and White, 2003) et plus récemment par l'introduction de microscopies innovantes à feuillet de lumière (Power and Huisken, 2017).

Les microscopes à feuillets lumineux illuminent un mince plan de l'échantillon (Figure 22). Un ou plusieurs objectifs orthogonaux au plan d'illumination, couplés à des caméras recueillent l'image du plan éclairé, offrant ainsi un gain de vitesse d'acquisition de plusieurs ordres de grandeur par rapport aux autres types de microscopies à balayage (Figure 22) (Power and Huisken, 2017; Vieites-Prado and Renier, 2021). Le développement révolutionnaire de cette microscopie à feuillets lumineux, associé à des marqueurs fluorescents pour les structures subcellulaires clés, a donc fait entrer le domaine de l'embryogenèse des vertébrés et des invertébrés dans une ère où le comportement de chaque cellule embryonnaire vivante et de ses descendants peut être imagé et suivi pendant des heures avec une résolution temporelle inférieure à la minute et une précision spatiale inférieure au micron (Khairy and Keller, 2011; Biasuz et al., 2018).

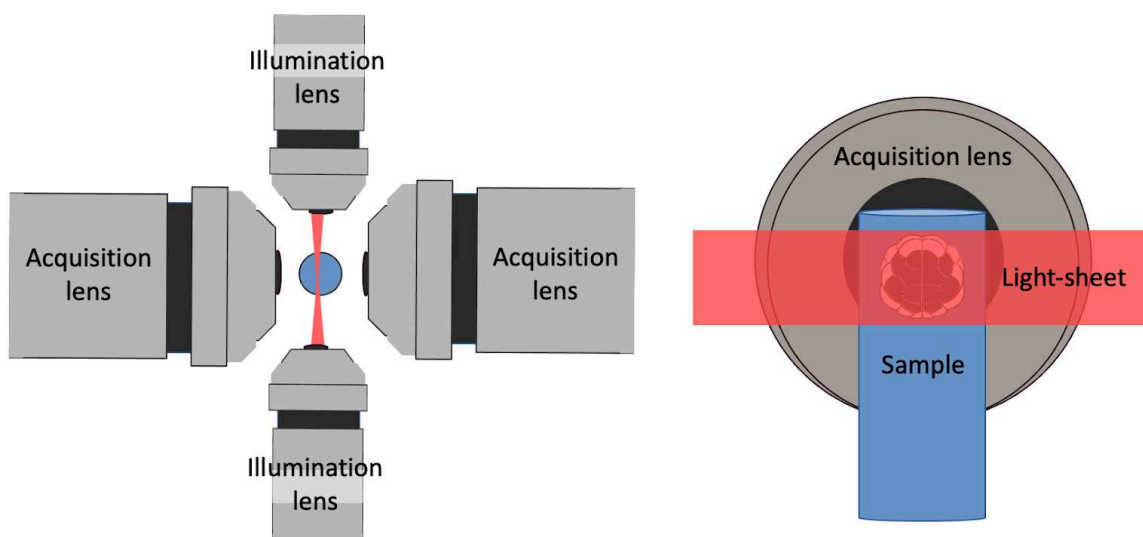


Figure 22 : Microscopie à fluorescence multi-vues à feuillets lumineux. Schéma montrant les caractéristiques du microscope multi-vues à feuillets lumineux : deux chemins optiques orthogonaux assurent l'illumination d'un plan sur l'échantillon (rouge) et la détection de la fluorescence.

Cependant, comme toute méthode d'imagerie, elle est soumise à des contraintes fondamentales entre la résolution et le champ de vision et la taille de l'embryon étudié (Vieites-Prado and Renier, 2021). Les grossissements typiques des systèmes commerciaux sont relativement faibles et produisent des images à faible résolution par rapport aux microscopes à épifluorescence ou confocaux et les chambres d'acquisitions sont souvent limitées en termes de taille d'échantillon (Vieites-Prado and Renier, 2021).

Dans ce contexte, les embryons d'ascidies offrent des possibilités uniques d'imagerie du développement avec la combinaison autrement difficile à atteindre d'une résolution subcellulaire et d'un champ de vision incluant l'ensemble de l'embryon (Veeman and Reeves, 2015).

En effet, bien qu'ils soient plus grands que les œufs particulièrement petits de *C. elegans*, les œufs d'ascidie sont beaucoup plus petits que ceux d'autres organismes modèles tels que le xénope, le poisson zèbre ou la drosophile (Figure 23 A) (Veeman and Reeves, 2015). Par exemple, l'œuf de *Ciona* d'un diamètre de 140 μm est 125 fois plus petit en volume que l'œuf de poisson zèbre d'un diamètre de 700 μm . Ainsi, si l'on compare par exemple la taille d'un embryon de *Ciona* au début de l'élongation de la queue à un stade comparable chez le poisson zèbre, le champ de vision capable de couvrir l'embryon entier de *Ciona* ne couvre l'équivalent que de quelques somites de l'embryon de poisson zèbre (Figure 23 B). Cela a des implications importantes pour l'utilisation de méthodes d'imagerie volumétrique où les embryons d'ascidie peuvent être imagés dans leur intégralité (*in toto*) avec une résolution suffisante pour résoudre finement les formes de toutes les cellules embryonnaires et l'arrangement de ces cellules.

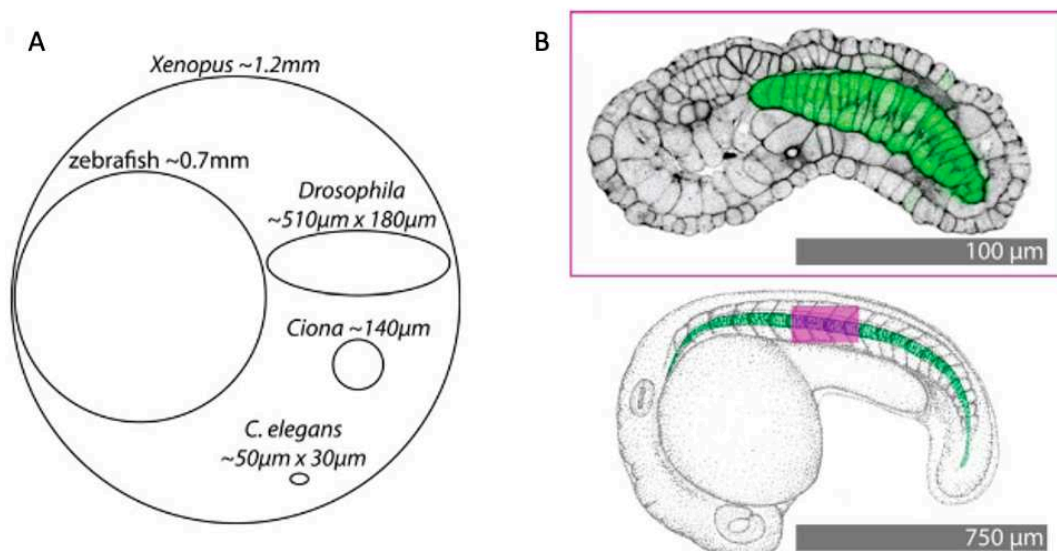


Figure 23 : Comparaison de la taille des embryons d'ascidies à ceux d'autres espèces modèles. A) Schéma montrant les tailles relatives des embryons de *Ciona* comparés à ceux de *C. elegans*, *Drosophila*, poisson zèbre (zebrafish) et *Xenopus*. B) Haut : image confocale avec coloration phalloïdine (noir, contraste inversé pour correspondre à l'image du poisson zèbre) et un transgène GFP spécifique de la notocorde (vert) dans un embryon de *Ciona intestinalis* au début de l'élongation de la queue. Bas : image confocale d'un embryon de poisson zèbre au même stade de développement, montrant la notocorde colorée en vert et les somites en noir.

stade 19; Bas : dessin à la plume d'un embryon de poisson zèbre à un stade comparable d'extension précoce de la queue. La notochorde est pseudo-colorée en vert. Le cadre violet autour de l'embryon de *Ciona intestinalis* est reporté en conservant l'échelle dans l'embryon de poisson zèbre montrant que le champ de vision capable de couvrir l'embryon entier de *Ciona* ne couvre l'équivalent que de quelques somites de l'embryon de poisson zèbre. D'après (Veeman and Reeves, 2015).

En plus de leur petite taille, les embryons d'ascidies solitaires se développent avec un faible nombre de cellules. Par exemple, leur larve ne compte qu'environ 2600 cellules (Lemaire, 2009). Malgré ce nombre remarquablement faible de cellules, l'anatomie sous-jacente du têtard a beaucoup en commun avec celle des vertébrés (Lemaire, 2011; Veeman and Reeves, 2015).

Enfin, les embryons d'ascidies présentent trois autres propriétés favorables à l'imagerie du développement de leurs embryons vivants : leur développement est externe et rapide (<18h pour atteindre le stade larvaire chez *Ciona* et *Phallusia* à 18°C) et les embryons de *Phallusia mammillata* sont optiquement transparents.

Une capacité unique de segmenter et suivre chaque cellule au cours du développement précoce des ascidies

Basé sur l'ensemble des propriétés avantageuses des embryons d'ascidies et afin de quantifier la reproductibilité embryonnaire chez l'ascidie *Phallusia mammillata*, notre équipe a développé, dès 2015, une capacité unique à segmenter et suivre chaque cellule au cours du développement précoce de ces embryons (Guignard, 2015). Tout d'abord, l'utilisation d'un microscope multi-vues à feuillet de lumière (de Medeiros et al., 2015), permet l'imagerie sur une longue période du développement précoce des embryons transparents de *Phallusia*, marqués au niveau des membranes plasmiques de chaque cellule, avec une résolution temporelle de 2 minutes (Figure 24 A). En parallèle, l'équipe a développé un algorithme de segmentation adapté à ces jeux de données issus de la microscopie à feuillet de lumière à haut débit appelé ASTEC (pour Adaptive Segmentation and Tracking of Embryonic Cells, segmentation et suivi adaptatifs des cellules embryonnaires). ASTEC segmente et suit simultanément des cellules entières en projetant itérativement les segmentations d'un point temporel sur le suivant pour produire une version numérisée de chaque embryon imagé (Figure 24 B). Les segmentations de toutes les cellules embryonnaires existant à chaque point temporel peuvent être ensuite intégrées dans des lignées cellulaires globales, auxquelles les destins connus sont associés et à partir desquelles des paramètres géométriques quantitatifs peuvent être extraits, ouvrant ainsi la voie à la description systématique de l'embryogenèse chez *Phallusia* avec une résolution cellulaire.

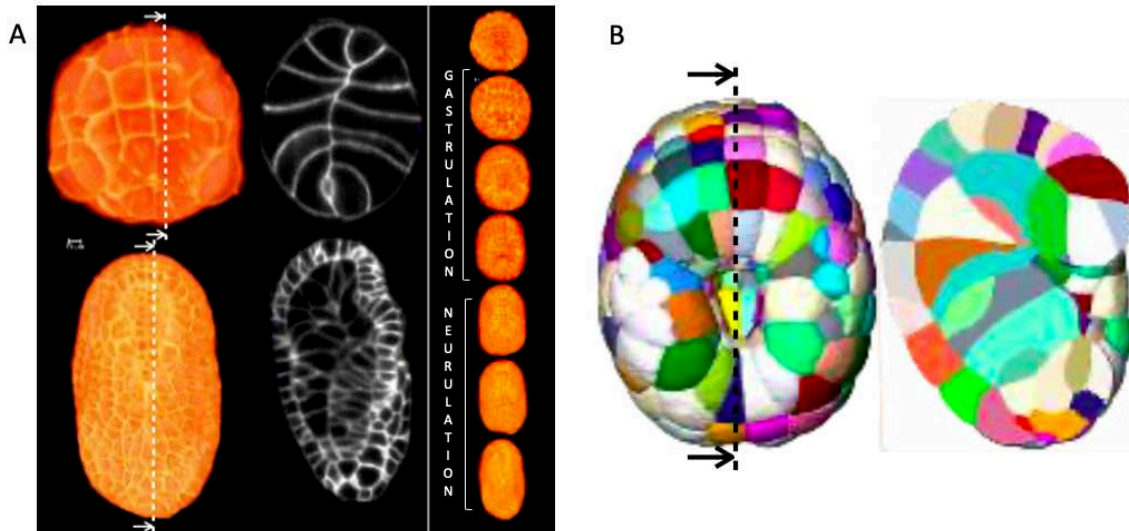


Figure 24 : Imagerie multi-vues à haute résolution et segmentation d'embryons de *Phallusia mammillata*. A) Vues végétale et dorsale du rendu 3D d'un embryon imagé, après fusion des images prises selon les quatre angles de vue. La ligne blanche pointillée indique les plans des coupes parasagittales. B) Vue dorsale de la segmentation du même embryon de *Phallusia mammillata* au stade de la gastrula tardive. Les couleurs sont arbitraires. La ligne noire pointillée indique le plan de coupe parasagittale. Adapté de (Guignard, 2015).

La voie de signalisation ERK est simplifiée chez les ascidies

En plus des propriétés des embryons d'ascidie favorables à l'imagerie d'événements du développement embryonnaire sur de longues périodes avec une haute résolution spatio-temporelle, les ascidies présentent des caractéristiques intéressantes pour l'étude de la voie ERK.

Tout d'abord, bien que les ascidies soient le groupe frère des vertébrés, elles ont conservé une plus grande compacité génomique. En effet, ces deux groupes ont divergé avant les deux séries d'événements de duplication du génome entier qui se sont produits dans la lignée des vertébrés et qui ont conduit à l'apparition de multiples paralogues pour chaque gène (Dehal and Boore, 2005), avec des activités potentiellement divergentes.

Ainsi, dans le génome des ascidies, les composants de la voie de signalisation ERK sont présents sous forme d'orthologues représentatifs uniques. Par exemple, *Phallusia* ERK1/2 représente à la fois ERK1 et ERK2 des vertébrés. Curieusement, le génome des ascidies est dépourvu de gènes codant pour le RAS classique, notamment H-RAS, K-RAS et N-RAS (Keduka et al., 2009). L'absence de RAS classique est compensée par M-RAS, qui agit comme médiateur de la voie de signalisation ERK dans les embryons de *Ciona*, comme il le fait également dans les cellules de mammifères (Keduka et al., 2009).

Les ascidies font un usage intensif de la voie FGF/RAS/ERK au cours de leur embryogenèse précoce

Les ascidies font un usage intensif de la voie FGF/RAS/ERK pour modeler leurs embryons précoces. Ainsi, 17 des 22 événements connus d'induction cellulaire jusqu'à la gastrulation impliquent cette voie (voir pour revue (Lemaire, 2009)). Au-delà de la gastrulation, de nombreux événements de

spécification de type cellulaire impliquent également la voie ERK. Par exemple, au cours du développement du système nerveux central (voir pour revue (Hudson, 2016)) ou de la spécification cardiaque (Razy-Krajka et al., 2018).

Activité de la signalisation ERK surface-de-contact dépendante au cours de l'induction neurale chez les ascidies

Certaines de ces inductions ERK dépendantes ont été caractérisées de manière assez détaillée, en particulier l'induction neurale dont l'étude précise a permis de mettre en évidence le rôle de la géométrie des contacts cellulaires dans la régulation de l'activité de la signalisation ERK.

Dans les embryons d'ascidies, la spécification du système nerveux commence au stade 32 cellules avec l'induction neurale, un processus développemental au cours duquel des cellules de l'ectoderme (le tissu cible) acquièrent un destin neuronal en réponse à des signaux provenant d'une région embryonnaire adjacente spécifique, le mésendoderme (le tissu inducteur) (Moreau et al., 2020). L'induction neurale précoce dans les embryons d'ascidie fait référence au début de la spécification d'une partie du système nerveux central, du cerveau antérieur et du tube neural dorsal (Hudson, 2016).

Elle a lieu au cours d'un seul cycle cellulaire d'environ 30 minutes au stade 32 cellules. À ce stade, l'embryon est constitué de 16 cellules ectodermiques dans l'hémisphère animal et de 16 cellules mésendodermiques dans l'hémisphère végétatif (Figure 25). Parmi les 16 cellules de l'ectoderme, deux paires de cellules uniquement : a6.5 et b6.5, sont induites pour former les progéniteurs neuraux (Figure 25).

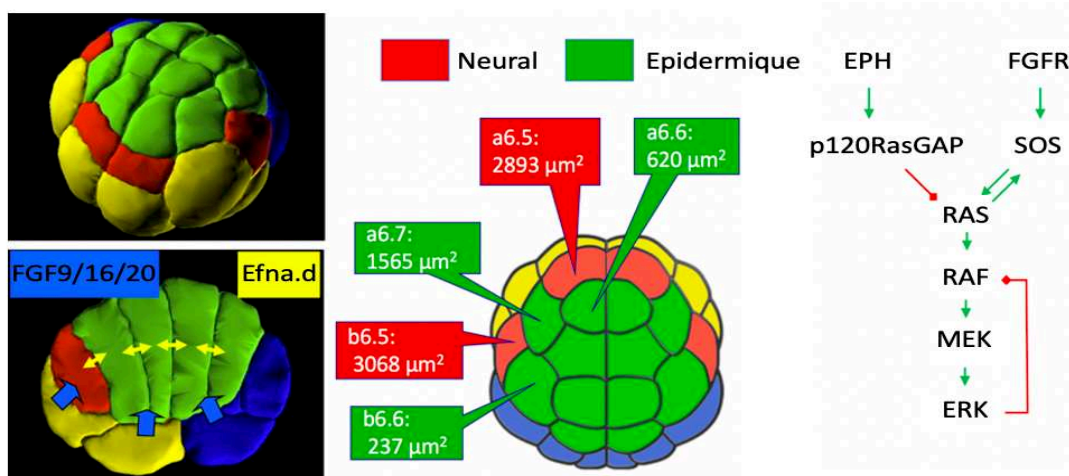


Figure 25 : Induction neurale surface-de-contact dépendante chez les ascidies via la voie ERK. A gauche : vues (haut : de côté, bas : coupe transversale) d'embryons de *Ciona intestinalis* au stade 32-cellules. Rouge : cellules animales neurales, vert : cellules animales épidermiques, bleu/jaune : cellules végétatives. Au centre : surfaces de contacts établies par chaque cellule animale individuelle avec les cellules végétatives émettant du FGF9/16/20. A droite : Schéma de la voie FGF-Eph/ERK impliquée dans l'induction neurale. Vert : activation, rouge : inhibition.

Un signal FGF, plus précisément le FGF9/16/20, a été identifié comme signal de l'induction neurale et il a été démontré qu'il agit via la cascade RAS/ERK (Figure 25) (Hudson and Lemaire, 2001; Bertrand

et al., 2003; Hudson et al., 2003). Le FGF9/16/20 commence à être exprimé dans l'hémisphère végétatif au stade 16-cellules et son expression est maintenue au stade 32-cellules. Lorsque la signalisation FGF/ERK est inhibée, toutes les cellules animales adoptent le destin épidermique. En revanche, lorsqu'elles sont traitées avec du FGF basique (bFGF recombinant), elles deviennent toutes neurales (Hudson and Lemaire, 2001; Hudson et al., 2003). Ces résultats révèlent donc que les cellules animales sont toutes compétentes pour répondre au signal FGF9/16/20 inducteur et adopter le destin neural et que la présence du ligand est le facteur limitant.

L'induction neurale médiée par le FGF9/16/20 et l'activité de la voie ERK a été directement liée à l'activation transcriptionnelle du gène *Otx*, qui joue un rôle essentiel dans la différenciation du système nerveux central antérieur chez les ascidies et de nombreuses espèces (Acampora et al., 2001; Bertrand et al., 2003). Lorsque la traduction du signal FGF9/16/20 est bloquée, l'expression d'*Otx* dans les quatre précurseurs neurales est perdue et ils expriment plutôt des gènes spécifiques de l'épiderme. Inversement, la surexpression de l'ARNm du FGF9/16/20 dans des explants ectodermiques conduit à une expression ectopique d'*Otx* dans une majorité de cellules de animales (Bertrand et al., 2003).

Ainsi, alors que toutes les cellules animales sont compétentes pour répondre au signal FGF9/16/20 et sont en contact direct avec les cellules végétatives exprimant le FGF9/16/20, seules quatre cellules animales (les paires a6.5 et b6.5) sur un total de 16 sont induites comme précurseurs neurales (Figure 25). Pour expliquer ce phénomène, une première réponse a été suggérée par les travaux de Tassy et collègues (Tassy et al., 2006). En effet, le développement et l'application d'un protocole de segmentation manuelle sur des images 3D d'embryons fixés (entiers ou coupés en deux afin de manipuler expérimentalement les surfaces de contact entre blastomères), leur a permis de reconstruire des embryons numériques 3D de *Ciona* et de quantifier précisément les surfaces de contact cellule-cellule dans ces embryons. Ces résultats ont montré que lors de l'induction neurale, la surface de contact entre les cellules inductrices et les cellules compétentes pour répondre au signal FGF est un déterminant majeur de l'induction. Dans tous les cas, dans les embryons entiers, a6.5 et b6.5 sont les cellules animales qui ont la plus grande surface de contact avec les cellules végétatives inductrices (Figure 25). De plus, dans les embryons entiers et coupés en deux, la surface de contact entre un blastomère animal et les cellules végétatives est très fortement corrélée, qualitativement et quantitativement, avec l'expression d'*Otx*. (Tassy et al., 2006). Toujours dans cette même étude, en supposant que le niveau de signalisation FGF/RAS/ERK d'une cellule est proportionnel à la zone de contact cellulaire avec les cellules végétatives exprimant le FGF, un mécanisme de seuil a été proposé dans lequel à partir d'une certaine surface de contact la signalisation ERK devient activée, conduisant à l'expression d'*Otx* dans ces cellules (Tassy et al., 2006). Cette hypothèse nécessite une supposition supplémentaire selon laquelle le FGF9/16/20 n'est pas diffusible et agit comme un signal à courte portée.

Ce premier modèle d'induction neurale chez les ascidies a ensuite été approfondi par la découverte de régulateurs négatifs du processus. Un ligand de l'éphrine, *Efna.d*, présent dans les cellules animales elles-mêmes, agit via *Eph3/p120RasGAP* pour atténuer la cascade RAS/ERK et prévenir l'activation ectopique d'*Otx* dans les cellules animales non neurales (Figure 25) (Haupaix et al., 2013; Ohta and Satou, 2013). Les cellules animales ayant la surface de contact la plus élevée avec les cellules végétatives exprimant le FGF9/16/20 ont, inversement, la surface de contact la plus plus

faible avec les cellules de animales exprimant l'éphrine (Tassy et al., 2006; Ohta and Satou, 2013). Lorsque la signalisation éphrine/Eph est bloquée, des cellules supplémentaires animales répondent aux signaux FGF9/16/20 activant ERK et exprimant *Otx* (Ohta and Satou, 2013). Ainsi, la signalisation antagoniste de la voie éphrine/Eph semble être essentielle pour réguler l'activité de la voie ERK et pour que les cellules animales des ascidies présentent un seuil correct de réponse aux signaux FGF. La sélection des précurseurs neuraux parmi les cellules animales semble dépendre du rapport entre leur surface de contact cellulaire avec les cellules végétatives inductrices (exprimant FGF9/16/20) et leur surface de contact cellulaire avec les cellules animales inhibitrices (exprimant *Efna.d*).

Cependant, la façon dont l'activité de ERK est régulée par ces signaux antagonistes FGF et éphrine pour conduire à une activation transcriptionnelle d'*Otx* spécifique dans les 4 cellules neurales n'est pas claire. En effet, bien qu'il semble qu'au cours de l'induction neurale, l'activité de ERK soit surface-de-contact dépendante, on ne sait pas, par exemple, si l'activation de ERK par le FGF9/16/20 est proportionnelle aux surfaces de contacts exposées au ligand ou si l'activité de ERK présente un seuil, c'est-à-dire que la voie ERK est activée uniquement au delà d'une certaine surface de contact exposée à l'inducteur.

OBJECTIFS DU TRAVAIL DE THÈSE

Au cours de mon projet de thèse, je me suis intéressé à la caractérisation de la dynamique spatio-temporelle quantitative de la voie ERK au cours du développement précoce de l'ascidie *Phallusia mammillata*, jusqu'en milieu de neurulation (~400 cellules).

Comme illustré dans cette introduction, en dépit de l'importance de la voie ERK nos connaissances actuelles de cette voie sont pour la plupart statiques et une compréhension intégrée de sa dynamique spatio-temporelle manque. Les développements récents de rapporteurs génétiquement codés de l'activité ERK ont ouvert la voie à sa mesure quantitative en temps réel à une résolution unicellulaire. Cependant, le nombre d'études de biologie du développement utilisant ces biosenseurs dans des contextes multi-cellulaires *in vivo* est encore limité et notre compréhension quantitative de la dynamique spatio-temporelle de la voie ERK est limitée à quelques éléments de compréhension non reliés entre eux.

Ainsi, j'ai étudié la voie FGF/Ras/ERK chez l'ascidie *Phallusia mammillata*, un petit invertébré marin membres des tuniciers, groupe frère des vertébrés, qui possède des gènes uniques FGFR, SOS, Ras, Raf, MEK et ERK et qui utilise cette voie pour contrôler de nombreux événements de communication cellulaire au cours de son embryogenèse précoce (Lemaire, 2009).

Sur ce modèle, notre équipe a également développé au cours des années précédant mon arrivée une capacité unique à identifier/segmenter/suivre les cellules sur de longues périodes de leur développement avec une haute résolution temporelle, une condition essentielle pour mesurer automatiquement la dynamique de l'activité de la voie ERK avec une résolution cellulaire au cours de l'embryogenèse. Cependant, la méthodologie permettant de mesurer automatiquement la dynamique de l'activité de ERK *in vivo* en temps réel au cours du développement des embryons d'ascidies restait à établir.

Afin de décrire le pattern spatio-temporel d'activation de la voie FGF/Ras/ERK, je me suis donc tout d'abord concentré sur le développement des méthodologies expérimentales nécessaires à la quantification manuelle *in vivo* en temps réel de l'activité ERK dans les embryons de *Phallusia mammillata* en utilisant le biosenseur ERK-KTR (Regot et al., 2014). Comme preuve de principe, nous avons utilisé cet outil dans le cadre de la révision d'un article visant à tester l'hypothèse selon laquelle les surfaces de contact entre les cellules qui émettent un signal et celle qui répondent fournissent les informations quantitatives qui déterminent le résultat des inductions au cours de l'embryogenèse chez les ascidies, un modèle alternatif au modèle classique de « Morphogène » où la concentration du ligand fournit cette information. Ces résultats sont présentés dans le chapitre 1.

Cependant, la méthodologie développée dans le cadre de cette étude présente certaines limitations, notamment le fait que les mesures effectuées manuellement sont très chronophages. Il faudrait l'équivalent d'un an de travail à la main pour mesurer le ERK-KTR dans les noyaux de toutes les cellules d'un seul embryon imagé toutes les 2 minutes entre le stade 32-cellules et le milieu de neurulation. Les mesures manuelles n'ont donc été effectuées que sur un stade donné, le stade 64-cellules et de façon discontinue, ne permettant pas d'accéder à la dynamique fine de l'activité ERK.

J'ai donc contribué au développement d'outils visant à automatiser certaines étapes de la méthodologie afin de produire un atlas spatio-temporel de l'activité ERK au cours de l'embryogenèse chez *Phallusia*.

Un des objectifs de cet atlas fonctionnel de la signalisation ERK, qui peut être combiné aux informations géométriques des cellules dans l'embryon, est de quantifier la variabilité de la signalisation ERK dans les cellules homologues d'un même embryon (génome identique) et d'embryons différents (polymorphisme), en lien avec la variabilité des géométries des contacts cellulaires. J'ai donc tout d'abord travaillé au développement d'un cadre théorique pour le nommage automatique et non biaisé des cellules équivalentes dans les embryons stéréotypés de l'ascidie *Phallusia mammillata* jusqu'en milieu de neurulation (~400 cellules à ce stade) (chapitre 2).

J'ai ensuite travaillé sur le développement des méthodes expérimentales et informatiques pour automatiser les mesures du ERK-KTR dans les noyaux et le cytoplasme de chaque cellule d'un embryon en développement (chapitre 3). Ce travail a permis de générer un premier atlas fonctionnel de la dynamique spatio-temporelle de l'activité ERK sur une large période du développement embryonnaire. Ces mesures de l'activité de ERK à partir du ERK-KTR ont été mises à disposition sous la forme d'un atlas interactif en ligne à l'aide d'un outil en ligne dédié appelé "MorphoNet" pour chacun des 4 embryons analysés. L'analyse de la mesure de l'activité ERK moyenne dans les différentes cellules a ensuite permis de montrer que l'on retrouve ce qui est connu des inductions ERK-dépendantes chez les ascidies, ce qui soutient la pertinence biologique des mesures automatiques de l'activité ERK de notre atlas. Enfin l'analyse de l'activité moyenne de ERK, en fournissant de nouveaux potentiels candidats pour des inductions différentielles ERK-dépendantes entre couples de cellules soeurs, suggère que nos connaissances actuelles ne sont pas exhaustives, un objectif qui pourrait être atteint par une analyse détaillée de notre atlas spatio-temporel systématique de l'activité ERK dans un plus grand nombre d'embryons.

RÉSULTATS

CHAPITRE 1 : L'ANALYSE QUANTITATIVE IN VIVO DE L'ACTIVITÉ ERK AU STADE 64-CELLULES SUPPORTE LE MODE DE COMMUNICATION CELLULAIRE SURFACE-DE-CONTACT DÉPENDANT DANS LES EMBRYONS D'ASCIDIES

Introduction

Le développement embryonnaire précoce des ascidies montre une invariance cellulaire globale jusqu'au stade de la gastrula (Lemaire, 2009, 2011) et les lignées cellulaires de plusieurs tissus semblent être stéréotypées jusqu'aux stades du bourgeon caudal (Figure 19) (Cole and Meinertzhagen, 2004; Razy-Krajka et al., 2014; Tolkin and Christiaen, 2016). Contrairement aux nématodes, dont la morphologie des embryons varie entre les espèces (Goldstein, 2001), la géométrie des embryons d'ascidies est elle restée essentiellement inchangée depuis l'émergence du groupe, il y a environ 400 millions d'années (Figure 19) (Delsuc et al., 2018), malgré une divergence génomique importante (Lemaire, 2011; Stolfi et al., 2014).

Afin de quantifier et mieux comprendre les mécanismes qui déterminent cette extrême reproductibilité du développement embryonnaire précoces des ascidies, notre équipe, en travaillant avec des collaborateurs de long terme issus de différentes disciplines : mathématiques appliquées, informatique, traitement d'images biomédicales et réseaux d'apprentissage profonds, a développé une capacité unique de segmenter et suivre chaque cellule toutes les 2 minutes pendant l'embryogenèse de *Phallusia mammillata*. (Guignard, 2015; Guignard et al., 2020). Ainsi, en combinant l'imagerie multi-vues à feuille de lumière (MuViSPIM) d'embryons vivants de l'ascidie *Phallusia mammillata*, exprimant un marquage membranaire fluorescent, avec un algorithme de segmentation développé dans l'équipe et adapté à ce type de données, appelé ASTEC, une version numérisée de chaque embryon imagé est produite (Figure 24). Cette méthode a permis la reconstruction de 10 embryons de *Phallusia mammillata* (Astec-pm1-10) entre les stade 32-cellules (st.6) et le début de l'élongation de la queue (st. 17) (Guignard et al., 2020). A partir de cet ensemble de données d'un nouveau type, de nombreux paramètres ont pu être précisément mesurés et suivis tels que : les volumes cellulaires, la position des cellules, le voisinage cellulaire et les surfaces de contacts entre cellules voisines (Guignard et al., 2020). L'analyse de la variabilité des ces différents paramètres cellulaires a révélé un développement embryonnaire "figé", sans croissance cellulaire, montrant une grande reproductibilité des positions cellulaires et avec une conservation des voisinages allant jusqu'à la conservation précises des mesures relatives des surfaces de contacts entre cellules voisines (Guignard et al., 2020).

Pour expliquer cette extrême reproductibilité de la géométrie du développement embryonnaire des ascidies (Figure 19) (Guignard et al., 2020), notre équipe a proposé un modèle selon lequel la mesure précise des surfaces de contact entre cellules émettrices et cellules réceptrices serait un déterminant essentiel de la communication cellulaire, imposant des contraintes favorisant l'invariance développementale. En effet, le développement des ascidies se caractérise par un nombre de cellules faibles et une restriction du destin plus précoce que chez la plupart des embryons animaux. Ainsi, 94 des 112 cellules de la gastrula précoce de l'ascidie ont un destin déterminé, chacune contribuant à un seul type de tissu larvaire (Figure 26) (Lemaire, 2009). La communications cellulaires à courte portée joue un rôle crucial dans la spécification de nombreux destins (Nishida, 2005; Lemaire, 2009) et des travaux antérieurs sur l'induction neurale au stade 32 cellules par la voie FGF/ERK ont montré

que cette induction ne se produit que dans les quatre cellules ectodermiques dont la surface de contact avec les cellules émettrices de FGF dépasse une valeur seuil (Figure 25) (Tassy et al., 2006).

Vegetal view of 112-cell stage embryo

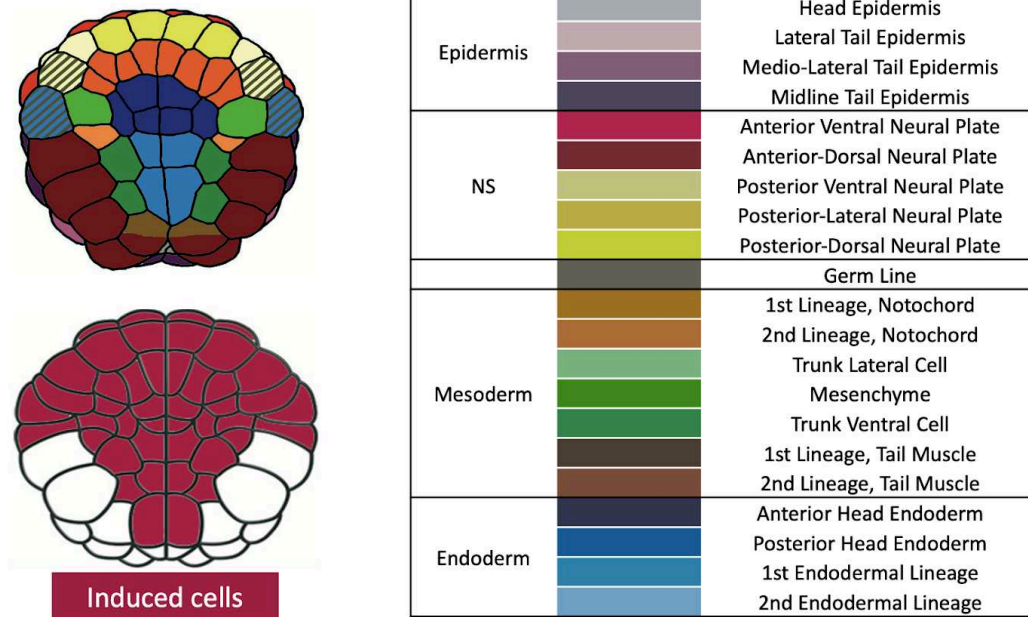


Figure 26 : Les mécanismes de communication cellulaire à courte portée peuvent imposer des contraintes favorisant l'invariance du développement des ascidies. Vue végétative d'un embryon au stade 112-cellules (st. 10). Haut : la coloration sur chaque blastomère indique son destin (tableau avec les noms des destins associés à droite). La majorité des cellules sont restreintes à un seul destin contribuant à un seul type de tissu larvaire. Bas : la coloration sur chaque blastomère indique les cellules induites par des mécanismes de communication cellulaire de courte portée. Adapté de (Lemaire, 2009).

Pour tester la généralité de ce modèle d'inductions cellulaires surfaces-de-contacts dépendantes, nous avons combiné imagerie quantitative *in vivo*, informatique, modélisation théorique et manipulations expérimentales.

A partir des atlas morphologiques quantitatifs, un cadre théorique modélisant le lien entre les contacts cellule-cellule et les inductions cellulaires a été développé. Ce modèle mathématique, basé sur les systèmes d'équations différentielles ordinaires et la loi d'action de masse, intègre la géométrie des contacts cellulaires en considérant que le nombre total de récepteurs pouvant être activés par un ligand est proportionnel à la surface de la cellule exposée au ligand. Ainsi, pour les six voies de signalisation (FGF, Ephrin, Wnt, Bmp, Nodal et Notch) impliquées dans toutes les inductions embryonnaires connues jusqu'en début de gastrulation (Lemaire, 2009), ce modèle simplifié d'inductions à courte distance produit une estimation semi-quantitative du niveau relatif de signalisation pour chaque voie dans chaque cellule.

Ma contribution à cette étude a été de développer l'approche expérimentale permettant de comparer les prédictions semi-quantitatives des niveaux d'activation des voies issues du modèle à des mesures quantitatives *in vivo* dans des embryons en cours développement. Je me suis concentré

sur la voie FGF/ERK, qui est la principale voie d'induction dans les embryons d'ascidies précoces (Lemaire, 2009) et pour laquelle un biosenseur fluorescent de l'activité ERK, composant le plus en aval de la cascade, avait été précédemment développé et utilisé dans d'autres organismes modèles : le ERK-KTR (Kinase Translocation Reporter), qui est exporté du noyau lors de l'activation de ERK (Regot et al., 2014). J'ai donc co-microinjecté dans des ovocytes de *Phallusia mammillata* des ARNm synthétiques codants pour un marqueur membranaire fluorescent et pour le ERK-KTR. Les embryons vivants exprimant ces deux marqueurs ont été imagés dans un microscope multi-vues à feuilles de lumière (MuViSPIM) puis digitalisé et les intensités ERK-KTR dans le noyau de chaque cellule ont été mesurées manuellement et normalisées pour pouvoir comparer les intensités entre embryons.

J'ai commencé par vérifier que le ERK-KTR, qui a été développé initialement pour des cellules de mammifères, reflète aussi l'activité ERK dans les embryons d'ascidies en : i) montrant que les signaux ERK-KTR sont hautement reproductibles entre embryons, ii) en contrôlant que l'activité ERK inférée d'après le ERK-KTR est réduite à un niveau basal après avoir bloqué l'activité de MEK (kinase en amont de ERK) grâce à un inhibiteur pharmacologique et iii) en comparant le pattern des valeurs ERK-KTR mesurées au pattern d'évènements d'inductions connus impliquant cette voie au stade 64-cellules.

Le niveau d'activité de ERK a été quantifié dans 5 embryons différents pour chacune des cellules, sur 5 pas de temps au cours du stade 64-cellules, stade précoce de l'embryogenèse chez les ascidies, au cours duquel plusieurs inductions ERK dépendantes ont lieu. Un modèle mathématique relativement simple permet d'inférer le niveau d'activation ERK à partir des mesures du ERK-KTR (Regot et al., 2014). Après la segmentation de cet ensemble de données et l'extraction des mesures de contacts cellule-cellule, nous avons lancé le modèle et montré que ses prédictions d'activation de la voie FGF/ERK sont fortement corrélées aux valeurs d'activité ERK inférées à partir des mesures expérimentales du ERK-KTR (corrélation de Spearman 0,9). Le modèle d'induction simplifié dépendant des surfaces-de-contacts fournit donc des estimations quantitatives fiables des niveaux d'activation de la signalisation ERK pour chaque cellule des embryons d'ascidie au stade 64-cellules.

Pour analyser plus en détail la réponse du niveau d'activité de la voie ERK aux variations de la surface d'exposition au signal, j'ai modifié expérimentalement les surfaces de contact entre les cellules en coupant un embryon (préalablement micro-injecté) au stade 2-cellules le long de son plan de symétrie bilatéral. Cette perturbation embryologique n'a pas empêché l'identification des cellules embryonnaires individuelles. Néanmoins, elle a suffisamment modifié les contacts cellulaires pour altérer les intensités de signalisation ERK inférées dans plusieurs cellules, le modèle prédisant correctement la direction de ces changements.

Ainsi, l'ensemble de ces résultats ont montré que la mesure des surfaces de contacts entre cellules inductrices et réceptrices est un bon prédicteur du niveau d'activité de la voie ERK et ils soutiennent la proposition selon laquelle les inductions surfaces-de-contacts dépendantes sont un mécanisme majeur chez les ascidies. L'utilisation intensive de ce mode d'inductions surfaces-de-contacts dépendants jouerait alors un rôle dans l'extrême reproductibilité du développement embryonnaire observé chez les ascidies en imposant des contraintes sur la géométrie des contacts individuels.

Ces résultats ainsi que l'étude complète sont publiés dans l'article : « Contact-area dependent cell communications and the morphological invariance of ascidian embryogenesis » publié en juillet 2020 dans le journal Science (Lien vers l'article : <https://www.science.org/doi/10.1126/science.aar5663>).

Un extrait du supplément d'information de cet article ciblant ma contribution est également présenté. Le supplément d'information complet peut quant à lui être téléchargé depuis le site du journal :

https://www.science.org/doi/suppl/10.1126/science.aar5663/suppl_file/aar5663_guignard_sm.pdf.

Manuscrit de l'article : Guignard, Léo et *al.* "Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis." Science (New York, N.Y.) Vol. 369,6500 (2020): EAAR5663.

RESEARCH ARTICLE SUMMARY

DEVELOPMENTAL BIOLOGY

Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis

Léo Guignard*, Ulla-Maj Fiúza*, Bruno Leggio, Julien Laussu, Emmanuel Faure, Gaël Michelin, Kilian Biasuz, Lars Hufnagel†, Grégoire Malandain†‡, Christophe Godin†‡, Patrick Lemaire†‡

INTRODUCTION: Within each animal species, embryonic development is highly reproducible, ensuring the faithful production of a complex organism with precisely arranged and shaped organs. In most animal embryos, reproducibility is found at the tissue scale, the behaviors of individual cells being stochastic beyond the first cell divisions. Ascidiaceans, a group of marine invertebrate chordates, show an extreme form of embryonic reproducibility: Homologous cells can be found across individual embryos, and early embryonic cell lineages are considered invariant. Embryonic geometries are even conserved between species, which diverged 400 million years ago and have very dissimilar genomes. Because of their evolutionary conservation of early embryonic development

and ability to buffer genetic divergence, ascidiaceans constitute attractive model systems to study the mechanisms driving cellular reproducibility.

RATIONALE: To quantify embryonic reproducibility in the ascidian *Phallusia mammillata*, we first built a high-resolution atlas of embryonic cell lineages, cell shapes, and cell interactions. We imaged 10 live embryos every 2 min up to the end of the neurula stages using multiview light-sheet microscopy. To systematically measure the developmental variability of a range of temporal and spatial cellular features, we developed a robust and scalable adaptive segmentation and tracking of embryonic cells procedure (ASTEC) compatible with high-throughput multiview light-sheet

imaging datasets. We related these features to cell fate specification, which in ascidiaceans is mainly controlled by differential sister cell inductions. Inspired by previous work indicating that the area of contact to signaling cells controls ascidian neural induction, we integrated our geometric description with a signaling gene expression atlas. This integration allowed us to test, through computational and experimental approaches, the hypothesis that contact area-dependent cell communication imposes constraints on embryonic geometries.

RESULTS: We found that, up to the neurula stages, *Phallusia* embryos develop without cell growth, programmed cell death, or cell neighbor exchanges. Beyond cell position, cell cycle duration, and cell lineages, we observed a

ON OUR WEBSITE

Read the full article at <https://dx.doi.org/10.1126/science.aar5663>

high reproducibility of cell arrangements: 75% of cells shared at least 80% of their neighbors in all 10 embryos studied. Furthermore, the areas of contact

between homologous cells varied by less than 20% across embryos. Mechanistically, we uncovered a tight link between the control of cell arrangements and asymmetric cell divisions, which give rise to sister cells of distinct fates. We then combined computational and experimental approaches to reveal that areas of cell contact between signaling and responding cells have sufficient encoding potential to explain all known early embryonic inductions, without the need to invoke gradients of ligand concentration. Finally, using geometrical perturbations of embryonic development we demonstrated that precise areas of cell-cell contact were important for mesendodermal and neural fate specification.

CONCLUSION: Our work establishes the highly reproducible ascidian embryo as a framework to bridge cell behaviors, morphogenesis, and the underlying regulatory program. The ASTEC pipeline allows systematic automated whole-cell segmentation and tracking across whole embryos in high-throughput light-sheet datasets. Second, we establish the geometric control of embryonic inductions as an alternative to classical morphogen gradients and suggest that the range of cell signaling events sets the scale at which embryonic reproducibility is observed. Finally, our study suggests that the high level of reproducibility of ascidian embryonic geometries may paradoxically lift constraints on the evolution of ascidian genomes, thereby contributing to rapid molecular evolution. ■

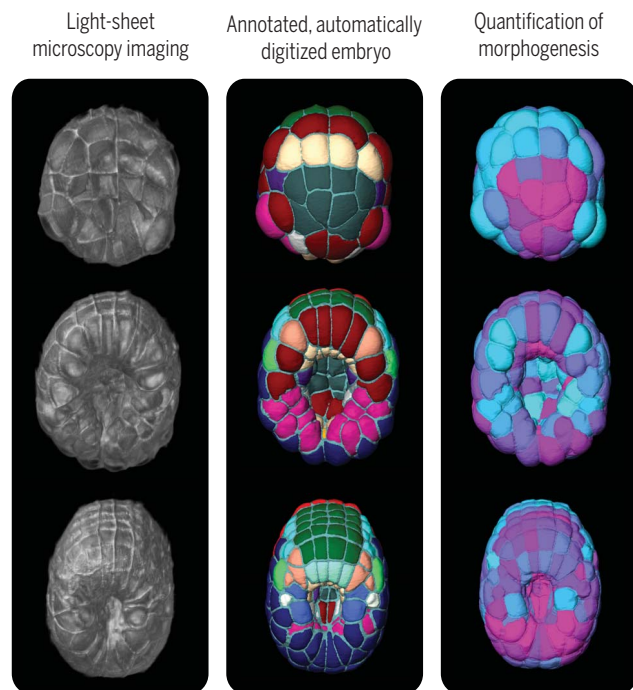
The list of author affiliations is available in the full article online.

*These authors contributed equally to this work.

†Corresponding author. Email: hufnagel@embl.de (L.H.); gregoire.malandain@inria.fr (G.Ma.); christophe.godin@inria.fr (C.G.); patrick.lemaire@crbm.cnrs.fr (P.L.)

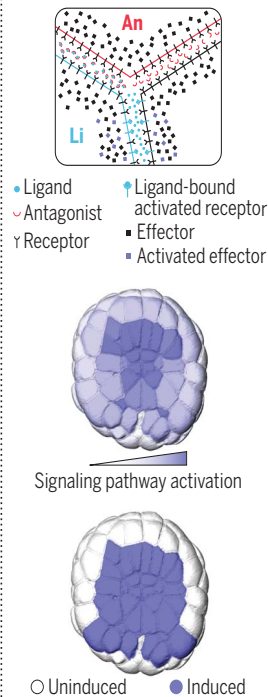
‡These authors contributed equally to this work.

Cite this article as L. Guignard et al., *Science* **369**, eaar5663 (2020). DOI: [10.1126/science.aar5663](https://doi.org/10.1126/science.aar5663)

Quantitative analysis of *Phallusia* embryogenesis

Reconstruction and modeling of *Phallusia* embryogenesis. (Left) Quantitative analysis of *Phallusia* embryogenesis. We combined live light-sheet imaging of cell membranes (left images) with automated cell segmentation and tracking with color-coded cell fates (center images) to extract quantitative cell morphological properties (right images, color-coded by cell compactness). From top to bottom: embryo at the 64-cell, mid-gastrula, and late gastrula stages. (Right) Cell signaling model. We first made simplifying assumptions concerning the distribution and diffusion of signaling pathway components (top) and then integrated cell contact geometry with gene expression profiles to predict pathway activation levels in single cells (center) and binarized induction status (bottom).

Cell signaling model



RESEARCH ARTICLE

DEVELOPMENTAL BIOLOGY

Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis

Léo Guignard^{1,2,3,*†}, Ulla-Maj Fiúza^{1,4,*}, Bruno Leggio^{1,2,5}, Julien Laussu^{1,†}, Emmanuel Faure^{1,2,6,§}, Gaël Michelin⁷, Kilian Biasuz¹, Lars Hufnagel^{4¶}, Grégoire Malandain^{7¶}, Christophe Godin^{2,5¶}, Patrick Lemaire^{1¶}

Marine invertebrate ascidians display embryonic reproducibility: Their early embryonic cell lineages are considered invariant and are conserved between distantly related species, despite rapid genomic divergence. Here, we address the drivers of this reproducibility. We used light-sheet imaging and automated cell segmentation and tracking procedures to systematically quantify the behavior of individual cells every 2 minutes during *Phallusia mammillata* embryogenesis. Interindividual reproducibility was observed down to the area of individual cell contacts. We found tight links between the reproducibility of embryonic geometries and asymmetric cell divisions, controlled by differential sister cell inductions. We combined modeling and experimental manipulations to show that the area of contact between signaling and responding cells is a key determinant of cell communication. Our work establishes the geometric control of embryonic inductions as an alternative to classical morphogen gradients and suggests that the range of cell signaling sets the scale at which embryonic reproducibility is observed.

The development of each animal species is highly reproducible, ensuring the faithful replication of a complex organism in which organs are precisely shaped, sized, and positioned. This fidelity of embryogenesis despite genetic polymorphism and fluctuating environmental conditions is critical for the perpetuation of species.

Reproducibility can be observed at different scales. Many vertebrate and invertebrate embryos initially display a stereotyped pattern of orientation and geometry of individual cell divisions, which can be explained by simple self-organizing mechanical rules (1). This initial cellular reproducibility is usually lost during early cleavage stages, as the number of cells increases. It is, however, maintained during

later developmental stages in nematodes (2) and in ascidians (3), a group of marine invertebrates closely related to vertebrates (4). Unlike those of nematodes (5), ascidian embryo geometries have even remained essentially unchanged since the emergence of the group, around 400 million years ago (3, 6), despite extensive genomic divergence and gene regulatory network rewiring (7). Because of the distinctive evolutionary conservation of their early embryonic development and because of their notable ability to buffer genetic divergence, ascidians constitute an attractive model

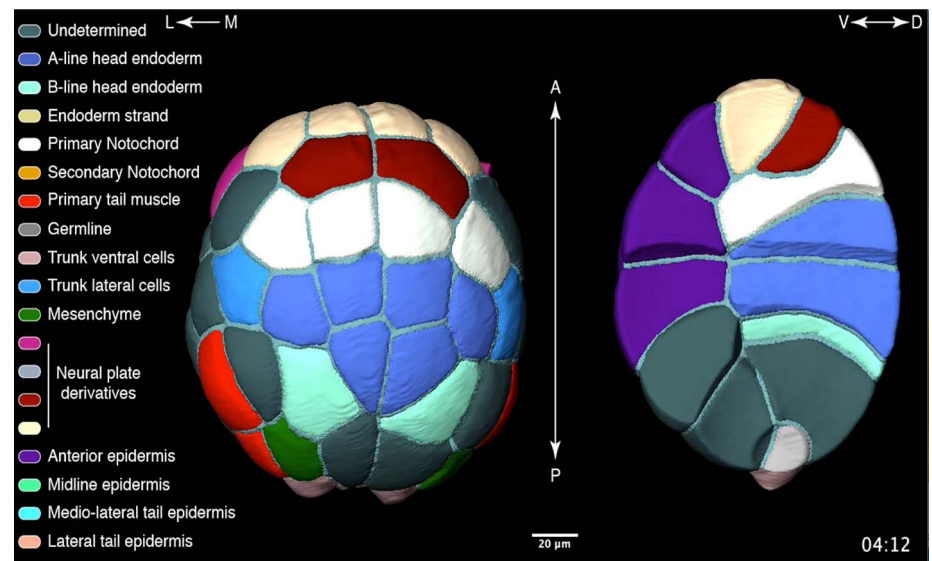
system to study the mechanisms driving and maintaining cellular reproducibility.

Development of ascidians and nematodes is characterized by lower cell number and earlier fate restriction than most animal embryos. Ninety four of the 110 early gastrula cells in the ascidian *Halocynthia roretzi* are fate restricted, each contributing to a single larval tissue type (8). Most mesendodermal tissues are generated by a very small, and specific, number of fate-restricted early gastrula progenitors: six for the notochord, four for the mesenchyme, and two for the heart or trunk lateral cells (8). Although reduced cellular complexity alone cannot explain invariance (9), the precision of the mechanisms needed to specify a small, specific number of fate-restricted progenitors of each tissue may impose constraints favoring developmental invariance.

Short-range cell communication plays a crucial role in the specification of most of the ascidian mesendodermal and neural lineages (10, 11). Previous work on the induction at the 32-cell stage of ascidian neural progenitors by fibroblast growth factor (FGF) revealed that this induction only occurs in the four ectodermal cells whose area of contact to FGF-emitting cells exceeds a threshold value (12). This contact area-dependent induction exemplifies how the specification of a precise number of tissue precursors can rely on the precise spatial arrangement of individual embryonic cells.

In this work, we tested the hypothesis that complex feedback between fate specification and morphogenetic processes, whose strength is amplified by the reduced cell number, promotes developmental invariance. We first developed methods to quantitatively study cellular variability in the ascidian *Phallusia mammillata*. We then used these methods in the chordate embryo to explore the relationships between

¹CRBM, Université de Montpellier, CNRS, 34293 Montpellier, France. ²Virtual Plants, Université de Montpellier, CIRAD, INRA, Inria, 34095 Montpellier, France. ³Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA. ⁴Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany. ⁵Laboratoire Reproduction et Développement des Plantes, Université de Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA, Inria, 69342 Lyon, France. ⁶Institut de Recherche en Informatique de Toulouse (IRIT), Universités Toulouse I et III, CNRS, INPT, ENSEEIHT, 31071 Toulouse, France. ⁷Morpheme, Université Côte d'Azur, Inria, CNRS, I3S, France. *These authors contributed equally to this work. †Present address: Max Delbrück Center for Molecular Medicine in the Helmutz Association, Berlin Institute of Health, 10115 Berlin, Germany. ‡Present address: Institut Clément Ader, Université Toulouse III, CNRS, INSA, ISAE-SUPAERO, Mines-Albi, 31400 Toulouse, France. §Present address: Equipe ICAR, LIRMM, Université de Montpellier, CNRS, 34095 Montpellier, France. ¶Corresponding author. Email: hufnagel@embl.de (L.H.); gregoire.malandain@inria.fr (G.Ma.); christophe.godin@inria.fr (C.G.); patrick.lemaire@crbm.cnrs.fr (P.L.) #These authors contributed equally to this work.



Movie 1. Digitized *P. mammillata* embryo. Vegetal view (left) and side view through a sagittal section (right) of the ASTEC-Pm1 segmented and fate-colored *P. mammillata* embryo. Anterior is to the top. The tissue fate color code is shown on the left-hand side.

the reproducibility of cell arrangements, the precision of fate specification processes, and the range of embryonic inductions.

A high-resolution geometric atlas of embryonic cell shapes and interactions

Using confocal multiview light-sheet microscopy (13), we imaged the development of 10 optically transparent *P. mammillata* embryos with fluorescently labeled plasma membranes, every 2 min from cleavage to neurula stages (Fig. 1A, figs. S1 to S4, table S1, and movies S1 and S2). Systematic segmentation and long-

term tracking of all membrane-labeled cells of a developing embryo remain challenging (14), despite recent advances (15–19). Analysis of the limitations of our previous two-pass multi-angle image acquisition, three-dimensional reconstruction and cell segmentation–automated lineage tracking (MARS-ALT) pipeline (20) (figs. S5 to S8) guided the development of an algorithm adapted to high-throughput light-sheet datasets called ASTEC, for adaptive segmentation and tracking of embryonic cells (Fig. 1, B and C). ASTEC is a single-pass algorithm that simultaneously segments and tracks

whole cells by iteratively projecting segmentations from one time point onto the next (figs. S9 to S14). ASTEC produces a digitized version of each embryo: a segmentation of all embryonic cells existing at each time point, integrated into global cell lineages, to which we associated their known larval tissue fates and from which quantitative geometrical parameters can be extracted (Fig. 1D, figs. S15 to S17, Movie 1, and movie S3). The 10 digitized *P. mammillata* embryos, ASTEC-Pm1 to ASTEC-Pm10, constitute a quantitative and dynamic atlas of cell positions, geometries, and ancestry over a large fraction of a metazoan developmental program (Movie 2 and movie S4), which can be interactively explored through the MorphoNet online morphological browser (21).

Figure 1C illustrates the ASTEC-Pm1 digitized embryo, which contains a total of 67,003 segmented three-dimensional (3D) cell “snapshots,” and describes the behavior over time of 1395 individual cells generated by 697 cell division events. Manual inspection revealed that the pipeline only missed 4/1399 cells and 3/700 cell divisions, and it assigned 99% of voxels to the correct cell (figs. S6 and S18 and tables S2 and S3). In a manner consistent with previous work in another ascidian species (22), no programmed cell death was detected, such that all cells but one were tracked until their division or the end of the movie (fig. S15). On this dataset, ASTEC outperformed RACE, a reference membrane-based segmentation algorithm adapted to embryonic datasets (16) (fig. S19). The segmentation and tracking accuracy of all 10 embryos was comparable to that of ASTEC-Pm1 (tables S2 and S3).

Ascidian development is reproducible down to the scale of cell-cell contacts

We systematically measured the developmental variability of a range of temporal and spatial cell features during gastrulation and neurulation. At the temporal level, the median variability of the cell cycle durations of homologous cells (equivalent bilateral cells within an embryo or across embryos) was on the order of one time point, within the image sampling error (figs. S20 to S22). The cell lineages of homologous cells were similar (Fig. 2A), as quantified with a normalized tree-edit distance capturing the average cost per cell of the transformation of one tree into the other (fig. S23). This high lineage similarity led to a quasi-invariant number of cells at any given developmental stage across the 10 embryos (2% average variation in cell number across all embryos) (Fig. 2B). At the spatial level, we observed up to a twofold variation in the total volume of individual embryos, reflecting egg size heterogeneity, as no cell growth occurred during the considered developmental period (fig. S24). The relative cell volumes (i.e., cell volume divided by total embryo volume) (fig. S25) and relative cell

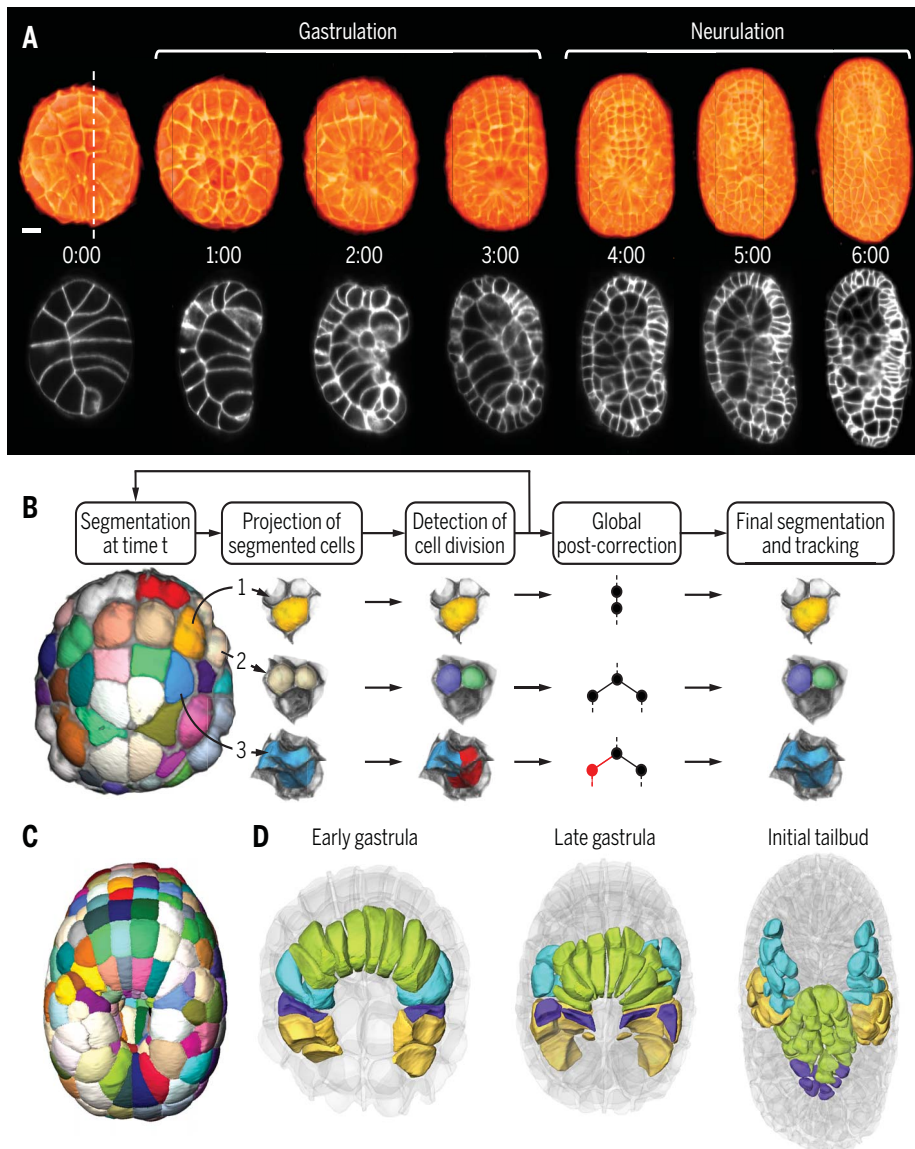


Fig. 1. High-resolution multiview light-sheet imaging and segmentation of *P. mammillata* embryos.

(A) (Top) Vegetal and dorsal views of 3D rendering at the indicated time points of the imaged embryo after fusion of the images taken along the four angles of view. The time scale indicates the duration of imaging in hours. (Bottom) Parasagittal sections along the plane shown with the vertical dashed white line. Scale bar, 20 μm . (B) The ASTEC pipeline. Different cases are illustrated: a nondividing cell (yellow), a dividing cell (white), an oversegmented nondividing cell (blue) corrected during postcorrection. For more details, see fig. S9. (C) Dorsal view of the segmentation of ASTEC-Pm1 at the late gastrula stage (time $t = 2$ hours 30 min). Colors are arbitrary. (D) Position of mesoderm progenitors in ASTEC-Pm1 at the early gastrula ($t = 1$ hour 10 min), late gastrula ($t = 2$ hours 40 min), and late neurula ($t = 4$ hours 40 min) stages. B-line mesenchyme, yellow; trunk lateral cells, cyan; primary notochord, green; secondary notochord, purple. Anterior is to the top. For more details, see fig. S16.

positions of matching cells within and between embryos (Fig. 2, C and D) had, however, a median variability below 5%, highlighting the reproducibility of cell geometries and positions as well as their scaling with egg size. The reproducibility of relative cell positions was higher in *Phallusia* than in a *Caenorhabditis elegans* dataset (23) (Fig. 2C and fig. S26). Cell neighborhoods were also highly conserved: Up

to the late neurula stage, 75% of matching cells shared at least 80% of their physical neighbors across all embryos (Fig. 2D and fig. S27). Most shared cell contacts lasted throughout the cell cycle (fig. S27), such that cell neighbor exchanges, a major source of stochastic variability during animal embryogenesis, were rare. The relative areas of shared cell-cell contacts varied by less than 20% across the 10 embryos (Fig. 2E).

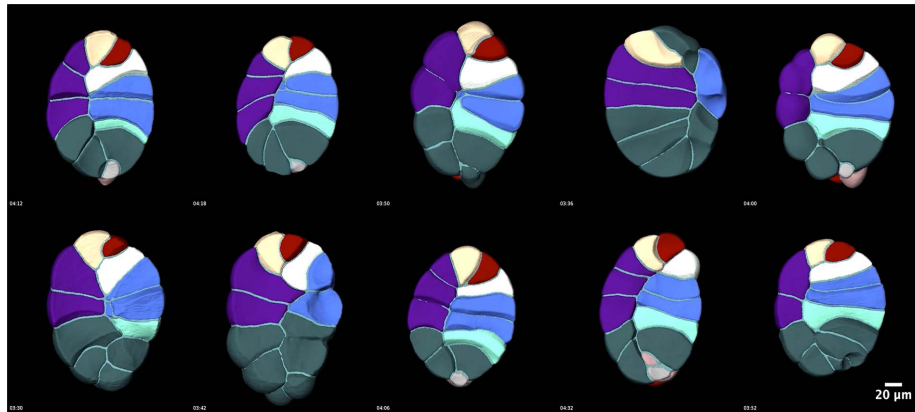
To assess the developmental impact of the high genetic diversity found in ascidians (24), we compared the variability observed between matching bilateral cells within an embryo, which share the same genotype, with that observed between homologous cells of embryos generated from different wild-caught adults of different genotypes. Variability in cell lineage trees (Fig. 2A) and relative cell volumes (fig. S25C) of homologous cells was not significantly different within and between embryos, whereas relative cell positions (Fig. 2C) and areas of cell contacts (Fig. 2E) were only slightly less reproducible between embryos. Consistent with a developmental control, as proposed in *C. elegans* (23), cell position variability differed between cell lineages (fig. S27).

Areas of cell contacts are modulated by asymmetric cell divisions

Cell contact areas are determined by cell growth, cell death, cell neighbor exchanges, cell shape changes, and the geometry of cell divisions. The first three processes are absent or rare in ascidian embryos. In this study, we focused on the impact of cell divisions on cell contacts.

We first classified cell divisions into symmetric and asymmetric divisions. Symmetric divisions are defined as giving rise to two daughters with the same larval tissue fate. By contrast, asymmetric divisions produce two daughter cells differing in their larval tissue fates. Because 64-cell-stage progenitors seeding similar lineage trees, according to the tree-edit distance, adopt the same larval tissue fate (figs. S28 and S29), we identified asymmetric divisions by comparing the lineages seeded by sister cells, as exemplified in Fig. 3B. This approach correctly classified 82 and 100% of a ground truth of 28 asymmetric and 30 symmetric divisions, respectively (tables S4 to S6). The same classification, applied to cell divisions of previously undetermined status, identified 12 novel asymmetric cell divisions (table S6). These divisions were found in the lineages of larval tissues made of a diversity of cell fates or giving rise to several juvenile mesodermal and endodermal tissues. More than 32% of divisions (44 out of 136) occurring during the period considered are therefore asymmetric.

Asymmetric cell divisions have been associated to the control of cell division geometry (25). Most *Phallusia* embryonic cell divisions until the neurula stages produced daughters of equal volumes (fig. S30A and tables S7 and S8) and equal lifespan (fig. S30B and tables S9 and S10). Thirty-two divisions (19%), mostly found in the mesendodermal and neural lineages, produced daughters with consistently unequal volumes across embryos (Fig. 3A and tables S7 and S8). These unequal divisions also affected the timing of division of daughter cells. The larger sibling generally divided earlier (fig. S30C), in agreement with a negative



Movie 2. Sagittal section of 10 digitized *P. mammillata* embryos. Side views through sagittal sections across the ASTEC-Pm1 to ASTEC-Pm10 segmented and fate-colored *P. mammillata* embryos. Anterior is to the top. Color code as in Movie 1.

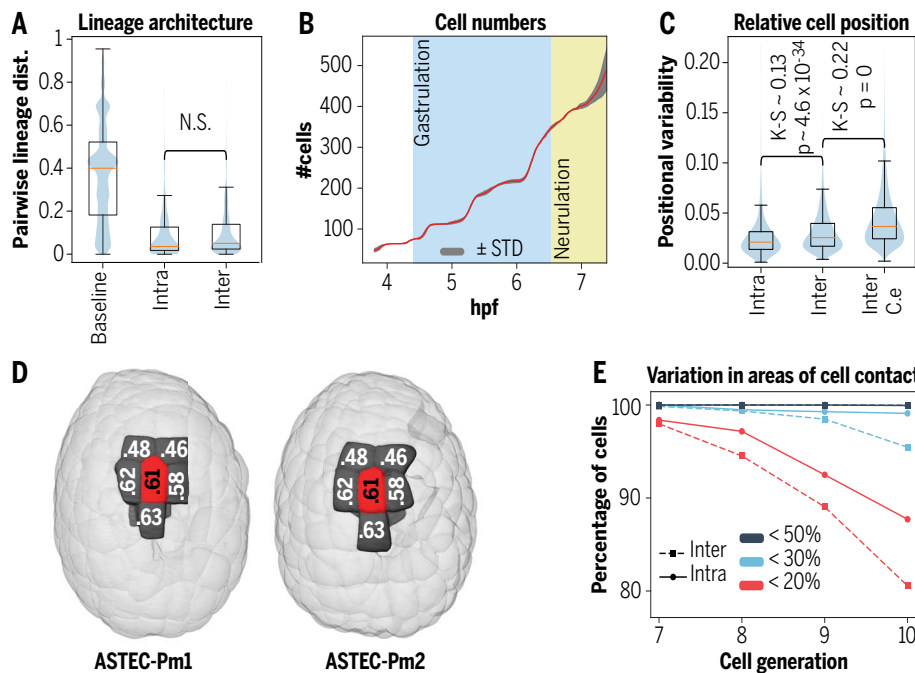


Fig. 2. Cellular reproducibility of ascidian development. (A) Distributions of pairwise distances between *Phallusia* cell lineage trees originated from individual cells at the 64-cell stage. Baseline: all trees; Intra: trees originating from matching bilateral left and right cells within each embryo; Inter: trees originating from homologous cells between embryos. In (A) and (C), boxes show the first, second, and third quartiles; whiskers show the range to 1.5 interquartile. N.S., not significant according to the Kolmogorov-Smirnov (K-S) test. (B) Comparison of the cell numbers across time between 10 individual *Phallusia* embryos after linear temporal rescaling (see supplementary materials). hpf, hours post fertilization at a temperature of 18°C. (C) Violin plots showing the distributions of the differences in relative cell position between matching bilateral cells in a *Phallusia* embryo (Intra), homologous *Phallusia* (Inter), or *C. elegans* (Inter C.e.) cells across embryos. (D) Conserved neighborhood and spatial position of the a9.61 cell in the ASTEC-Pm1 (tp73) and ASTEC-Pm2 (tp53) embryos. Ventral view. (E) Evolution in time (for cellular generations 7 to 10) of the percentage of cells showing a variability in areas of contact to shared neighbors smaller than the indicated thresholds between left and right cells of the same embryo (solid lines) or between homologous cells of distinct embryos (dashed lines).

correlation between cell volume and cell cycle duration observed throughout ascidian embryogenesis (fig. S30D), as in *C. elegans* (26).

Inequality and asymmetry of divisions were often associated in *Phallusia*. Seventy-three percent of asymmetric cell divisions gave rise to sister cells that showed inequality in their volumes or cell cycle durations (Fig. 3C and tables S7 to S10). Conversely, 78% of geometrically unequal divisions and 92% of divisions producing sisters with asynchronous cell cycles were asymmetric (Fig. 3C and tables S6 to S10). The mid-gastrula neural plate exemplifies a near-perfect association between asymmetric divisions, geometrically unequal divisions, and asynchronously dividing sister cells (Fig. 3, D to F).

Asymmetric divisions were also preferentially associated with oriented cell divisions. By default, cells divide at 90° from the division orientation of their mother (27), such that a doublet of sister cells divides to form a square or tetrahedral clone of four cells (fig. S31A). We could score the orientation of 14 known symmetric, 11 known asymmetric, and 6 candidate asymmetric cell divisions. Whereas all known symmetric divisions followed the default orientation pathway, eight known and

five candidate asymmetric divisions deviated from it (fig. S31, B and C).

Asymmetric cell divisions in *Phallusia* are thus associated with an active control of the volumes or relative positions of sister cells and thereby of the contact areas they establish with their neighbors. This geometrical control may be important for the differential fate specification of sister cells.

A model of geometric control of cell inductions

In ascidian embryos, most asymmetric divisions involve cell communication. Two non-exclusive mechanisms have been described. First, sister cells can be differentially induced after the division as a result of their exposure to distinct signals. Second, the mother cell can be polarized by extracellular signals before its division, causing the inheritance of activated signaling pathway components by only one of its daughters (11). Neural induction at the 32-cell stage involves sister differential cell induction and is controlled by the area of contact established with inducing cells (12), providing a proof of principle of a functional connection between cell arrangements and fate specifica-

tion. To test the generality of this mechanism, we augmented our geometric description of embryogenesis with a representation of the spatiotemporal distribution of signaling pathway components (Fig. 4, A and B) and developed a theoretical framework to analyze the impact of the range of signaling proteins on inductive processes.

Given the conservation of the expression of signaling genes between *Phallusia* and *Ciona* (28, 29), we inferred the expression of *Phallusia* signaling genes with cellular resolution from extensive *Ciona* signaling gene expression atlases (30). We focused our attention on six developmental signaling pathways (FGF, ephrin, Wnt, Bmp, Nodal, and Notch) showing patterned expression of extracellular ligands or antagonists up to the early gastrula stages (fig. S32). These pathways were sufficient to explain all known inductions up to this stage (11).

We next introduced a series of simplifying assumptions relative to both the competence of cells and the availability of inducing extracellular ligands. Receptors and intracellular components of the six candidate pathways were maternally expressed and their mRNAs uniformly distributed throughout the egg cytoplasm (30). We therefore estimated that, up to the gastrula stages, all cells were equally competent to respond to all pathways and that the receptors were uniformly distributed on cell surfaces. Extracellular ligand availability was therefore the main driver of pathway activation. When ligands and antagonists were colocalized at a cell-cell interface, secreted antagonists were assumed to fully inactivate their cognate ligands (Fig. 4D and fig. S33). Finally, we tested two scenarios for the range of secreted ligands and antagonists (Fig. 4, C and D, blue and red, respectively). This range was either limited to first-order neighbors contacting emitting cells (Fig. 4C, left) or extended to second-order neighbors (Fig. 4C, right). In both cases, potential concentration gradients were neglected, free ligands being either absent or present at a uniform concentration at the interface between two cells.

Interactions between ligands, receptors, and intracellular effectors were then modeled using the law of mass action (see supplementary materials). The cell geometry was introduced into the equations by proportionally linking the total number of receptors that could be activated by a ligand with the area of the cell exposed to free ligand molecules (Fig. 4D). For each pathway, we modeled differential sister inductions by computing the fraction of activated intracellular effectors as a function of the surface of each sister cell exposed to the inducer (Fig. 4E). In the case of polarization of the mother cell, activated effectors in each sister cell were computed from the portion of the exposed mother's surface inherited by the cell under consideration. The fraction of activated

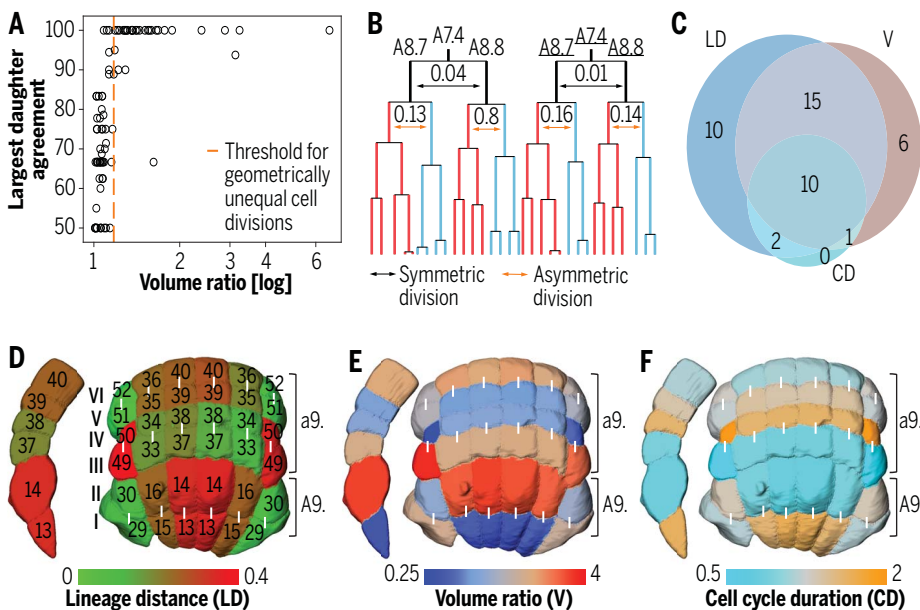


Fig. 3. Relationships between unequal cleavage, unequal cell cycle duration, and asymmetric divisions. (A) Unequal cell divisions between the 32-cell and neurula stages were identified as divisions in which the largest daughter is the same across >90% of all instances of this division in the dataset. (B) Example of symmetric and asymmetric divisions in the A7.4 neural progenitors lineages in ASTEC-Pm1 (by convention, right cells are underlined, left ones are not). The numbers indicate the tree-edit distance between the lineage trees seeded from the daughters of A8.7 and A8.8 cells. The divisions of A7.4 (right and left) are symmetric (lineage distance < 0.051), while those of A8.7 and A8.8 on both sides of the embryo are asymmetric (lineage distance > 0.051). Note the similarity of the lineages originating from the left and right A7.4 progenitors, slight division heterochronies accounting for differences in the final cell number. (C) Venn diagram showing the overlap of cell divisions leading to asymmetric divisions (LD), unequal sister cell cycle duration (CD), and geometrically unequal divisions (V). (D to F) Dorsal and lateral view of the a- and A-line derived neural plate, color-coded for lineage tree distance (D), volume ratios (E), and cell cycle duration ratios (F) between sister cells, linked by a white bar. The cell names are indicated in (D). Lineage tree distances are features of a cell pair, and ratios are features of individual cells.

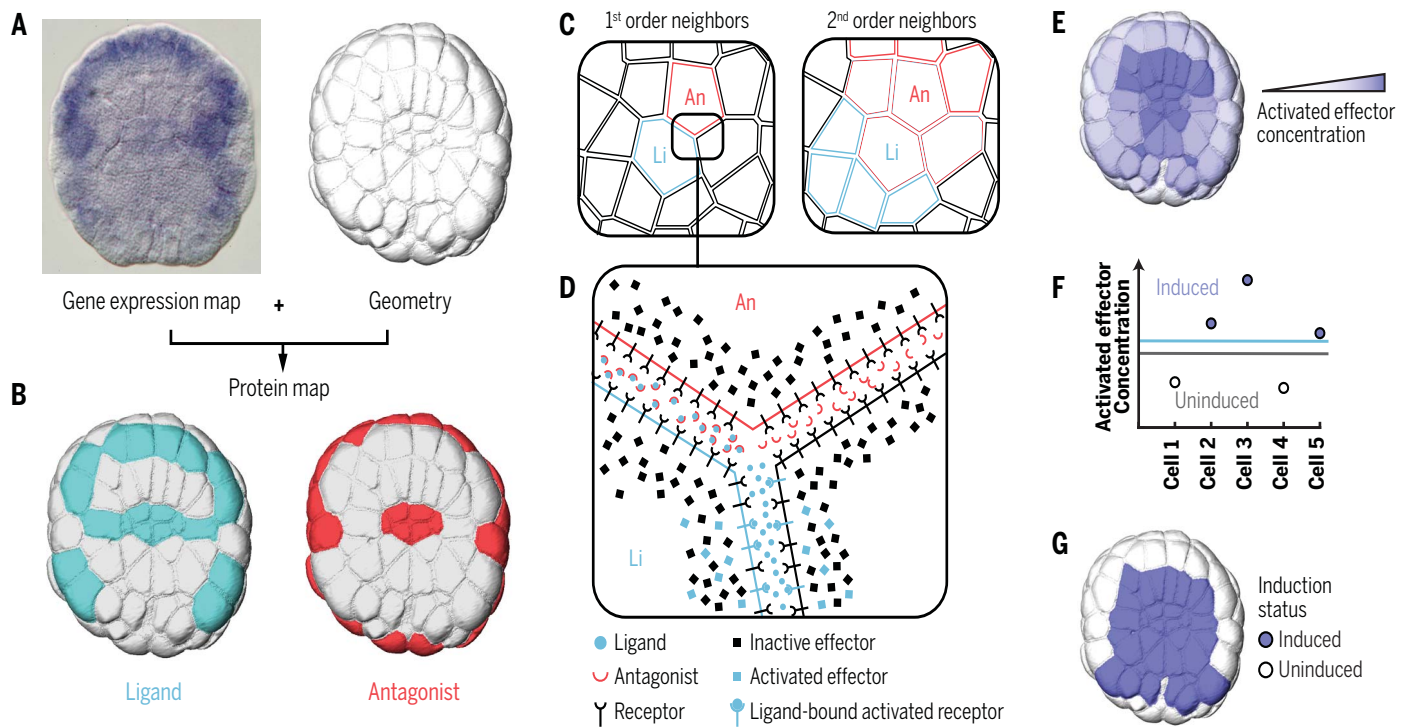


Fig. 4. Differential sister cell induction model. (A and B) Integration of gene expression information by in situ hybridization from the ANISEED database [(A), left] and geometrical information from ASTEC segmentation [(A), right] generated a map of cells expressing extracellular ligands and antagonists for the pathways of interest (B). (C) Schematic representation of the distribution of extracellular ligands (blue) and antagonists (red) in the vicinity of ligand-expressing (Li) or antagonist-expressing (An) cells, according to the diffusion range of proteins. (D) Illustration of the localization of ligands, antagonists,

free or ligand-bound receptors, and active or inactive effectors at the interface of three cells expressing a ligand Li (blue cell) or its antagonist An (red cell). (E) Activated effector concentration map computed from the mass action law. (F) Theoretical schema illustrating the identification of biologically motivated induction thresholds (horizontal lines) for activated effector concentrations (see supplementary materials). (G) Induced cell map obtained by combining activated effector concentrations shown in (E) and activated effector concentration thresholds computed according to the principle shown in (F).

intracellular effectors is well approximated by a linearly increasing function of the area of exposure to free ligands when the areas of cell contact and the biochemical parameters remain within the range found in the literature (fig. S34).

Finally, to convert continuous levels of effector activation into a binary map of induced and uninduced cells (Fig. 4, F and G, and figs. S35 and S36), we implemented induction thresholds (Fig. 4F) governed by tunable free parameters that defined both the minimum effector activity levels for an induction and a minimum ratio of activity between induced and uninduced sister cells. This simplified differential induction model thus provides both a semi-quantitative estimate of the relative level of signaling for each pathway in each cell and a qualitative prediction of differential inductions.

Contact area-dependent inductions drive ascidian fate specification

We next explored the parameter space to adjust induction thresholds and find the best fit of the model to a ground truth of sister cells known to be differentially induced (14) or not (49). When the range of secreted protein was limited to first-order neighbors, the model, run on ASTEC-Pm1, successfully recapitulated all 14 known differential induction events with their correct inducers, with only 11 false pre-

dictions (Fig. 5A; figs. S37, and S40 to S57; and table S13). This fit was reached with a relatively large set of possible combinations of parameters values (fig. S58), which were consistent with the scientific literature [e.g., time necessary for an induction was consistent with the one found for FGF in (37)] (table S12). In agreement with the high rate of molecular evolution in ascidians, the affinity and concentration of ligands and receptors appeared least constrained (fig. S58). Training the model on two other embryos revealed that the predictions were robust to natural variability (figs. S38 and S39) and were obtained with similar optimal parameter sets (tables S12 and S13). The quality of the predictions, however, strongly decreased when the model was trained either without taking into account cell contact areas (Fig. 5A) or with surfaces approximated from a Voronoi tessellation of the cell barycenters (table S13). Retraining the model with a range of action of secreted ligands and inhibitors increased to second-order neighbors degraded its performance (Fig. 5A) and strongly decreased its robustness to parameter variations (fig. S58), which is consistent with the notion that secreted proteins act by direct contact to emitting cells.

The model run with short-range inducers also correctly predicted events it was not trained on. For instance, the model trained to predict

differential sister induction also predicted the simultaneous induction by FGF of both daughters of the A6.1 and B6.1 endoderm precursors (II) (figs. S40 and S41). Running the model with optimal parameters for the wild-type (WT) situation but after virtual inactivation of ephrin receptor (Eph) signaling phenocopied the morphant Eph phenotype (II) and affected all FGF-mediated differential inductions except that of the daughters of the B6.4 mesenchyme and muscle precursor (32) (fig. S59).

Finally, we compared semiquantitative predictions of pathway activation with in vivo measurements in live embryos. We focused on the Eph-FGF/extracellular signal-regulated kinase (ERK) pathway, the main inducing pathway in early ascidian embryos (table S11) and imaged a live ERK activity biosensor, ERK-kinase translocation reporter (KTR) (33) (fig. S60). Experimentally measured values of ERK activity were highly reproducible between embryos (figs. S61 and S62), depended on MEK kinase activity (fig. S63), and were highly correlated to in silico predicted FGF/ERK pathway activation levels (Spearman correlation: 0.9) (Fig. 5B). Our simplified contact area-dependent induction model thus provides both reliable qualitative predictions of inductions and quantitative estimates of pathway activation levels in the different cells of WT embryos.

Robustness of ascidian fate specification to in vivo and in silico perturbations

To further analyze the sensitivity of fate specification to changes in surface of exposure to free ligand, we experimentally modified cell arrangements. Figure 5, B to D, and fig. S64 illustrate the case of a half embryo obtained by bisecting an embryo at the two-cell stage along its plane of cleavage. As expected (12), ectopic neural induction of epidermal precursors was observed at the 32-cell stage. The area of contact established between FGF-emitting vegetal cells and the a6.6 head epidermis precursor was strongly increased, as was the re-

sulting ERK signaling intensity in this cell, leading to neural induction (fig. S64).

Embryo halving also affected FGF-dependent mesodermal fate decisions at the 64-cell stage (Fig. 5, B to D). In vivo ERK signaling activity was modified in several cells, with the model correctly predicting the direction of the changes (Fig. 5B). Surface changes observed in the half embryo could even alter cell fate. During normal embryogenesis, the A6.3 mother cell divides asymmetrically to produce the A7.5 endodermal progenitor and the A7.6 trunk lateral cell progenitor. A7.5 is induced by FGF signaling, whereas the induction of A7.6 is antagonized by

its contact with ephrin-expressing b-line animal cells (Fig. 5C, left: contact between the green A7.6 and the dark blue b7.9 cell) (11). In the half embryo, the repositioning of the B-line vegetal blastomeres sheltered A7.6 from ephrin signaling (Fig. 5C, right: contact between the green A7.6 and the ephrin-negative light blue B7.5 and B7.7 cells). As a result, A7.6 received a level of ERK signaling comparable to the induced A7.1, A7.2, A7.5, and B7.1 endodermal precursors in the half embryo and accordingly seeded an endoderm-like lineage tree (Fig. 5D and fig. S65).

Finally, we explored in silico the effect of systematic geometrical perturbations that cannot easily be generated experimentally. We ran the model with an optimal set of parameters for the WT situation but applied $\pm 30\%$ variation in the surface of contact to FGF-expressing cells during the 64-cell stage, without change in the contacts to ephrin-expressing cells or in the total surface of each cell. These conditions altered the predicted induction status of 8 of 32 bilateral cell pairs (Fig. 6A). Extending this approach to all pathways, 22 of 32 cell pairs changed induction status for at least one pathway upon a $\pm 30\%$ variation in the surface of contacts they established with signaling cells (Fig. 6A). Similar conclusions were drawn at the 112-cell stage (fig. S66). A high level of precision and reproducibility of cell-cell contacts is thus crucial for early ascidian fate specification.

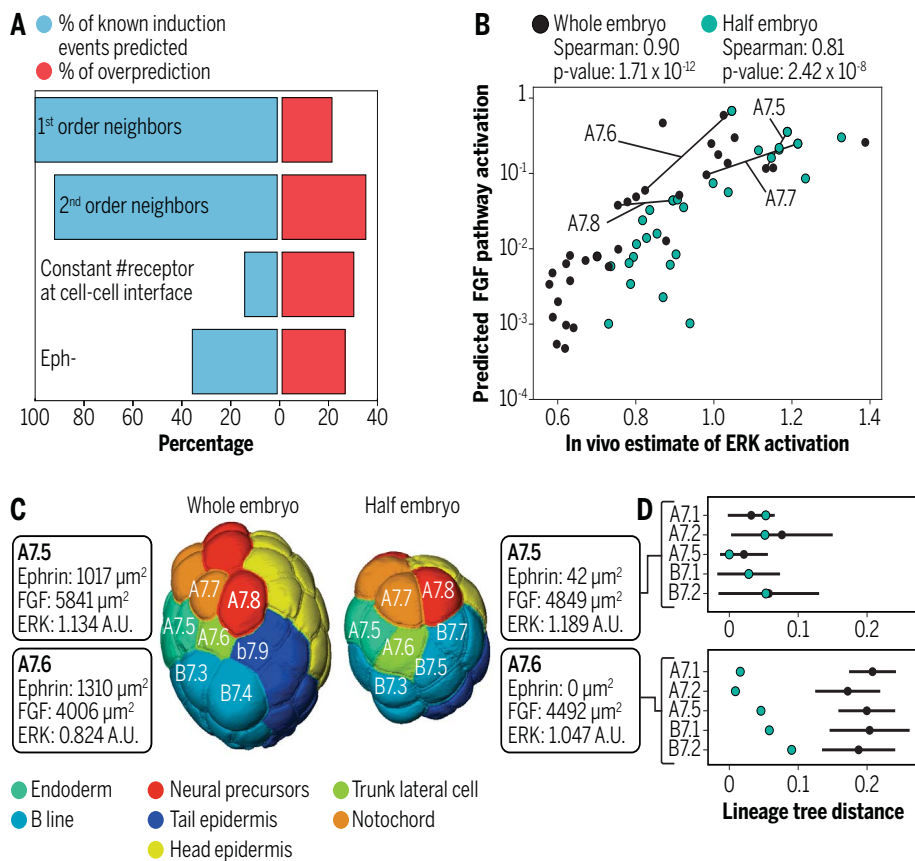


Fig. 5. Areas of contact mediate cell-cell communication. (A) Percentage of known (blue) or overpredicted (red) differential cell inductions predicted by the model on ASTEC-Pm1 when the range of secreted proteins is limited to first- or second-order neighbors, when the number of receptors exposed to ligands is uncoupled from the areas of contact, or when the Eph receptor is virtually inactivated. (B) Comparison of predicted and in vivo estimated FGF/ERK pathway activation at the 64-cell stage for a whole embryo (ASTEC-Pm13, black dots) and a half embryo (ASTEC-half-Pm1, blue-green dots). Predicted FGF pathway activity (y axis, log value, dimensionless) was obtained by running the model without diffusion. In vivo ERK activation (x axis, [fluorescence intensity] $^{-1}$) was estimated by imaging ERK-KTR-Clover-expressing embryos and determining the normalized nuclear amount of reporter, whose inverse provides an estimate of the log value of ERK activation (see supplementary materials section 10.1.3). The Spearman's correlation coefficient (S) and P value that this correlation is obtained by chance are indicated for both datasets. (C) Lateral views of mid-64-cell stage whole and half embryos with indication of cell fates, cell names, cell neighborhood, and areas of contact to FGF- and ephrin-expressing cells and ERK signaling activity (inverse normalized nuclear KTR values) of the A7.5 (blue-green) and A7.6 (green) cells. A.U., arbitrary units. (D) (Top) Distances between the lineage trees seeded by A7.5 and the indicated endodermal precursors in ASTEC-half-Pm1 (blue-green dots). Black dots indicate the average lineage distance, in whole embryos (ASTEC-Pm1 to -Pm8), of A7.5 to the indicated cells. The black line represents one standard deviation. (Bottom) Same analysis, but for the trunk lateral cell precursor A7.6 [green in (C)]. Embryo halving does not substantially affect the lineage distance of A7.5 to endoderm but shifts the lineage seeded by A7.6 to an endoderm-like one.

Concluding remarks

In this work, we systematically quantified and analyzed the behavior and fate decisions of individual embryonic cells across gastrulation and neurulation in *Phallusia*. We then enriched this geometric description with measurements of signaling activities and gene expression activity. Combined with high throughput single-cell genomics data (34), emerging computational methods to map such data onto reference geometrical descriptions (35), and the determination of mechanical properties by inference from the geometry or by direct force measurements (36, 37), our work establishes the ascidian embryo as a system to bridge cell behaviors, the forces that drive morphogenesis, and the underlying regulatory program.

Descriptions of cell inductions during animal embryogenesis often highlight the role of long-range morphogen gradients (38). In this scenario, the outcome of a cell communication event is determined by local morphogen concentrations, whereas variations in individual cell contacts are considered to play a minor part (Fig. 6B, top). Here, we show that areas of contact between communicating cells have a sufficient encoding potential to explain most fate specification events in a chordate embryo (Fig. 6B, bottom), extending previous work in ascidians and vertebrates (12, 39). Our work thus reveals that quantitative cues can shift from ligand concentration

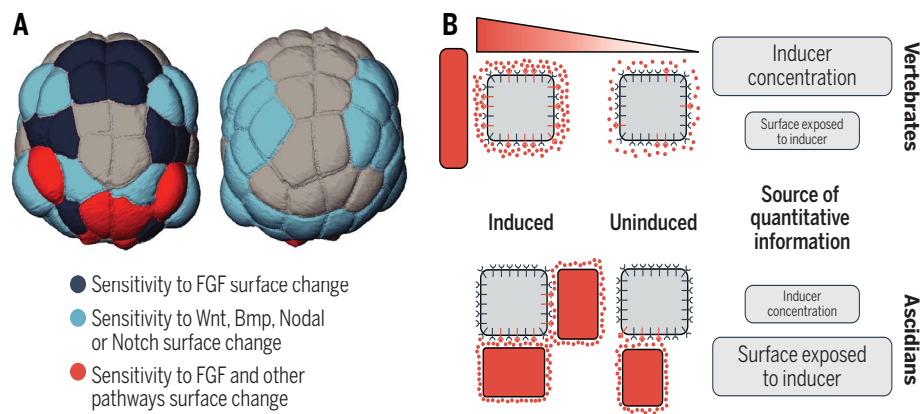


Fig. 6. Contact area-dependent cell inductions constrain embryonic geometries. (A) Projection onto a 64-cell-stage embryo of cells whose predicted induced state changed when exposed to a $\pm 30\%$ in silico change in their area of contact to cells expressing FGF (dark blue), ligands or inhibitors of other pathways (light blue), or both (red). See also fig. S66 for corresponding data at the 112-cell stage. (B) Comparison of two modes of cell induction: diffusible morphogen gradients, as proposed for vertebrates, and contact area-dependent cell inductions, as proposed for ascidians. In the former, secreted inducer concentration is the major determinant of induction; in the latter, the area exposed to the ligand plays a major role.

to areas of cell contacts. Consistent with an equivalence of processes, variations in cell contact areas or in ligand concentrations have comparable effects on ascidian cell inductions (fig. S67).

The prevalent use of contact area-dependent inductions may also causally explain the counterintuitive association between the cellular reproducibility of the ascidian embryonic program and a high level of genetic divergence. Precisely and robustly shaping a morphogen gradient involves the coordinated recruitment of multiple cellular functions to control the production, degradation, transport, or endocytosis of ligands and receptors (38). Because of the shallowness and noisiness of such gradients, additional layers of regulatory cross-talk between target transcription factors (38) or subsequent cell sorting (40) are needed to transform the coarse information provided by the gradient into sharp boundaries. The precise response to long-range morphogen gradients thus involves sophisticated layers of regulation, which may constrain the architecture of regulatory networks and global genome evolution. By contrast, the surface-of-contact-dependent signaling we described here ensures that sister cells are exposed to sufficiently different signals to directly define boundaries, thereby limiting the need for subsequent transcriptional refinements. Contact area-dependent inductions are therefore likely to involve fewer layers of regulation than long-range morphogen gradients, thereby relaxing constraints on genome evolution.

Methods summary

All methods used in this study are recapitulated in detail in the supplementary materials, which include the following 14 sections: (i) the imaging protocol, (ii) the preprocessing of the movies of *P. mammillata* development, (iii)

the limitation of the state-of-the-art methods for cell segmentation and tracking, (iv and v) our segmentation and tracking algorithm (ASTE) and its postprocessing step, (vi) the quantification of the quality of the segmentations and trackings, (vii and viii) the analysis of developmental stereotypy and fate restriction events, (ix) the mathematical model of cell-cell signaling, (x) the validation of the model through the in vivo measurement of ERK signaling, and (xi) the validation of the model through the analysis of the effect of experimental perturbations of surfaces of contacts. Sections xii to xiv describe the authors' contributions, list the tables, and indicate how to access software and datasets.

Note added in proof: A reference (106) was added at proof stage to a relevant new resource on single-cell transcriptional profiles covering *P. mammillata* early embryogenesis.

REFERENCES AND NOTES

- A. Pierre, J. Sallé, M. Wühr, N. Minc, Generic theoretical models to predict division patterns of cleaving embryos. *Dev. Cell* **39**, 667–682 (2016). doi: [10.1016/j.devcel.2016.11.018](https://doi.org/10.1016/j.devcel.2016.11.018); pmid: [27997824](https://pubmed.ncbi.nlm.nih.gov/27997824/)
- J. E. Sulston, E. Schierenberg, J. G. White, J. N. Thomson, The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983). doi: [10.1016/0012-1606\(83\)90201-4](https://doi.org/10.1016/0012-1606(83)90201-4); pmid: [6684600](https://pubmed.ncbi.nlm.nih.gov/6684600/)
- P. Lemaire, Evolutionary crossroads in developmental biology: The tunicates. *Development* **138**, 2143–2152 (2011). doi: [10.1242/dev.048975](https://doi.org/10.1242/dev.048975); pmid: [21558365](https://pubmed.ncbi.nlm.nih.gov/21558365/)
- F. Delsuc, H. Brinkmann, D. Chourrout, H. Philippe, Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **439**, 965–968 (2006). doi: [10.1038/nature04336](https://doi.org/10.1038/nature04336); pmid: [16495997](https://pubmed.ncbi.nlm.nih.gov/16495997/)
- B. Goldstein, On the evolution of early development in the Nematoda. *Philos. Trans. R. Soc. London Ser. B* **356**, 1521–1531 (2001). doi: [10.1098/rstb.2001.0977](https://doi.org/10.1098/rstb.2001.0977); pmid: [11604120](https://pubmed.ncbi.nlm.nih.gov/11604120/)
- F. Delsuc et al., A phylogenomic framework and timescale for comparative studies of tunicates. *BMC Biol.* **16**, 39 (2018). doi: [10.1186/s12915-018-0499-2](https://doi.org/10.1186/s12915-018-0499-2); pmid: [29653534](https://pubmed.ncbi.nlm.nih.gov/29653534/)
- A. Stolfi et al., Divergent mechanisms regulate conserved cardiopharyngeal development and gene expression in

- distantly related ascidians. *eLife* **3**, e03728 (2014). doi: [10.7554/eLife.03728](https://doi.org/10.7554/eLife.03728); pmid: [25209999](https://pubmed.ncbi.nlm.nih.gov/25209999/)
- H. Nishida, Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue restricted stage. *Dev. Biol.* **121**, 526–541 (1987). doi: [10.1016/0012-1606\(87\)90188-6](https://doi.org/10.1016/0012-1606(87)90188-6); pmid: [3582738](https://pubmed.ncbi.nlm.nih.gov/3582738/)
- D. A. Voronov, Y. V. Panchin, Cell lineage in marine nematode *Enoplus brevis*. *Development* **125**, 143–150 (1998). pmid: [9389672](https://pubmed.ncbi.nlm.nih.gov/9389672/)
- H. Nishida, Specification of embryonic axis and mosaic development in ascidians. *Dev. Dyn.* **233**, 1177–1193 (2005). doi: [10.1002/dvdy.20469](https://doi.org/10.1002/dvdy.20469); pmid: [15973692](https://pubmed.ncbi.nlm.nih.gov/15973692/)
- P. Lemaire, Unfolding a chordate developmental program, one cell at a time: Invariant cell lineages, short-range inductions and evolutionary plasticity in ascidians. *Dev. Biol.* **332**, 48–60 (2009). doi: [10.1016/j.ydbio.2009.05.540](https://doi.org/10.1016/j.ydbio.2009.05.540); pmid: [19433085](https://pubmed.ncbi.nlm.nih.gov/19433085/)
- O. Tassy, F. Daian, C. Hudson, V. Bertrand, P. Lemaire, A quantitative approach to the study of cell shapes and interactions during early chordate embryogenesis. *Curr. Biol.* **16**, 345–358 (2006). doi: [10.1016/j.cub.2005.12.044](https://doi.org/10.1016/j.cub.2005.12.044); pmid: [16488868](https://pubmed.ncbi.nlm.nih.gov/16488868/)
- G. de Medeiros et al., Confocal multiview light-sheet microscopy. *Nat. Commun.* **6**, 8881 (2015). doi: [10.1038/ncomms9881](https://doi.org/10.1038/ncomms9881); pmid: [26602977](https://pubmed.ncbi.nlm.nih.gov/26602977/)
- V. Ulman et al., An objective comparison of cell-tracking algorithms. *Nat. Methods* **14**, 1141–1152 (2017). doi: [10.1038/nmeth.4473](https://doi.org/10.1038/nmeth.4473); pmid: [29083403](https://pubmed.ncbi.nlm.nih.gov/29083403/)
- E. Faure et al., A workflow to process 3D+time microscopy images of developing organisms and reconstruct their cell lineage. *Nat. Commun.* **7**, 8674 (2016). doi: [10.1038/ncomms9674](https://doi.org/10.1038/ncomms9674); pmid: [26912388](https://pubmed.ncbi.nlm.nih.gov/26912388/)
- J. Stegmaier et al., Real-time three-dimensional cell segmentation in large-scale microscopy data of developing embryos. *Dev. Cell* **36**, 225–240 (2016). doi: [10.1016/j.devcel.2015.12.028](https://doi.org/10.1016/j.devcel.2015.12.028); pmid: [26812020](https://pubmed.ncbi.nlm.nih.gov/26812020/)
- Y. Azuma, S. Onami, Biologically constrained optimization based cell membrane segmentation in *C. elegans* embryos. *BMC Bioinformatics* **18**, 307 (2017). doi: [10.1186/s12859-017-1717-6](https://doi.org/10.1186/s12859-017-1717-6); pmid: [28629355](https://pubmed.ncbi.nlm.nih.gov/28629355/)
- P. Villoutreix et al., An integrated modelling framework from cells to organism based on a cohort of digital embryos. *Sci. Rep.* **6**, 37438 (2016). doi: [10.1038/srep37438](https://doi.org/10.1038/srep37438); pmid: [27910875](https://pubmed.ncbi.nlm.nih.gov/27910875/)
- M. A. Luengo-Oroz et al., 3D+t morphological processing: Applications to embryogenesis image analysis. *IEEE Trans. Image Process.* **21**, 3518–3530 (2012). doi: [10.1109/TIP.2012.2197007](https://doi.org/10.1109/TIP.2012.2197007); pmid: [22562755](https://pubmed.ncbi.nlm.nih.gov/22562755/)
- R. Fernandez et al., Imaging plant growth in 4D: Robust tissue reconstruction and lineage at cell resolution. *Nat. Methods* **7**, 547–553 (2010). doi: [10.1038/nmeth.1472](https://doi.org/10.1038/nmeth.1472); pmid: [20543845](https://pubmed.ncbi.nlm.nih.gov/20543845/)
- B. Leggio et al., MorphoNet: An interactive online morphological browser to explore complex multi-scale data. *Nat. Commun.* **10**, 2812 (2019). doi: [10.1038/s41467-019-10668-1](https://doi.org/10.1038/s41467-019-10668-1); pmid: [31249294](https://pubmed.ncbi.nlm.nih.gov/31249294/)
- W. R. Jeffery, Programmed cell death in the ascidian embryo: Modulation by FoxA5 and Manx and roles in the evolution of larval development. *Mech. Dev.* **118**, 111–124 (2002). doi: [10.1016/S0925-4773\(02\)00236-8](https://doi.org/10.1016/S0925-4773(02)00236-8); pmid: [12351175](https://pubmed.ncbi.nlm.nih.gov/12351175/)
- X. Li et al., Systems properties and spatiotemporal regulation of cell position variability during embryogenesis. *Cell Rep.* **26**, 313–321.e7 (2019). doi: [10.1016/j.celrep.2018.12.052](https://doi.org/10.1016/j.celrep.2018.12.052); pmid: [30625313](https://pubmed.ncbi.nlm.nih.gov/30625313/)
- E. M. Leffler et al., Revisiting an old riddle: What determines genetic diversity levels within species? *PLoS Biol.* **10**, e1001388 (2012). doi: [10.1371/journal.pbio.1001388](https://doi.org/10.1371/journal.pbio.1001388); pmid: [22984349](https://pubmed.ncbi.nlm.nih.gov/22984349/)
- X. Morin, Y. Bellaïche, Mitotic spindle orientation in asymmetric and symmetric cell divisions during animal development. *Dev. Cell* **21**, 102–119 (2011). doi: [10.1016/j.devcel.2011.06.012](https://doi.org/10.1016/j.devcel.2011.06.012); pmid: [21763612](https://pubmed.ncbi.nlm.nih.gov/21763612/)
- Y. Arata, H. Takagi, Y. Sako, H. Sawa, Power law relationship between cell cycle duration and cell volume in the early embryonic development of *Caenorhabditis elegans*. *Front. Physiol.* **5**, 529 (2015). doi: [10.3389/fphys.2014.00529](https://doi.org/10.3389/fphys.2014.00529); pmid: [25674063](https://pubmed.ncbi.nlm.nih.gov/25674063/)
- S. Strome, Determination of cleavage planes. *Cell* **72**, 3–6 (1993). doi: [10.1016/0092-8674\(93\)90041-N](https://doi.org/10.1016/0092-8674(93)90041-N); pmid: [8422680](https://pubmed.ncbi.nlm.nih.gov/8422680/)
- A. Madgwick et al., Evolution of embryonic cis-regulatory landscapes between divergent *Phallusia* and *Ciona* ascidians. *Dev. Biol.* **448**, 71–87 (2019). doi: [10.1016/j.ydbio.2019.01.003](https://doi.org/10.1016/j.ydbio.2019.01.003); pmid: [30661644](https://pubmed.ncbi.nlm.nih.gov/30661644/)
- U.-M. Fiuza, T. Negishi, A. Rouan, H. Yasuo, P. Lemaire, A Nodal/Eph signalling relay drives the transition from apical constriction to apico-basal shortening in ascidian endoderm

- invagination. *bioRxiv* 418988 [Preprint]. 1 December 2019. <https://doi.org/10.1101/418988>.
30. J. Dardailon *et al.*, ANISEED 2019: 4D exploration of genetic data for an extended range of tunicates. *Nucleic Acids Res.* **48**, D668–D675 (2019). doi: [10.1093/nar/gkz955](https://doi.org/10.1093/nar/gkz955); pmid: [31680137](https://pubmed.ncbi.nlm.nih.gov/31680137/)
 31. G. J. Kim, A. Yamada, H. Nishida, An FGF signal from endoderm and localized factors in the posterior-vegetal egg cytoplasm pattern the mesodermal tissues in the ascidian embryo. *Development* **127**, 2853–2862 (2000). pmid: [10851130](https://pubmed.ncbi.nlm.nih.gov/10851130/)
 32. G. J. Kim, G. Kumano, H. Nishida, Cell fate polarization in ascidian mesenchyme/muscle precursors by directed FGF signaling and role for an additional ectodermal FGF antagonizing signal in notochord/nerve cord precursors. *Development* **134**, 1509–1518 (2007). doi: [10.1242/dev.02825](https://doi.org/10.1242/dev.02825); pmid: [17360771](https://pubmed.ncbi.nlm.nih.gov/17360771/)
 33. C. de la Cova, R. Townley, S. Regot, I. Greenwald, A real-time biosensor for ERK activity reveals signaling dynamics during *C. elegans* cell fate specification. *Dev. Cell* **42**, 542–553.e4 (2017). doi: [10.1016/j.devcel.2017.07.014](https://doi.org/10.1016/j.devcel.2017.07.014); pmid: [28826819](https://pubmed.ncbi.nlm.nih.gov/28826819/)
 34. C. Cao *et al.*, Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* **571**, 349–354 (2019). doi: [10.1038/s41586-019-1385-y](https://doi.org/10.1038/s41586-019-1385-y); pmid: [31292549](https://pubmed.ncbi.nlm.nih.gov/31292549/)
 35. M. Nitzan, N. Karaiskos, N. Friedman, N. Rajewsky, Gene expression cartography. *Nature* **576**, 132–137 (2019). doi: [10.1038/s41586-019-1773-3](https://doi.org/10.1038/s41586-019-1773-3); pmid: [31748748](https://pubmed.ncbi.nlm.nih.gov/31748748/)
 36. K. Sugimura, P.-F. Lenne, F. Graner, Measuring forces and stresses in situ in living tissues. *Development* **143**, 186–196 (2016). doi: [10.1242/dev.119776](https://doi.org/10.1242/dev.119776); pmid: [26786209](https://pubmed.ncbi.nlm.nih.gov/26786209/)
 37. K. Sherrard, F. Robin, P. Lemaire, E. Munro, Sequential activation of apical and basolateral contractility drives ascidian endoderm invagination. *Curr. Biol.* **20**, 1499–1510 (2010). doi: [10.1016/j.cub.2010.06.075](https://doi.org/10.1016/j.cub.2010.06.075); pmid: [20691592](https://pubmed.ncbi.nlm.nih.gov/20691592/)
 38. A. Sagner, J. Briscoe, Morphogen interpretation: Concentration, time, competence, and signaling dynamics. *WIREs Dev. Biol.* **6**, e271 (2017). doi: [10.1002/wdev.271](https://doi.org/10.1002/wdev.271); pmid: [28319331](https://pubmed.ncbi.nlm.nih.gov/28319331/)
 39. O. Shaya *et al.*, Cell-cell contact area affects notch signaling and notch-dependent patterning. *Dev. Cell* **40**, 505–511.e6 (2017). doi: [10.1016/j.devcel.2017.02.009](https://doi.org/10.1016/j.devcel.2017.02.009); pmid: [28292428](https://pubmed.ncbi.nlm.nih.gov/28292428/)
 40. F. Xiong *et al.*, Specified neural progenitors sort to form sharp domains after noisy Shh signaling. *Cell* **153**, 550–561 (2013). doi: [10.1016/j.cell.2013.03.023](https://doi.org/10.1016/j.cell.2013.03.023); pmid: [23622240](https://pubmed.ncbi.nlm.nih.gov/23622240/)

ACKNOWLEDGMENTS

We thank the members of the Lemaire, Godin, and Hufnagel teams for discussion and advice throughout this project; P. Kristofori (CRBM) for contributions to the analysis of KTR experiments; T. Rasse (EMBL) for running RACE on ascidian datasets; J. Stegmaier (RWTH Aachen University, Germany) for advice; A. McDougall (LBDV, Villefranche/mer, France) for the generous gift of the PH-GFP expression construct; and Y. Hitoyoshi (LBDV, Villefranche/mer, France) for the PH-Tdtomato expression construct. We thank T. Schneid (EMBL) for help with building the pGEM-PH-citrine construct and P. Neveu (EMBL) for discussions on cloning methods. We thank the Montpellier Image Resources (MRI) facility for hosting the MuViSPIM used in this study and S. de Rossi (MRI) for providing expert help with the MuViSPIM. We are grateful to P.-J. Keller and A. Pavlopoulos (Janelia Research Labs, USA), to F. Fagotto (CRBM, Montpellier, France), and to P. Ōnal (NYU) for critical reading of the manuscript. We also thank the Montpellier S12C2 service for IT support and P. Richard and M. Plays (CRBM) for careful animal husbandry. **Funding:** This work was funded by core support from CNRS to P.L.; by Inria (core support and IPL Morphogenetics) to C.G. and G.Ma.; by the Geneshape project (ANR-SYSC-018-02) to P.L. and C.G.; and by the Dig-Em project (ANR-14-CE11-0013-01) to P.L., C.G., and G.Ma. L.H. was supported by the European Molecular Biology Laboratory and the Center of Modeling and Simulation in the Biosciences (BIOMS) of the University of Heidelberg. L.G. was supported by a doctoral contract from the CBS2 doctoral school of the University of Montpellier 2, by the Fondation pour la Recherche Médicale (FRM) (FDT20140931061), and by the Morphoscope2 Equipex project (ANR-11-EQPX-0029). U.-M.F. was

supported by the Geneshape and Dig-Em projects, by the FRM (SPF20120523969), and by the EMBL Interdisciplinary Postdoc Programme under Marie Curie Actions. The Institut de Biologie Computationnelle de Montpellier (IBC; ANR-11-BINF-0002) supported E.F. and B.L. B.L. and J.L. were supported by the Dig-Em project. G.Mi. was supported by Inria (IPL Morphogenetics) and the Dig-Em project. K.B. was supported by a Ph.D. contract from the EpiGenMed laboratory of excellence (ANR-10-LABX-12-01). The Fondation Bettencourt-Schueller funded the Montpellier MuViSPIM (with a contribution from the University of Montpellier). **Author contributions:** Author contributions are listed in detail in the supplementary materials. **Data and materials availability:** See supplementary materials. The software, standard parameter files, and tutorials have been deposited to Github (<https://github.com/astec-segmentation>). All digitized embryos (fused intensity movies, segmentations, and meshes) have been deposited to figshare (https://figshare.com/projects/Phallusia_mammillata_embryonic_development/64301) and can be interactively explored through MorphoNet (www.morphonet.org/). All instructions and links are detailed and kept up to date in the ASTEC webpage of our institute (www.crblm.cnrs.fr/en/astec/).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6500/eaar5663/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S67
Tables S1 to S13
References (41–106)
MDAR Reproducibility Checklist
Data S1 to S18
Movies S1 to S4

[View/request a protocol for this paper from Bio-protocol.](#)

4 December 2019; accepted 29 April 2020
10.1126/science.aar5663

Contact area–dependent cell communication and the morphological invariance of ascidian embryogenesis

Léo Guignard, Ulla-Maj Fiúza, Bruno Leggio, Julien Laussu, Emmanuel Faure, Gaël Michelin, Kilian Biasuz, Lars Hufnagel, Grégoire Malandain, Christophe Godin and Patrick Lemaire

Science **369** (6500), eaar5663.
DOI: 10.1126/science.aar5663

Cell-cell contacts specify cell fate

Ascidians, or sea squirts, are marine invertebrate filter feeders with highly reproducible cellular events and invariant embryonic cell lineages. Guignard *et al.* studied the ascidian embryo to address the determinants of this cellular reproducibility. They introduce computational methods for the robust and automated segmentation, tracking, and analysis of whole-cell behaviors in high-throughput light-sheet microscopy datasets. This work shows that cell induction can be controlled by the contact area among cells. The range of cell signaling is proposed to set the scale at which animal embryonic reproducibility is observed. A high level of reproducibility of embryonic geometries may also counter-intuitively lift constraints on genome evolution, thereby contributing to the rapid molecular evolution observed in ascidians.

Science, this issue p. eaar5663

ARTICLE TOOLS

<http://science.sciencemag.org/content/369/6500/eaar5663>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2020/07/08/369.6500.eaar5663.DC1>

REFERENCES

This article cites 97 articles, 26 of which you can access for free
<http://science.sciencemag.org/content/369/6500/eaar5663#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

Extraits du supplément d'information : "Validation biologique des prédictions du modèle à partir de mesures de l'activité ERK *in vivo* pour chaque cellule au stade 64-cellules."

10 Biological validation of the model output using live single cells measurements of ERK activity at the 64-cell stage (Supplement to Figure panels 5B)

To qualitatively compare the amount of intra-cellular effector E^* computed by our model with single-cell live measurements of FGF/Ras/ERK signaling activities, we used the ERK-Kinase Translocation Reporter, a synthetic real-time biosensor, which translates ERK activity into a nucleo-cytoplasmic shuttling event (46). We made the assumption that the amount of Ras-GTP at the plasma membrane, computed by our model, is monotonously related to ERK activity, measured by the ERK-KTR sensor.

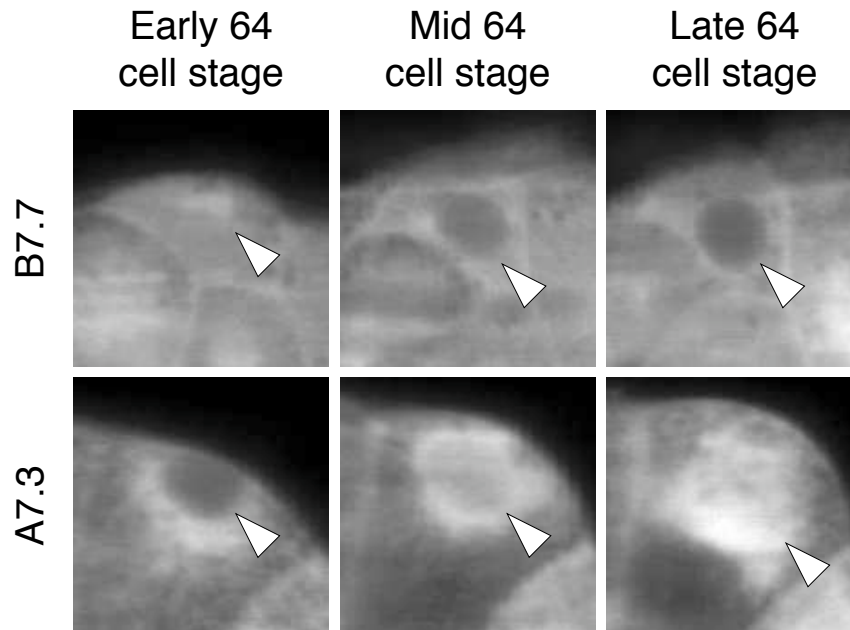


Figure S60: Examples of dynamic ERK-KTR nucleo-cytoplasmic shuttling in two example cells The figure shows the dynamic ERK-KTR nucleo-cytoplasmic shuttling in two example cells at the 64-cell stage in *ASTECPm13*. The ERK-KTR sensor is progressively exported out of the nucleus of the B7.7 mesenchyme precursor as the cell cycle progresses, indicative of increased ERK signalling. The A7.3 notochord precursor shows the opposite behavior, consistent with the notion that the notochord fate results from the polarization of the A6.3 mother cell (80).

10.1 KTR signal analysis

10.1.1 Processing of two-channel PH-tdTomato/ERK-KTR-Clover image datasets

At each time point, four two-channels (PH-tdTomato and ERK-KTR-Clover) acquisitions are acquired. Fusion is first performed on PH-tdTomato channel (see section 2.2): resulting computed transformations as well as image weights are used to fuse the ERK-KTR-Clover channel. This results in two-channel fusion images.

To extract surface of contact areas, the fused PH-tdTomato channel was segmented using ASTEC and manually named according to the Conklin nomenclature. The fused ERK-KTR channel was then used to infer ERK activity levels in each named cell.

10.1.2 Quantification strategy for the ERK-KTR signal

As the ERK-KTR reporter is exported out of the nucleus upon phosphorylation of the reporter by active ERK kinase, we used Fiji/ImageJ (99) to measure the mean ERK-KTR fluorescence intensity in a manually-placed disk enclosed in the equatorial plane of each nucleus (we verified that the precise position and size of the disk did not significantly affect the measurements). For each embryo, we measured the ERK-KTR in all nuclei at 5 time points, spaced by 4-6 minutes. These measures were made blindly without knowledge of the cell identities. The cell naming from the segmentation was then associated with the ERK-KTR values thus obtained, revealing highly similar values in matching bilaterally

symmetric cells.

To compare the values obtained in different embryos, which may not have received the exact same amount of synthetic ERK-KTR-Clover mRNA, we normalized each nuclear intensity by the mean ERK-KTR fluorescence intensity within the whole embryo at each developmental time analyzed. These mean intensities were automatically computed as the average of the intensity in the ERK-KTR channel located within the boundary of the embryo, given by the cell segmentation of the membrane channel. Attempts to normalize the signal by the total cytoplasmic intensity of each cell gave results inconsistent with known inductions, suggesting that only part of the highly heterogeneous cytoplasmic ERK-KTR can shuttle to and from the nucleus.

10.1.3 Time registration of ERK-KTR signals in different embryos

The precise time points at which the ERK-KTR measurements were performed may differ between embryos. To compare the dynamics and the magnitude of normalised ERK-KTR signals in different embryos, a non-linear time registration of ERK-KTR signal was performed. Having measured the nuclear KTR signal for all cells in each embryo at five distinct time points, the problem consisted in identifying the temporal relation between measurements in different embryos. To this end, we employed a multi-step iterative process. At first, assuming all N_e embryos were measured at exactly the same five equally spaced timepoints 1, 2, 3, 4, 5, we calculated the average KTR signal per cell c per timepoint t as

$$\eta^{(0)}(c, t) = \frac{1}{N_e} \sum_{i=1}^{N_e} \nu_i^{(0)}(c, t), \quad (10.1)$$

where $\nu_i^{(0)}(c, t)$ is the KTR value measured for cell c at time t in embryo i . Based on this, we defined the distance in KTR signal between embryo i and the average as

$$\mathcal{D}_i^{(0)} = \frac{1}{N_c} \sum_{c=1}^{N_c} \left(\frac{1}{N_t} \sum_{t=1}^{N_t} (\nu_i^{(0)}(c, t) - \eta^{(0)}(c, t))^2 \right) = \left\langle \left\langle (\nu_i^{(0)}(c, t) - \eta^{(0)}(c, t))^2 \right\rangle_t \right\rangle_c, \quad (10.2)$$

where N_t is the number of measured timepoints per cell and N_c the number of cells in the embryo, $\langle \cdot \rangle_t$ stands for the time average for each cell and $\langle \cdot \rangle_c$ stands for the average over all cells.

The five equally spaced timepoints $t_n = n, n = 1, \dots, 5$ will serve as a reference time template for inter-embryonic comparisons of KTR signal.

Lifting now the assumption that all embryos were measured at the exact same timepoints, the set of KTR signal measurements on embryo i becomes a function of five ordered parameters $\tau_{i,1} < \tau_{i,2} < \tau_{i,3} < \tau_{i,4} < \tau_{i,5}$, corresponding at the actual time instants at which embryo i has been measured. By assuming a quasi-linear time evolution of the KTR signal between any two consecutive time instants $\tau_{i,n}, \tau_{i,n+1}$, the estimated KTR values $\nu_i^{\text{est}}(c, t_n | \tau_{i,1}, \dots, \tau_{i,5})$ per embryo and per cell at the five reference timepoints $t_n = n, n = 1, \dots, 5$ is obtained by linear interpolation between the two time instants $\tau_{i,n}, \tau_{i,n+1}$ closest to t_n , and is thus a function of the five time parameters $\tau_n, n = 1, \dots, 5$.

This also means that both the average KTR signal, Eq. (10.1), and the KTR distance between each embryo and the average, Eq. (10.2), become functions of all τ s:

$$\eta(c, t | \{\tau\}) = \frac{1}{N_e} \sum_{i=1}^{N_e} \nu_i^{\text{est}}(c, t | \tau_{i,1}, \dots, \tau_{i,5}), \quad (10.3)$$

$$\mathcal{D}_i(\{\tau\}) = \frac{1}{N_c} \sum_{c=1}^{N_c} \left(\frac{1}{N_t} \sum_{t=1}^{N_t} (\nu_i^{\text{est}}(c, t | \tau_{i,1}, \dots, \tau_{i,5}) - \eta(c, t | \{\tau\}))^2 \right), \quad (10.4)$$

where $\{\tau\}$ is a short-hand notation for $\{\tau_{1,1}, \dots, \tau_{1,5}, \tau_{2,1}, \dots, \tau_{2,5}, \dots, \tau_{N_e,5}\}$.

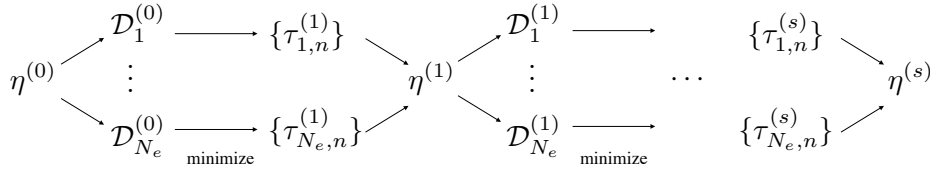


Figure S61: Schematic flow of the iterative process of time registration of KTR signal.

In order to estimate the τ parameters for each embryo, starting from the zero-th order values $\mathcal{D}_i^{(0)}$, we performed a non-linear numerical optimisation by iteratively minimizing the function $\mathcal{D}_i(\tau_1, \dots, \tau_5)$ in Eq. (10.4) to find the approximation of the τ s at each step s . These approximations, one per embryo, are then used to calculate the next estimation of the average $\eta^{(s)}$ at step s from Eq. (10.3). This procedure was performed until $\max |\eta^{(s)} - \eta^{(s-1)}| \leq 10^{-2}$, with the max performed over all cells and timepoints. This procedure is schematically represented in Fig. S61.

After convergence we obtained a reliable estimation of both the average KTR signal $\eta_{\text{conv}}(c, t_n)$ and the signal for each embryo $\nu_i^{\text{conv}}(c, t_n)$, $t_n = n, n = 1, \dots, 5$ per cell and per reference timepoint. This process brings all embryos onto the same time scaffold and allows for meaningful comparison of their KTR signal at corresponding timepoints.

Note that, mutant or experimentally manipulated embryos can be included in the pool of N_e embryos to which the analysis is applied, provided the perturbation does not affect the majority of cells: since Eq. (10.4) gives the average agreement per embryo, if only a small fraction of cells is impacted by the perturbation, the procedure can still be applied with relatively high confidence, as localised differences will be averaged out of the distance function. Interestingly, this procedure can efficiently single out the cells affected by the perturbation as those farthest away from the average after time registration.

10.1.4 The subcellular localization of ERK-KTR correlates with ERK activity in ascidian embryos

To confirm that the ERK-KTR-mClover sensor initially developed for mammalian cells (46) was also a reporter of ERK activity in ascidian embryos, and that our measurement strategy was suitable, we first compared the pattern of ERK-KTR nuclear accumulation with known ERK activity in ascidians. We focused our analysis at the 64-cell stage and used our census of known FGF-induced cells, which includes the progeny of 5 vegetal mother cells giving rise to differentially-induced daughters (A6.2, A6.3, A6.4, B6.2, B6.4) (Table S11) and that of the A6.1 endoderm precursor (98).

Analysis of the temporally-registered patterns of inferred ERK activity revealed a high level of reproducibility between embryos in all cells, and a clean separation of the cells known to be induced from uninduced cells (Fig. S62). Consistent with previous work in *Halocynthia roretzi* (100), the B7.1 and B7.2 posterior head endoderm progenitors strongly activate the ERK pathway (Fig. S62).

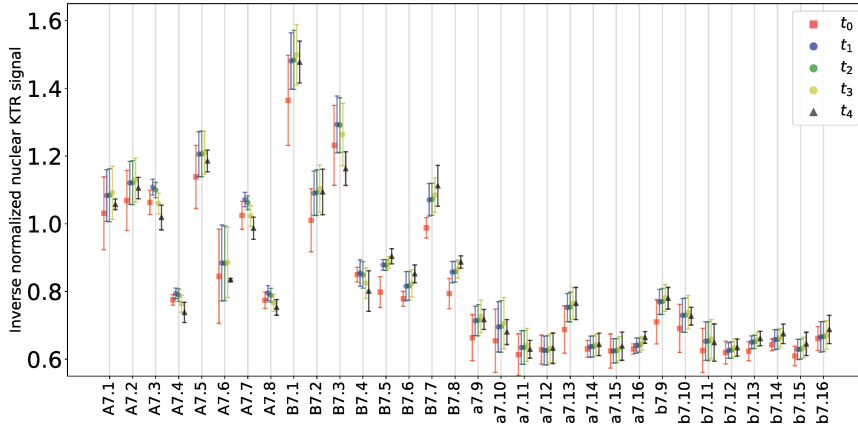


Figure S62: Average time-registered inverse normalized ERK-KTR value for wild-type embryos. The figure shows, for each cell at the 64-cell stage, the average inverse normalized value of nuclear KTR, obtained after time registration of 5 wild-type embryos. Each cell at the 64-cell stage was measured at five distinct time points (shown in different colors after time registration). Each point corresponds to the average value after time registration, with its standard deviation indicated by vertical bars. Time registration was performed according to the scheme of Fig. S61

Consistent with the proposal that ERK activation in early ascidian embryos mostly results from the activation of the Ras/Raf/MEK cascade, pharmacological inhibition of MEK activity from the early 16-cell stage strongly reduced the activation of ERK activation (Fig. S63).

10.2 Comparison of ERK-KTR signal to model predictions

To compare the predictions of the differential induction model with the experimentally-inferred activation of ERK, we ran the differential induction model on these embryos. We first determined a *best-consensus* model parametrization. As detailed in Table S12, the optimal values for the model parameters obtained after extensive parameter-space exploration for three wild-type embryos have a narrow yet non-zero variance. In addition, more than one optimal parameter configuration was found after each parameter exploration. To determine an average parametrization of the model we looked, among all optimal configurations found for the three wild-type embryos of Table S12, for the most similar ones. To this end, each parameter configuration p for embryo $i = ASTEC - Pm1, ASTEC - Pm5, ASTEC - Pm8$ was translated into a n -dimensional vector V_{p^i} after normalising the values of each parameter in the range $[0, 1]$. We then selected, among all possible optimal configurations for the three embryos, the three configurations $q^1 = q^{ASTEC-Pm1}, q^5 = q^{ASTEC-Pm5}, q^8 = q^{ASTEC-Pm8}$ that minimize $d(V_{q^1}, V_{q^5}) + d(V_{q^1}, V_{q^8}) + d(V_{q^5}, V_{q^8})$, $d(a, b)$ being the euclidean distance between two vectors a and b . We finally defined the *best-consensus* model parameterization by taking, for each parameter, the average of its values in q^1, q^5 and q^8 . This produced the set of values given in Table S12 that were used as parameter values when running our model on embryos for which KTR signal was measured.

We then performed a correlation analysis between the registered ERK-KTR values and the model predictions of the state of activation of RasGTP, assumed to be a monotonic function of ERK_{Model}^* . Since both ERK_{Model}^* and v_i^{conv} are information given for each cell, we quantified how much the two

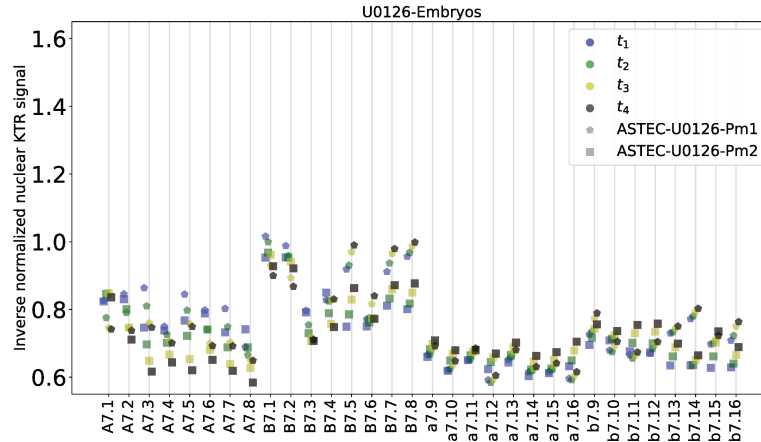


Figure S63: Time-registered inverse normalized ERK-KTR value for two embryos treated with the U0126 MEK inhibitor from the early 16-cell stage. The figure shows, for each cell at the 64-cell stage of two U0126-MEK-inhibitor-treated embryos, the time-registered inverse normalized ERK-KTR nuclear value. In both embryos and within the 64-cell stage and for each cell, the nuclear KTR was measured at 4 distinct time points (shown in different colors after time registration). Time registration was performed between the two U0126-MEK-inhibitor-treated datasets according to the scheme of Fig. [S61](#). Note the high level of reproducibility of the measurements between embryos.

sets of data are monotonic functions of one another by computing the Spearman correlation between $ERK_{Model}^*(c)$ and $v_i^{conv}(c)$ as functions of embryonic cells.

11 Analysis of the effect of changes in the cell-cell contact areas (Supplementary information to Figure 5B-D)

11.1 Production, naming and analysis of half embryos

To analyze the effects of changes in the surface of contact to FGF- or Ephrin-expressing cells, we halved embryos micro-injected with mRNA for PH-tdTomato and ERK-KTR-Clover at the 2-cell stage with a thin glass filament. Embryos were imaged in the MuViSPIM for both channels (488nm and 561nm) from the 16-cell stage onwards at 18°C for 151 consecutive time points, separated by 2 minutes intervals. The membrane and KTR channels were fused, segmented, manually curated as before. Due to the perturbation, naming was more difficult than in whole embryos, and proceeded iteratively.

Overall, we considered that the halving procedure does not globally alter fate specification processes or cell identities and chose the naming compatible with the smallest number of respecified cells. Animal and vegetal hemispheres were first identified thanks to the earlier cleavage of vegetal cells at the 24- and 44-cell stages. The antero-posterior axis was then identified by looking for anterior cells in the animal and vegetal hemisphere seeding a lineage consistent with a6.5 (anterior neural plate) and with A6.2 and A6.4 (notochord and tail nerve cord identity). The identity of these cells was further confirmed by the analysis of the KTR signal and by the analysis of their relative positions (a6.5 is expected to contact A6.2 and A6.4). These cells were used as anchors to identify the halving plane of the embryo and name

the other cells following the Conklin rule: two cells sharing a mother have consecutive indices, the lower index corresponding to the cell closest to the contact point between the vegetal-most cells and the halving plane.

Having identified each cell, we measured the nuclear KTR signal (see above) at the 32- and 64-cell stages in each cell at approximately 50% of its cell cycle (Fig. 5B, C, [S65](#) and [S64](#)). We also ran the model with the parameters of Table [reftab:mod-param-ctr](#) and compared the results to ERK activation values inferred from KTR nuclear signal measurements (Fig. 5B). Changes in fate specification processes were inferred from the comparative analysis of cell lineage trees within each analysed embryo or half embryo.

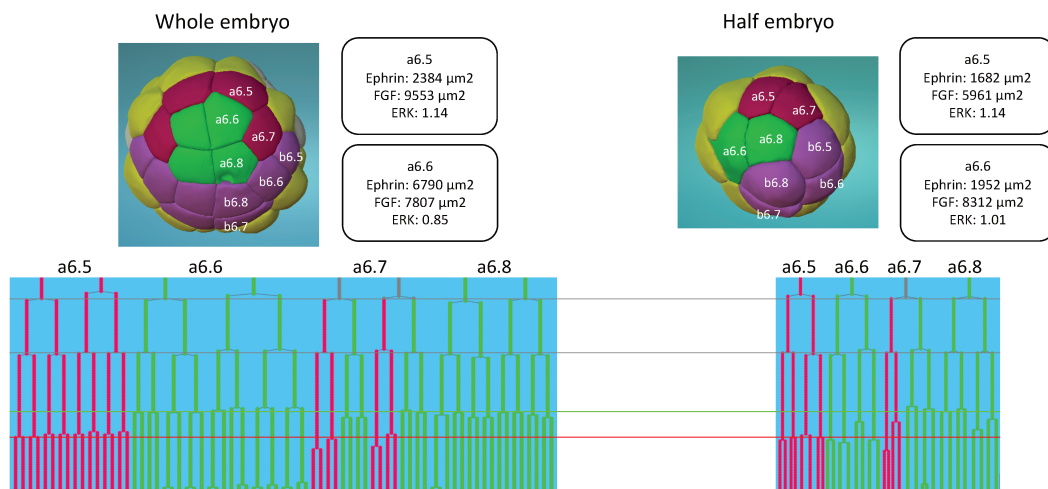


Figure S64: Effect of halving embryos on the neighborhood, ERK signalling intensity and cell lineage of the a6.5 and a6.6 blastomeres. View of ASTEC-Pm13 and ASTEC-half-Pm1 at the 32-cell stage. Green cells are future head epidermis, red future neural cells, mauve b-line ectoderm and yellow FGF-expressing vegetal cells. Bottom panels: a-line cell lineage trees originating from the indicated cells. Grey lines correspond to cells not yet fate restricted, green branches correspond in Whole embryos to future head epidermis, red to future neural cells. The horizontal lines indicate the timing of division of neural (red line), epidermal (green line) and undifferentiated (grey line). Note the characteristic delay in the 3rd cleavage of the neural precursors compared to epidermal precursors, and the difference in the cell lineage tree seeded by a6.6 in the whole embryo (epidermal) and in the half embryo (neural).

Discussion

L'intégration des mesures précises des géométries cellulaires dans un cadre théorique modélisant le lien entre les contacts cellule-cellule et les inductions cellulaires, combiné à des manipulations expérimentales, établit donc la surface des contacts entre cellules communicantes comme déterminant général de la spécification du destin dans les embryons précoces d'ascidies. Ce travail introduit le contrôle géométrique des inductions embryonnaires comme une alternative aux gradients morphogènes classiques et suggère que la nécessité de réduire la variabilité des zones de contacts cellulaires pour assurer une signalisation inductive robuste pourrait être une cause majeure de l'extrême reproductibilité développementale observée chez les ascidies et qui pourrait paradoxalement lever les contraintes sur l'évolution de leurs génomes. En effet, alors que la réponse précise aux gradients morphogènes à longue distance nécessite des couches de régulation sophistiquées (Balaskas et al., 2012; Sagner and Briscoe, 2017), qui peuvent contraindre l'architecture des réseaux de régulation et l'évolution globale du génome, la signalisation dépendante de la surface de contact garantit que deux cellules sœurs sont exposées à des signaux suffisamment différents pour définir directement les frontières, limitant ainsi le besoin de raffinements transcriptionnels ultérieurs.

Le développement d'un nouvel algorithme adapté aux données d'imagerie de microscope à feuille de lumière MuViSPIM, ASTEC a permis de générer dix embryons digitalisés de *Phallusia mammillata* qui constituent un atlas morphologique quantitatif et dynamique sans précédent. Cependant, la solution de segmentation actuelle, basée sur un algorithme faisant appel à la technique de partage des eaux (watershed), peut encore être difficile dans sa prise en main, avec notamment un large éventail de paramètres libres, ce qui pourrait limiter son utilisation sur d'autres types de données. Le large ensemble de données 3D+temps de l'atlas morphologique de *Phallusia mammillata* (images fusionnées des membranes et segmentations corrigées) ouvre néanmoins maintenant la porte aux méthodes d'apprentissage profond fondées sur les réseaux neuronaux convolutifs qui commencent à être appliquées à la biologie du développement et qui requièrent un grand nombre de données pour l'apprentissage (Tokuoka et al., 2020). Il serait intéressant de voir si ces nouvelles approches, seules ou en combinaison avec les outils classiques de segmentation et de suivi, contribueront à ce que le suivi cellulaire et le phénotypage systématiques à long terme deviennent une analyse standard en biologie du développement.

Enfin, l'ensemble de ce travail constitue une preuve de principe quant à notre capacité de combiner la production d'atlas morphologiques des embryons de *Phallusia mammillata* avec des mesures quantitatives de l'activité de la voie FGF/ERK à partir du biosenseur ERK-KTR. Cependant, la méthodologie développée pour cette étude présente certaines limites de passage à l'échelle, rendant difficile la production d'un atlas fonctionnel de l'activité de la signalisation ERK dans les embryons d'ascidies. En effet, l'identification et le nommage des cellules équivalentes pour chaque embryon reconstruit est réalisé à la main, une méthode très chronophage et source potentielle d'erreurs. Les mesures des signaux ERK-KTR sont également effectuées manuellement et il faudrait l'équivalent d'un an de travail à la main pour mesurer le ERK-KTR dans les noyaux de toutes les cellules d'un seul embryon imagé toutes les 2 minutes entre le stade 32-cellules et la fin de la neurulation. Les mesures n'ont donc été effectuées que sur un stade donné, le stade 64-cellules et de façon discontinue, ne permettant pas d'accéder à la dynamique fine de l'activité ERK.

Ainsi, afin de réaliser un atlas quantitatif de la signalisation ERK au cours de l'embryogenèse précoce chez *Phallusia mammillata*, je me suis donc ensuite concentré sur i) l'automatisation du nommage des cellules des embryons d'ascidies reconstruits (chapitre 2) et ii) sur l'automatisation des mesures du signal ERK-KTR au cours du développement et l'étude de ces données pour 4 embryons de *Phallusia mammillata* entre le stade 32-cellules (St.6) et le stade de neurula tardive (st. 16) (chapitre 3).

CHAPITRE 2 : L'IDENTIFICATION ET LE NOMMAGE SYSTÉMATIQUE ET AUTOMATIQUE DES CELLULES HOMOLOGUES DANS LE EMBRYONS STÉRÉOTYPÉS D'ASCIDIES ÉTABLI LES ASCIDIES COMME UN CADRE NUMÉRIQUE POUR LA CONSTRUCTION ET L'INTÉGRATION D'ATLAS 4D FONCTIONNELS MULTI-ÉCHELLES À LA RÉOLUTION CELLULAIRE

Introduction

L'identification systématique de cellules équivalentes dans des embryons distincts est une condition préalable pour la réalisation d'un atlas fonctionnel de l'activité de ERK. En effet, cette étape est essentielle pour d'étudier la dynamique de l'activité de la signalisation de ERK à la résolution cellulaire dans des cellules identifiées et pour intégrer, dans un référentiel commun, une cohorte d'embryons individuels afin d'étudier la variabilité de la signalisation ERK entre cellules homologues d'embryons différents.

Dans la plupart des organismes modèles, des points de repère ou d'ancrage communs, tels que des protéines ou des ARNm à motifs spatiaux, doivent être définis et systématiquement inclus pour recalibrer des ensembles de données distincts sur un individu "standard" ou prototypique (Tomer et al., 2010; Castro-González et al., 2012; Knowles and Biggin, 2013). Cependant, certains taxons, dont les ascidies, présentent une telle reproductibilité cellulaire du développement que chaque cellule a des homologues dans tous les embryons d'une espèce. Il a donc été possible de définir des classes de cellules et de leur attribuer des noms canoniques uniques.

Les cellules embryonnaires des ascidies sont traditionnellement nommées selon un schéma proposé par Conklin (Figure 21) (Conklin, 1905). Alors que la syntaxe du nommage proposé par Conklin et qui assure la cohérence des noms de cellules est largement applicable au-delà des stades de la gastrula, sa règle de décision pour l'attribution des noms entre deux cellules sœurs, qui repose sur une distance géodésique, ne peut pas être étendue au-delà du début de la gastrulation, et ne peut donc pas être utilisée de manière générale pour automatiser la dénomination *de novo* d'un nouvel embryon. D'autres règles inspirées de Conklin ont été utilisées pour nommer les cellules de la plaque neurale situées à l'extérieur à des stades ultérieurs (Nicol and Meinertzhagen, 1988; Cole and Meinertzhagen, 2004). En revanche, la localisation interne des cellules précurseurs cardiogéniques identifiées expérimentalement a empêché leur nommage selon la syntaxe de Conklin, les noms qui leur ont été attribués reflétant leur évolution cellulaire larvaire mais pas leur origine clonale (Stolfi et al., 2010).

Ainsi, pour pallier l'absence de règles de décision claires permettant de nommer *de novo* toutes les cellules d'un embryon après le stade 76 cellules, stade au cours duquel la première division interne a lieu, nous avons précédemment choisi de nommer manuellement un embryon de référence, en utilisant des décisions *ad hoc* inspirées de Conklin, et de transférer manuellement cette information aux nouveaux embryons (Guignard et al., 2020). Cette tâche qui a permis de constituer un atlas morphologique de 10 embryons de *Phallusia mammillata* avec plus de 6000 cellules nommées jusqu'à la 10^{ème} génération (la 11^{ème} génération apparaît en fin de neurulation), a donc été largement tributaire d'une expertise manuelle chronophage et sujette aux erreurs.

Pour établir pleinement les embryons d'ascidies comme un cadre numérique permettant de construire, intégrer et analyser des atlas 4D fonctionnels avec une résolution cellulaire, il était donc nécessaire d'automatiser l'identification et la nomenclature systématique des cellules embryonnaires pour les embryons reconstruits d'ascidies.

En étroite collaboration avec Grégoire Malandain (INRIA, Sophia Antipolis), spécialiste en analyses d'images biologiques, analyses topologiques et construction d'atlas, un cadre théorique et les méthodes algorithmiques permettant l'identification automatisée et impartiale des cellules dans les embryons stéréotypés de l'ascidie *Phallusia mammillata* ont été développées.

En nous appuyant sur la syntaxe historique du nommage de Conklin et sur la disponibilité de la population d'embryons de référence entièrement segmentés, suivis et nommés manuellement (Guignard et al., 2020), nous avons établi une stratégie de nommage basée sur le transfert d'information des noms de cellules depuis cette collection d'embryons. En effet, nous avons tout d'abord montré que les mesures des surfaces de contact entre cellules voisines sont suffisamment invariantes à travers les embryons de *Phallusia* pour représenter de manière unique chaque cellule d'un embryon par le vecteur de ses surfaces de contact avec les cellules voisines directes et que ces mesures précises des voisinages sont suffisamment différentes entre cellules sœurs pour les distinguer sans ambiguïté. Nous avons donc établi une nouvelle règle de décision basée sur cette mesure du voisinage. D'après la syntaxe de Conklin chaque cellule $F_{p,q}$ se divise en $F_{(p+1).(2q-1)}$ et $F_{(p+1).(2q)}$. A partir d'un nommage initial manuel (classiquement réalisé au stade 64-cellules), le nom $F_{p,q}$ pour chaque cellule est propagé jusqu'à sa division grâce à l'information du lignage. Ainsi, quand une cellule se divise, la seule décision consiste à choisir à quelle cellule fille attribuer le nom $F_{(p+1).(2q-1)}$ et à quelle cellule fille attribuer le nom $F_{(p+1).(2q)}$. Pour prendre cette décision, nous avons implémenté une mesure de distance du voisinage division-à-division qui permet de nommer automatiquement les deux cellules filles d'un nouvel embryon par héritage depuis un embryon de *Phallusia* déjà nommé. Afin de rendre notre procédure de nommage automatique plus robuste et de tenir compte de la variabilité biologique du développement, nous avons décidé d'établir toute la population comme embryons de référence, pour lesquels nous avons d'abord développé des méthodes informatiques afin de détecter et de corriger systématiquement les erreurs résiduelles issues du nommage manuel initial. Enfin nous avons montré que notre méthode, qui permet en quelques minutes de nommer entièrement un embryon (jusqu'à la 11^{ème} génération), est très précise à la fois sur les données de référence et sur les nouveaux ensembles de données.

Dans le cadre de ce travail de recherches, avec notre collaborateur Grégoire Malandain et mon directeur de thèse Patrick Lemaire, j'ai participé à la conceptualisation du cadre théorique qui comprend : l'analyse de l'invariance spatiale et temporelle au cours du développement chez *Phallusia mammillata*; le développement des méthodes permettant d'identifier de potentielles erreurs dans le nommage manuel initial (erreurs de non respect de la syntaxe de Conklin et inversions des noms de cellules sœurs) et de nommer automatiquement chaque cellule fille après une division; le développement d'un score de divergence qui identifie les cellules qui présentent une configuration de voisinage cellulaire la plus inhabituelle dans un embryon nouvellement segmenté par rapport à notre base de données d'embryons de références. J'ai également été en charge des analyses ayant conduit à i) l'établissement d'une version corrigée et cohérente du nommage des embryons de référence, ii) à la quantification de la précision de la méthode du nommage

automatique pour chaque embryon référence, et iii) à l'établissement d'une population d'embryons de références pour tenir compte de la variabilité biologique du développement. Enfin, j'ai généré un nouvel embryon de *Phallusia* reconstruit, Astec-Pm16, pour lequel j'ai analysé qualitativement et quantitativement la qualité du nommage afin de valider la stratégie du nommage automatique sur un nouveau jeu de données.

L'étude complète est présentée dans le manuscrit de l'article "Systematic automated cell naming of segmented ascidian embryos establishes ascidians as a digital framework to build, analyse and integrate multiscale functional 4D atlases at cellular resolution".

Manuscrit de l'article : Biasuz, Kilian et al. "Systematic automated cell naming of segmented ascidian embryos establishes ascidians as a digital framework for constructing, analysing and integrating multi-scale functional 4D atlases at cellular resolution."

Systematic automated cell naming of segmented ascidian embryos establishes ascidians as a digital framework for constructing, analysing and integrating multi-scale functional 4D atlases at cellular resolution

Kilian Biasuz^{1,#}, Patrick Lemaire¹ and Grégoire Malandain^{2,#}

1) CRBM, Université de Montpellier, CNRS, France

2) Université Côte d'Azur, Inria, CNRS, I3S, France

#: Corresponding authors

ABSTRACT

Systematic identification of equivalent cells across distinct embryos is a prerequisite for the integration, into a common referential, of cohorts of individual embryos. This task is largely reliant on time-consuming and error-prone manual expertise. Here, we present a theoretical framework for unbiased and automated cell identification in stereotyped embryos of the ascidian, *Phallusia mammillata*. Building upon the historic Conklin naming syntax and the availability of a population of 7 fully-segmented, tracked and manually named reference embryos, we present a naming strategy based on the transfer of the naming of cells from this collection of embryos. We show that the measurements of contact areas between neighbouring cells is sufficiently invariant across *Phallusia* embryos to uniquely represent each cell of an embryo by the vector of its contact surfaces with first order cell neighbours and we implemented a division-to-division distance measurements to automatically name the two daughter cells after each division. To render our automatic naming procedure more robust and account for developmental biological variability, we decided to establish the whole population as reference embryos, for which we first developed computational methods to systematically detect and correct residual manual naming errors. Finally, we show that our technique is highly accurate on both ground-truth and novel datasets and provide first evidence supporting that the comparison of the precise measurements of the cell neighbourhood between *Phallusia* embryos can be used to quantify the natural developmental variability in ascidian.

INTRODUCTION

How a single fertilised cell progressively transforms into a complex organism, with tissues composed of a multitude of cell types in correct numbers and position, is the central question of developmental biology (Wolpert, 1994). Decades of research in this field have sketched a broad portrait of the fundamental processes involved from gene transcription to the precise control and coordination of the behaviour of cells during morphogenesis. Developmental mechanisms have, however, traditionally been studied at the tissue level, in a qualitative manner and with low temporal resolution (Biasuz et al., 2018).

The recent breakthrough development of light-sheet microscopy (Khairy and Keller, 2011) and dedicated computational tools now allow the digital reconstruction of animal and plant development from 3D + time live imaging data (4D) (McDole et al., 2018; Duncan et al., 2019; Guignard et al., 2020). In the same way as geographic information systems (GIS) atlases allow to register, project and

query heterogeneous datasets onto a spatial reference database, these dynamical single-cell resolution anatomical reconstruction of cell properties can be augmented by the integration of the transcriptomes (Karaiskos et al., 2017; Briggs et al., 2018; Farrell et al., 2018; Wagner et al., 2018; Cao et al., 2019; Packer et al., 2019; Sladitschek et al., 2020; Zand and Ruan, 2021) or proteomes (Budnik et al., 2018; Ma et al., 2021) of individual cells. Such multimodal and multiscale digital atlases of animal development provide a quantitative description of the genetic and cellular underpinnings of morphogenesis, which can support a system-level modelling of development (Biasuz et al., 2018; Pastor-Escuredo and del Álamo, 2020).

One of the most challenging aspects of this strategy is that the heterogeneous data (e.g. transcriptomes, cell shapes...) to be integrated are generally captured in different embryos, by different experimentalists, in different labs (Di Gregorio and Hadjantonakis, 2006; Knowles and Biggin, 2013; Armit et al., 2017). The integration of heterogeneous data into a digital atlas of wild-type development thus requires the definition of a common referential onto which cohorts of individuals can be mapped, despite natural or experimental phenotypic variability (Oates et al., 2009).

In most model organisms, common landmarks or anchors, such as spatially-patterned proteins or mRNAs, need to be defined and systematically included to register distinct datasets onto a “standard” or prototypical individual (Tomer et al., 2010; Castro-González et al., 2012; Knowles and Biggin, 2013). Some taxa however, including nematodes and ascidians, show such a cellular reproducibility of development, that each cell has homologs across all embryos of a species. It is thus possible to define cell classes and to assign them unique canonical names. For instance, (Guignard et al., 2020) defined 354 cell classes between the 64-cell and late neurula stages of the ascidian *Phallusia mammillata*. Each cell of a particular embryo being an instance of a cell class, is then named by inheritance of the name of the class (Conklin, 1905; Sulston and Horvitz, 1977). Such model organisms are particularly favourable to the high resolution integration of heterogeneous multiscale datasets.

Using the optically transparent embryos of the ascidian *Phallusia mammillata*, we recently reported the accurate and automated systematic whole-cell segmentation and tracking over 5 cell generations of plasma membrane-labelled cells with a 2-min temporal resolution, up to the neurula stage (Guignard et al., 2020). The integration of this quantitative and dynamic description of the positions, contacts and division patterns of all cells in 10 embryos revealed that development was indeed reproducible down to the subcellular scale over this period of development. Reaching this conclusion, however, required the manual naming of more than 6000 cells, a time-consuming and potentially error prone process. To fully establish a high-throughput digital framework to build, analyse and integrate multiscale functional 3D + time ascidian atlases with cellular resolution, there is thus a need to automate the systematic identification and naming of embryonic cells.

Ascidian embryonic cells are traditionally named according to a scheme proposed by Conklin (Conklin, 1905). Each embryonic cell is given a name of the form Fp.q (e.g. A7.6), where F is a letter that indicates the founding lineage at the 8-cell stage, from which the cell originates (A is anterior vegetal, a anterior animal, B posterior vegetal and b posterior animal), p is the rank of the mitotic cell cycle (e.g. p = 3 means that the cell is in its 3rd mitotic cell cycle since fertilisation) and q is an integer reflecting the relative position of the cell with respect to the vegetal pole of the embryo or to the

midline, compared to its sibling. To distinguish bilaterally symmetrical left and right cells (from vegetal view), right cells are given a “*” suffix (e.g. A7.6*), while the matching left cells receive either no suffix or, more recently, an “_” suffix (e.g. A7.6_) (Figure 1).

As development proceeds and cells divide, the Conklin naming syntax specifies that cell Fp.q divides into two daughters, named F(p + 1).(2q) and F(p + 1).(2q – 1). The two daughters are then individually named according to their relative position in the embryo: F(p + 1).(2q – 1) is given to the daughter whose apical surface is geodesically closer to the vegetal pole of the embryo or to the midline, if the geodesic distance to the vegetal pole is identical between the two daughter cells. This use of the geodesic distance limits the applicability of the Conklin naming scheme to stages where all cells have a contact with the external medium. From the 76-cell stage onwards, the axis of some cell divisions is oriented perpendicular to the apical surface of the embryo, generating an internal cell whose geodesic distance to the vegetal pole is not defined.

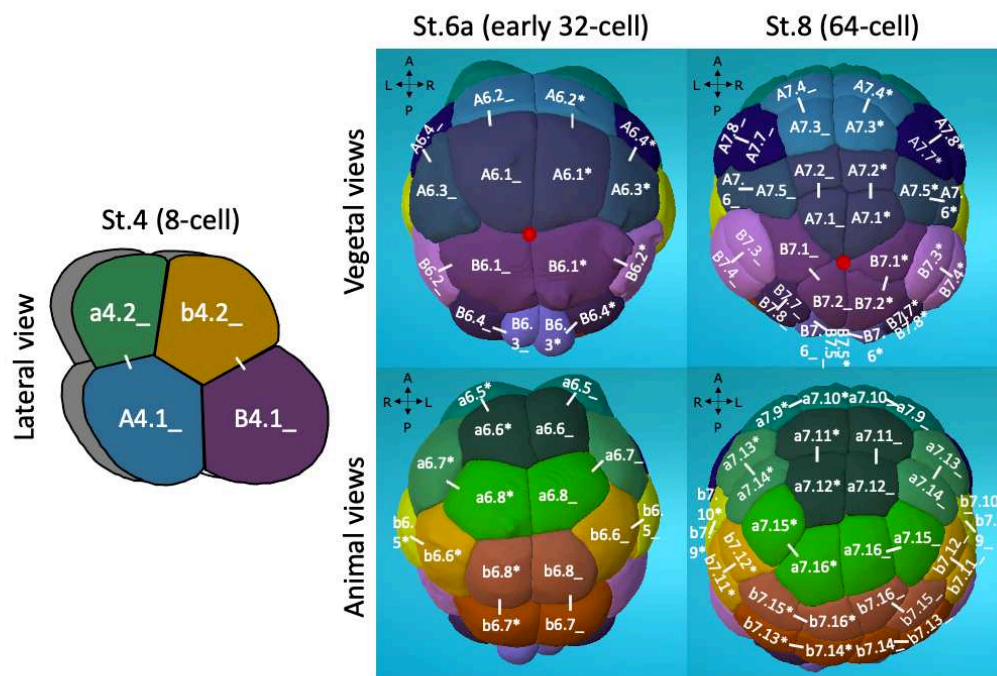


Figure 1 : Cell naming of early ascidian embryos according to the Conklin nomenclature. Left: lateral schema of an 8-cell embryo, showing the position of the 4 founding cell lineages. Middle and right: vegetal (top) and animal (down) views of a 32-cell (middle) and a 64-cell (right) stage *Phallusia mammillata* embryo. Each bilateral cell pair at the 32-cell stage is assigned a unique colour, which is propagated at the 64-cell stage. Names (in white) are according to Conklin. White lines link sister cells. Antero-Posterior (AP) and Left-Right (LR) axes are represented. The red disks materialise the vegetal pole of the embryo, used as a reference for the measurement of the geodesic distance for the Conklin naming. Middle, left : snapshots of the Astec-Pm-Stages embryo at the indicated stage (Guignard et al., 2020) visualised using MorphoNet (Leggio et al., 2019).

Thus, while the Conklin naming syntax ensures the consistency of cell names and is broadly applicable beyond the gastrula stages, its decision rule for assigning a name to sister cells cannot be extended beyond the onset of gastrulation, and thereby cannot be generally used to automate the *de novo* naming of a new embryo. Additional Conklin-inspired rules have been used to name the

externally located neural plate cells at later stages (Nicol and Meinertzhagen, 1988; Cole and Meinertzhagen, 2004). By contrast, the internal location of experimentally-identified cardiogenic precursor cells has precluded their individual naming according to the Conklin syntax, their assigned names reflecting their larval cell fates but not their clonal origin (Stolfi et al., 2010).

To overcome the absence of clear decision rules allowing to *de novo* name all cells of an embryo past the 76-cell stage, we previously chose to manually name a reference embryo, using *ad hoc* decisions loosely inspired by Conklin, and to manually transfer this information to new embryos (Guignard et al., 2020). Here, we present the algorithmic methods we developed to automate this transfer process. We kept the Conklin syntax, and show that following the division of a cell, the surface of contact each of its daughter cells establishes with its first order neighbours (see glossary table S1) provides sufficient information to unambiguously name it. To render our automatic naming procedure more robust and account for developmental biological variability, we decided to establish a population of reference embryos, for which we first developed tools to detect and correct residual manual naming errors. Finally, we show that our technique is highly accurate on both ground-truth and novel datasets and provide first evidence supporting that the comparison of the precise measurements of the cell neighbourhood between *Phallusia* embryos can be used to quantify the variability of the ascidian development.

RESULTS

Temporal and spatial invariance of local cell neighbourhoods during *Phallusia* development

The traditional Conklin naming strategy illustrates the difficulty of identifying sister cells by their relative position with respect to a global embryonic landmark. We thus explored whether the first order neighbourhood (see glossary table S1) of sister cells would be sufficient to unambiguously distinguish them. Figure 2A shows that at each time point in the 7 segmented Astec-Pm embryos (Guignard et al., 2020) with the most canonical early development, each embryonic cell contacts on average 5 to 6 first order neighbours contact, over more than 5% of its surface.

Comparison of the neighbourhoods between two embryos E and E' can be hampered by slight heterochronies in cell divisions, causing cell C_i to contact cell C_j in E , and the daughters of C_j in E' . This difficulty can be offset by defining a common virtual neighbourhood of closest common ancestors, in which the surface of contact between C_i and C_j in E' is the sum of the surfaces of contact of C_i with the daughters of C_j (Figure S1A).

In this common referential, the neighbourhood of each cell snapshot (see glossary table S1) $C_i^{E,t}$, the instance of cell C_i in embryo E at time t , can be represented by the vector of its contact surfaces with all cells present at this time point, where $S_{i,j}^{E,t}$ is the contact surface of $C_i^{E,t}$ with $C_j^{E,t}$:

$$C_i^{E,t} = \begin{pmatrix} S_{i,1}^{E,t} \\ \dots \\ S_{i,j}^{E,t} \\ \dots \end{pmatrix}$$

The similarity of neighbourhoods between two cell snapshots $C_i^{E,t}$ and $C_k^{E',t'}$, originating from the same or different embryos, can then be estimated by the fraction of the surface of $C_i^{E,t}$ and $C_k^{E',t'}$ that contacts different cells, computed with the Manhattan distance:

$$Nb(C_i^{E,t}, C_k^{E',t'}) = \frac{1}{2} \sum_j \left| \frac{S_{i,j}^{E,t}}{\sum_l |S_{i,l}^{E,t}|} - \frac{S_{k,j}^{E',t'}}{\sum_l |S_{k,l}^{E',t'}|} \right|$$

Using this distance, we first analysed the temporal stability of neighbourhoods of a given cell at successive time points (Figure 2B, Figure S2). While the neighbourhood distances evolve rapidly between the first two time points after each cell division, they stabilise from the 2nd time point onwards (Figure 2B, C), consistent with our previous analysis (Guignard et al., 2020). In all subsequent analyses, contact areas with neighbouring cells, computed 4 time points after division (called subsequently time point T_4), were considered representative of the contacts established by a cell over most of its lifetime.

Having established a method to quantitatively compare local cell neighbourhoods, we next assessed the similarity of local neighbourhoods of various classes of cells. Figure 2D illustrates that the distributions of neighbourhood distances between pairs of matching cells (see glossary table S1), sister cell pairs and other possible cell pair combinations of the same generation, within and across embryos, show little overlaps. As expected, the local neighbourhoods of pairs of cells of the same generation that are neither matching nor sister cells (see glossary table S1) were most dissimilar. On average, matching cell pairs contacted different cells over approximately 20% of their surface, consistent with the reported reproducibility of cell contacts during ascidian embryogenesis (Guignard et al., 2020). By contrast, sister cell pairs contacted different cells over 60% of their surface on average.

Taken together, these results confirm that the surfaces of contacts established by a cell with its first order neighbours become temporally stable shortly after completion of mitosis, are highly conserved between embryos, and differ significantly between sisters.

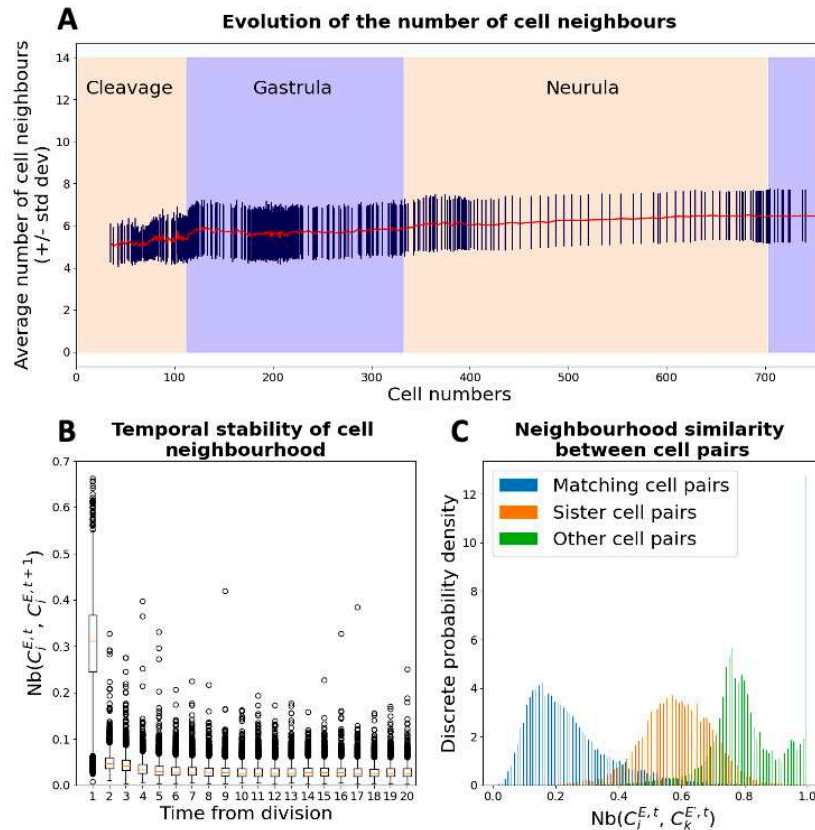


Figure 2 : Temporal and spatial reproducibility of cell neighbourhoods during *Phallusia mammillata* development. A) Evolution of the average number (red line) of cell neighbours touching a cell over >5% of its surface (+/- standard deviation, dark blue) during early *Phallusia* development. B) Stability of neighbourhood analysis. For each cell, the stability of its neighbourhood was determined by calculating the neighbourhood distance between two successive time points throughout the entire lifetime of the cell. Top, middle, and bottom lines in the box plot indicate 75th percentile, median, and 25th percentile of data, respectively. C) Neighbourhood distance measurements between pairs of cells: matching cells (blue), sister cells (orange), other cells (green, includes all cells of the same generation except matching and sister cells) within and across embryos. A-C) n=7 embryos.

Automated cell naming using the local invariance of surface of contact with neighbouring cells

We will now explore how local cell neighbourhood information can be used to automatically name cells. Let us first consider how cells, prior to their naming, are identified during the segmentation process, using the ASTEC pipeline as an example (Guignard et al., 2020). Thanks to the high temporal resolution of lightsheet imaging, each cell C_i^E of an embryo E is imaged at several time points. During the segmentation process, the cell snapshot $C_i^{E,t}$ is identified by an arbitrary cell label, which in ASTEC has the syntax $[t][k]$, where $[k]$ is a unique 4-digit numerical identifier indicating the segmentation rank of the cell at time point t (e.g. label 890152 indicates the 152nd cell segmented at time point 89). The different snapshots of a cell C_i^E thus receive different numerical identifiers within a segmentation, or between different segmentation runs of the same or different embryos.

The automated naming of an embryo E consists, for each cell C_i^E , in associating to every cell snapshot $C_i^{E,t}$ the unique Conklin cell name, $Fp.q$, corresponding to the C_i cell class. The automated naming process we designed includes four steps, of which step 3 is the most challenging: 1) initiation by associating a $Fp.q$ name to each cell snapshot present at an early time point, either manually based

on its position in the embryo and annotator knowledge or by transfer from an named embryo of same stage (Michelin et al., 2015); 2) propagation of this $F_{p,q}$ name to all $C_i^{E,t}$ snapshots, using cell lineage information ; 3) upon division of C_i^E , association of one of the only two possible names, $F_{(p+1).(2q-1)}$ and $F_{(p+1).(2q)}$, to each of its daughters, $d_0(C_i^E)$ and $d_1(C_i^E)$, by comparison with a reference embryo R ; 4) iteration of steps 2 and 3 until all cells are named.

To perform step 3, let us first represent the division of cell C_i^R in a reference embryo R as the vector $DIV(C_i^R) = (d_0(C_i^R), d_1(C_i^R))$, where $d_0(C_i^R)$ and $d_1(C_i^R)$ are, by construction, the cells named $F_{(p+1).(2q-1)}$ and $F_{(p+1).(2q)}$, respectively (Figure S3).

Let us now consider the division of C_i^E , which can be represented by either $DIV(C_i^E) = (d_0(C_i^E), d_1(C_i^E))$ or, after permutation of the daughters, by $DIV^P(C_i^E) = (d_1(C_i^E), d_0(C_i^E))$ (Figure S3), where $d_0(C_i^E)$ is by construction the daughter cell with the lowest cell label in the segmented image at T_4 after the division of C_i^E . Naming the daughters of C_i^E by transfer from the reference embryo is equivalent to deciding which of $DIV(C_i^E)$ or $DIV^P(C_i^E)$ is more similar to $DIV(C_i^R)$. Having established in the previous section that the first order neighbourhoods of sister cells differ, we thus compared the neighbourhoods of the daughters of C_i^E and C_i^R at T_4 (Figure S3). To this aim, we introduced a division-to-division distance (D2D), built from the neighbourhood distances between daughter cells in E and R , and which is minimal when the neighbourhoods of the d_0 and d_1 daughters of C_i^E and C_i^R are in perfect agreement:

$$D2D(DIV(C_i^E), DIV(C_i^R)) = 1/2(Nb(d_0(C_i^{E,T4}), d_0(C_i^{R,T4})) + Nb(d_1(C_i^{E,T4}), d_1(C_i^{R,T4})))$$

Comparison of $D2D(DIV(C_i^E), DIV(C_i^R))$ and $D2D(DIV^P(C_i^E), DIV(C_i^R))$ is then sufficient to name $d_0(C_i^E)$ and $d_1(C_i^E)$: if $D2D(DIV(C_i^E), DIV(C_i^R)) < D2D(DIV^P(C_i^E), DIV(C_i^R))$, $d_0(C_i^E)$ inherits the biological name of $d_0(C_i^R)$ and $d_1(C_i^E)$ that of $d_1(C_i^R)$. Otherwise, $d_0(C_i^E)$ inherits the name of $d_1(C_i^R)$ and $d_1(C_i^E)$ that of $d_0(C_i^R)$.

The potential of this naming strategy relies on the expectation that matched (Figure 3A) and permuted (Figure 3B) C_i^E divisions have sufficiently different D2D distances to C_i^R . As expected, the two histograms showed little overlaps (Figure 3C) and the matched-to-mismatched distance difference between the D2D distances of the normal (DIV) and permuted (DIV^P) divisions had a high mean (0.35) and a low standard deviation (0.14) (Figure 3D). We note, however, that for a small minority of cell divisions, the pairwise D2D distance of the permuted division is smaller than that of the normal one, which suggests that the name of some sister cells had been mistakenly permuted in our original manual reference naming.

To further evaluate the accuracy of our automated naming procedure, we next renamed each segmented Astec-Pm embryo *de novo* by automated transfer from Astec-Pm8, as previously done manually (Guignard et al., 2020). The automated and manual procedures agreed on the names of 86,3-95,9% of cells, depending on the embryo (Table 1). Our procedure is thus able to correctly name a large majority of cells.

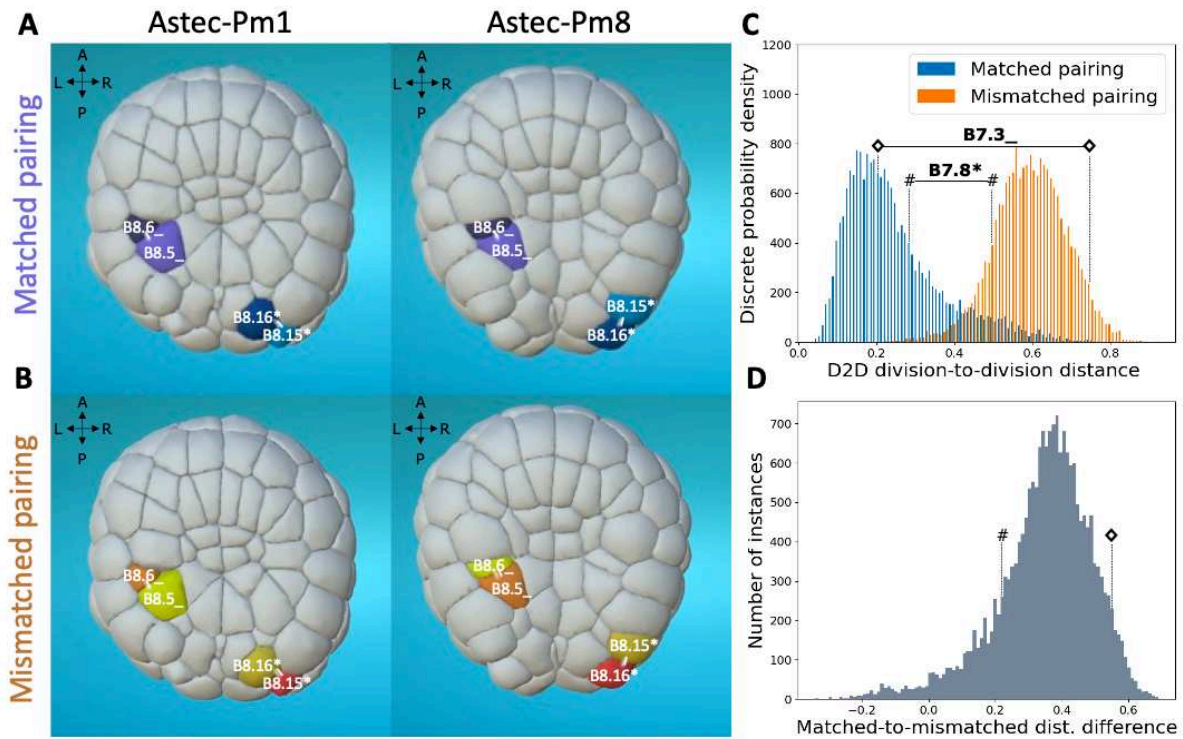


Figure 3 : Pairwise division-to-division distances distinguish matching and permuted sister cells.

A-B) Illustrations of the matched (A; $D2D(DIV(C_i^E), DIV(C_i^E))$) and mismatched (B; $D2D(DIV(C_i^E), DIV^P(C_i^E))$) pairing of the daughters of B7.3_ or B7.8* between Astec-Pm1 and Astec-Pm8. In both examples, the pairwise D2D division-to-division distances are computed between sister cells represented with the same colour code. C) Histograms of the pairwise division-to-division distances between either matched (blue) or mismatched (orange) sister cells (blue), within and across the 7 reference embryos. (\diamond) and ($\#$) highlight the positions on the histograms of the D2D division-to-division distance values for B7.3_ (\diamond) and B7.8* ($\#$) divisions between Astec-Pm1 and Astec-Pm8 with matched or mismatched pairings with $D2D(DIV(B7.3_Astec-Pm1), DIV(B7.3_Astec-Pm8))=0.20$, $D2D(DIV(B7.3_Astec-Pm1), DIV^P(B7.3_Astec-Pm8))=0.74$, $D2D(DIV(B7.8*Astec-Pm1), DIV(B7.8*Astec-Pm8))=0.28$, $D2D(DIV(B7.8*Astec-Pm1), DIV^P(B7.8*Astec-Pm8))=0.50$. D) Histogram of the differences between mismatched and matched D2D division-to-division distances ($D2D(DIV(C_i^E), DIV^P(C_i^E)) - D2D(DIV(C_i^E), DIV(C_i^E))$). (\diamond) and ($\#$) highlight the positions on the histogram of the D2D division-to-division distance differences values for B7.3_ (\diamond ; value=0.54) and B7.8* ($\#$; value=0.22).

Embryo ID	Manual naming	Automated naming by transfer from Astec-Pm8		Differences between automated & manual namings	
	# Names	# Names	# Unnamed cells	# Different names*	% Agreement
Astec-Pm1	725	713	12	71	88,6
Astec-Pm3	596	584	12	56	88,6
Astec-Pm4	507	495	12	34	90,9
Astec-Pm5	604	592	12	71	86,3
Astec-Pm7	389	387	2	16	95,4
Astec-Pm8	723	711	12	18	95,9
Astec-Pm9	518	506	12	41	89,8
All embryos	4062	3988	74	307	90,6

Table 1 : Comparison of the manual and automated *de novo* naming of Astec-Pm embryos by transfer from Astec-Pm8. (#) = "Number of". *: Includes both primary and induced differences (see glossary table S1).

Correction of manual errors in the reference dataset

The disagreement between the original manual and *de novo* automated namings of 5 to 15% of cells, and the negative values in Figure 3D suggested that our original manual naming may be partially erroneous. We indeed identified two types of discrepancies between namings: syntax errors and sister cell name permutations, which in most cases could be traced back to annotation errors in the manual naming. In this section we present how we identified and resolved these discrepancies to generate a corrected naming for our reference embryo collection.

First, we detected a limited number of errors affecting the Conklin syntax. For example, the left b10.91* cell in Astec-Pm1 was mistakenly named as its right counterpart, b10.91_, while the daughters of b9.58* in Astec-Pm4 were named b10.113* and b10.114*, instead of the expected b10.115* and b10.116*. Such syntactic errors were automatically detected by parsing the cell lineage and cell name information (see methods), resulting in the correction of 36 cell names (Table S2), including 19 in Astec-Pm8.

The second source of differences consisted in the permutation of the names of sister cells. Figure 4 A-B illustrates that the daughters of the left and right b7.12 cells in Astec-Pm8 are permuted compared to their Astec-Pm5 counterparts. Clustering the normal and permuted divisions of the b7.12* and b7.12_ cells across all embryos using our D2D division-to-division distance revealed that all embryos behave consistently except Astec-Pm8 (Figure 4C). We thus considered that sister cell names were permuted in the original naming if they differ from what is found in the majority of reference embryos.

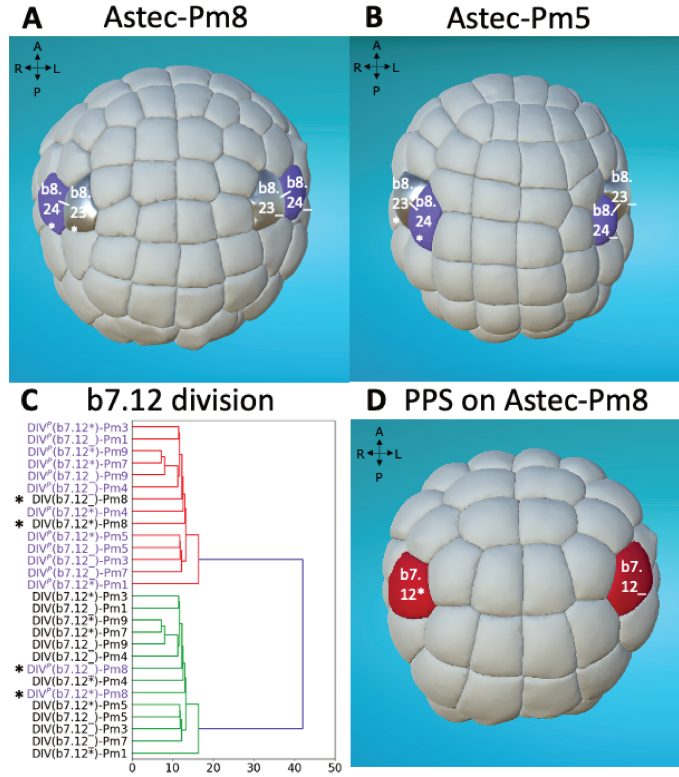


Figure 4: Systematic identification of sister cells with inverted manual names in a minority of embryos. A-B) Visualisation using MorphoNet (Leggio et al., 2019) of manually assigned names for daughters of b7.12_ and b7.12* divisions in Astec-Pm8 (A) and Astec-Pm5 (B). Cell names are displayed and cells with the same name between the two embryos are colour coded (blue and highlighted). C) Dendrogram of the b7.12 divisions. The ‘single’ linkage method is used to compute the distance between two clusters. Asterisks (*) highlight the positions of the b7.12 divisions of Astec-Pm8 in the dendrogram. D) Visualisation of the b7.12 cells (in red) with a negative value of the permutation proposal score (PPS) on Astec-Pm8 using MorphoNet. B-C-D) Animal views, Antero-Posterior (AP) and Left-Right (LR) axes are represented.

To systematically identify permuted naming among the 4082 named cells of the reference embryos, we reasoned that if the names of the daughters of cell C_i^E are inverted compared to most reference embryos R , the sum of the pairwise D2D division-to-division distances of the permuted division of C_i^E to matching unpermuted divisions in reference embryos should be smaller than the sum of the pairwise D2D division-to-division distances of its unpermuted division to other unpermuted divisions. We thus constructed a Permutation Proposal Score:

$$PPS(C_i^E) = \sum_{R \in \mathbb{R}, R \neq E} D2D(DIV^P(C_i^E), DIV(C_i^R)) - \sum_{R \in \mathbb{R}, R \neq E} D2D(DIV(C_i^E), DIV(C_i^R))$$

which is expected to have a negative value in case of a naming permutation. Fifty-four divisions with negative PPS values were identified across the dataset and manually checked by visualising with MorphoNet the divisions of mother cells with negative PPS values in each segmented embryo (Figure 4C). The permutations proposed by the PPS were either obvious according to Conklin’s rules, suggestive of a manual error during the original naming or more ambiguous because the cells were in an internal position where Conklin’s rules were not applicable, or because their division angle was atypical. In such cases, we accepted the proposed permutation to maximise the consistency of the

namings between embryos. Overall, all proposals were accepted and 150 (3,6%) cell names were corrected: 108 daughters of cells whose division was permuted, and 42 granddaughters indirectly affected by this permutation (Table S2).

This corrected database can now be considered the reference to name with a Conklin syntax all embryonic *Phallusia* cells up to the 10th generation, including those that could not previously be assigned a Conklin name as illustrated for 9th generation cells in the B7.5 cardiopharyngeal lineage (Razy-Krajka et al., 2014) (Figure S4).

Naming with respect to a collection of reference embryos increases accuracy and robustness to natural variability

The above corrections changed the names of 38 cells in the Astec-Pm8 embryo initially used as a reference to manually name the other embryos. As expected, the *de novo* automated naming of the 6 other reference embryos with respect to the corrected Astec-Pm8 significantly increased the agreement between automated and corrected reference namings (Table 2). Some discrepancies, however, persisted which were affected by the choice of the reference embryo from which the names were transferred. This is illustrated with the *de novo* automated naming of all 6 other reference embryos from the corrected Astec-Pm5 (Table 3). Figure S5 illustrates that as a result of the different cleavage orientations of the A7.2 cells in Astec-Pm8 and Astec-Pm5, using either of these embryos as a naming reference led to different namings of the A7.2 daughters of Astec-Pm3. Naming by transfer from a single embryo is thus sensitive to natural variability.

To increase the robustness of our automatic naming procedure, we extended it to name an embryo *E* by inheritance from a population of reference embryos \mathcal{R} by considering the average division-to-division distances between a division in *E* and its homologs in \mathcal{R} (see methods). Automated renaming of each Astec-Pm embryo *de novo* by transfer from the population of the other reference embryos showed higher level of accuracy compared to the naming from the corrected Astec-Pm8 (Table 2) or Astec-Pm5 (Table 3), achieving 99,9% of agreement between automated and corrected reference namings (Table 4), including divisions with variable orientations, such as that of A7.2 (Figure S5). Using a reference embryo population thus increases the robustness of our automatic naming procedure to natural variability.

Embryo ID	Corrected manual naming	Automated naming by transfer from the corrected Astec-Pm8		Differences between automated & corrected manual namings	
	# Names	# Names	# Unnamed cells	# Different names*	% Agreement
Astec-Pm1	732	732	0	10	98,6
Astec-Pm3	598	598	0	0	100
Astec-Pm4	509	509	0	2	99,6
Astec-Pm5	606	606	0	10	98,3
Astec-Pm7	389	389	0	0	100
Astec-Pm8	730	730	0	0	100
Astec-Pm9	518	518	0	2	99,6
All embryos except Astec-Pm8	3352	3352	0	24	99,3

Table 2: Comparison of the corrected manual and *de novo* automated naming of Astec-Pm embryos by transfer from the corrected Astec-Pm8. (#)="Number of". *: Includes both primary and induced differences (see glossary table S1).

Embryo ID	Corrected manual naming	Automated naming by transfer from the corrected Astec-Pm5		Differences between automated & corrected manual namings	
	# Names	# Names	# Unnamed cells ^{\$}	# Different names*	% Agreement
Astec-Pm1	732	620	112	22	96,5
Astec-Pm3	598	584	14	18	96,9
Astec-Pm4	509	509	0	4	99,2
Astec-Pm5	606	606	0	0	100
Astec-Pm7	389	389	0	4	99
Astec-Pm8	730	620	110	12	98,1
Astec-Pm9	518	516	2	14	97,3
All embryos except Astec-Pm5	3844	3844	238	74	97,9

Table 3: Comparison of the corrected manual and *de novo* automated naming of Astec-Pm embryos by transfer from the corrected Astec-Pm5. (#)="Number of". (\$)="Cells that are not named because they do not exist and are not named in the reference embryo, here Astec-Pm5. These unnamed cells are therefore not counted as differences between automated and manual naming. *: Includes both primary and induced differences (see glossary table S1).

Embryo ID	Corrected manual naming	Automated naming from the corrected reference population (except the tested one)		Differences between automated & corrected manual namings	
	# Names	# Names	# Unnamed cells	# Different names*	% Agreement
Astec-Pm1	732	732	0	2	99,7
Astec-Pm3	598	598	0	0	100
Astec-Pm4	509	509	0	0	100
Astec-Pm5	606	606	0	4	99,3
Astec-Pm7	389	389	0	0	100
Astec-Pm8	730	730	0	0	100
Astec-Pm9	518	518	0	0	100
All embryos	4082	4082	0	6	99,9

Table 4: Comparison of the manual corrected and automated *de novo* naming of Astec-Pm embryos by transfer from the corrected reference population (except the tested one). (#) = "Number of". Reference = all Astec-Pm embryos except the tested one. *: Includes both primary and induced differences (see glossary table S1).

***De novo* naming of a new embryo**

Finally, we applied our method to name *de novo* the 764 cells up to the 10th generation of a new embryo, Astec-Pm16, imaged between the 32-cell and mid-neurula stages, segmented with ASTEC and manually curated (see methods).

To qualitatively assess the accuracy of the automatic naming on this new embryo, we assigned a different colour to each pair of bilateral cells in the Astec-Pm8 reference embryo and applied this same colour map to Astec-Pm16. Comparison of the position of like-colored cells in the two embryos using MorphoNet (Leggio et al., 2019) suggested a correct overall assignment of cell names in Astec-Pm16 (Figure 5, A-B).

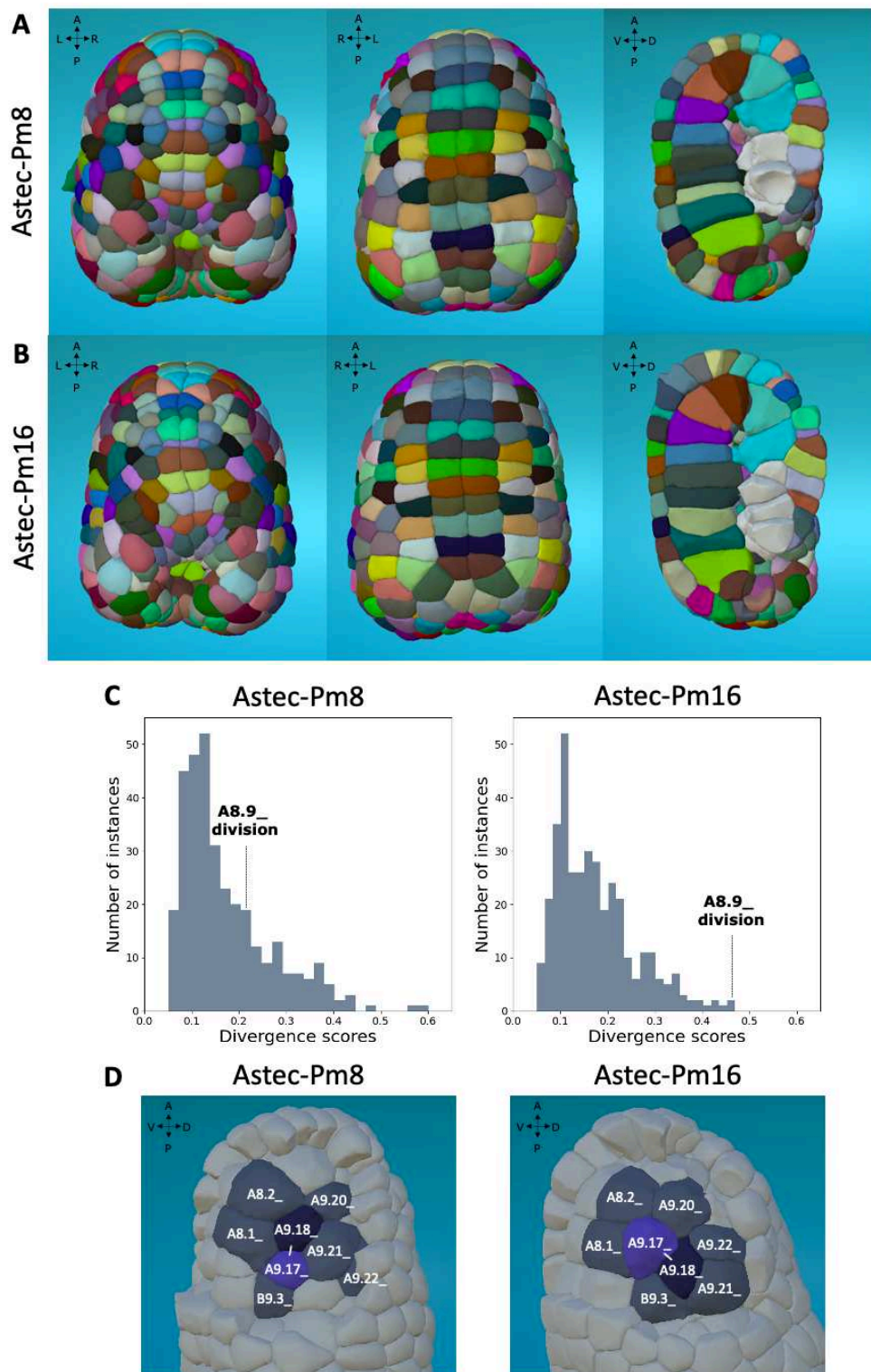


Figure 5: Assessment of the *de novo* naming of a new segmented embryo : Astec-Pm16. A-B) Each pair of symmetrically matching cells is colour coded based on the cell names, allowing a qualitative comparison between Astec-Pm8 (A) and Astec-Pm16 (B). Uncoloured cells correspond to cells of the 11th generation which is the generation from which the cells are no longer named. Left: dorsal views, middle : ventral views, right: sagittal views of late gastrula embryos (St. 13) visualised via MorphoNet. C) Histogram of the divergence scores for Astec-Pm8 and Astec-Pm16. Comparison of the two distributions was performed according to the Kolmogorov-Smirnov (K-S) test showing that the two distributions are identical (The null hypothesis is that the two distributions are identical, p-value=0.06). The position on the histograms of the A8.9_ division scores are presented for Astec-Pm8 (divergence score value=0.21) and Astec-Pm16 (divergence score

value=0.46) D) Illustration of the atypical division angle of the A8.9_ division in Astec-Pm16 compared to the other Astec-Pm, represented here by Astec-Pm8, showing a 90° shift in Astec-Pm16. A8.9_ daughter cells are in blue. White lines link A8.9_ daughter cells. Neighbouring cells visible on both embryos are grey. Cell names are in white. Antero-Posterior (AP) and Left-Right (LR) axes are represented.

We next evaluated quantitatively the quality of the automated *de novo* naming of Astec-Pm16, by reasoning that cells that have the most unusual pattern of cell neighbours are most likely to be misnamed. We thus used as a divergence score the average division-to-division distances between a cell division whose daughters' names are to be assessed and the closest two reference embryos. This score can be affected by two sources of variability: the orientation of the cell division considered and that of its immediate neighbours. To increase the sensitivity of the score to the orientation of the cell division of interest, we computed the D2D distance in a common neighbourhood referential in which neighbouring cells of the same generation as the cells to be assessed are virtually merged (Figure S1B). Figure 5D illustrates a representative case of a cell division whose divergence score is small in Astec-Pm8 and large in Astec-Pm16, as a result of different division orientations. This example illustrates that cell divisions with large divergence scores correspond to ambiguous cases where manual naming would have been challenging.

We then computed the divergence score for each division of Astec-Pm16 and compared it to that of Astec-Pm8, automatically renamed *de novo* by transfer from the population of the other reference embryos. The score distributions in the two embryos were not statistically different (Figure 5C) according to the Kolmogorov-Smirnov (K-S) test (see methods), suggesting that cell neighbourhoods in Astec-Pm16 are as canonical as those of Astec-Pm8, and thus that the qualities of the naming of Astec-Pm8 (Table 4) and Astec Pm16 are similar.

Taken together, the qualitative and quantitative assessments suggested an overall correct assignment of the cell names in Astec-Pm16 (Figure 5), supporting that our naming strategy is accurate for both ground-truth and novel datasets. We also used division-to-division distances to identify cells that have the most unusual pattern of cell neighbours in the new segmented embryo compared to our database of reference embryos. This highlighted that comparison of the precise cell neighbourhood measurements between *Phallusia* embryos can be used to study and quantify the natural variability of ascidian development.

DISCUSSION

The systematic identification of equivalent cells in separate embryos is a prerequisite for the integration of cohorts of individual embryos and multi-scale datasets into a common referential. In ascidians, the development shows such cellular invariance that each cell has homologs across all embryos and it is possible to assign to each cell classe a unique canonical name, traditionally according to a scheme proposed by Conklin (Figure 1) (Conklin, 1905). This task of cell identification was, however, largely reliant on time-consuming and error-prone manual expertise. Based on the historical Conklin naming syntax and the availability of a population of 7 fully segmented and manually named reference embryos of the ascidian *Phallusia mammillata* (Guignard et al., 2020), we developed a theoretical framework to transfer cell naming from this embryo collection towards new reconstructed embryos. We showed that contact areas between neighbouring cells are sufficiently invariant across *Phallusia* embryos to uniquely represent each cell of an embryo by the vector of its contact surfaces with first order cell neighbours and we implemented a division-to-division distance

measurements to automatically name the two daughter cells after each division. To ensure that our automatic naming procedure is robust to natural biological developmental variability, we established the entire collection as reference embryos, for which we first developed computational methods to systematically detect and correct residual manual naming errors. This corrected database of optimally named embryos can now be considered the reference to name with a Conklin syntax all embryonic *Phallusia* cells up to the 10th generation, including those that could not previously be assigned a Conklin name. We showed that our technique is highly accurate on both ground-truth and novel datasets and provided first evidence supporting that the comparison of the precise measurements of the cell neighbourhood between *Phallusia* embryos can be used to quantify the variability of the ascidian development. To conclude, by automating cell naming and connecting embryos into a common referential, this study establishes the highly reproducible, simple and transparent embryos of ascidians as a digital framework of choice to build, analyse and integrate multiscale functional 4D atlases with cellular resolution. Such multimodal and multiscale digital atlases of a chordate development will ultimately support a system-level modelling of development (Biasuz et al., 2018; Pastor-Escuredo and del Álamo, 2020).

The stereotypical development of ascidians provides a rigorous framework for systematic cell naming

Given the high reproducibility of ascidian embryonic development, already in the early 20th century Edwin G. Conklin proposed a scheme to systematically identify each cell of early ascidian embryos (Conklin, 1905). This scheme is composed of both a naming syntax that ensures the consistency of cell names and of decision rules for assigning a name to sister cells, based on their relative position with respect to the geodesic distance to the vegetal pole of the embryo or to the midline (Conklin, 1905). The use of these rules, even by a non-expert, should result in unambiguous naming. However, his proposal presents several limitations to name each cell of an ascidian embryo from the strict application of his rules. Indeed, while the naming syntax is largely applicable, the use of geodesic distance limits the applicability of Conklin's decision rules to stages where all cells have contact with the external environment, a rule that no longer applies beyond the onset of gastrulation. From the 76-cell stage onwards, the axis of some cell divisions is oriented perpendicular to the apical surface of the embryo, generating an internal cell whose geodesic distance to the vegetal pole is not defined. In addition, in some situations, the naming proposed by Conklin seems to deviate from the stated rules. This is the case, for example, for the daughter cells of a6.5 (Figure 1), none of which appear to be closer to the vegetative pole, and whose names should be permuted in the original naming established by Conklin for *Styela (Cynthia) partita* embryos (Conklin, 1905).

To overcome these difficulties in applying the Conklin rules to name each ascidian embryo *de novo*, we chose an alternative approach based on the transfer of the naming between distinct ascidian embryos. A first approach would have been to rely on the stereotypy of the ascidian embryonic development, which induces an identical relative position of homologous cells between embryos, to spatially register segmented images between two embryos and transfer the name of one reference embryo to another. This method is a popular approach used for automatic cell annotation in the nematode *C. elegans* (Long et al., 2009; Aerni et al., 2013; Toyoshima et al., 2019) and has already been developed to register 64-cell (St.8) and 112-cell (St.10) stages *Phallusia* embryos (Michelin et al., 2015). However, with this method, the difficulty is transferred to the registration method which

should be able to handle whole embryos that need to be recorded both temporally and spatially, which is computationally intensive and difficult to implement with multiple reference embryos. Here, we showed that the precise measurements of contact areas between neighbouring cells are sufficiently invariant across *Phallusia* embryos between homologous cells and sufficiently different between sister cells to distinguish them unambiguously. This allowed for the compact representation of each cell of a segmented *Phallusia* embryo by the vector of its contact surfaces with first order cell neighbours which, coupled with the use of the cell lineage information, is sufficient to decide the names to transfer for a lower cost, rendering our method to name cells extremely computationally efficient on large 3D + time datasets. Moreover, a local decision allows to manage naturally the division heterochronies.

Limitations

Naming new segmented *Phallusia mammillata* embryos by transfer from a population of already segmented and named embryos is subject to certain limitations. Firstly, it is important to ensure that the reference embryos are named consistently with each other. This consistency can be a posteriori assessed and ensured by dedicated tools. Secondly, cells of a test embryo can only be named if their homologues are named in the reference database. With our current Astec-Pm collection of embryos, the cell naming ends at the 10th generation included, a period which covers already a significant part of the whole embryonic development. Most of the tissues present in the larva are composed of cells that are of the next generation (11th) and only up to 13 generations are observed for the others (Dardaillon et al., 2020). Moreover, it includes cell names for all tissue fates specification events (Lemaire, 2009) and numerous cell types specification events, notably within the neural plate (reviewed in (Hudson, 2016)). Finally, the work we accomplished here to automate the naming of segmented *Phallusia* embryos has allowed us to come across interesting cases of biological variability. For example, we identified an atypical division orientation of A8.9_ in Astec-Pm16. We also noticed the bi-modal division of A7.2 for which we observed two clear orientations each represented in several reference embryos. This revealed that ascidian development is not strictly determined. However, in our current strategy, each couple of daughter cells is named by considering the average division-to-division distance measured between the new embryo and the reference dataset, meaning that an atypical division may be misnamed depending on the number of representations of this division in the reference database. To address this issue, it will be essential to increase the population of reference embryos and to distinguish between natural variation (small variation in the measurement of contact area between embryos) and variability (different orientation of divisions, e.g. A7.2) within the reference embryo population.

This raised the issue of whether these cells resulting from non-canonical or bi-modal divisions should be named? For atypical divisions that are symmetric, i.e. giving rise to two equipotential daughter cells with the same larval tissue fate, it can be assumed that these cells can be named automatically using the neighbourhood to ensure a reproducible and consistent naming between embryos. This applies, for example, for the bi-modal A7.2 division where the two daughter cells are known to be both already restricted to the anterior endodermal tissue (Shi and Levine, 2008)). For asymmetric divisions which, by contrast, produce two daughter cells differing in their larval tissue fates, their fate specification mechanisms may affect the decision to name them automatically or not. If cells autonomously segregate determinants depending on their division orientation, naming them using

their neighbours rather than their division orientation can be misleading. If however cells are induced by signals from their neighbours, our automated naming is more likely to respect inductive clues. However, this becomes more complicated for atypical or bi-modal divisions for which the fate specification mechanisms are not known. A first step would be to identify among them the symmetric divisions which can be distinguished from asymmetric divisions by comparing the lineage seeded by sister cells (Guignard et al., 2020). However, the choice of naming will always remain questionable for a number of divisions. For this reason, we propose to always couple the name of a cell with its divergence score in anatomical ontologies. This allows each cell name to be preserved while keeping track of the cells that have the most unusual cell neighbourhood pattern in each newly named embryo.

Future directions

The corrected database of optimally named embryos that we have established can now be considered as the reference for Conklin naming of *Phallusia* embryonic cells, including those that could not previously be assigned a Conklin name. However, in the current reference database cells are not named beyond 10th generation. Nevertheless, few cells in later generations are precisely identified according to the Conklin syntax and it would therefore be possible to extend the reference database and assign a name with a Conklin syntax to each cell from the 11th generation to the last.

Here, we have provided the first evidence supporting that the theoretical framework we developed to represent each cell of a segmented *Phallusia* embryo by the vector of its contact surfaces with first-order neighbouring cells can be used to quantify and study the variability of ascidian development. However, to go further in this direction, it will be necessary to increase the population size of segmented and named *Phallusia* embryos.

Because of the extreme level of reproducibility of the ascidian early embryonic development, only slight differences in the geometry of embryos of different species are observed (Lemaire, 2011). We therefore expect our automatic naming method to be broadly applicable to accurately name other solitary ascidian species directly from our collection of 7 optimally named *Phallusia* embryos. This will open the door to the study of variability between ascidian species, a good candidate being, for example, the fully transparent embryos of the ascidian *Ascidiella aspersa*.

Finally, given that the developmental characteristics of the nematodes which develop with invariant cell lineage and high reproducibility of relative cell position (Li et al., 2019) are very similar to those of ascidians, we also expect our approach to be transferable to these organisms. However, there is currently a lack of data from *C. elegans* embryos imaged with labelled membranes, whose imaging for cell identification is based mainly on nuclear labelling (Long et al., 2009; Aerni et al., 2013; Toyoshima et al., 2019; Chaudhary et al., 2021). Nevertheless, cell surface can be approximated using the Voronoï segmentation algorithm, a method applied to automatically infer cell-cell contact up to the ~350-cell *C. elegans* embryo (Chen et al., 2018) and it would be interesting to try to name this dataset with our method.

MATERIALS AND METHODS

Astec-Pm16 *Phallusia mammillata* segmented embryo

Phallusia mammillata was provided by the Centre National de Ressources Biologiques Marines in Roscoff (France). Embryo was obtained by *in vitro* fertilisation and dechorionated as in (Sardet et al., 2011). *Phallusia mammillata* unfertilized eggs were microinjected under a stereoscope, with approximately 45pg of PH-GFP membrane labelling mRNA construct as in (Guignard et al., 2020).

Embryo was deposited by gravity without embedding at the bottom of a 1% Low Melt Agarose (Carl Roth, 6351.1) cone-shaped well and imaged at a constant temperature of ~18°C in artificial seawater using a multiview lightsheet microscope (MuViSPIM, Luxendo). PH-GFP was excited with a 488 nm laser (LuxX 488-60, Omicron) with simultaneous two-sided illumination. The emitted light was collected by two opposing detection arms positioned perpendicularly to the illumination plane, resulting in simultaneous dual acquisition. Each detection arm was equipped with a 20× Olympus water dipping objective (NA 1.0) combined with a tube lens leading to a 33.3 fold image magnification. The filter used was the band-path BrightLine 525/30 filter (Semrock). Filtered emitted light was collected by Hamamatsu V2 Flash 4 SCMOS cameras. An electronic confocal slit detection (eCSD) mode was used during image acquisition to minimise the capture of scattered light and improve contrast. The electronic slit size was matched in accordance to the diameter of the illumination beams (60 pixels). At each time point, two acquisitions were sequentially performed, the second one orthogonal to the first, generating four whole-embryo stacks, i.e. two sets of two matching 3D intensity image stacks (0, 90, 180, 270 degrees) each covering the whole imaged embryo. The embryo was imaged every two minutes with an image voxel size of $0.195 \times 0.195 \times 1 \mu\text{m}^3$. The intensity values at each voxel were encoded in 16 bits.

The four images from each camera at a given time point were then fused to build a 3D isotropic image of the embryo using an updated version of the ASTEC suite described in (Guignard et al., 2020) and available on the INRIA gitlab (<https://gitlab.inria.fr/astec/astec>)

All membrane-labelled embryonic cells of the developing Astec-Pm16 embryo were then segmented at each time point using the updated ASTEC pipeline (Guignard et al., 2020), producing a reconstructed version of the embryo integrating cell lineages and quantitative geometrical properties including cell contact surfaces, stored in an .xml file.

Segmentation errors (over- and under-segmentations or missing cells) were manually corrected using the MorphoNet 'MorphoPlot' 3D curation package (<https://morphonet.org/helpfiles/API/morphonet.plot.html>).

Initial naming at the 64-cell stage of Astec-Pm16 was performed by associating manually a name to each cell snapshot present at the 28th time point of the acquisition based on its position in the embryo and annotator knowledge. This step was performed by editing the .xml properties file of Astec-Pm16 using the sublime-text file editor and adding a "cell_name" key containing the association between each cell label to its name (e.g. `<cell cell-id="280001">'a7.0001_'</cell>`). For more information on the properties file, see the astec documentation (https://astec.gitlabpages.inria.fr/astec/astec_embryoproperties.html).

Collection of data from the literature

The Astec-Pm1/3/4/5/7/8 & 9 *Phallusia mammillata* segmented embryos (MuViSPIM fused images, segmentation files, cell lineage and cell informations including cell names) (Guignard et al., 2020) were downloaded from FigShare (<https://figshare.com/s/765d4361d1b073beedd5>) and explored via MorphoNet (<http://www.morphonet.org>). Among the 10 Astec-Pm embryos available, we excluded Astec-Pm2 because it partially exited the field of view during imaging leading to segmentation errors, and ASTEC-Pm6 and -Pm10 because of the presence of uncanonical cell divisions prior to the 64-cell stage.

Detection of manual errors affecting the Conklin syntax in the reference dataset

The “astec_contact_atlas” script (for more informations, see https://astec.gitlabpages.inria.fr/astec/astec_contact_atlas.html), with the “diagnosis_properties” parameter sets to “true”, automatically parses the cell lineage and cell name information of the .xml properties file for each input .xml “atlasFiles” to check whether the cell names agrees with the Conklin syntax. This diagnosis script reports:

- Cases of inaccurate propagation of cell names through the life of a cell:
 - cell snapshots that are not named despite having a named non-dividing cell snapshot predecessor;
 - cell snapshots that are named despite having an unnamed cell snapshot predecessor;
 - cell snapshots that are named differently from their non-dividing cell snapshot predecessor;
- Cases of inaccurate propagation of cell names across a cell division:
 - cell snapshots just after a division whose name does not follow the Conklin syntax ($Fp.q \rightarrow [F(p+1).2q ; F(p+1).(2q-1)]$);
 - cell snapshots just after a division with a name compatible with the Conklin syntax, but with an unnamed sister cell;
 - cell snapshots just after a division with a name compatible with the Conklin syntax, but with a sister cell that has the same name;
 - cell snapshots just after a division with a name compatible with the Conklin syntax, but with a sister cell name that is incompatible with the Conklin syntax.

This diagnosis was launched on each Astec-Pm reference embryo. All detected syntax errors were corrected by manually editing the relevant .xml properties file with the sublime-text file editor. Repeating this diagnostic procedure after each round of correction ensured that no indirect syntax errors were introduced

Detection and correction of manually permuted sister cell names in the reference dataset

We used the Permutation Proposal Score (PPS) to identify sister cells whose name was likely to be permuted in the original manual naming.

As sister name permutations can affect the calculation of the PPS on neighbouring sister cells and also affect the names of cells of later generations, we corrected the permuted names in the reference dataset iteratively, generation by generation, from the 7th generation to the 10th. Starting

from 7th generation cells, we first computed the PPS for each Astec-Pm reference embryo using the “astec_contact_atlas script” (for more informations, see https://astec.gitlabpages.inria.fr/astec/astec_contact_atlas.html). We then used Morphonet to manually inspect cells with negative PPS values in each segmented embryo. When appropriate, sister cell names were permuted manually by editing in the .xml properties files of each Astec-Pm embryo using the sublime-text file editor. The names of the daughter cells of renamed cells, were first manually deleted and then automatically renamed according to the other reference embryos using the “astec_contact_naming” script (https://astec.gitlabpages.inria.fr/astec/astec_contact_naming.html). After all 7th generation cells were corrected, the PPS was recomputed and the following generations were iteratively analysed in the same way.

Automated cell naming of an embryo E by comparison with a population of reference embryo \mathcal{R}

To name an embryo E by inheritance from a population of reference embryos \mathcal{R} , we considered the average division-to-division distances (D2D) between a division in E and its homologs in \mathcal{R} . In this context, the automated naming process compares $\sum_{R \in \mathcal{R}, R \neq E} \text{D2D}(\text{DIV}(C_i^E), \text{DIV}(C_i^R))$ and $\sum_{R \in \mathcal{R}, R \neq E} \text{D2D}(\text{DIV}^P(C_i^E), \text{DIV}(C_i^R))$ to name $d_0(C_i^E)$ and $d_1(C_i^E)$:

if $\sum_{R \in \mathcal{R}, R \neq E} \text{D2D}(\text{DIV}(C_i^E), \text{DIV}(C_i^R)) < \sum_{R \in \mathcal{R}, R \neq E} \text{D2D}(\text{DIV}^P(C_i^E), \text{DIV}(C_i^R))$, $d_0(C_i^E)$ inherits the biological name of $d_0(C_i^R)$ and $d_1(C_i^E)$ that of $d_1(C_i^R)$. Otherwise, $d_1(C_i^E)$ inherits the name of $d_0(C_i^R)$ and $d_0(C_i^E)$ that of $d_1(C_i^R)$.

To increase the number of reference cell divisions taken into account, each hemi-embryos (left or right) were considered as distinct reference embryos.

Statistical analysis

The two-sample Kolmogorov-Smirnov (K-S) statistical analysis was performed using the `scipy.stats.ks_2samp` function of the `scipy.stats` module of the python `scipy` library (https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ks_2samp.html#scipy.stats.ks_2samp). This test compares the underlying continuous distributions $F(x)$ and $G(x)$ of two independent samples. The null hypothesis is that the two distributions are identical, $F(x)=G(x)$ for all x . If the KS statistic is small or the p-value is high (>0.05), then we cannot reject the null hypothesis.

Data and code availability

Software scripts for reproducing the analyses and to perform automated *de novo* naming are available via the gitlab repository (<https://gitlab.inria.fr/astec/astec>). A documentation can be found at astec.gitlabpages.inria.fr/astec/ : see the “astec_contact_atlas” section for more details on the scripts used to assess the quality/consistency of the set of the 7 Astec-Pm embryos already named and see the “astec_contact_naming” section for more informations on the scripts used to name an embryo and to assess its quality from the divergence score. The corrected cell namings of Astec-Pm1/3/4/5/7/8 & 9 and the corrected fate maps are deposited and available from FigShare (XX) and have been uploaded in MorphoNet (new names and fates have a “V2” suffix in the infos

menu, e.g. “Cell Name V2”). A tutorial can be found at gitlab.inria.fr/astec/astec-tutorial. The fused images, segmentations and embryo properties (including cell lineage, cell names and tissue fates) of Astec-Pm16 can also be downloaded from XXXX. Its segmentations and information can be interactively explored via MorphoNet (<https://morphonet.org/RB9nb4LZ>).

REFERENCES

- Aerni, S.J., Liu, X., Do, C.B., Gross, S.S., Nguyen, A., Guo, S.D., Long, F., Peng, H., Kim, S.S., Batzoglou, S., 2013. Automated cellular annotation for high-resolution images of adult *Caenorhabditis elegans*. *Bioinformatics* 29, i18–i26. <https://doi.org/10.1093/bioinformatics/btt223>
- Armit, C., Richardson, L., Venkataraman, S., Graham, L., Burton, N., Hill, B., Yang, Y., Baldock, R.A., 2017. eMouseAtlas: An atlas-based resource for understanding mammalian embryogenesis. *Dev. Biol.* 423, 1–11. <https://doi.org/10.1016/j.ydbio.2017.01.023>
- Biasuz, K., Leggio, B., Faure, E., Lemaire, P., 2018. The “computable egg”: Myth or useful concept? *Curr. Opin. Syst. Biol.* 11, 91–97. <https://doi.org/10.1016/j.coisb.2018.09.003>
- Briggs, J.A., Weinreb, C., Wagner, D.E., Megason, S., Peshkin, L., Kirschner, M.W., Klein, A.M., 2018. The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* 360, eaar5780. <https://doi.org/10.1126/science.aar5780>
- Budnik, B., Levy, E., Harmange, G., Slavov, N., 2018. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* 19, 161. <https://doi.org/10.1186/s13059-018-1547-5>
- Cao, C., Lemaire, L.A., Wang, W., Yoon, P.H., Choi, Y.A., Parsons, L.R., Matese, J.C., Wang, W., Levine, M., Chen, K., 2019. Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* 571, 349–354. <https://doi.org/10.1038/s41586-019-1385-y>
- Castro-González, C., Ledesma-Carbayo, M.J., Peyriéras, N., Santos, A., 2012. Assembling models of embryo development: Image analysis and the construction of digital atlases. *Birth Defects Res. Part C Embryo Today Rev.* 96, 109–120. <https://doi.org/10.1002/bdrc.21012>
- Chaudhary, S., Lee, S.A., Li, Y., Patel, D.S., Lu, H., 2021. Graphical-model framework for automated annotation of cell identities in dense cellular images. *eLife* 10, e60321. <https://doi.org/10.7554/eLife.60321>
- Chen, L., Ho, V.W.S., Wong, M.-K., Huang, X., Chan, L., Ng, H.C.K., Ren, X., Yan, H., Zhao, Z., 2018. Establishment of Signaling Interactions with Cellular Resolution for Every Cell Cycle of Embryogenesis. *Genetics* 209, 37–49. <https://doi.org/10.1534/genetics.118.300820>
- Cole, A.G., Meinertzhagen, I.A., 2004. The central nervous system of the ascidian larva: mitotic history of cells forming the neural tube in late embryonic *Ciona intestinalis*. *Dev. Biol.* 271, 239–262. <https://doi.org/10.1016/j.ydbio.2004.04.001>
- Conklin, E.G., 1905. Mosaic development in ascidian eggs. *J. Exp. Zool.* 2, 145–223. <https://doi.org/10.1002/jez.1400020202>
- Dardaillon, J., Dauga, D., Simion, P., Faure, E., Onuma, T.A., DeBiasse, M.B., Louis, A., Nitta, K.R., Naville, M., Besnardeau, L., Reeves, W., Wang, K., Fagotto, M., Guérault-Bellone, M., Fujiwara, S., Dumollard, R., Veeman, M., Volff, J.-N., Roest Crollius, H., Douzery, E., Ryan, J.F., Davidson, B., Nishida, H., Dantec, C., Lemaire, P., 2020. ANISEED 2019: 4D exploration of genetic data for an extended range of tunicates. *Nucleic Acids Res.* 48, D668–D675. <https://doi.org/10.1093/nar/gkz955>
- Di Gregorio, A., Hadjantonakis, A.-K., 2006. The multidimensionality of cell behaviors underlying morphogenesis: a case study in ascidians. *BioEssays* 28, 874–879. <https://doi.org/10.1002/bies.20453>
- Duncan, L.H., Moyle, M.W., Shao, L., Sengupta, T., Ikegami, R., Kumar, A., Guo, M., Christensen, R., Santella, A., Bao, Z., Shroff, H., Mohler, W., Colón-Ramos, D.A., 2019. Isotropic Light-Sheet Microscopy and Automated Cell Lineage Analyses to Catalogue *Caenorhabditis elegans* Embryogenesis with Subcellular Resolution. *J. Vis. Exp. JoVE* 10.3791/59533.

- <https://doi.org/10.3791/59533>
- Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., Schier, A.F., 2018. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* 360, eaar3131. <https://doi.org/10.1126/science.aar3131>
- Guignard, L., Fiúza, U.-M., Leggio, B., Laussu, J., Faure, E., Michelin, G., Biasuz, K., Hufnagel, L., Malandain, G., Godin, C., Lemaire, P., 2020. Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis. *Science* 369. <https://doi.org/10.1126/science.aar5663>
- Hudson, C., 2016. The central nervous system of ascidian larvae. *WIREs Dev. Biol.* 5, 538–561. <https://doi.org/10.1002/wdev.239>
- Karaiskos, N., Wahle, P., Alles, J., Boltengagen, A., Ayoub, S., Kipar, C., Kocks, C., Rajewsky, N., Zinzen, R.P., 2017. The *Drosophila* embryo at single-cell transcriptome resolution. *Science* 358, 194–199. <https://doi.org/10.1126/science.aan3235>
- Khairy, K., Keller, P.J., 2011. Reconstructing embryonic development. *genesis* 49, 488–513. <https://doi.org/10.1002/dvg.20698>
- Knowles, D.W., Biggin, M.D., 2013. Building quantitative, three-dimensional atlases of gene expression and morphology at cellular resolution. *Wiley Interdiscip. Rev. Dev. Biol.* 2, 767–779. <https://doi.org/10.1002/wdev.107>
- Leggio, B., Laussu, J., Carlier, A., Godin, C., Lemaire, P., Faure, E., 2019. MorphoNet: an interactive online morphological browser to explore complex multi-scale data. *Nat. Commun.* 10, 2812. <https://doi.org/10.1038/s41467-019-10668-1>
- Lemaire, P., 2011. Evolutionary crossroads in developmental biology: the tunicates. *Development* 138, 2143–2152. <https://doi.org/10.1242/dev.048975>
- Lemaire, P., 2009. Unfolding a chordate developmental program, one cell at a time: Invariant cell lineages, short-range inductions and evolutionary plasticity in ascidians. *Dev. Biol.* 332, 48–60. <https://doi.org/10.1016/j.ydbio.2009.05.540>
- Li, X., Zhao, Z., Xu, W., Fan, R., Xiao, L., Ma, X., Du, Z., 2019. Systems Properties and Spatiotemporal Regulation of Cell Position Variability during Embryogenesis. *Cell Rep.* 26, 313–321.e7. <https://doi.org/10.1016/j.celrep.2018.12.052>
- Long, F., Peng, H., Liu, X., Kim, S.K., Myers, E., 2009. A 3D digital atlas of *C. elegans* and its application to single-cell analyses. *Nat. Methods* 6, 667–672. <https://doi.org/10.1038/nmeth.1366>
- Ma, X., Zhao, Z., Xiao, L., Xu, W., Kou, Y., Zhang, Y., Wu, G., Wang, Y., Du, Z., 2021. A 4D single-cell protein atlas of transcription factors delineates spatiotemporal patterning during embryogenesis. *Nat. Methods* 18, 893–902. <https://doi.org/10.1038/s41592-021-01216-1>
- McDole, K., Guignard, L., Amat, F., Berger, A., Malandain, G., Royer, L.A., Turaga, S.C., Branson, K., Keller, P.J., 2018. In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level. *Cell* 175, 859–876.e33. <https://doi.org/10.1016/j.cell.2018.09.031>
- Michelin, G., Guignard, L., Fiuza, U.-M., Lemaire, P., Godine, C., Malandain, G., 2015. Cell pairings for ascidian embryo registration, in: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). Presented at the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI 2015), IEEE, Brooklyn, NY, USA, pp. 298–301. <https://doi.org/10.1109/ISBI.2015.7163872>
- Nicol, D., Meinertzhagen, I.A., 1988. Development of the Central Nervous System of the Larva of the Ascidian, *Ciona intestinalis* L. 30.
- Oates, A.C., Gorfinkiel, N., González-Gaitán, M., Heisenberg, C.-P., 2009. Quantitative approaches in developmental biology. *Nat. Rev. Genet.* 10, 517–530. <https://doi.org/10.1038/nrg2548>
- Packer, J.S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E., Dueck, H., Stefanik, D., Tan, K., Trapnell, C., Kim, J., Waterston, R.H., Murray, J.I., 2019. A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science* 365, eaax1971. <https://doi.org/10.1126/science.aax1971>

- Pasini, A., Amiel, A., Rothbacher, U., Roure, A., Lemaire, P., Darras, S., 2006. Formation of the Ascidian Epidermal Sensory Neurons: Insights into the Origin of the Chordate Peripheral Nervous System. *PLOS Biol.* 4, e225. <https://doi.org/10.1371/journal.pbio.0040225>
- Pastor-Escuredo, D., del Álamo, J.C., 2020. How Computation Is Helping Unravel the Dynamics of Morphogenesis. *Front. Phys.* 8, 31. <https://doi.org/10.3389/fphy.2020.00031>
- Razy-Krajka, F., Lam, K., Wang, W., Stolfi, A., Joly, M., Bonneau, R., Christiaen, L., 2014. Collier/OLF/EBF-Dependent Transcriptional Dynamics Control Pharyngeal Muscle Specification from Primed Cardiopharyngeal Progenitors. *Dev. Cell* 29, 263–276. <https://doi.org/10.1016/j.devcel.2014.04.001>
- Sardet, C., McDougall, A., Yasuo, H., Chenevert, J., Pruliere, G., Dumollard, R., Hudson, C., Hebras, C., Le Nguyen, N., Paix, A., 2011. Embryological Methods in Ascidians: The Villefranche-sur-Mer Protocols, in: Pelegri, F.J. (Ed.), *Vertebrate Embryogenesis, Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp. 365–400. https://doi.org/10.1007/978-1-61779-210-6_14
- Shi, W., Levine, M., 2008. Ephrin signaling establishes asymmetric cell fates in an endomesoderm lineage of the *Ciona* embryo. *Development* 135, 931–940. <https://doi.org/10.1242/dev.011940>
- Sladitschek, H.L., Fiuza, U.-M., Pavlinic, D., Benes, V., Hufnagel, L., Neveu, P.A., 2020. MorphoSeq: Full Single-Cell Transcriptome Dynamics Up to Gastrulation in a Chordate. *Cell* 181, 922–935.e21. <https://doi.org/10.1016/j.cell.2020.03.055>
- Stolfi, A., Gainous, T.B., Young, J.J., Mori, A., Levine, M., Christiaen, L., 2010. Early Chordate Origins of the Vertebrate Second Heart Field. *Science* 329, 565–568. <https://doi.org/10.1126/science.1190181>
- Sulston, J.E., Horvitz, H.R., 1977. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* 56, 110–156. [https://doi.org/10.1016/0012-1606\(77\)90158-0](https://doi.org/10.1016/0012-1606(77)90158-0)
- Tomer, R., Denes, A.S., Tessmar-Raible, K., Arendt, D., 2010. Profiling by Image Registration Reveals Common Origin of Annelid Mushroom Bodies and Vertebrate Pallium. *Cell* 142, 800–809. <https://doi.org/10.1016/j.cell.2010.07.043>
- Toyoshima, Y., Wu, S., Kanamori, M., Sato, H., Jang, M.S., Oe, S., Murakami, Y., Teramoto, T., Park, C., Iwasaki, Y., Ishihara, T., Yoshida, R., Iino, Y., 2019. An annotation dataset facilitates automatic annotation of whole-brain activity imaging of *C. elegans* (preprint). *Neuroscience*. <https://doi.org/10.1101/698241>
- Wagner, D.E., Weinreb, C., Collins, Z.M., Briggs, J.A., Megason, S.G., Klein, A.M., 2018. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* 360, 981–987. <https://doi.org/10.1126/science.aar4362>
- Wolpert, L., 1994. Do We Understand Development? *Science* 266, 571–572. <https://doi.org/10.1126/science.7939707>
- Yasuo, H., Hudson, C., 2007. FGF8/17/18 functions together with FGF9/16/20 during formation of the notochord in *Ciona* embryos. *Dev. Biol.* 302, 92–103. <https://doi.org/10.1016/j.ydbio.2006.08.075>
- Zand, M., Ruan, J., 2021. Spatial mapping of single cells in the *Drosophila* embryo from transcriptomic data based on topological consistency. *F1000Research* 9, 1014. <https://doi.org/10.12688/f1000research.24163.2>

ACKNOWLEDGEMENTS

We thank the members of the Tunicate team at CRBM for fruitful discussions throughout the course of this work and more particularly Benjamin Gallean for contribution for the visualisation of the data on MorphoNet and his help for statistical analysis. We gratefully acknowledge Marc Plays' and Philippe Richard's expert animal husbandry. P.L. and G.M. were CNRS and INRIA staff scientists, respectively. K.B. was funded by an Epigenmed (ANR-10-LABX-12-01) doctoral contract and by a University of Montpellier contract extension in compensation for Covid-19 restrictions and lab

closure . This work was supported by the Dig-Em (ANR-14-CE11-0013-01), Cell-whisper (ANR-19-CE13-0020) and Morphoscope (ANR-11EQPX-0029) grants of the Agence Nationale de la recherche (ANR). We acknowledge the support and help of the Montpellier Resource Imaging (MRI) facility and in particular of Sylvain de Rossi. MRI is a core facility of the France-BioImaging infrastructure, funded by the ANR (ANR-10-INBS-04).

SUPPLEMENTAL INFORMATION

Terms	Definitions
First order neighbourhood	Neighbouring cells with which a cell has direct physical contacts.
Cell snapshot	Thanks to the high temporal resolution of lightsheet imaging, each cell C_i^E of an embryo E is imaged at several time points. A cell snapshot corresponds to C_i^E at a given time point. All snapshots of a cell are linked by cell lineage relationships.
Matching cells	Cells with the same Fp.q names within an Astec-Pm embryo (bilateral cells) or across Astec-Pm embryos (homologous cells).
Cell generation	A cell generation includes all cells that are produced after the same number of divisions since the egg.
Primary and induced cell name differences	A primary cell name difference corresponds to the first time that a name difference is observed in a lineage. According to Conklin's syntax, the name of a cell determines the name of its descendants, so the induced differences correspond to the observed name differences resulting from a primary cell name difference.

Table S1 : Glossary table

Embryo ID	# Misnamed cells due to manual errors in the naming syntax*	# PPS	# PPS Accepted	# Renamed cells following PPS analysis*
Astec-Pm1	9	10	10	24
Astec-Pm3	2	9	9	22
Astec-Pm4	2	3	3	6
Astec-Pm5	2	17	17	44
Astec-Pm7	0	1	1	2
Astec-Pm8	19	7	7	38
Astec-Pm9	2	7	7	14
Total	36	54	54	150

Table S2: Summary of the manual errors corrected using the Permutation Proposal score.

(#)="Number of". *: Includes both primary and induced errors (see glossary table S1).

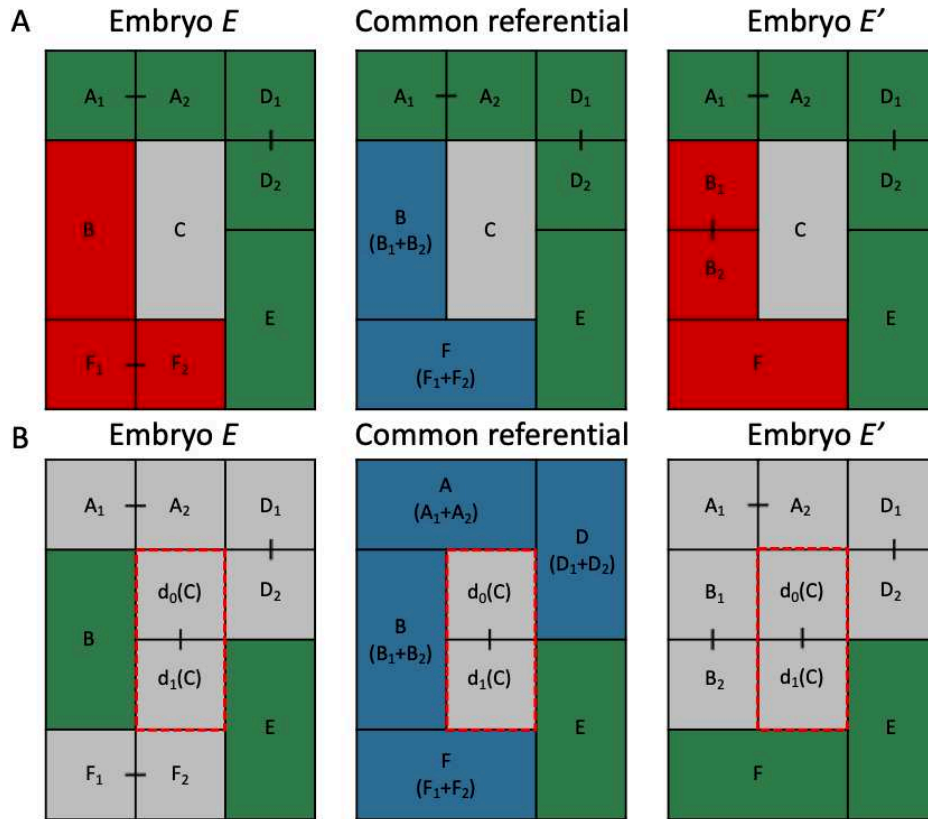


Figure S1: Taking into account division heterochronies to compute local neighbourhoods. A) Illustration showing the neighbourhoods of a cell C (grey) between two embryos E and E' . Sister cells present in only one of the two embryos (red) are virtually merged, to define a common referential (blue). Sister cells present in the two embryos are in green. Black lines link sister cells. B) Illustration showing the neighbourhoods of the two daughter cells of a cell C ($d_0(C), d_1(C)$, surrounded by red dotted lines) between two embryos E and E' . Cells of the same generation of $d_0(C), d_1(C)$ are in grey and cells of the previous generation belonging to the generation of their mother cell C are in green. To increase the sensitivity of the divergence score to the orientation of the cell division $DIV(C)=d_0(C), d_1(C)$, we computed the D2D distance in a common neighbourhood referential in which neighbouring cells of the same generation of $d_0(C), d_1(C)$ (in grey) are virtually merged (in blue)

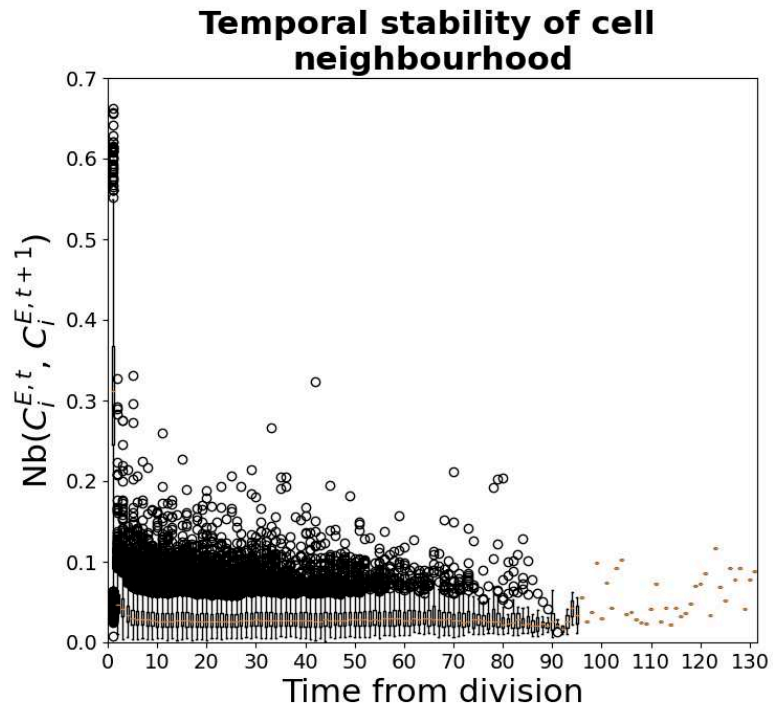


Figure S2: temporal stability of cell neighbourhoods. For each cell, the stability of its neighbourhood was determined by calculating the neighbourhood distance between two successive time points throughout the entire lifetime of the cell. Top, middle, and bottom lines in the box plot indicate 75th percentile, median, and 25th percentile of data, respectively.

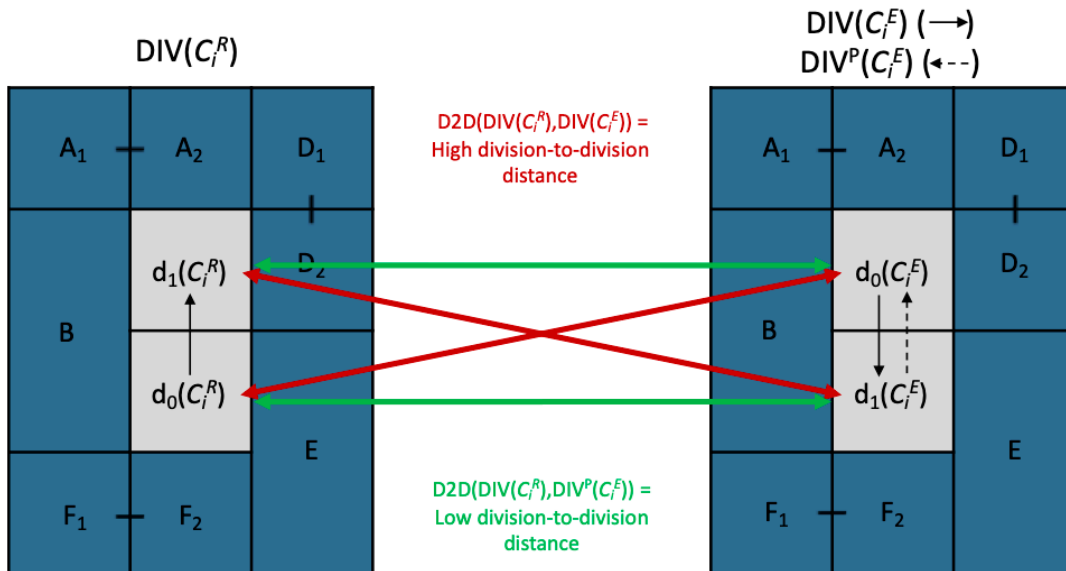


Figure S3 : Automated cell naming of C_i^E daughter cells by transfer from a reference embryo R using comparison of division-to-division distances $D2D(DIV(C_i^R), DIV(C_i^E))$ and $D2D(DIV(C_i^R), DIV^P(C_i^E))$. The division of cell C_i^R in a reference embryo (left panel) is represented by $DIV(C_i^R) = (d_0(C_i^R), d_1(C_i^R))$. The same division of C_i^E in an E embryo to be named (right panel) can be represented by $DIV(C_i^E) = (d_0(C_i^E), d_1(C_i^E))$ (black arrow) or by permutation of the daughter cells $DIV^P(C_i^E) = (d_1(C_i^E), d_0(C_i^E))$ (dotted arrow). For both divisions, d_0 is by construction the daughter cell with the lowest cell label in the segmented image at T_4 after the division. The divisions are placed in their common referential of neighbouring cells (dark blue cells). Black lines link sister cells. In this example, $D2D(DIV^P(C_i^E), DIV(C_i^R)) < D2D(DIV(C_i^E), DIV(C_i^R))$ and therefore $d_0(C_i^E)$ inherits its biological name from $d_1(C_i^R)$ and $d_1(C_i^E)$ from $d_0(C_i^R)$.

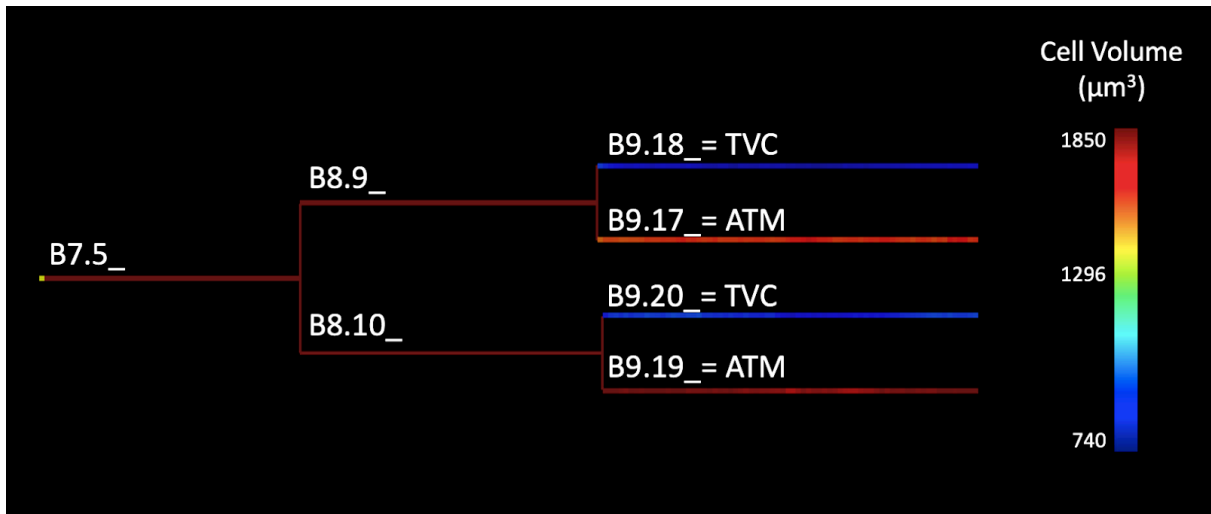


Figure S4 : Conklin name assignment to the cardiopharyngeal lineages. Schematic lineage tree of the Astec-Pm1 B7.5_ cell with colour code on branches indicating the cell volumes (scale bar on the right). The unequal divisions of B8.9_ in B9.18_ (smaller) and B9.17_ (bigger) and of B8.10_ in B9.20_ (smaller) and B9.19_ (bigger) allowed us to assign a name to the B8.9 and B8.10 daughter cells which were previously called based on differential fate specification of the cardiopharyngeal progenitors (called Trunk Ventral Cells (TVC)) versus the Anterior Tail Muscles (ATM) (Razy-Krajka et al., 2014).

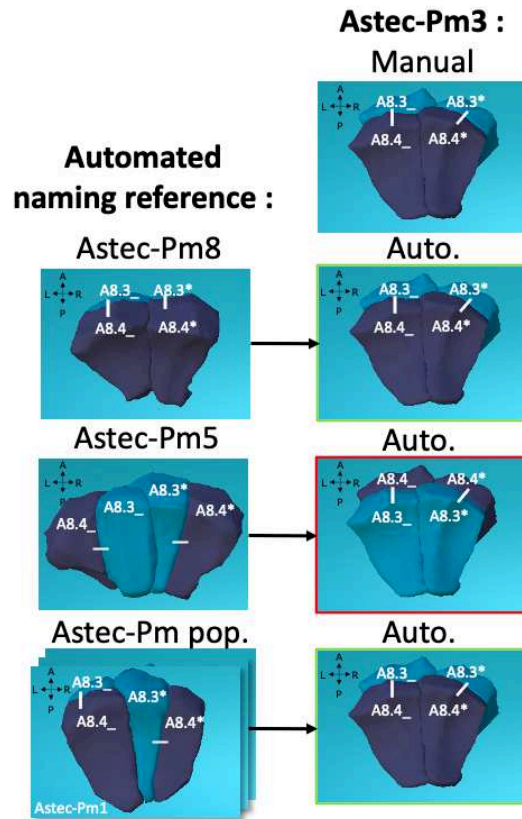


Figure S5 : Effect of the choice of the reference on the automated naming output. A7.2 daughters of Astec-Pm3 (top right, manual corrected naming) are named differently when Astec-Pm8 or Astec-Pm5 are used as reference. The result of the automated naming from the Astec-Pm population of reference embryos (except Astec-Pm3) is also shown on the bottom. Cell names are displayed and cells with the same name are colour coded. Green and red boxes illustrate either a correct (green) or a wrong (red) naming of the A7.2 daughters of Astec-Pm3 from the indicated reference. Vegetal views, Antero-Posterior (AP) and Left-Right (LR) axes are represented.

Discussion

L'identification systématique de cellules équivalentes dans des embryons distincts est une condition préalable à l'intégration de cohortes d'embryons individuels et d'ensembles de données multi-échelles dans un référentiel commun. Chez les ascidies, le développement présente une telle invariance cellulaire que chaque cellule a des homologues dans tous les embryons et qu'il est possible d'attribuer à chaque classe cellulaire un nom canonique unique, traditionnellement selon un schéma proposé par Conklin (Figure 1) (Conklin, 1905). Cette tâche d'identification des cellules était cependant largement tributaire d'une expertise manuelle longue et sujette aux erreurs.

Sur la base de la syntaxe historique de Conklin et de la disponibilité d'une population de 7 embryons de référence entièrement segmentés et nommés manuellement de l'ascidie *Phallusia mammillata* (Guignard et al., 2020), nous avons développé un cadre théorique pour transférer le nommage des cellules de cette collection d'embryons vers de nouveaux embryons reconstruits. Nous avons montré que les surfaces de contact entre cellules voisines sont suffisamment invariantes à travers les embryons de *Phallusia* pour représenter de façon unique chaque cellule d'un embryon par le vecteur de ses surfaces de contact avec les cellules voisines de premier ordre et nous avons implémenté une mesure de distance division-division pour nommer automatiquement les deux cellules filles après chaque division.

Pour s'assurer que notre procédure de nommage automatique est robuste à la variabilité naturelle du développement biologique, nous avons établi la collection entière comme embryons de référence, pour lesquels nous avons d'abord développé des méthodes computationnelles pour détecter et corriger systématiquement les erreurs résiduelles issues du nommage manuel.

Cette base de données corrigée d'embryons nommés de façon optimale peut maintenant être considérée comme la référence pour nommer avec une syntaxe de Conklin toutes les cellules des embryons de *Phallusia* jusqu'à la 10^{ème} génération, y compris celles qui n'ont pas pu recevoir de nom de Conklin auparavant (Razy-Krajka et al., 2018).

Nous avons montré que notre technique est très précise à la fois sur les données de référence et sur de nouveaux ensembles de données et nous avons fourni les premières preuves que la comparaison des mesures précises du voisinage des cellules entre les embryons de *Phallusia* peut être utilisée pour quantifier la variabilité du développement de l'ascidie.

Pour conclure, en automatisant l'identification et le nommage des cellules et en connectant les embryons dans un référentiel commun, cette étude établit les embryons stéréotypés d'ascidies, simples et transparents, comme un cadre numérique de choix pour construire et intégrer des atlas 4D fonctionnels multi-échelles avec une résolution cellulaire.

CHAPITRE 3 : ANALYSE QUANTITATIVE ET SYSTÉMATIQUE DE LA DYNAMIQUE SPATIO-TEMPORELLE DE L'ACTIVITÉ DE LA SIGNALISATION ERK AU COURS DU DÉVELOPPEMENT PRÉCOCE DE *PHALLUSIA MAMMILLATA*

Introduction

Après avoir établi la méthodologie expérimentale permettant de mesurer l'activité de ERK dans les embryons de *Phallusia mammillata*, grâce aux mesures manuelles du rapporteur ERK-KTR, et participé au développement du cadre théorique et des outils algorithmiques permettant de nommer automatiquement chaque cellule des embryons reconstruits, je me suis concentré sur le développement des méthodologies permettant de mesurer automatiquement et systématiquement l'activité de ERK dans toutes les cellules d'embryons qui se développent.

Je me suis tout d'abord demandé si le ERK-KTR (Regot et al., 2014) était le meilleur rapporteur pour suivre l'activité de ERK. Pour cela, j'ai comparé les mesures manuelles de l'activité de ERK obtenues à partir du ERK-KTR ou obtenues à partir du SKARS (Durandau et al., 2015) dans un embryon au stade 64-cellules. En effet, tout comme le ERK-KTR, le SKARS, qui a été initialement désigné pour suivre l'activité de ERK chez la levure (Fus3 et Kss1), subit un export nucléaire en réponse à l'activation de ERK. Cependant, contrairement au ERK-KTR qui dispose de séquences d'export et d'import nucléaire, le SKARS possède uniquement une séquence d'export nucléaire, suggérant un comportement potentiellement différent de celui du ERK-KTR en réponse à l'activité de ERK. La comparaison des mesures de ces deux outils ayant montré un comportement comparable des deux rapporteurs, aucun ne présentant une meilleure gamme dynamique de réponse, j'ai continué avec le ERK-KTR.

Je me suis ensuite concentré sur la mise en place de l'approche expérimentale permettant d'imager le ERK-KTR, les membranes et les noyaux sur des canaux séparés. Alors que l'imagerie de protéines marquées avec des protéines fluorescentes vertes et rouges est couramment utilisée chez l'ascidie, pour pouvoir imaginer sur un troisième canal il m'a fallu adapter pour la première fois l'imagerie de protéines marquées en infrarouge chez les ascidies. En effet, toutes les iRFP (infraRed Fluorescent Protein, protéine fluorescente infrarouge) requièrent le co-facteur biliverdine pour émettre de la lumière. Ce co-facteur est présent dans de nombreux types cellulaires (Filonov et al., 2011), cependant, dans les embryons d'ascidie, sans l'ajout de biliverdine dans l'eau de mer dans laquelle les embryons se développent, l'iRFP n'émet pas de lumière. A partir de ce constat, j'ai mis en place un protocole permettant d'imager des embryons incubés dans de l'eau de mer + 25µm biliverdine. J'ai donc micro-injecté des ovocytes avec des ARNm codants pour un marqueur de membranes vert (PH-GFP), un marqueur des noyaux rouge (H2B-mCherry) et le ERK-KTR en rouge lointain (ERK-KTR-iRFP), les ovocytes étant incubés dans la biliverdine depuis la micro-injection jusqu'à la fin de la période d'imagerie. Le ERK-KTR étant un marquage plus étendu dans la cellule que ceux des membranes ou des noyaux, l'utilisation du ERK-KTR en rouge lointain permettait de diminuer au maximum l'émission des photons les plus énergétiques pouvant altérer le développement embryonnaire. A partir de ce protocole expérimental, 4 embryons imagés toutes les 2 minutes pour ces trois marquages ont été choisis car présentant les meilleurs développement embryonnaire : Astec-Pm16, 17, 18, 19.

A partir de cet ensemble de données, chaque embryon à tout d'abord été reconstruit avec ASTEC en

utilisant le marquage des membranes (Guignard et al., 2020) et nommé automatiquement à partir des outils présentés dans le chapitre 2. En collaboration avec Benjamin Gallean, ingénieur informaticien dans notre équipe, nous avons ensuite développé de nouvelles méthodes algorithmiques permettant de mesurer automatiquement les signaux nucléaires et cytoplasmiques du ERK-KTR-iFRP à partir du marquage des noyaux H2B-mCherry. Ensemble nous avons mis au point une pipeline d'analyse d'image pleinement adaptée : i) capable de mesurer à l'intérieur des noyaux sur une large période du développement embryonnaire de *Phallusia*, ii) qui tient compte de l'hétérogénéité du signal ERK-KTR dans le cytoplasme pour ne mesurer que la région péri-nucléaire comme valeur du cytoplasme pour le calcul du ratio "ERK-KTRcytoplasme/ERK-KTRnoyau", et iii) qui tient compte de la disparition de l'enveloppe nucléaire et de la reprise de l'import nucléaire pendant la mitose.

Ce travail a permis de générer le premier atlas de la dynamique spatio-temporelle de l'activité de ERK à la résolution cellulaire sur une grande période de leur développement, depuis le stade 32-cellules jusqu'en fin de neurulation. Ces mesures de l'activité de ERK à partir du ERK-KTR, réalisées sur 1750 cellules identifiées, ont été mises à disposition sous la forme d'un atlas interactif en ligne à l'aide du navigateur morphodynamique MorphoNet pour chacun des 4 embryons analysés.

Après avoir mesuré l'activité ERK pour chaque cellule de nos 4 embryons de *Phallusia mammillata*, nous avons ensuite cherché à évaluer comment nos données quantitatives sur la dynamique spatio-temporelle de ERK se rapportent à ce qui est connu de l'activité ERK chez les ascidies. En effet, le pattern spatio-temporel d'activation de ERK a été bien étudié au cours de l'embryogenèse des ascidies *Halocynthia roretzi* et *Ciona intestinalis* (types A et B), en utilisant un anticorps spécifique de ERK phosphorylé et activé (dp-ERK) (Nishida, 2003 ; Hudson et al, 2007 ; Yasuo et Hudson, 2007 ; Shi et Levine, 2008 ; Haupaix et al., 2013, 2014 ; Razy-Krajka et al., 2018 ; Williaume et al., 2021). L'analyse des patterns spatiaux d'activation de ERK à l'aide de ces colorations immunohistochimiques a permis d'établir une liste de cellules dp-ERK positives ou négatives. Ainsi, l'analyse de la mesure de l'activité ERK moyenne calculée à partir des mesures du ERK-KTR pour cette liste de cellules présentes dans notre dataset a permis de retrouver le statut correct pour 91,6% des cellules de la liste, ce qui soutient la pertinence biologique des mesures automatiques de l'activité ERK de notre atlas et suggère également que l'activité ERK moyenne est un bon prédicteur des inductions cellulaires dépendantes de l'ERK chez les ascidies. Nous avons alors ensuite étudié l'activité moyenne de ERK pour toutes les cellules, ce qui révèle de nouveaux potentiels candidats pour des inductions différentielles ERK-dépendantes entre couples de cellules soeurs et suggère que nos connaissances actuelles ne sont pas exhaustives, un objectif qui pourrait être atteint par une analyse détaillée de notre atlas spatio-temporel systématique de l'activité ERK dans un plus grand nombre d'embryons.

L'ensemble de ce travail est présenté dans le manuscrit de l'article "Systematic and quantitative monitoring of the spatio-temporal dynamic of the ERK signalling at single cell resolution during early ascidian embryogenesis".

Manuscrit de l'article : Biasuz, Kilian et al. "Systematic and quantitative monitoring of the spatio-temporal dynamic of the ERK signalling at single cell resolution during early ascidian embryogenesis."

Systematic and quantitative monitoring of the spatio-temporal dynamic of the ERK signalling at single cell resolution during early ascidian embryogenesis

Kilian Biasuz^{1,#}, Benjamin Gallean¹, Ingrid Mormin¹, and Patrick Lemaire^{1,#}

1) CRBM, Université de Montpellier, CNRS, France

#: Corresponding authors

ABSTRACT

The extracellular signal-regulated kinase (ERK) signalling plays pivotal roles in a broad variety of developmental processes. To explain how this single pathway can play such diverse roles, it has been proposed that the dynamic spatial and temporal characteristics of ERK activity signalling may encode the diversity of ERK-dependent outcomes. Recent advances in genetically-encoded fluorescent biosensors have revealed previously unknown ERK activation dynamics in space and time and their functional importance, mainly in cultured cells. However, ERK dynamics during embryonic development have still only been visualised in a limited number of model organisms. We studied the ERK signalling in ascidians, marine invertebrate chordates, which develop with a fixed cell lineage, so that each cell can be named and found in the same position in all embryos, providing a rigorous framework for single cell resolution analyses. In addition, light-sheet imaging of membrane-marked live embryos of the ascidian *Phallusia mamillata* coupled with recently developed algorithms, allows the reconstruction of their embryonic development and the automated identification of each cell, making ascidian embryos a unique context for building and integrating multi-scale atlases. Based on these advantageous properties, we imaged embryos expressing membrane and nucleus labelling and the ERK-kinase translocation reporter (KTR) that enables live quantification of ERK activity. We developed a computational workflow to automatically quantify the nuclear and cytoplasmic signal intensity of the ERK-KTR every 2 minutes in each cell and generated the first functional atlas of the spatio-temporal dynamic of the ERK activity over a wide period of embryonic development. We show that the measured ERK activity recapitulates what is known about ERK-dependent inductions in ascidians and identifies novel candidates of ERK-dependent inductions.

INTRODUCTION

Successful embryogenesis requires the differentiation of the right cell types, in the right numbers and in the right positions (Wolpert et al., 2015). In most cases, the decisions made by individual cells are guided by signals from their neighbours. A surprisingly small set of signalling pathways act in combination to guide patterning and tissue morphogenesis in a time- and space-specific manner (Perrimon et al., 2012; Li and Elowitz, 2019). We now have a list of essential components of these signalling systems, which are involved in embryonic development and are also known to be reused in physiological and pathological processes such as wound healing, regeneration, cancer cell invasion and metastasis (Li and Elowitz, 2019).

The evolutionary-conserved extracellular signal-regulated kinase (ERK) signalling pathway is a key regulator of animal development. Hyperactivating mutations in individual components of the ERK pathway are also responsible for a class of human developmental syndromes, called RASopathies

(Figure 1) (Rauen, 2013; Patel and Shvartsman, 2018) and are also present in many types of cancer (Lavoie and Therrien, 2015; Sanchez-Vega et al., 2018). When activated by extracellular ligands, transmembrane receptors such as receptor tyrosine kinases (RTKs) trigger the initiation of the ERK signalling (Figure 1) (Lavoie et al., 2020). Upon receptor activation, cytoplasmic guanine exchange factors (GEFs, SOS) move to the plasma membrane and activate the small GTPase RAS, which in turn initiates the RAF/MEK/ERK serine-threonine kinase cascade. As a result, ERK catalyses the phosphorylation of various substrates. Transcription factors phosphorylated by ERK induce gene expression/repression, while other proteins such as cytoskeletal regulators function directly through phosphorylation, causing a wide range of cellular phenotypic changes, such as differentiation, proliferation, cell motility and tumorigenesis (Lavoie et al., 2020). Upon stimulation, positive and negative feedback mechanisms act at several levels of the pathway, depending on the stimulus and the cell type (Shin et al., 2009). How this sophisticated architecture shapes the dynamic response to input stimuli remains unclear (Patel and Shvartsman, 2018).

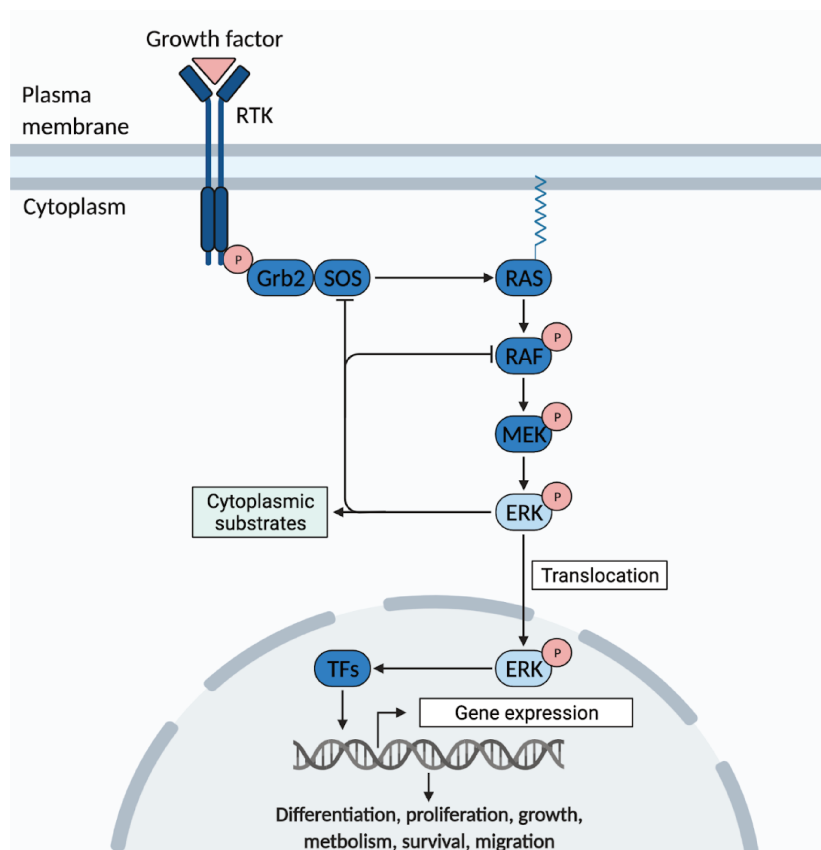


Figure 1 : The ERK signalling pathway. Upon ligand-dependent activation of a receptor tyrosine kinase (RTK), the small GTPase RAS is activated, leading to the subsequent activation of the RAF-MEK-ERK serine-threonine kinase cascade. ERK in turn phosphorylates cytoplasmic and nuclear substrates, including transcription factors.

ERK signalling has been extensively studied in various fields of research (Patel and Shvartsman, 2018; Lavoie et al., 2020; Nakamura et al., 2021). Many findings have been clarified using immunolabeling, inhibitors or gain- and loss-of-function experiments, representing the multifunctional aspect of ERK (Shilo and Barkai, 2017). However, we still do not fully understand why and how the ERK signalling participates in such diverse functions during development processes. It has been proposed that the spatial and temporal dynamic features of ERK activity signalling may in themselves encode the diversity of ERK-dependent outcomes (Antebi et al., 2017; Li and Elowitz, 2019). In a classic example,

rat PC12 cells use different positive or negative feedback mechanisms to regulate the nerve growth factor receptor (TrkA) and the epidermal growth factor receptor (EGFR), resulting in either sustained ERK activity and differentiation in response to TrkA activation, or transient ERK activity and proliferation in response to EGFR activation (Marshall, 1995; Santos et al., 2007).

Recent advances in live quantitative imaging of ERK signalling using genetically encoded fluorescent biosensors, which allow visualisation of signal transduction systems at the single cell level, have helped to verify this concept of the dynamical control of the ERK signalling both in cultured cells and in developmental contexts and have revealed further levels of complexity (Nakamura et al., 2021). A variety of ERK activity kinetics are now observed in a range of cell types or in response to different stimuli; these include transient, sustained and pulsatile ERK responses and differences in the temporal dynamics, amplitude or cumulative load of the ERK response have been shown to promote distinct cellular responses (Santos et al., 2007; Albeck et al., 2013; Aoki et al., 2013; de la Cova et al., 2017; Hiratsuka et al., 2015; Lim et al., 2015; Mayr et al., 2018; Johnson and Toettcher, 2019; Pokrass et al., 2020; Simon et al., 2020).

The ERK signalling system provides an important and fascinating example of how signalling pathways dynamically encode input signals to induce distinct cellular responses. Quantifying the ERK activity dynamics in the context of single cells during development using rapidly responding live reporters is therefore essential to further our understanding of signalling dynamics and its impact on cellular behaviours such as tissue formation *in vivo*. However, single-cell live *in vivo* quantitative studies of ERK activity during embryonic development are only beginning to emerge (de la Cova et al., 2017; Johnson and Toettcher, 2019; Moreno et al., 2019; Pokrass et al., 2020; Simon et al., 2020). The combined application of live embryo imaging methods at the cell resolution with dedicated computational tools to track individual cells and quantify ERK biosensor is still challenging. Consequently, our quantitative understanding of the spatiotemporal dynamics of the ERK pathway is still limited to a few scattered insights (Biasuz et al., 2018).

Here we studied the ERK pathway during embryogenesis of the ascidian *Phallusia mammillata*, a member of the tunicates, the sister group of vertebrates (Delsuc et al., 2006, 2018). Because ascidians diverged from the vertebrate lineage before the two rounds of genome duplication, they have unique genes for each component of the cascade, whereas vertebrates have multiple paralogues with divergent activities (Dehal and Boore, 2005; Lavoie et al., 2020). Embryos of the ascidian *Phallusia mammillata* have two other favourable properties: they are small (130 μm) and fully transparent, thus easy to image, and they develop with invariant cell lineages, so that each cell can be named and found in the same position and with conserved neighbourhoods in all embryos, providing a rigorous framework for single-cell resolution analyses (Conklin, 1905; Nicol and Meinertzhagen, 1988; Cole and Meinertzhagen, 2004). This extreme reproducibility of ascidian embryonic development is associated with a particular mode of fate specification. Indeed, cell fates in ascidians are mainly controlled by differential sister cell inductions which are proposed to be largely cell contact area dependent (Guignard et al., 2020).

Among the developmental signalling pathways used to shape their early embryos, ascidians make extensive use of the fibroblast growth factor (FGF)/ERK signalling pathway (Lemaire, 2009). Seventeen of the twenty-two known cell induction events up to gastrulation involve this pathway

(Lemaire, 2009) and, beyond gastrulation many cell type specification events also involve the FGF/ERK pathway. This is the case, for example, during central nervous system development (Hudson, 2016) or cardiac specification (Razy-Krajka et al., 2018). Although much is known about the signalling molecules and transcriptional regulators involved (Lemaire, 2009; Hudson, 2016; Williaume et al., 2021)), there is still a lack of a comprehensive and quantitative understanding of the dynamics of the FGF/ERK signalling activity.

Using the transparent embryos of the ascidian *Phallusia mammillata*, we recently reported accurate and automated whole-cell segmentation and tracking of membrane-labelled cells (Guignard et al., 2020), which can now be combined with accurate and automated systematic cell naming for each new embryo segmented (Biasuz et al., 2022, *in prep.*). We also successfully manually quantified the ERK activity in each cell of living 64-cell stage embryos using the ERK-Kinase translocation reporter (ERK-KTR), a genetically encoded fluorescent sensor of ERK kinase activity that converts a phosphorylation event into a nucleo-cytoplasmic shuttle event of the synthetic reporter (Figure 2 A) (Regot et al., 2014; Guignard et al., 2020).

Here we developed a systematic automated pipeline to quantitatively describe the spatio-temporal dynamics of ERK activity with cellular resolution and 2 minutes temporal resolution during the *Phallusia mammillata* embryogenesis. We first imaged 4 embryos, from the late 32-cell stage to the late-neurula stage, expressing membrane and nucleus labelling and the ERK-KTR (Table S1). We then developed a computational workflow to automatically quantify the nuclear and cytoplasmic signal intensity of ERK-KTR every 2 minutes in each cell and generated the first functional atlas of the spatio-temporal dynamics of ERK activity over a broad period of embryonic development. The results were made available in the form of an interactive online atlas using the morphodynamic browser MorphoNet (Leggio et al., 2019). The analysis of the dataset revealed that the measurement of average ERK activity recapitulates what is known about ERK-dependent inductions in ascidians and identifies novel candidates of ERK-dependent inductions.

RESULTS

Development of an automated ERK-KTR quantification strategy for live ERK activity during the early embryogenesis of the ascidian *Phallusia mammillata*

To monitor ERK activity during embryonic development of *Phallusia mammillata*, we used an ERK-kinase translocation reporter (ERK-KTR) (Regot et al., 2014). ERK-KTR consists of an ERK docking site (fragment of human Elk1, an ERK substrate), nuclear localisation (NLS) and export (NES) signals, and the iRFP fluorescent protein (Figure 2A). When the ERK signalling is inactive, the sensor is not phosphorylated and nuclear import predominates, resulting in increased nuclear enrichment of iRFP fluorescence (Figure 2A). Conversely, when ERK signalling is active, activated ERK phosphorylates the sensor triggering nuclear export (Figure 2A). The cytoplasm to nucleus ratio of ERK-KTR can thus serve as a quantitative readout of ERK activity (Regot et al., 2014; Mayr et al., 2018; Pokrass et al., 2020; Simon et al., 2020; Williaume et al., 2021).

The ERK-KTR sensor, which was originally developed for mammalian cells (Regot et al., 2014), has previously been reported to quantify ERK activity in the individual a-line cells of late 32-cell stage and

in all the cells of 64-cell stage embryos of the ascidian *Phallusia mammillata* (Guignard et al., 2020; Williaume et al., 2021). Consistent with the proposal that ERK activation in early ascidian embryos mostly results from the activation of the FGF/FGFR/RAS/RAF/MEK cascade, *dnFGFR* injected embryos and embryos treated with a pharmacological MEK inhibitor, U0126, from the early 16-cell stage show strongly reduced ERK activity (Guignard et al., 2020; Williaume et al., 2021). Analysis of the inferred ERK activity patterns revealed in both studies a high reproducibility of the level of ERK activity between embryos and supported that the level of ERK nuclear activation is proportional to the area of cell contact with FGF-expressing cells (Guignard et al., 2020; Williaume et al., 2021).

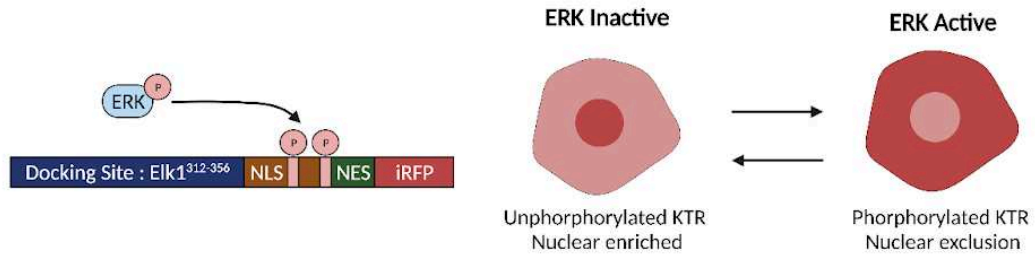
Nuclear and cytoplasmic intensity measurements of ERK-KTR in these studies were performed manually, an exhaustive work not allowing access to the precise spatio-temporal dynamics of ERK activity over a high number of cells and a large period of embryonic development. To systematically monitor ERK signalling dynamics in each cell of *Phallusia mammillata* embryos, we developed experimental and computational methods to automatically quantify ERK-KTR reporter nuclear and cytoplasmic signals. Using multi-view light-sheet microscopy, 4 living *Phallusia mammillata* embryos, called Astec-Pm16 to Astec-Pm19 (Table S1), produced from eggs microinjected with mRNA encoding fluorescent markers for plasma membranes (PH-GFP), nuclei (H2B-mCherry) as well as the ERK-KTR (ERK-KTR-iRFP) were imaged every 2 minutes from the late 32-cell stage (st.6b) to the late-neurula stage (st.16) (Figure 2B) (Table S1). Fluorescently-labelled plasma membranes were used to segment each cell of the developing embryo with the ASTEC pipeline, as previously described (Figure 2B) (Guignard et al., 2020). To compare the ERK activity between embryos with cellular resolution, the cells of the reconstructed embryos were automatically named according to a reference population of optimally named embryos (Biasuz et al., 2022, *in prep.*).

To automatically quantify the ERK-KTR nuclear and cytoplasmic signals, we developed an image analysis workflow based on nuclei labelling (Figure 2B). First, in each cell snapshot, the instance of a cell in an embryo at a time t , the H2B-mCherry signal is segmented using the multi-Otsu thresholding method (see methods). However, we noted that the very bright H2B labelling extends beyond the nucleus and therefore does not allow segmentation of the precise boundaries of the nuclei. Thus, to be sure to measure in each nucleus over a wide development window, based on the segmentations of the nuclei, a sphere is placed at the barycentre of each nucleus segmentation (Figure 2B; Figure S1) (see methods). The compact shape of nuclei ensures that each sphere is fully included in the nucleus (Figure S1). Finally, this method is applied to the ERK-KTR-iRFP signal channel to provide nuclear measurements of the ERK-KTR signal (Figure 2B).

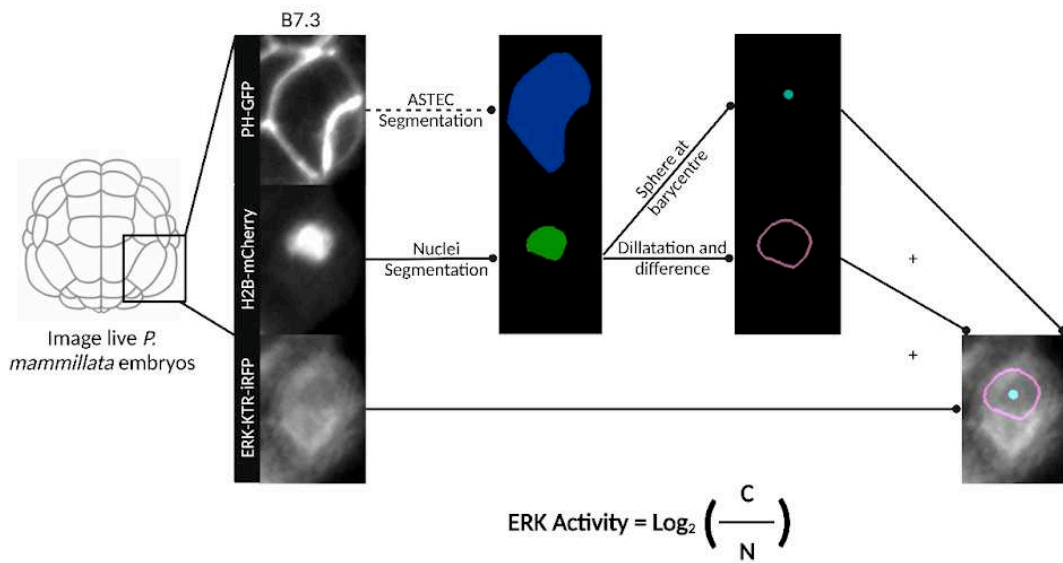
As the amount of microinjected ERK-KTR mRNA may vary between embryos and its inheritance may differ between cells within an embryo (Figure S2A), we next measured cytoplasmic ERK-KTR. However, attempts to normalise the signal by the total cytoplasmic intensity of each cell yielded results inconsistent with known inductions, suggesting that only a portion of the highly heterogeneous cytoplasmic ERK-KTR can shuttle to and from the nucleus (Figure S2B). We reasoned that the part of the cytoplasmic ERK-KTR signal involved in nucleo-cytoplasmic exchange is located in the proximity of the nucleus. The C/N ratio is then calculated by dividing the median cytoplasmic ERK-KTR intensity in the immediate vicinity of the nucleus (see methods) by the median nuclear ERK-KTR intensity. The ERK activity is then considered proportional to $\text{Log}_2(\text{C/N ratios})$ (Figure 2B) as in (Pokrass et al., 2020; Williaume et al., 2021).

Since during mitosis the nuclear envelope breaks down, abolishing the separation between nucleoplasm and cytoplasm, the dynamics of the ERK kinase activity that occurs during this period cannot be measured with the ERK-KTR (Regot et al., 2014; Kudo et al., 2018). In animal cells, the nuclear envelope breaks down during the prometaphase and forms again during telophase. The nucleus recovers its import capacity very rapidly upon nuclear envelope formation, a few minutes after the onset of anaphase (Dultz et al., 2008, 2009; Ferreira and Maiato, 2021). To determine the time window over which nuclear envelope breakdown extends for the cells in our *Phallusia mamillata* embryos dataset, we reasoned that the disruption of the nuclear envelope should result in a rapid change in the inferred dynamics of ERK, which should be observed in each cell of the dataset. We analysed the average dynamics of the C/N ratio around the cell divisions (Figure 2C). This allowed us to identify a change in the dynamics of the C/N ratio 8 time step (16 min.) before each cell division; most likely as a result of entry into mitosis (Figure 2C). This was further supported by analysis of the mitotic phases using H2B-mCherry labelling (Figure 2C). Indeed, Figure 2C clearly shows, for a representative cell, its entry into prophase 7 time step before division (14 min.). These two observations suggest that ERK-KTR measurements for the 8 time steps before division should be excluded from the analyses. In Figure 2C, one can also observe the anaphase 4 time step (8 min.) before the cell divides, suggesting a recovery of nuclear transport over the next few time steps. Looking again at the average dynamics of the C/N ratio, we noticed that it stabilised rapidly after the cell divisions (Figure 2C). Based on these observations we started measuring ERK activity from the second time step of the life of each cell.

A



B



C

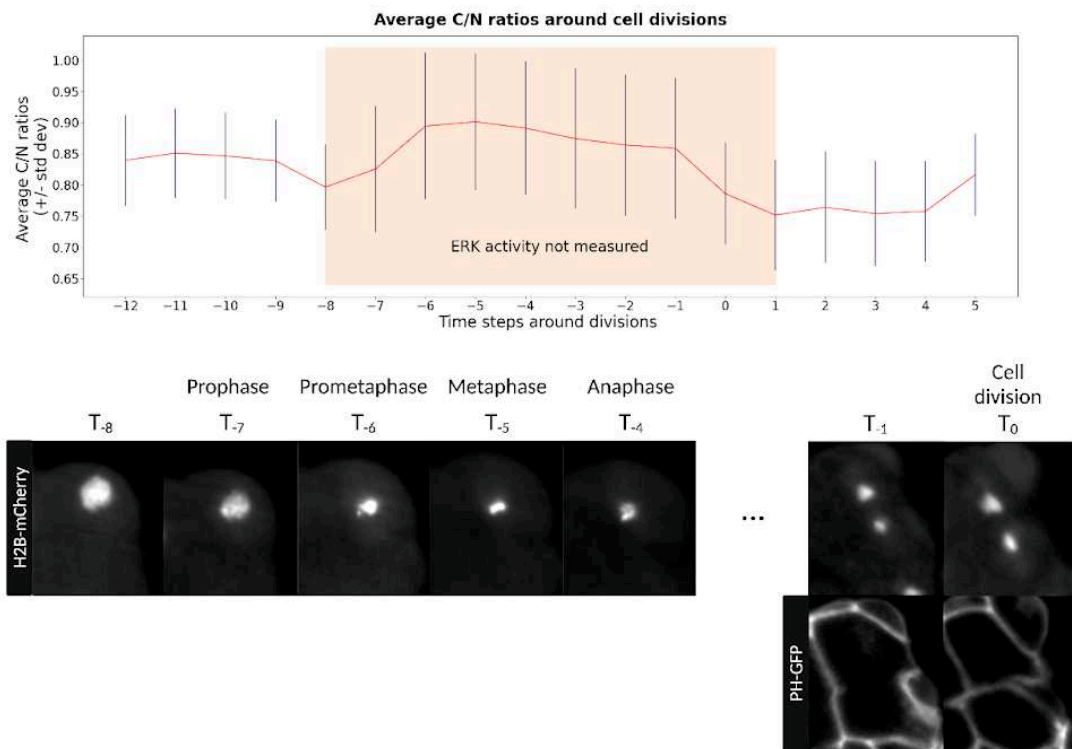


Figure 2 : Pipeline of automated measurements of nuclear and cytoplasmic ERK-KTR signals. A) Left : the ERK-KTR sensor is composed of peptides from Elk1 (Elk1³¹²⁻³⁵⁶) which promote docking of ERK, followed by a module containing NLS and NES sequences containing residues for phosphorylation by ERK. Right : the ERK-KTR principle (Regot et al., 2014). When ERK kinase activity is low, the ERK-KTR sensor protein is not phosphorylated and is enriched at the nuclear level. When ERK kinase activity is high, the ERK-KTR sensor is phosphorylated and excluded from the nucleus. B) Automated pipeline of ERK-KTR quantification. Living *Phallusia mammillata* embryos are imaged every 2 minutes using multi-views light-sheet microscopy. The embryos express membrane labelling (PH-GFP), nuclear labelling (H2B-mCherry) and the ERK-KTR reporter (ERK-KTR-iRFP). The membrane channel is used to reconstruct each cell using the already established cell segmentation and tracking tool, ASTEC (dotted line). To automatise the nuclear and cytoplasmic signals of the ERK-KTR reporter, we developed a new workflow based on the nuclei channel (black lines). First, in each cell for each time point, the H2B-mCherry signal is segmented. Then, on the basis of these segmentations, a small diameter sphere is positioned at the barycentre of each segmented nucleus. Applied to the ERK-KTR-iRFP signal channel this provides nuclear (N) measurements of the ERK-KTR signal. The nucleus boundary is also used to measure the cytoplasmic signal of ERK-KTR in the vicinity of the nucleus : the nucleus boundary is first extended a few pixels in each direction to ensure that it is outside the nucleus, and then extended a second time. The peri-nuclear area formed by the difference of these two extensions constitutes the cytoplasmic (C) measurement of the ERK-KTR reporter when applied to the ERK-KTR-iRFP channel. The C/N ratio is calculated by dividing the median intensity of the cytoplasmic ERK-KTR (C) by the median intensity of the nuclear ERK-KTR. The inferred ERK activity is then calculated using Log_2 (C/N ratios). C) Taking into account the nuclear envelope breakdown during mitosis. Top : analysis of the dynamique of the C/N ratio around cell divisions. We extracted the dynamics of the C/N ratio for each of the cells in our dataset, from the cells of the last generations back to the previous generations. Using lineage information, all these dynamics were aligned to the time of division (T_0 = appearance of the two daughter cells). The average (red line) of all these dynamics (+/- standard deviation, dark blue) is shown. Down : mitotic phases analysis using H2B-mCherry nuclear labelling for a representative cell (B6.1*). T_0 : division of the cell as illustrated by the formation of the membrane between the two daughter cells (PH-GFP channel). The names of the phases of the mitotic cycle are indicated.

This dataset of *Phallusia* embryos constitutes the first atlas of the spatio-temporal dynamics of ERK activity for every single cell in living embryos over a large period of their embryogenesis (Figure 3). The development of experimental and image analysis methods allowed the automation of nuclear and cytoplasmic measurements of ERK-KTR and the quantitative monitoring of ERK activity for 1750 cells (Table S1). To facilitate exploration and sharing of this atlas, the results were made available in the form of an interactive online atlas using the morphodynamic browser MorphoNet (Leggio et al., 2019) (Figure 3; Table S1; Movies S1-S12) (see methods).

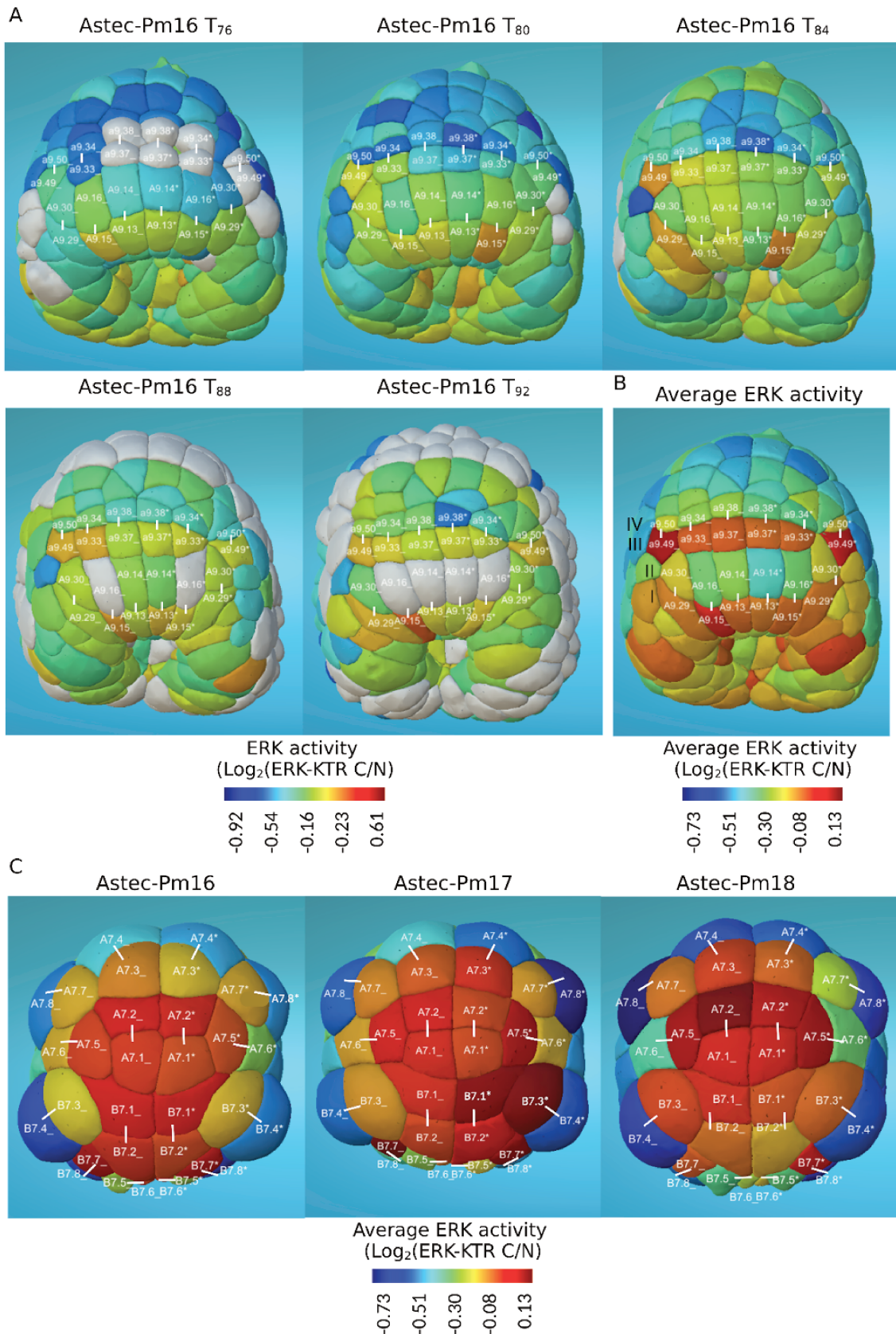


Figure 3 : Atlas of the spatio-temporal dynamics of ERK activity with cellular resolution A) Spatio-temporal dynamics of ERK activity measured during the 6-row neural plate stage on Astec-Pm16. The ERK activity is colour-coded and the scale bar is indicated. Uncoloured cells correspond to mitotic cells in which the ERK-KTR C/N ratio is not measured. The cell names for rows I (lower row) /II/III and IV (upper row) of the neural plate are shown. The acquisition

time steps are indicated (one time step every two minutes). B) Average ERK activity measured in each cell and plotted on Astec-Pm16 T₈₄. The average ERK activity is colour-coded and the scale bar is indicated. The names of cells in rows I/II/III and IV of the neural plate are indicated, revealing a higher average ERK activity in cells in rows I and III compared to rows II and IV. White lines link sister cells. Dorsal views, anterior is to the top. C : Reproducibility of the average ERK activity patterns between embryos. Vegetal views of Astec-Pm16, Astec-Pm17, Astec-Pm18 embryos at the 64-cell stage with average ERK activity colour-coded. Names of the vegetal cells are shown, with the lines link sister cells. Visualised using MorphoNet.

Automated measurement of ERK-KTR C/N ratio throughout early *Phallusia mammillata* embryogenesis recapitulates previous knowledge

Having measured ERK activity for each cell in our 4 *Phallusia mammillata* embryos (Astec-Pm16-19; Table S1), we then sought to assess how our quantitative data on the spatio-temporal dynamics of ERK relate to what is known about ERK activity in ascidians.

The spatio-temporal pattern of activation of ERK was investigated during embryogenesis of the ascidians *Halocynthia roretzi* and *Ciona intestinalis* (types A and B), using an antibody specific to diphosphorylated activated ERK (dp-ERK) (Nishida, 2003; Hudson et al., 2007; Yasuo and Hudson, 2007; Shi and Levine, 2008; Haupaix et al., 2013, 2014; Razy-Krajka et al., 2018; Guillaume et al., 2021). Analysis of the spatial patterns of ERK activation using these immunohistochemical stains allowed to establish a list of dp-ERK positive cells, from the appearance of the first signals at the 32-cell stage (st. 6) to the tail elongation stages. ERK activation in these cells has been linked to the embryonic fates inductions (Nishida, 2003; Hudson et al., 2007; Yasuo and Hudson, 2007; Shi and Levine, 2008; Lemaire, 2009; Haupaix et al., 2013, 2014; Razy-Krajka et al., 2018; Guillaume et al., 2021). A remarkable feature of the ERK activation pattern in ascidians is that ERK is frequently activated in only one of the sister cells following asymmetric cell divisions (i.e. resulting in two sister cells with different larval tissue fate). This is observed, for example, during neural plate specification where ERK activation alternates between positive and negative cells along the anterior-posterior axis (Nishida, 2003; Hudson et al., 2007; Haupaix et al., 2013, 2014).

Overall, ERK activation patterns are consistent between *Halocynthia roretzi* and *Ciona intestinalis* (types A and B) (Nishida, 2003; Yasuo and Hudson, 2007; Haupaix et al., 2013, 2014). When differences are observed (Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008) or dp-ERK data is only available in one species, we built our ground truth of ERK activation patterns for ascidians, by considering only patterns consistent in both species. Using this approach, we established a list of 14 dp-ERK positive cells in both species (Table S2) and a list of 23 asymmetric divisions resulting in one dp-ERK positive and one dp-ERK negative sister cell (Table S3). For simplicity, when cells show a pattern of ERK activation that varies over time (Nishida, 2003; Yasuo and Hudson, 2007), we consider that a dp-ERK-positive cell at one point in its life is positive over its entire life span.

To compare the imaging of dp-ERK signals in fixed ascidian embryo with our spatio-temporal measurements of the dynamics of ERK activity of our dataset, in which the choice of analysis time step may influence the inferred ERK activity observed (Figure 3A), we first calculated the average ERK activity for each cell (Figure 3B, C) (see methods). This showed a high reproducibility of the average ERK activity measured between symmetric left-right (L/R) cells (Figure 3B, C) (Figure S3) that is a first indication that the automatic measurement of the ratio (C/N) of the ERK-KTR reporter reflects ERK

activity. This also suggests that to reduce the number of cells to be analysed, each cell can be represented by the mean of the average ERK activity measurement between the two symmetric L/R cells.

Next, to compare the qualitative measurements of dp-ERK signals, which resulted in a list of positive or negative cells, with our quantitative and continuous measurements of ERK activity, we sought to determine a threshold that separates induced ERK-positive cells from uninduced ERK-negative cells in our dataset. For this, we plotted the average ERK activity measured for each pair of L/R cells belonging to the ground truth lists (Tables S2-S3), in each Astec-Pm (Figures S4-S15). A lower threshold, below which a cell is considered ERK-negative, is then set at the level of the cell known to be ERK-negative in the literature with the highest average measured ERK activity. An upper threshold, above which a cell is considered ERK-positive, is set at the level of the cell known to be ERK-positive with the lowest average measured ERK activity and which is above the lower threshold to maintain consistency between thresholds (Figures S4-S15). By averaging the values of these two thresholds, we obtain a single threshold that allows us to infer an ERK positive or ERK negative status for each cell measured. We first noted that when applying a single all-generation threshold for each embryo, a large number of cells known to be dp-ERK positive would be considered negative (Figures S4; S9; S12). Indeed, some cells from later generations known to be dp-ERK negative show a higher average level of ERK activity than many cells from earlier generations known to be positive. We therefore calculated and applied one set of upper and lower thresholds for each generation of cells (Figures S5-S8; S10-S11; S13-S15; Table S4). Using this simple method, 109/119 measured L/R cell pairs were assigned to the correct status (Table S4). This is the case, for example, for all cells in rows I and II of the neural plate which are positive while their sister cells, in rows II and IV respectively, are negative (Figure 3B, Table S4). The status of 10 cells differs, however. The cells for which the status differs only in 1 out of 4 embryos (a7.10 in Astec-Pm18, A8.6 in Astec-Pm18, A8.14 in Astec-Pm17) could be attributed to a problem with the automatic measurement process or to developmental variability. Cells with different status from what is known in several embryos (b7.9 in Astec-Pm16/17/18, b7.10 in Astec-Pm17/18, A8.13 in Astec-Pm17/18) reflect either a difference in ERK activity in these cells in *Phallusia* compared to what is known in *Halocynthia* and *Ciona*, which is unlikely, a normalisation that is not adapted in these cells or could reflect differences between ERK activation, measured from dp-ERK antibodies, and ERK activity, inferred from ERK-KTR.

Overall, both the left-right symmetry in inferred ERK activity and the correct identification of 91.6% of the known ERK-positive or ERK-negative cells we analysed confirm the biological relevance of ERK activity inference from automated measurements of nuclear and cytoplasmic ERK-KTR signals over several generations of embryonic *Phallusia mammillata* cells. These data also suggest that the mean ERK activity is a good predictor of ERK-dependent cell inductions in ascidians.

Measurement of mean ERK activity between sister cells identifies potential new ERK-dependent inductions

We next asked whether the induction thresholds, determined above for each generation, could identify divisions giving rise to an ERK positive and an ERK negative sister cell that are not in our ground truth (Table S3).

This strategy first allowed the correct identification, within our dataset of 4 embryos, of 12/12 ERK-dependent asymmetric divisions known in at least one of the two reference species *Halocynthia roretzi* or *Ciona intestinalis* (Table S5). The application of this strategy then resulted in a list of 21 divisions with one ERK-positive and one ERK-negative sister cell, in at least one of the embryos, for which no ERK-dependent induction event had previously been reported (Table S6). Among these asymmetric divisions, 5 divisions (b7.8, b7.9, b7.10, b7.11, b7.13; Table S6) are found in at least two of the embryos, making them good candidates for potential ERK dependent induction events. In some cases, the variability observed could be explained by variability in the orientation of these divisions, which caused sister cells to differ in the cell contacts with neighbouring cells expressing an FGF ligand (Figure 4). 16 divisions were also identified in only one embryo, mainly reflecting the absence of representatives of these divisions in the other embryos. Among these divisions, 8 are 9th generation divisions of A-line and a-line neural cells (Figure 4; Table S6). An example is the division of a9.52, where it is known that only the daughter cell a10.103 expresses the transcription factor *Six1/2* and whose progeny will form a different type of neuron from those of a10.104 (Abitua et al., 2015), suggesting a potential role of ERK signalling in this induction event, which will require experimental confirmation.

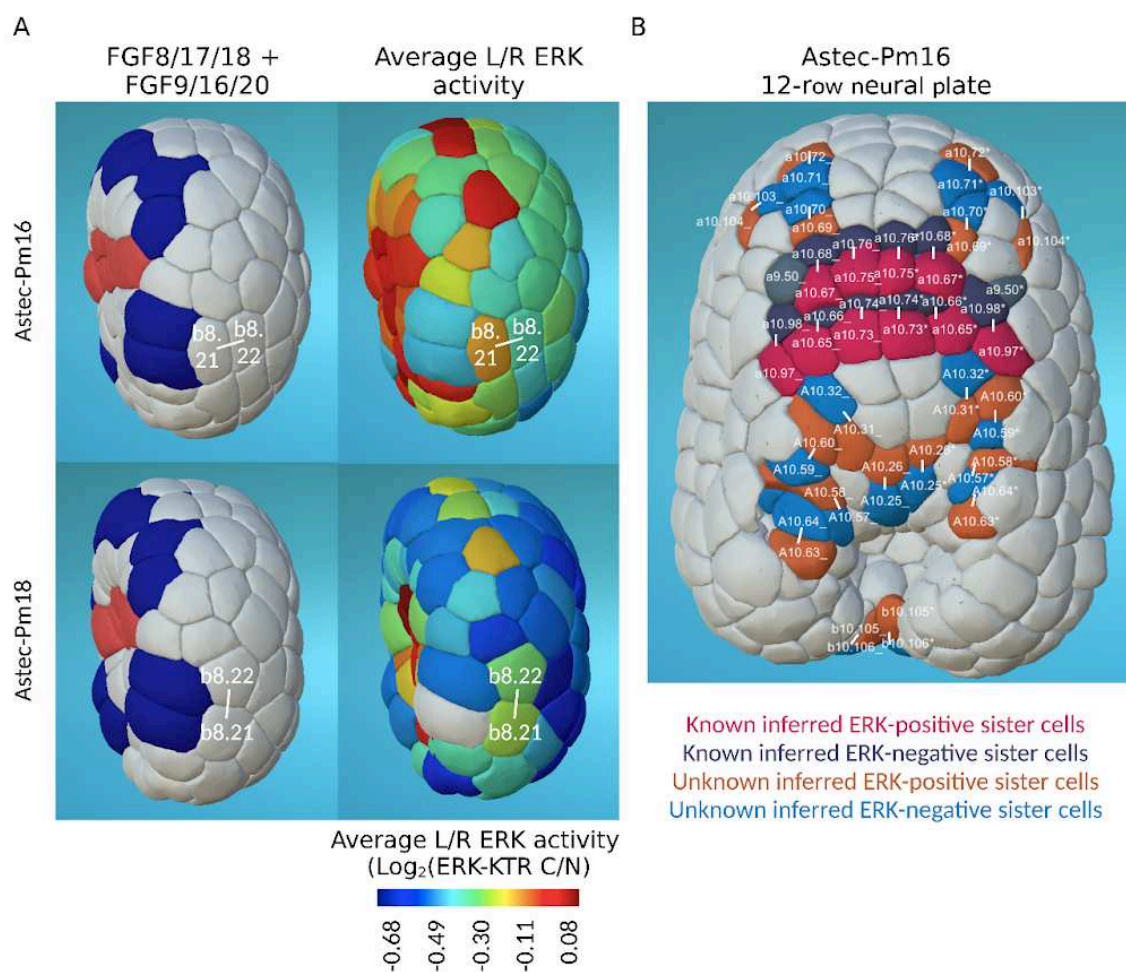


Figure 4 : Identification of novel cell divisions giving rise to one ERK-positive and one ERK-negative sister cell. A : Example of the b7.11 division giving rise to b8.21 with a positive inferred ERK status and b8.22 with a negative status in Astec-Pm16, 17, 19 but not in Astec-Pm18. This can be explained by the atypical orientation of the b7.11 division in Astec-Pm18 that gives rise to two sister cells contacting the FGF9/16/20 emitting cells. Lateral views, anterior is

to the top. Expression profiles of FGF8/17/18/19 (in red) and FGF9/16/20 (in blue) as well as the average L/R ERK activity are shown. B : Many new sister cell pairs of the 10th generation are proposed to have differential ERK activity, especially in the nervous system. Dorsal view, anterior is to the top. Cell name for selected cells are shown, with the lines link sister cells.

Overall, this analysis suggests that our current knowledge of the role of ERK signalling during ascidian development, although well documented, is not yet fully comprehensive and that systematic measurement of mean ERK activity can identify potential ERK-dependent asymmetric divisions.

DISCUSSION

Our knowledge of the role of ERK dynamics during embryonic development has so far remained limited (Biasuz et al., 2018; Patel and Shvartsman, 2018; Li and Elowitz, 2019; Nakamura et al., 2021). This is mainly due to the difficulty of monitoring the activity of the ERK signalling pathway at the cellular level in developing embryos. Here, we studied the ERK pathway during ascidian embryogenesis. These embryos develop in a stereotyped manner with a fixed cell lineage, so that each can be identified. They make extensive use of the ERK signalling pathway during embryonic fate specification (Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008; Lemaire, 2009; Hudson, 2016; Tolkin and Christiaen, 2016; Razy-Krajka et al., 2018). Furthermore, the recent development of lightsheet imaging of the transparent embryos of the ascidian *Phallusia mammillata*, combined with dedicated algorithmic tools allowing digital reconstruction of the embryos and automatic naming of each cell, have made this model a framework of choice for constructing multi-modal atlases with cellular resolution (Guignard et al., 2020) (Biasuz et al., 2022, *in prep*).

To study the spatio-temporal dynamic of ERK activity with single-cell resolution, we first imaged embryos co-expressing membrane and nucleus fluorescent markers with ERK-KTR, a live reporter for ERK activity (Regot et al., 2014; Guignard et al., 2020; Williaume et al., 2021). We then developed a computational workflow to automatically quantify the nuclear and cytoplasmic signal intensity of the ERK-KTR every 2 min in each cell. We applied this strategy to generate the first functional atlas of the spatio-temporal dynamics of ERK activity over a broad period of embryonic development in 4 different embryos. The results are made available in the form of an interactive online atlas using the MorphoNet morphodynamic browser (Table S1) (Leggio et al., 2019). We showed that the measurement of average ERK activity recapitulates what is known about ERK-dependent inductions in ascidians, which supports the biological relevance of our atlas. We finally provide examples showing that these measurements identify novel candidates of ERK-dependent inductions.

Thus, although many ERK-dependent inductions have been described in recent decades, notably through the use of anti-dp-ERK antibodies (Nishida, 2003; Hudson et al., 2007; Yasuo and Hudson, 2007; Shi and Levine, 2008; Haupaix et al., 2013, 2014; Razy-Krajka et al., 2018; Williaume et al., 2021), this study suggests that our knowledge is not exhaustive, a goal that could be achieved by detailed analysis of our systematic spatio-temporal atlas of ERK activity. Indeed, we showed that measuring the average ERK activity compared to the threshold set for each generation is a good predictor of ERK-dependent inductions. The question now arises whether other quantitative parameters of the dynamics, such as the maximum activity of ERK signalling, its amplitude or integrated activity, as observed in other developmental contexts (Johnson and Toettcher, 2019; Pokrass et al., 2020), could also explain the inductions and which ones is the better predictor and more reproducible between embryos. However, in order to better characterise the spatio-temporal

dynamics of the ERK pathway in ascidian inductions and its variability, one of the main limitations so far is the lack of sufficient numbers of imaged embryos for divisions beyond the 8th generation. This can be explained by the difficulty of imaging well developing *Phallusia mammillata* embryos due to the phototoxicity inherent in the acquisition of 3 imaging channels. A solution to this phototoxicity issue would be to simultaneously image the membranes and nuclei in a single channel and to separate these two sources of information subsequently using dedicated deconvolution tools based on deep learning networks (Hallou et al., 2021).

In addition to allowing a comprehensive description of ERK-dependent inductions and characterising the informative role of the spatio-temporal dynamics of the ERK pathway in these events, the automation of measurements of ERK activity paves the way to a better understanding of the surface-of-contact-dependent behaviour of the ERK pathway. Indeed, while this contact-dependent mode of ERK signalling is beginning to be well documented for the early stages of the ascidian development with quantitative measurements of ERK activity *in vivo* (Guignard et al., 2020; Williaume et al., 2021), the lack of measurements for the later stages has not allowed to test directly the generality of this model. The data we have generated allows us to accurately test this model up to 10th generation cells, however, the lack of FGF ligand expression data for the late stages makes this task complicated. An alternative would be to study the covariance between precise measurements of contact areas and measurements of ERK activity for these stages for different embryos. Furthermore, the automation of ERK pathway activity monitoring allows its combination with pathway perturbations at different levels of the ERK cascade, which is essential for building a mechanistic model of the surface-of-contact dependent ERK activity. For example, recent quantitative measurements of ERK activity in a-line cells at the 32-cell stage coupled with inhibition of the Ephrin/Eph pathway have revealed the role of this pathway in dampening the ERK signalling, thereby regulating the spatial accuracy of the surface-of-contact with FGF expressing cells ERK-dependent inductions (Williaume et al., 2021).

Finally, it is important to mention that the methods that have been developed to automate nuclear and cytoplasmic measurements of ERK-KTR can be deployed to measure any other process based on a nuclear relocation event enhancing the attractiveness of the ascidian model for building multi-modal functional atlases at cellular resolution.

MATERIALS AND METHODS

Animals and ascidian embryos culture

Phallusia mammillata were provided by the Centre National de Ressources Biologiques Marines in Roscoff (France). The embryos were obtained by *in vitro* fertilisation and dechorionated as in (Sardet et al., 2011).

Unfertilised eggs of *Phallusia mammillata* were microinjected under a stereoscope with approximately 45pg of synthetic mRNA as in (Guignard et al., 2020).

To record near infra-red fluorescence, injected embryos were incubated with the cofactor biliverdin hydrochloride (Merck : 30891) in artificial seawater, at concentrations of 25µM as described in (Filonov et al., 2011). The dye was applied from microinjection to the end of the imaging period.

Constructs and synthetic mRNA

The plasmid pSpe3-ERK-KTR-iRFP was constructed by recombining the gateway construct pENTR-ERK-KTR-iRFP (a gift from Jared Toettcher; Addgene plasmid #111510; (Dine et al., 2018)) with the mRNA synthesis vector pDEST/Spe3-Rfa, as described in (deposited in addgene, plasmid #xxxx) (Roure et al., 2007).

The plasmid pSpe4-H2B-mCherry (deposited in addgene, plasmid #xxxx) was constructed by recombining the gateway construct pENTR-H2B (Roure et al., 2007) with the mRNA synthesis vector pDEST/Spe4-Rfa-mCherry (deposited in addgene, plasmid #xxxx). The pDEST/Spe4-Rfa-mCherry plasmid was produced by gene synthesis using pDEST-Spe3-RfA as template to which we added a BglIII-ZraI-AatII-StuI-SnaBI (agatctgacgtcaggccttagta) polylinker before the attR1 sequence and an EcoRV-SacII-AfeI-BstEII-AvrII (gatatcgcgccggaagcgctggttagctagg) polylinker after the attR2 sequence (pDEST/Spe4-Rfa deposited in addgene, plasmid #xxxx). Afterwards, mCherry was inserted by cloning using the EcoRV and AvrII restriction sites.

The mRNAs for microinjection were synthesised using the mMessage mMachine T3 Invitrogen transcription kits using as templates the pRN3-PH-GFP construct (a kind gift from Alex McDougall, Observatoire Océanologique de Villefranche-sur-Mer, France; (Carroll et al., 2003)), pSpe4-H2B-mCherry and pSpe3-ERK-KTR-iRFP, linearised with the KpnI (for pRN3-PH-GFP and pSpe4-H2B-mCherry) or SbfI (for Spe3-ERK-KTR-iRFP) restriction enzymes.

Imaging and reconstructing *Phallusia mammillata* embryos

Embryos were deposited by gravity without embedding at the bottom of a 1% Low Melt Agarose (Carl Roth, 6351.1) cone-shaped well and imaged at a constant temperature of ~18°C in artificial seawater using a multiview lightsheet microscope (MuViSPIM, Luxendo) (Table S1).

PH-GFP was excited with a 488 nm laser, H2B-mCherry with 561 nm laser and ERK-KTRiRFP with a 640 nm laser, with simultaneous two-sided illumination. The emitted light was collected by two opposing detection arms positioned perpendicularly to the illumination plane, resulting in simultaneous dual acquisition. Each detection arm was equipped with a 20× Olympus water dipping objective (NA 1.0) combined with a tube lens leading to a 33.3 fold image magnification. The filter used was the band-path BrightLine 525/30 filter (Semrock). Filtered emitted light was collected by Hamamatsu V2 Flash 4 sCMOS cameras using the electronic confocal slit detection (eCSD) mode. The electronic slit size was matched in accordance to the diameter of the illumination beams (60 pixels).

At each time point, for each channel, two acquisitions were sequentially performed, the second one orthogonal to the first, following this sequence : membrane then nuclei then ERK-KTR; 90° rotation; membrane then nuclei then ERK-KTR and generating four whole-embryo stacks, i.e. two sets of two matching 3D intensity image stacks (0, 90, 180, 270 degrees) per channel each covering the whole imaged embryo. Embryos were imaged every two minutes with an image voxel size of $0.195 \times 0.195 \times 1 \mu\text{m}^3$. The intensity values at each voxel were encoded in 16 bits. For each one of the three-channels (PH-GFP, H2B-mCherry and ERK-KTRiRFP) the four image stacks at a given time point

are fused based on the PH-GFP channel using an updated version of the ASTEC suite described in (Guignard et al., 2020) and available on the INRIA gitlab (<https://gitlab.inria.fr/astec/astec>). The resulting computed transformations as well as image weights are used to fuse the H2B-mCherry and ERK-KTR-iRFP channels. This results in three-channels fusion image stacks.

All membrane-labelled embryonic cells of the imaged embryos were then segmented at each time point using an updated version of the ASTEC pipeline (Guignard et al., 2020), producing a reconstructed version of each embryo integrating cell lineages and quantitative geometrical properties including cell contact surfaces, stored in an .xml file. Segmentation errors (over- and under-segmentations or missing cells) were manually corrected using the MorphoNet 3D curation package (<https://morphonet.org/helpfiles/API/morphonet.plot.html>).

Initial naming was performed at the 64-cell stage for each embryo by associating manually a name to each cell snapshot based on its position in the embryo and annotator knowledge. This step was performed by editing the .xml properties file of each embryo using the sublime-text file editor and adding a "cell_name" key containing the association between each cell label to its name (e.g <cell cell-id="280001">'a7.0001_'</cell>). For more information on the properties file, see the ASTEC documentation (https://astec.gitlabpages.inria.fr/astec/astec_embryoproperties.html).

Automated ERK-KTR signals quantification

For each time point, we start with the corrected cell segmentation. The 3D segmentation image for each cell is then extracted to generate a mask. This mask is applied to the image of the nuclei, allowing us to apply the Otsu thresholding nuclei segmentation method locally in each cell. If this method fails, i.e. if the threshold obtained for a cell = 0 or if after applying the threshold, the segmented nucleus obtained covers more than 50% of the mask of the cell, we apply a multi Otsu thresholding. This method splits the image of the nuclei into three groups, the first corresponding to the background of the image, the second to the nucleus and the third to a potential artefact. These steps therefore generate a segmentation of the H2B nucleus labelling for each cell snapshot. To be sure to measure within the nucleus, we then apply two successive erosions of one voxel each for each segmented nucleus. Then, the number of unconnected objects produced in each cell is determined using the connected component analysis (3 connectivity). If we have multiple objects, we need to choose one as being the nucleus. For each connected region, we calculate its volume and the largest object is chosen. Now we have a mask with one connected region: the nucleus. We compute the barycenter of the segmented nucleus and apply a sphere with a radius of 6 voxels. The sphere volume is determined by counting the number of voxels within the nucleus (voxel size = 0.3x0.3x0.3µm). Applied to the ERK-KTR image the median ERK-KTR intensity is measured in each sphere and constitutes the ERK-KTR nuclear measurement. Applied to the segmented image of the nuclei, this allows us to compute the median segmented nuclei intensity and assess the quality of our automatic sphere positioning strategy (0 = background, 1 = nucleus). To measure the ERK-KTR cytoplasmic signal in the perinuclear region, the uneroded segmented nucleus is first expanded 3 times (1 voxel expansion each time), to ensure that no measurement is made in the nucleus. We then apply a new dilation of 4 voxels and it is this region thus formed that constitutes the cytoplasm measurement. If a voxel falls outside the cell segmentation mask, it is not used for the measurement.

The volume of each region formed is determined by counting the number of voxels within the region (voxel size = $0.3 \times 0.3 \times 0.3 \mu\text{m}$). Applied to the ERK-KTR image, the median intensity of the ERK-KTR reporter is computed constituting the ERK-KTR cytoplasmic measurements. The whole ERK-KTR cell signal is obtained by computing the median ERK-KTR signal intensity within the cell coordinates using the cell segmentations. All these measurements are stored separately into the xml file containing all the information related to each reconstructed embryo and are then post-processed using the lineage information to i) compute new informations without the 2 first and the 8 last time points of the life of each cell; ii) compute the cytoplasmic to nuclear ratio for each snapshot; iii) compute the ERK activity ($\text{Log}_2(\text{C/N ratio})$), iv) compute the average ERK activity in each cell or for each pair of matching L/R cells, using the cell names. All this information is stored in the xml file.

Integration of the atlas of ERK activity into MorphoNet

For the online display of the ERK activity measured in each Astec-Pm embryo, Morphonet was used (<http://www.morphonet.org>). Morphonet is an interactive web-based platform for the visualisation and sharing of complex morphological data and metadata (Leggio et al., 2019). Leveraging its 3D Unity visual engine (<https://unity.com>), it offers a wide range of possible interactions with 2, 3 and 4D datasets. With a flexible hierarchical representation of biological structures and dedicated formats for associated metadata, users can track the dynamics of biological forms, onto which associated quantitative and qualitative properties can be projected.

Cells are represented by meshes in Morphonet, and the meshes generated by cell segmentation have been converted to *obj* format and uploaded to Morphonet together with the following information contained in each Astec-Pm xml file: lineage information, cell names, cell fates, mean ERK-KTR intensity per cell, ERK activity, mean ERK activity, mean L/R ERK activity.

Data and code availability

The automated nuclear and cytoplasmic ERK-KTR measurements workflow has been developed using python3. The images are stored and manipulated using the Numpy library. The image processing algorithms are from Scipy library and Skimage library. The xml file is manipulated using the xml.etree library. All the codes and documentation can be found on a Git repository : <https://gite.lirmm.fr/bgallean/ktr-analysis>.

Statistical analysis

The two-sample Kolmogorov-Smirnov (K-S) statistical analysis was performed using the `scipy.stats.ks_2samp` function of the `scipy.stats` module of the python scipy library (https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ks_2samp.html#scipy.stats.ks_2samp). This test compares the underlying continuous distributions $F(x)$ and $G(x)$ of two independent samples. The null hypothesis is that the two distributions are identical, $F(x)=G(x)$ for all x . If the KS statistic is high or the p-value is low (<0.05), then we can reject the null hypothesis.

REFERENCES

- Abitua, P.B., Gainous, T.B., Kaczmarczyk, A.N., Winchell, C.J., Hudson, C., Kamata, K., Nakagawa, M., Tsuda, M., Kusakabe, T.G., Levine, M., 2015. The pre-vertebrate origins of neurogenic placodes. *Nature* 524, 462–465. <https://doi.org/10.1038/nature14657>
- Albeck, J.G., Mills, G.B., Brugge, J.S., 2013. Frequency-Modulated Pulses of ERK Activity Transmit Quantitative Proliferation Signals. *Mol. Cell* 49, 249–261. <https://doi.org/10.1016/j.molcel.2012.11.002>
- Antebi, Y.E., Nandagopal, N., Elowitz, M.B., 2017. An operational view of intercellular signaling pathways. *Curr. Opin. Syst. Biol.* 1, 16–24. <https://doi.org/10.1016/j.coisb.2016.12.003>
- Aoki, K., Kumagai, Y., Sakurai, A., Komatsu, N., Fujita, Y., Shionyu, C., Matsuda, M., 2013. Stochastic ERK Activation Induced by Noise and Cell-to-Cell Propagation Regulates Cell Density-Dependent Proliferation. *Mol. Cell* 52, 529–540. <https://doi.org/10.1016/j.molcel.2013.09.015>
- Biasuz, K., Leggio, B., Faure, E., Lemaire, P., 2018. The “computable egg”: Myth or useful concept? *Curr. Opin. Syst. Biol.* 11, 91–97. <https://doi.org/10.1016/j.coisb.2018.09.003>
- Carroll, M., Levasseur, M., Wood, C., Whitaker, M., Jones, K.T., McDougall, A., 2003. Exploring the mechanism of action of the sperm-triggered calcium-wave pacemaker in ascidian zygotes. *J. Cell Sci.* 116, 4997–5004. <https://doi.org/10.1242/jcs.00846>
- Cole, A.G., Meinertzhagen, I.A., 2004. The central nervous system of the ascidian larva: mitotic history of cells forming the neural tube in late embryonic *Ciona intestinalis*. *Dev. Biol.* 271, 239–262. <https://doi.org/10.1016/j.ydbio.2004.04.001>
- Conklin, E.G., 1905. Mosaic development in ascidian eggs. *J. Exp. Zool.* 2, 145–223. <https://doi.org/10.1002/jez.1400020202>
- de la Cova, C., Townley, R., Regot, S., Greenwald, I., 2017. A Real-Time Biosensor for ERK Activity Reveals Signaling Dynamics during *C. elegans* Cell Fate Specification. *Dev. Cell* 42, 542–553.e4. <https://doi.org/10.1016/j.devcel.2017.07.014>
- Dehal, P., Boore, J.L., 2005. Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. *PLOS Biol.* 3, e314. <https://doi.org/10.1371/journal.pbio.0030314>
- Delsuc, F., Brinkmann, H., Chourrout, D., Philippe, H., 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439, 965–968. <https://doi.org/10.1038/nature04336>
- Delsuc, F., Philippe, H., Tsagkogeorga, G., Simion, P., Tilak, M.-K., Turon, X., López-Legentil, S., Piette, J., Lemaire, P., Douzery, E.J.P., 2018. A phylogenomic framework and timescale for comparative studies of tunicates. *BMC Biol.* 16, 39. <https://doi.org/10.1186/s12915-018-0499-2>
- Dine, E., Gil, A.A., Uribe, G., Brangwynne, C.P., Toettcher, J.E., 2018. Protein Phase Separation Provides Long-Term Memory of Transient Spatial Stimuli. *Cell Syst.* 6, 655–663.e5. <https://doi.org/10.1016/j.cels.2018.05.002>
- Dultz, E., Huet, S., Ellenberg, J., 2009. Formation of the Nuclear Envelope Permeability Barrier Studied by Sequential Photoswitching and Flux Analysis. *Biophys. J.* 97, 1891–1897. <https://doi.org/10.1016/j.bpj.2009.07.024>
- Dultz, E., Zanin, E., Wurzenberger, C., Braun, M., Rabut, G., Sironi, L., Ellenberg, J., 2008. Systematic kinetic analysis of mitotic dis- and reassembly of the nuclear pore in living cells. *J. Cell Biol.* 180, 857–865. <https://doi.org/10.1083/jcb.200707026>
- Ferreira, L.T., Maiato, H., 2021. Prometaphase. *Semin. Cell Dev. Biol.* 117, 52–61. <https://doi.org/10.1016/j.semcdb.2021.06.004>
- Filonov, G.S., Piatkevich, K.D., Ting, L.-M., Zhang, J., Kim, K., Verkhusha, V.V., 2011. Bright and stable near infra-red fluorescent protein for in vivo imaging. *Nat. Biotechnol.* 29, 757–761. <https://doi.org/10.1038/nbt.1918>
- Guignard, L., Fiúza, U.-M., Leggio, B., Laussu, J., Faure, E., Michelin, G., Biasuz, K., Hufnagel, L., Malandain, G., Godin, C., Lemaire, P., 2020. Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis. *Science* 369.

<https://doi.org/10.1126/science.aar5663>

- Hallou, A., Yevick, H.G., Dumitrascu, B., Uhlmann, V., 2021. Deep learning for bioimage analysis in developmental biology. *Development* 148, dev199616. <https://doi.org/10.1242/dev.199616>
- Haupaix, N., Abitua, P.B., Sirour, C., Yasuo, H., Levine, M., Hudson, C., 2014. Ephrin-mediated restriction of ERK1/2 activity delimits the number of pigment cells in the *Ciona* CNS. *Dev. Biol.* 394, 170–180. <https://doi.org/10.1016/j.ydbio.2014.07.010>
- Haupaix, N., Stolfi, A., Sirour, C., Picco, V., Levine, M., Christiaen, L., Yasuo, H., 2013. p120RasGAP mediates ephrin/Eph-dependent attenuation of FGF/ERK signals during cell fate specification in ascidian embryos. *Dev. Camb. Engl.* 140, 4347–4352. <https://doi.org/10.1242/dev.098756>
- Hiratsuka, T., Fujita, Y., Naoki, H., Aoki, K., Kamioka, Y., Matsuda, M., 2015. Intercellular propagation of extracellular signal-regulated kinase activation revealed by in vivo imaging of mouse skin. *eLife* 4, e05178. <https://doi.org/10.7554/eLife.05178>
- Hudson, C., 2016. The central nervous system of ascidian larvae. *WIREs Dev. Biol.* 5, 538–561. <https://doi.org/10.1002/wdev.239>
- Hudson, C., Lotito, S., Yasuo, H., 2007. Sequential and combinatorial inputs from Nodal, Delta2/Notch and FGF/MEK/ERK signalling pathways establish a grid-like organisation of distinct cell identities in the ascidian neural plate. *Development* 134, 3527–3537. <https://doi.org/10.1242/dev.002352>
- Johnson, H.E., Toettcher, J.E., 2019. Signaling Dynamics Control Cell Fate in the Early *Drosophila* Embryo. *Dev. Cell* 48, 361–370.e3. <https://doi.org/10.1016/j.devcel.2019.01.009>
- Kudo, T., Jeknić, S., Macklin, D.N., Akhter, S., Hughey, J.J., Regot, S., Covert, M.W., 2018. Live-cell measurements of kinase activity in single cells using translocation reporters. *Nat. Protoc.* 13, 155–169. <https://doi.org/10.1038/nprot.2017.128>
- Lavoie, H., Gagnon, J., Therrien, M., 2020. ERK signalling: a master regulator of cell behaviour, life and fate. *Nat. Rev. Mol. Cell Biol.* 21, 607–632. <https://doi.org/10.1038/s41580-020-0255-7>
- Lavoie, H., Therrien, M., 2015. Regulation of RAF protein kinases in ERK signalling. *Nat. Rev. Mol. Cell Biol.* 16, 281–298. <https://doi.org/10.1038/nrm3979>
- Leggio, B., Laussu, J., Carlier, A., Godin, C., Lemaire, P., Faure, E., 2019. MorphoNet: an interactive online morphological browser to explore complex multi-scale data. *Nat. Commun.* 10, 2812. <https://doi.org/10.1038/s41467-019-10668-1>
- Lemaire, P., 2009. Unfolding a chordate developmental program, one cell at a time: Invariant cell lineages, short-range inductions and evolutionary plasticity in ascidians. *Dev. Biol.* 332, 48–60. <https://doi.org/10.1016/j.ydbio.2009.05.540>
- Li, P., Elowitz, M.B., 2019. Communication codes in developmental signaling pathways. *Development* 146, dev170977. <https://doi.org/10.1242/dev.170977>
- Lim, B., Dsilva, C.J., Levario, T.J., Lu, H., Schüpbach, T., Kevrekidis, I.G., Shvartsman, S.Y., 2015. Dynamics of Inductive ERK Signaling in the *Drosophila* Embryo. *Curr. Biol.* CB 25, 1784–1790. <https://doi.org/10.1016/j.cub.2015.05.039>
- Marshall, C.J., 1995. Signaling: Transient versus Sustained Extracellular Signal-Regulated Kinase Activation. *Cell* 80, 179–185.
- Mayr, V., Sturtzel, C., Stadler, M., Grissenberger, S., Distel, M., 2018. Fast Dynamic in vivo Monitoring of Erk Activity at Single Cell Resolution in DREKA Zebrafish. *Front. Cell Dev. Biol.* 6. <https://doi.org/10.3389/fcell.2018.00111>
- Moreno, E., Valon, L., Levillayer, F., Levayer, R., 2019. Competition for Space Induces Cell Elimination through Compaction-Driven ERK Downregulation. *Curr. Biol.* 29, 23–34.e8. <https://doi.org/10.1016/j.cub.2018.11.007>
- Nakamura, A., Goto, Y., Kondo, Y., Aoki, K., 2021. Shedding light on developmental ERK signaling with genetically encoded biosensors. *Development* 148, dev199767. <https://doi.org/10.1242/dev.199767>
- Nicol, D., Meinertzhagen, I.A., 1988. Development of the Central Nervous System of the Larva of the Ascidian, *Ciona intestinalis* L. 30.

- Nishida, H., 2003. Spatio-temporal pattern of MAP kinase activation in embryos of the ascidian *Halocynthia roretzi*. *Dev. Growth Differ.* 45, 27–37. <https://doi.org/10.1046/j.1440-169X.2003.00672.x>
- Patel, A.L., Shvartsman, S.Y., 2018. Outstanding questions in developmental ERK signaling. *Dev. Camb. Engl.* 145, dev143818. <https://doi.org/10.1242/dev.143818>
- Perrimon, N., Pitsouli, C., Shilo, B.-Z., 2012. Signaling Mechanisms Controlling Cell Fate and Embryonic Patterning. *Cold Spring Harb. Perspect. Biol.* 4, a005975–a005975. <https://doi.org/10.1101/cshperspect.a005975>
- Pokrass, M.J., Ryan, K.A., Xin, T., Pielstick, B., Timp, W., Greco, V., Regot, S., 2020. Cell-Cycle-Dependent ERK Signaling Dynamics Direct Fate Specification in the Mammalian Preimplantation Embryo. *Dev. Cell* 55, 328–340.e5. <https://doi.org/10.1016/j.devcel.2020.09.013>
- Rauen, K.A., 2013. The RASopathies. *Annu. Rev. Genomics Hum. Genet.* 14, 355–369. <https://doi.org/10.1146/annurev-genom-091212-153523>
- Razy-Krajka, F., Gravez, B., Kaplan, N., Racioppi, C., Wang, W., Christiaen, L., 2018. An FGF-driven feed-forward circuit patterns the cardiopharyngeal mesoderm in space and time. *eLife* 7, e29656. <https://doi.org/10.7554/eLife.29656>
- Razy-Krajka, F., Lam, K., Wang, W., Stolfi, A., Joly, M., Bonneau, R., Christiaen, L., 2014. Collier/OLF/EBF-Dependent Transcriptional Dynamics Control Pharyngeal Muscle Specification from Primed Cardiopharyngeal Progenitors. *Dev. Cell* 29, 263–276. <https://doi.org/10.1016/j.devcel.2014.04.001>
- Regot, S., Hughey, J.J., Bajar, B.T., Carrasco, S., Covert, M.W., 2014. High-Sensitivity Measurements of Multiple Kinase Activities in Live Single Cells. *Cell* 157, 1724–1734. <https://doi.org/10.1016/j.cell.2014.04.039>
- Roure, A., Rothbacher, U., Robin, F., Kalmar, E., Ferone, G., Lamy, C., Missero, C., Mueller, F., Lemaire, P., 2007. A multicassette Gateway vector set for high throughput and comparative analyses in *Ciona* and vertebrate embryos. *PloS One* 2, e916. <https://doi.org/10.1371/journal.pone.0000916>
- Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W.K., Luna, A., La, K.C., Dimitriadoy, S., Liu, D.L., Kantheti, H.S., Saghafeina, S., Chakravarty, D., Daian, F., Gao, Q., Bailey, M.H., Liang, W.-W., Foltz, S.M., Shmulevich, I., Ding, L., Heins, Z., Ochoa, A., Gross, B., Gao, J., Zhang, H., Kundra, R., Kandoth, C., Bahceci, I., Dervishi, L., Dogrusoz, U., Zhou, W., Shen, H., Laird, P.W., Way, G.P., Greene, C.S., Liang, H., Xiao, Y., Wang, C., Iavarone, A., Berger, A.H., Bivona, T.G., Lazar, A.J., Hammer, G.D., Giordano, T., Kwong, L.N., McArthur, G., Huang, C., Tward, A.D., Frederick, M.J., McCormick, F., Meyerson, M., Van Allen, E.M., Cherniack, A.D., Ciriello, G., Sander, C., Schultz, N., 2018. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* 173, 321–337.e10. <https://doi.org/10.1016/j.cell.2018.03.035>
- Santos, S.D.M., Verveer, P.J., Bastiaens, P.I.H., 2007. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nat. Cell Biol.* 9, 324–330. <https://doi.org/10.1038/ncb1543>
- Sardet, C., McDougall, A., Yasuo, H., Chenevert, J., Pruliere, G., Dumollard, R., Hudson, C., Hebras, C., Le Nguyen, N., Paix, A., 2011. Embryological Methods in Ascidians: The Villefranche-sur-Mer Protocols, in: Pelegri, F.J. (Ed.), *Vertebrate Embryogenesis, Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp. 365–400. https://doi.org/10.1007/978-1-61779-210-6_14
- Shi, W., Levine, M., 2008. Ephrin signaling establishes asymmetric cell fates in an endomesoderm lineage of the *Ciona* embryo. *Development* 135, 931–940. <https://doi.org/10.1242/dev.011940>
- Shilo, B.-Z., Barkai, N., 2017. Lighting Up ERK Activity. *Dev. Cell* 40, 115–116. <https://doi.org/10.1016/j.devcel.2016.12.016>
- Shin, S.-Y., Rath, O., Choo, S.-M., Fee, F., McFerran, B., Kolch, W., Cho, K.-H., 2009. Positive- and negative-feedback regulations coordinate the dynamic behavior of the Ras-Raf-MEK-ERK

- signal transduction pathway. *J. Cell Sci.* 122, 425–435. <https://doi.org/10.1242/jcs.036319>
- Simon, C.S., Rahman, S., Raina, D., Schröter, C., Hadjantonakis, A.-K., 2020. Live Visualization of ERK Activity in the Mouse Blastocyst Reveals Lineage-Specific Signaling Dynamics. *Dev. Cell* 55, 341–353.e5. <https://doi.org/10.1016/j.devcel.2020.09.030>
- Tolkin, T., Christiaen, L., 2016. Rewiring of an ancestral Tbx1/10-Ebf-Mrf network for pharyngeal muscle specification in distinct embryonic lineages. *Development* 143, 3852–3862. <https://doi.org/10.1242/dev.136267>
- Williaume, G., de Buyl, S., Sirour, C., Haupaix, N., Bettoni, R., Imai, K.S., Satou, Y., Dupont, G., Hudson, C., Yasuo, H., 2021. Cell geometry, signal dampening, and a bimodal transcriptional response underlie the spatial precision of an ERK-mediated embryonic induction. *Dev. Cell.* <https://doi.org/10.1016/j.devcel.2021.09.025>
- Wolpert, L., Tickle, C., Arias, A.M., 2015. *Principles of Development*. Oxford University Press.
- Yasuo, H., Hudson, C., 2007. FGF8/17/18 functions together with FGF9/16/20 during formation of the notochord in *Ciona* embryos. *Dev. Biol.* 302, 92–103. <https://doi.org/10.1016/j.ydbio.2006.08.075>

ACKNOWLEDGEMENTS

We thank the members of the Tunicate team at CRBM for fruitful discussions throughout the course and more particularly Mathieu Bonrepaux for his contribution to the development of algorithmic methods to segment the nuclei. We thank Gregoire Malandain and Emmanuel Faure for their help in reconstructing the embryos and correcting the segmentations respectively, and their very enriching discussions. We gratefully acknowledge Marc Plays' and Philippe Richard's expert animal husbandry. K.B. was funded by an Epigenmed (ANR-10-LABX-12-01) doctoral contract and by a University of Montpellier contract extension in compensation for Covid-19 restrictions and lab closure. This work was supported by the Dig-Em (ANR-14-CE11-0013-01), Cell-whisper (ANR-19-CE13-0020) and Morphoscope (ANR-11EQPX-0029) grants of the Agence Nationale de la recherche (ANR). We acknowledge the support and help of the Montpellier Resource Imaging (MRI) facility and in particular of Sylvain de Rossi and Volker Baecker. MRI is a core facility of the France-BioImaging infrastructure, funded by the ANR (ANR-10-INBS-04).

SUPPLEMENTAL INFORMATION

Embryo ID	Labels	#Time points	#Cells (begin)	#Cells (end)	#Total cells with measured ERK activity	#Cell snapshots with measured ERK activity*	Dev. Stage (begin)	Dev. Stage (end)	Direct link MorphoNet
Astec-Pm16	PH-GFP H2B-mCherry ERK-KTR-iRFP	150	32	698	764	26480	St.6b late 32-cell	St.16 late neurula	https://morphonet.org/RB9nb4LZ
Astec-Pm17	PH-GFP H2B-mCherry ERK-KTR-iRFP	60	32	204	360	2707	St.6b late 32-cell	St.12 mid gastrula	https://morphonet.org/vnYp4afL
Astec-Pm18	PH-GFP H2B-mCherry ERK-KTR-iRFP	50	44	180	324	2408	St.7 44-cell	St.12 mid gastrula	https://morphonet.org/KO0jbXhX
Astec-Pm19	PH-GFP H2B-mCherry ERK-KTR-iRFP	35	64	183	302	1669	St.8 64-cell	St.12 mid gastrula	https://morphonet.org/XuF5QCcC

Table S1 : *Phallusia mammillata* dataset. The column "Labels" indicates the fluorescent proteins expressed by each embryo. # = number. *: The snapshots corresponding to the first two and the last eight time points of the life of each cell are not counted.

Cell names	References
A7.1	(Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008)
A7.2	(Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008)
B7.1	(Nishida, 2003; Yasuo and Hudson, 2007)
B7.2	(Nishida, 2003; Yasuo and Hudson, 2007)
B7.7	(Nishida, 2003; Shi and Levine, 2008)
a7.9	(Nishida, 2003; Hudson, 2016)
a7.10	(Nishida, 2003; Hudson, 2016)
b7.9	(Nishida, 2003; Hudson, 2016)
b7.10	(Nishida, 2003; Hudson, 2016)
A8.5	(Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008)
A8.6	(Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008)
A8.13	(Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008)
A8.14	(Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008)
B8.6	(Nishida, 2003; Yasuo and Hudson, 2007)

Table S2 : List of dp-ERK positive cells experimentally determined both in *Halocynthia roretzi* and *Ciona*.

Mother cell name	Sister cells name	dp-ERK signal	References
a5.3	a6.5	positive	(Nishida, 2003; Williaume et al., 2021)
	a6.6	negative	
a5.4	a6.7	positive	(Nishida, 2003; Williaume et al., 2021)
	a6.8	negative	
b5.3	b6.5	positive	(Nishida, 2003; Williaume et al., 2021)
	b6.6	negative	
A6.2	A7.3	positive	(Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008; Haupaix et al., 2013)
	A7.4	negative	
A6.4	A7.7	positive	(Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008; Haupaix et al., 2013)
	A7.8	negative	
B6.2	B7.3	positive	(Nishida, 2003; Yasuo and Hudson, 2007; Shi and Levine, 2008)
	B7.4	negative	
a6.7	a7.13	positive	(Nishida, 2003; Hudson, 2016)
	a7.14	negative	
a7.9	a8.17	positive	(Nishida, 2003; Hudson, 2016)
	a8.18	negative	
a7.10	a8.19	positive	(Nishida, 2003; Hudson, 2016)
	a8.20	negative	
a7.13	a8.25	positive	(Nishida, 2003; Hudson, 2016)
	a8.26	negative	
A8.7	A9.13	positive	(Nishida, 2003; Haupaix et al., 2014; Hudson, 2016)
	A9.14	negative	
A8.8	A9.15	positive	(Nishida, 2003; Haupaix et al., 2014; Hudson, 2016)
	A9.16	negative	
A8.15	A9.29	positive	(Nishida, 2003; Haupaix et al., 2014; Hudson, 2016)
	A9.30	negative	
A8.16	A9.31	positive	(Cole and Meinertzhagen, 2004; Hudson, 2016)
	A9.32	negative	
a8.17	a9.33	positive	(Nishida, 2003; Haupaix et al., 2014; Hudson, 2016)
	a9.34	negative	
a8.19	a9.37	positive	(Nishida, 2003; Haupaix et al., 2014; Hudson, 2016)
	a9.38	negative	
a8.25	a9.49	positive	(Nishida, 2003; Haupaix et al., 2014; Hudson, 2016)
	a9.50	negative	
a9.33	a10.65	positive	(Haupaix et al., 2014; Hudson, 2016)
	a10.66	negative	
a9.34	a10.67	positive	(Haupaix et al., 2014; Hudson, 2016)
	a10.68	negative	
a9.37	a10.73	positive	(Haupaix et al., 2014; Hudson, 2016)
	a10.74	negative	

a9.38	a10.75	positive	(Haupaix et al., 2014; Hudson, 2016)
	a10.76	negative	
a10.49	a10.97	positive	(Haupaix et al., 2014; Hudson, 2016)
	a10.98	negative	
a10.50	a10.99	positive	(Haupaix et al., 2014; Hudson, 2016)
	a10.100	negative	

Table S3 : List of differential dp-ERK positive/dp-ERK negative sister cells experimentally determined both in *Halocynthia roretzi* and *Ciona*.

		Inferred ERK status from average ERK activity measurements			
Cell name	dp-ERK signal from literature (Table S2; Table S3)	Astec-Pm16	Astec-Pm17	Astec-Pm18	Astec-Pm19
a6.5	positive	NA	NA	NA	NA
a6.7	positive	NA	NA	NA	NA
b6.5	positive	NA	NA	NA	NA
A7.1	positive	positive	positive	positive	NA
A7.2	positive	positive	positive	positive	NA
B7.1	positive	positive	positive	positive	NA
B7.2	positive	positive	positive	positive	NA
B7.7	positive	positive	positive	positive	NA
a7.9	positive	positive	positive	positive	NA
a7.10	positive	positive	positive	negative	NA
b7.9	positive	negative	negative	negative	NA
b7.10	positive	positive	negative	negative	NA
A8.5	positive	positive	positive	positive	positive
A8.6	positive	positive	positive	negative	positive
A8.13	positive	positive	negative	negative	positive
A8.14	positive	positive	negative	positive	positive
B8.6	positive	positive	positive	positive	positive
A7.3	positive	positive	positive	positive	NA
A7.4	negative	negative	negative	negative	NA
A7.7	positive	positive	positive	positive	NA
A7.8	negative	negative	negative	negative	NA
B7.3	positive	positive	positive	positive	NA
B7.4	negative	negative	negative	negative	NA

a7.13	positive	positive	positive	positive	NA
a7.14	negative	negative	negative	negative	NA
a8.17	positive	positive	positive	positive	positive
a8.18	negative	negative	negative	negative	negative
a8.19	positive	positive	positive	positive	positive
a8.20	negative	negative	negative	negative	negative
a8.25	positive	positive	positive	positive	positive
a8.26	negative	negative	negative	negative	negative
A9.13	positive	positive	NA	NA	NA
A9.14	negative	negative	NA	NA	NA
A9.15	positive	positive	NA	NA	NA
A9.16	negative	negative	NA	NA	NA
A9.29	positive	positive	NA	NA	NA
A9.30	negative	negative	NA	NA	NA
A9.31	positive	positive	NA	NA	NA
A9.32	negative	negative	NA	NA	NA
a9.33	positive	positive	NA	NA	NA
a9.34	negative	negative	NA	NA	NA
a9.37	positive	positive	NA	NA	NA
a9.38	negative	negative	NA	NA	NA
a9.49	positive	positive	NA	NA	NA
a9.50	negative	negative	NA	NA	NA
a10.65	positive	positive	NA	NA	NA
a10.66	negative	negative	NA	NA	NA
a10.67	positive	positive	NA	NA	NA
a10.68	negative	negative	NA	NA	NA
a10.73	positive	positive	NA	NA	NA
a10.74	negative	negative	NA	NA	NA

a10.75	positive	positive	NA	NA	NA
a10.76	negative	negative	NA	NA	NA
a10.97	positive	positive	NA	NA	NA
a10.98	negative	negative	NA	NA	NA
a10.99	positive	NA	NA	NA	NA
a10.100	negative	NA	NA	NA	NA

Table S4 : Comparison of ERK activity status between what is known in the literature and *in vivo* measurements of ERK activity from the ERK-KTR.

The list of known positive and negative dp-ERK cells is from tables S2 and S3. In our dataset, ERK status was inferred from measurements of average ERK activity in each left/right cell pair by comparing its value to thresholds we established in each embryo for each generation (Figures S4-S15). NA: indicate cells for which there is no measurement: either because they have not been imaged or because they have been imaged over too short a period, less than 14 time steps, not allowing enough values to establish a reliable average measurement. These 14 time steps were specified to include the first 2 time steps after cell appearance and the 8 time steps before cell division and at least 4 time steps for averaging. Bolded are cells for which the status inferred from the measurement of ERK activity *in vivo* is different from what is known from dp-ERK imaging.

Mother cell name	Sister cell name	Inferred ERK status from average ERK activity measurements				References
		Astec-Pm16	Astec-Pm17	Astec-Pm18	Astec-Pm19	
A7.6	A8.11	Negative	Negative	Negative	Negative	(Tolkin and Christiaen, 2016)
	A8.12	Positive	Positive	Positive	Positive	
A8.12	A9.23	Negative	NA	NA	NA	
	A9.24	Positive	NA	NA	NA	
A9.23	A10.45	Positive	NA	NA	NA	
	A10.46	Negative	NA	NA	NA	
B6.4	B7.7	Positive	Positive	Positive	NA	(Nishida, 2003)
	B7.8	Negative	Negative	Negative	NA	
B8.7	B9.13	Positive	NA	NA	NA	(Nishida, 2003)
	B9.14	Negative	NA	NA	NA	
B8.9	B9.17	Negative	NA	NA	NA	(Razy-Krajka et al., 2014) (Biasuz et al., 2022, <i>in prep</i>)
	B9.18	Positive	NA	NA	NA	
B8.10	B9.19	Negative	NA	NA	NA	
	B9.20	Positive	NA	NA	NA	

Table S5 : Sister cells with differential inferred ERK status from measurements of mean ERK activity and thresholding method for which status is known in *Halocynthia roretzi* or in *Ciona intestinalis*.

The status inferred from the measurement of ERK activity *in vivo* is consistent from what is known from dp-ERK imaging for all the cells of the list. ERK status was inferred from measurements of average ERK activity in each left/right cell pair by comparing its value to thresholds we established in each embryo for each generation (Figures S4-15). NA: indicate cells for which there is no measurement: either because they have not been imaged or because they have been imaged over too short a period, less than 14 time steps, not allowing enough values to establish a reliable average measurement. These 14 time steps were specified to include the first 2 time steps after cell appearance and the 8 time steps before cell division and at least 4 time steps for averaging.

		Inferred ERK status from average ERK activity measurements			
Mother cell name	Sister cell name	Astec-Pm16	Astec-Pm17	Astec-Pm18	Astec-Pm19
a6.8	a7.15	negative	ND	ND	NA
	a7.16	positive	ND	ND	NA
b6.8	b7.15	negative	ND	ND	NA
	b7.16	positive	ND	ND	NA
a7.14	a8.27	ND	Negative	ND	Positive
	a8.28	ND	Positive	ND	Negative
B7.8	B8.15	ND	Positive	Positive	Positive
	B8.16	ND	Negative	Negative	Negative
b7.9	b8.17	ND	Positive	ND	Positive
	b8.18	ND	Negative	ND	Negative
b7.10	b8.19	Positive	Positive	ND	Positive
	b8.20	Negative	Negative	ND	Negative
b7.11	b8.21	Positive	Positive	ND	Positive
	b8.22	Negative	Negative	ND	Negative
b7.13	b8.25	ND	Positive	Positive	Positive
	b8.26	ND	Negative	Negative	Negative
b7.14	b8.27	ND	ND	Positive	NA
	b8.28	ND	ND	Negative	NA
A8.4	A9.7	Negative	NA	NA	NA
	A9.8	Positive	NA	NA	NA
B8.5	B9.9	Positive	NA	NA	NA
	B9.10	Negative	NA	NA	NA
A9.13	A10.25	Negative	NA	NA	NA
	A10.26	Positive	NA	NA	NA
A9.16	A10.31	Positive	NA	NA	NA

	A10.32	Negative	NA	NA	NA
A9.29	A10.57	Negative	NA	NA	NA
	A10.58	Positive	NA	NA	NA
A9.30	A10.59	Negative	NA	NA	NA
	A10.60	Positive	NA	NA	NA
A9.32	A10.63	Positive	NA	NA	NA
	A10.64	Negative	NA	NA	NA
a9.35	a10.69	Positive	NA	NA	NA
	a10.70	Negative	NA	NA	NA
a9.36	a10.71	Negative	NA	NA	NA
	a10.72	Positive	NA	NA	NA
a9.52	a10.103	Negative	NA	NA	NA
	a10.104	Positive	NA	NA	NA
B9.10	B10.19	Positive	NA	NA	NA
	B10.20	Negative	NA	NA	NA
b9.53	b10.105	Positive	NA	NA	NA
	b10.106	Negative	NA	NA	NA

Table S6 : Sister cells with differential inferred ERK status from measurements of mean ERK activity and thresholding method identifies novel candidates of ERK-dependent inductions.

ERK status was inferred from measurements of average ERK activity in each left/right cell pair by comparing its value to thresholds we established in each embryo for each generation (Figures S4-15). NA: cells for which there is no measurement: either because they have not been imaged or because they have been imaged over too short a period, less than 14 time steps, not allowing enough values to establish a reliable average measurement. These 14 time steps were specified to include the first 2 time steps after cell appearance and the 8 time steps before cell division and at least 4 time steps for averaging. ND : sister cells with no differential inferred ERK status.

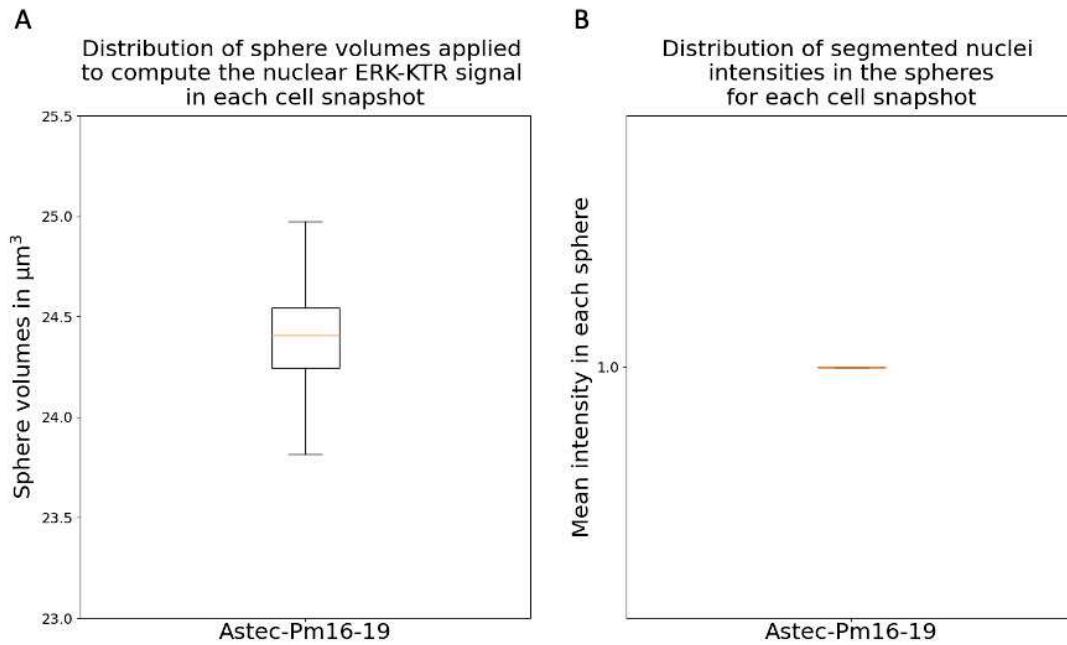
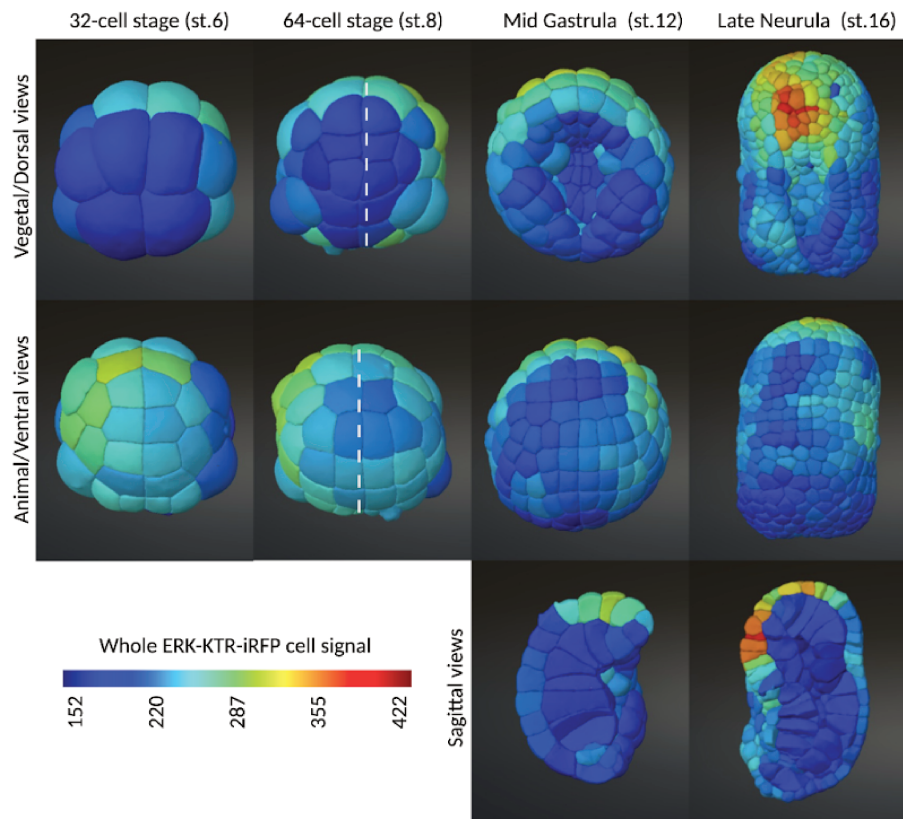


Figure S1 : High accuracy of the nuclear measurement of the ERK-KTR signal using a sphere of suitable volume.

A : After verifying that the ERK-KTR signal is homogeneous in the nucleus, we determined that a sphere with a radius of 6 voxels, equivalent to a volume of $24.42\mu\text{m}^3$ in our image, is large enough (it represents between 8 and 25% of the volumes of segmented nuclei present in our dataset) to reliably estimate the nuclear intensity of the ERK-KTR while ensuring that it is always within the entire development of the phallusia embryos. However, due to artefacts or malformed nuclei, during the sphere placement process, the spheres may have to be cut by the nucleus boundaries in order to ensure a nuclear measurement. Analysis of the volume distribution confirms that the majority of spheres remain broadly unchanged in size. The median value is $24.4\mu\text{m}^3$ and 99% of the 33264 snapshots in the dataset have a volume greater than $21\mu\text{m}^3$. As the sphere is made up of pixels, it is not surprising to observe some volumes slightly larger than $24.42\mu\text{m}^3$.

B : To systematically evaluate the accuracy of our sphere placement the automated sphere nuclear measurement was applied on the segmented nuclei images themselves. Indeed, a measurement of 1 is expected if the sphere is completely inside the segmented nucleus and 0 if the sphere is completely outside. As expected, the median value of the intensity measurement of the segmented nuclei in each sphere is 1 for the whole dataset of 33264 measured snapshots, confirming the high accuracy of the nuclear measurement with our strategy. For both analyses, the snapshots corresponding to the first two and the last eight time points of the life of each cell were excluded prior to the analyses. Top, middle, and bottom lines in the box plot indicate 75th percentile, median, and 25th percentile of data, respectively.

A



B

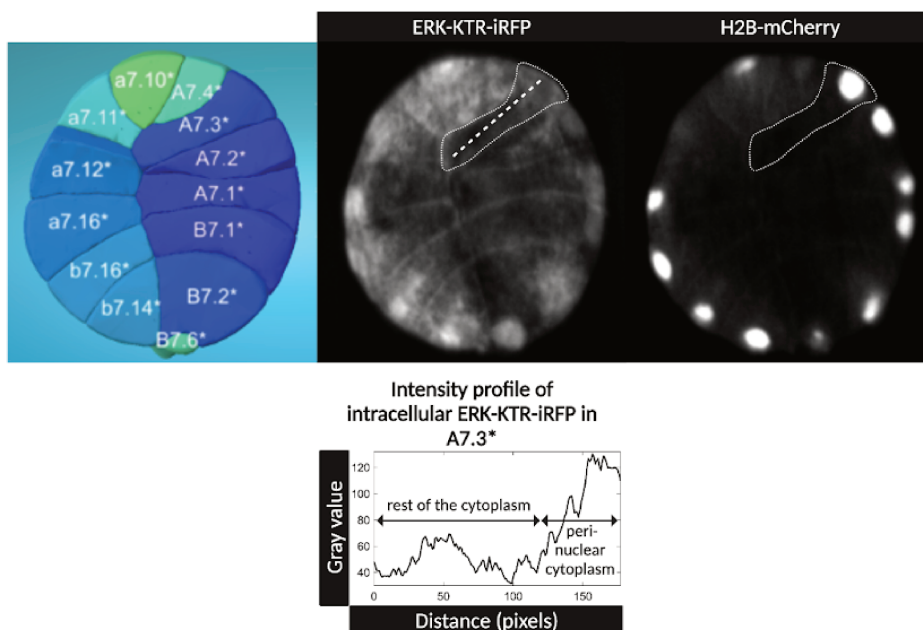


Figure S2 : Heterogeneity of cellular and intracellular distribution of ERK-KTR signal in *Phallusia mammillata* embryos. A) Different expression levels of ERK-KTR-iRFP between cells. Snapshots of the representative

Astec-Pm16 embryo showing the quantification of the mean ERK-KTR-iRFP cellular signal intensity for each cell at different developmental stages. The ERK-KTR cell signal is colour-coded and the scale bar is indicated. The white dotted lines on the 64-cell stage snapshots indicate the cross-section of the cells shown in B. B) Highly heterogeneous distribution of the ERK-KTR-iRFP in the cytoplasm. Left : cross-section of the reconstructed Astec-Pm16 embryo at the 64-cell stage with the colour-coded mean ERK-KTR-iRFP cellular intensities and cell names. Right : same sections showing ERK-KTR-iRFP and H2B-mCherry channels, revealing highly heterogeneous distribution of the ERK-KTR-iRFP in the cytoplasm. This is illustrated for the cell A7.3* (cell outline in dotted line): the ERK-KTR-iRFP intensity profile along the apical-basal major axis (white dotted line) shows a perinuclear region with a higher ERK-KTR-iRFP signal than the rest of the cytoplasm.

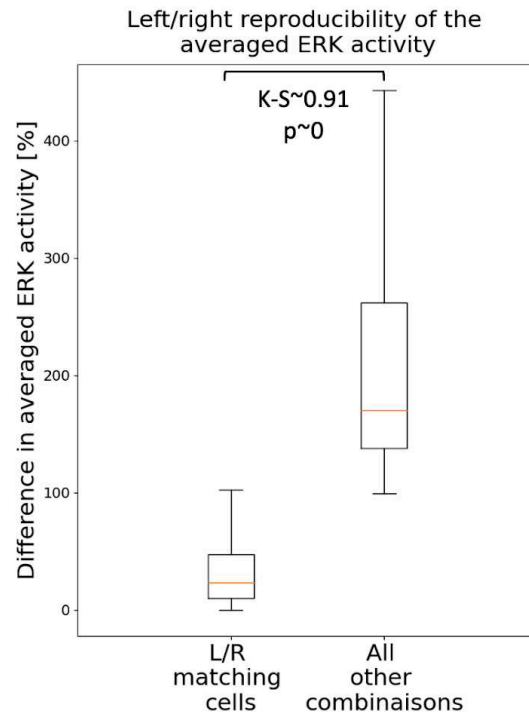


Figure S3 : Highly similar average ERK activity in bilaterally matching cells. In each embryo, the percentage difference between the mean ERK activities were calculated between matching left and right symmetric cells and all other possibilities. The distributions of the percentage differences in these two conditions for all 4 embryos are shown. Top, middle, and bottom lines in the box plot indicate 75th percentile, median, and 25th percentile of data, respectively. Comparison of the two distributions was performed according to the Kolmogorov-Smirnov (K-S) test showing that the two distributions are statistically different (the null hypothesis is that the two distributions are identical, K-S statistic = 0.91; p-value=0.0) (see methods).

Astec-Pm16 all generations

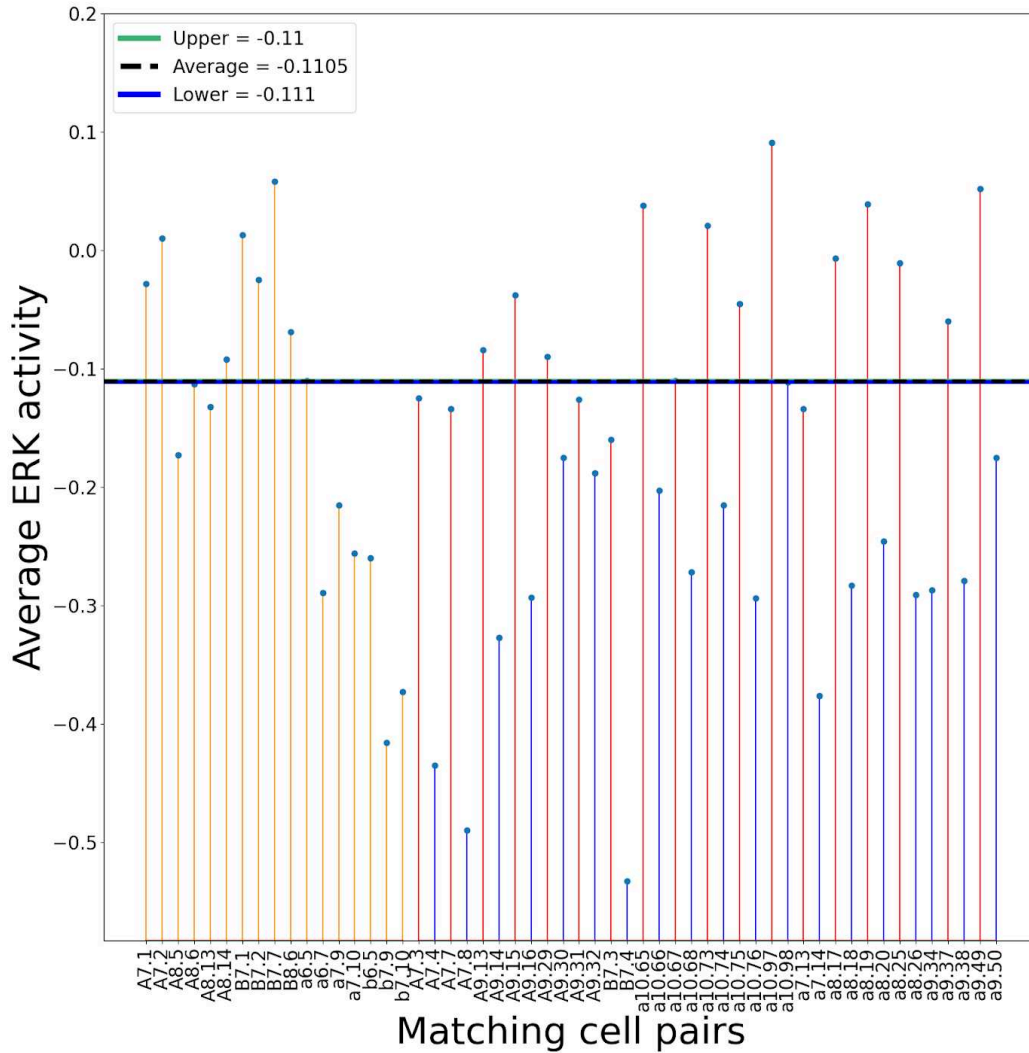


Figure S4 : Average ERK activity measured in Astec-Pm16 for cells belonging to the ground truth lists and determination of ERK status based on thresholding, all generations. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists all cell generations combined : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm16 7th generation

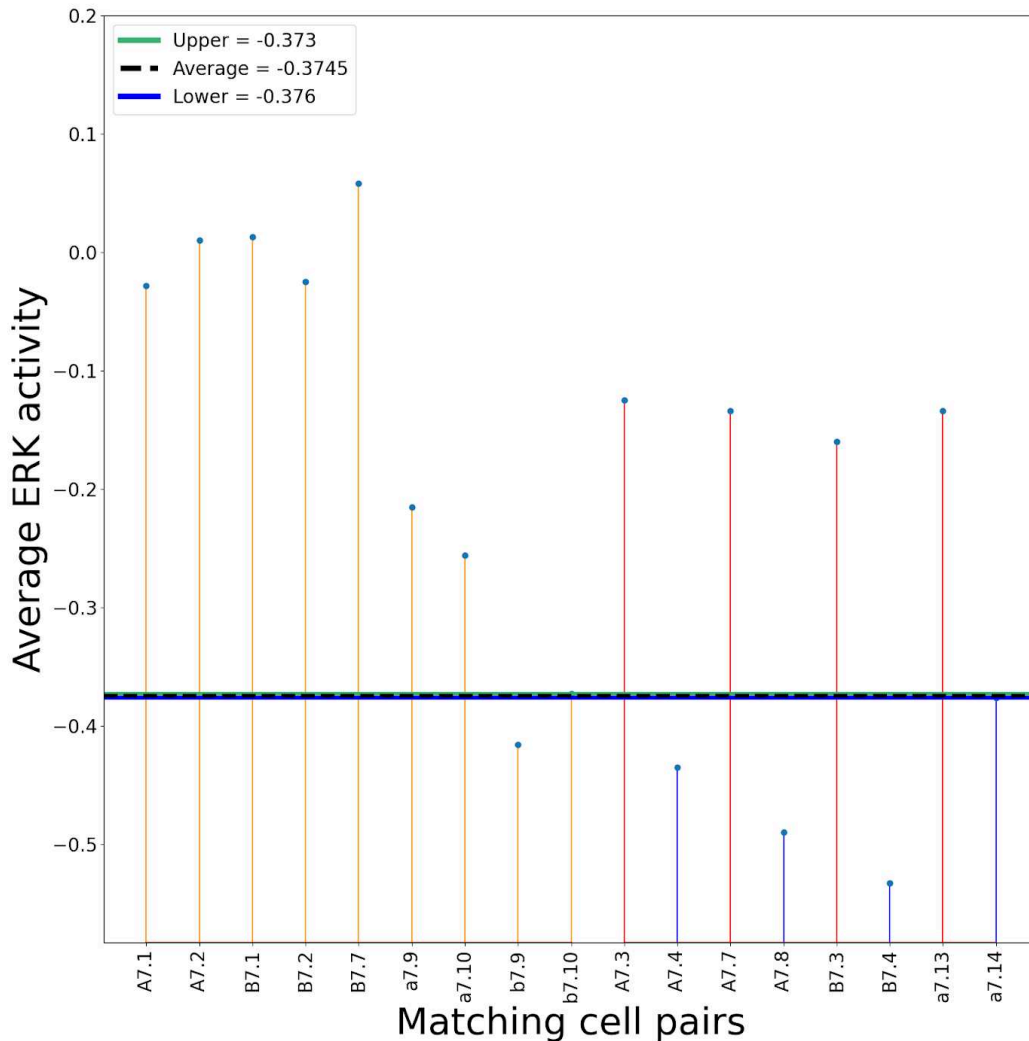


Figure S5 : Average ERK activity measured in Astec-Pm16 only for cells of the 7th generation belonging to the ground truth lists and determination of ERK status based on thresholding. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists of the 7th generation : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm16 8th generation

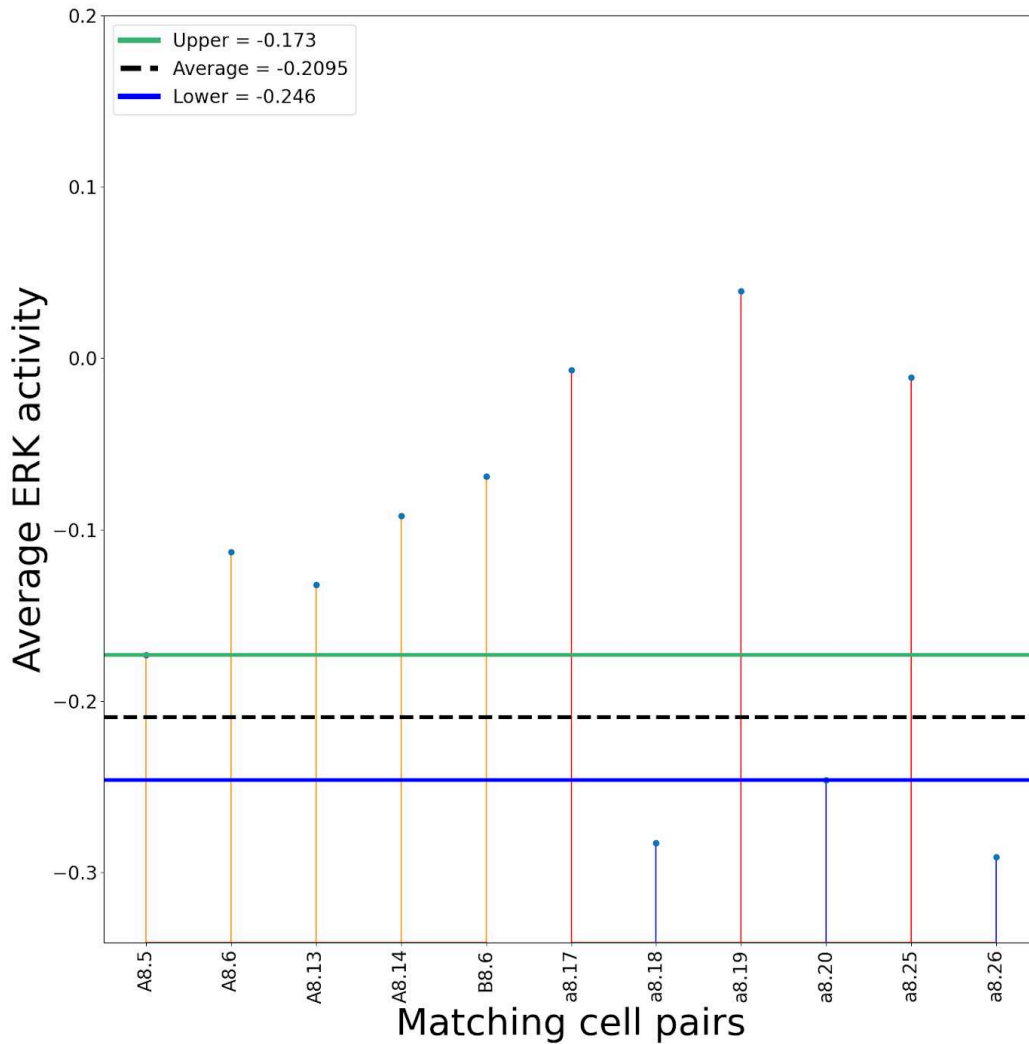


Figure S6 : Average ERK activity measured in Astec-Pm16 only for cells of the 8th generation belonging to the ground truth lists and determination of ERK status based on thresholding. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists of the 8th generation : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm16 9th generation

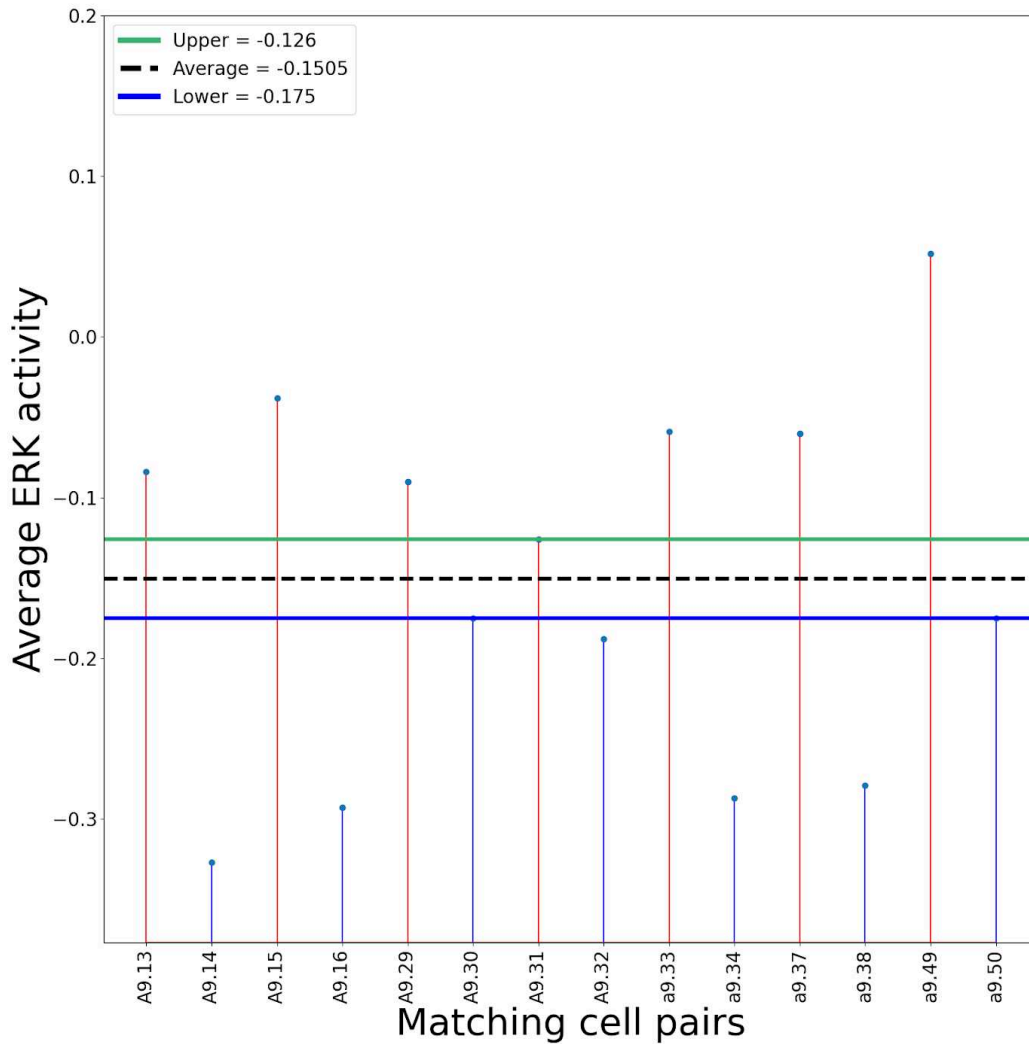


Figure S7 : Average ERK activity measured in Astec-Pm16 only for cells of the 9th generation belonging to the ground truth lists and determination of ERK status based on thresholding. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists of the 9th generation : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm16 10th generation

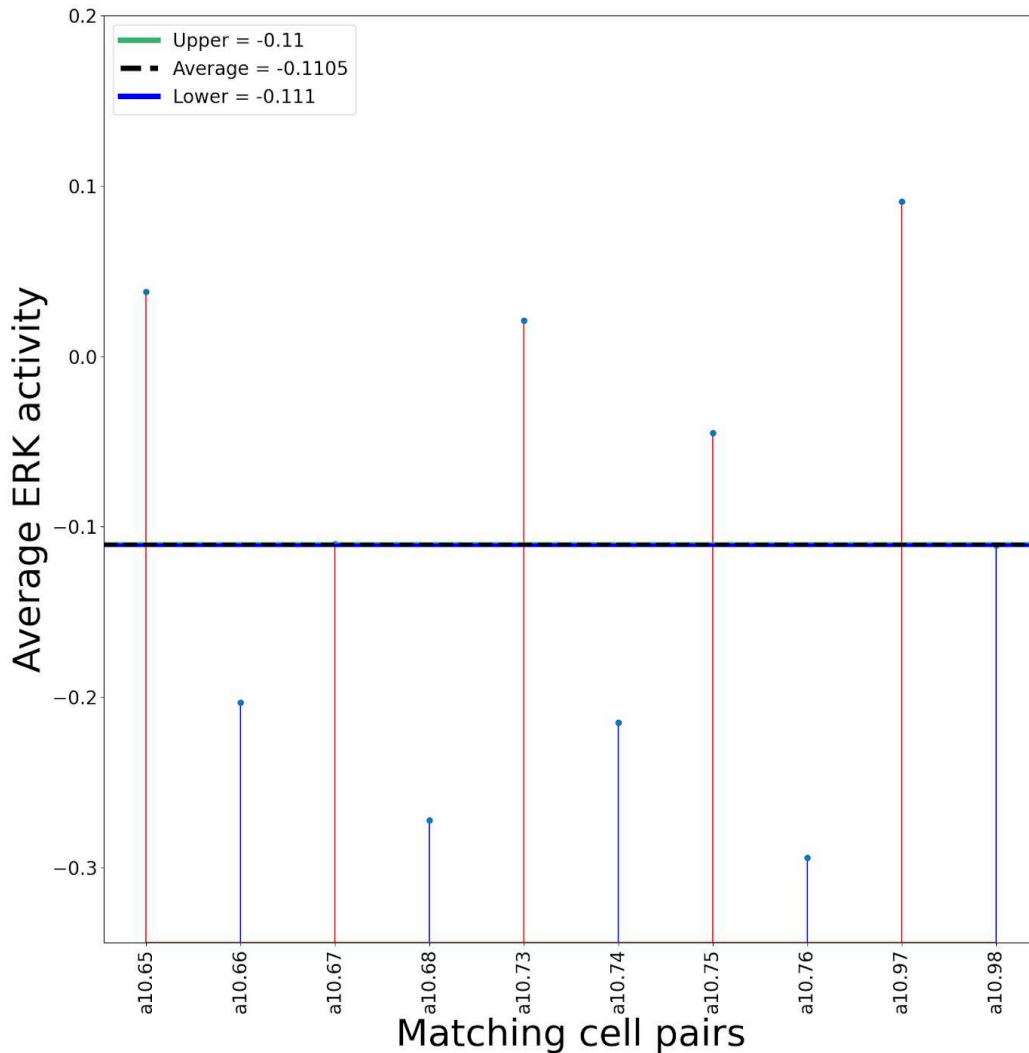


Figure S8 : Average ERK activity measured in Astec-Pm16 only for cells of the 10th generation belonging to the ground truth lists and determination of ERK status based on thresholding.

Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists of the 10th generation : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm17 all generations

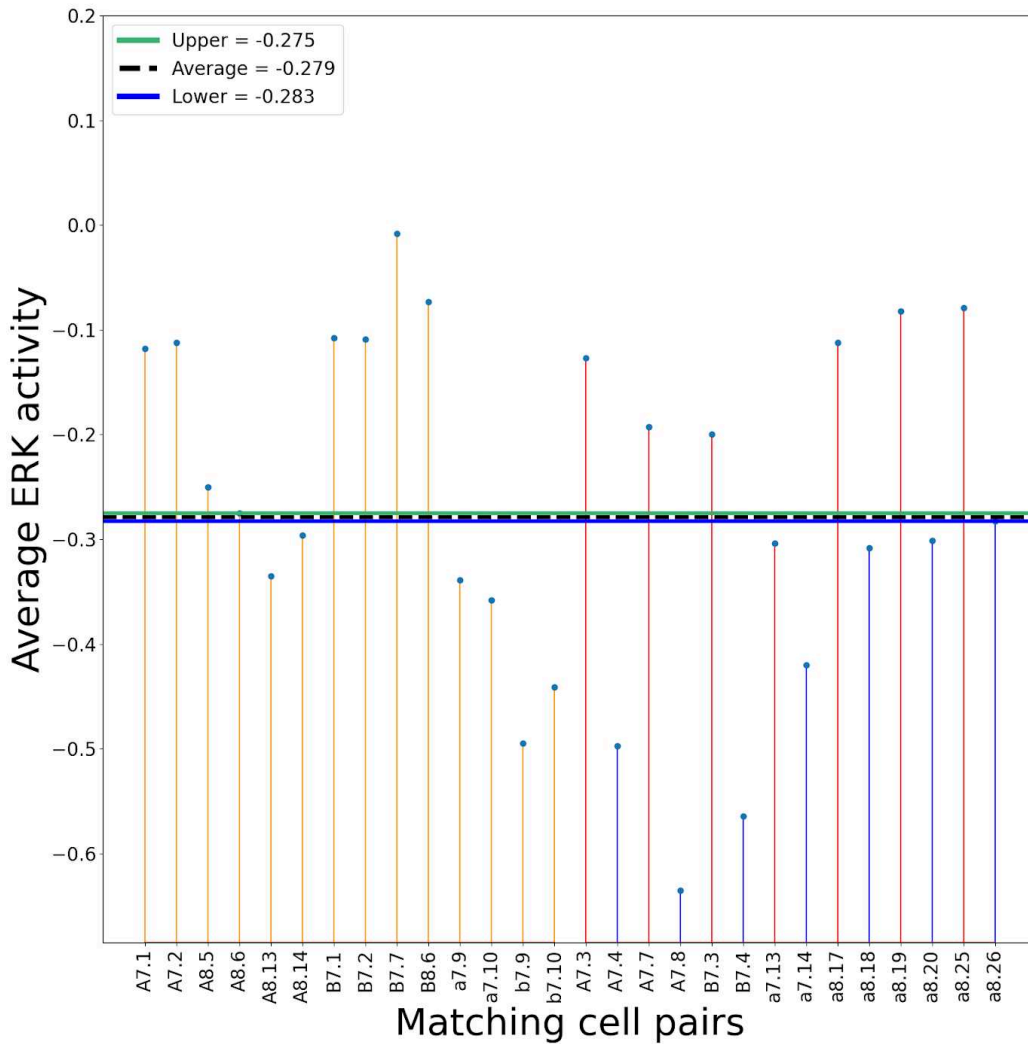


Figure S9 : Average ERK activity measured in Astec-Pm17 for cells belonging to the ground truth lists and determination of ERK status based on thresholding, all generations. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists all cell generations combined : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm17 7th generation

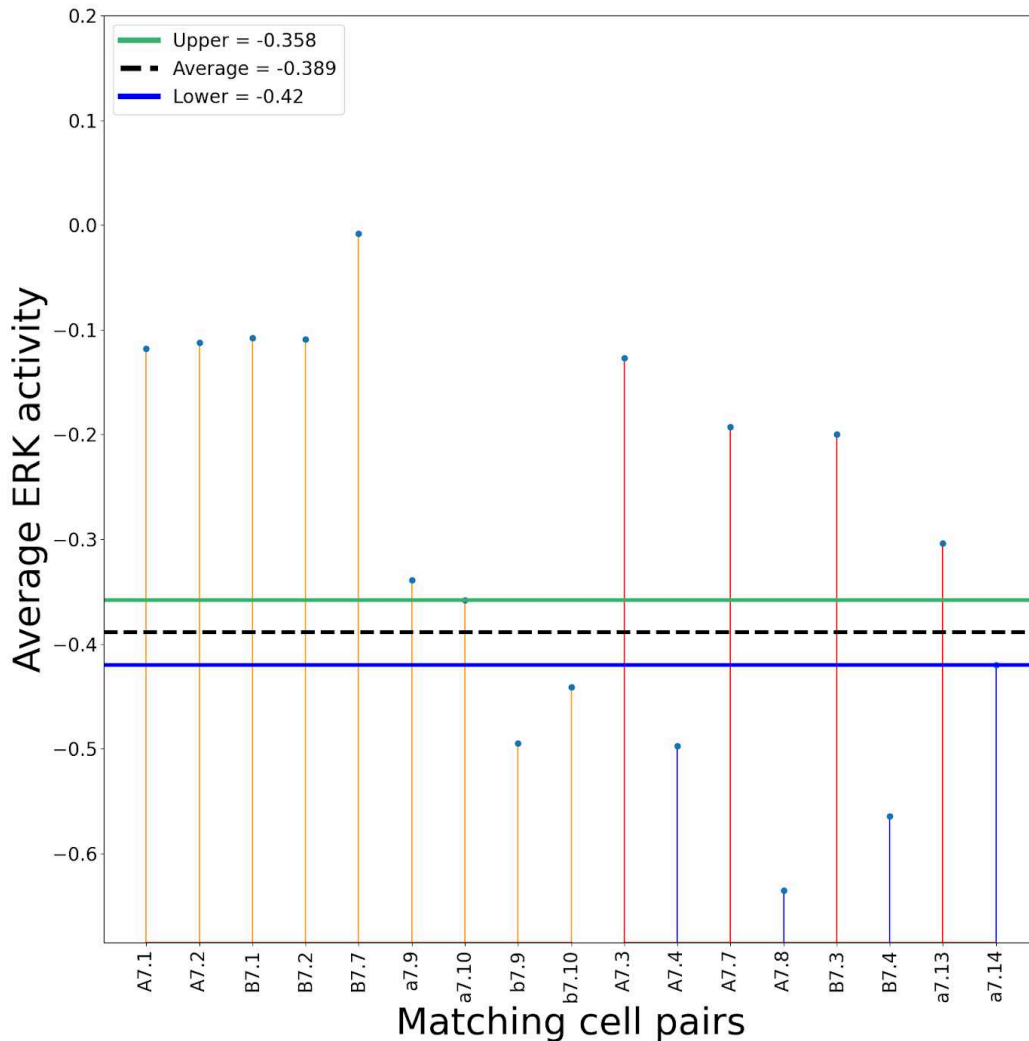


Figure S10 : Average ERK activity measured in Astec-Pm17 only for cells of the 7th generation belonging to the ground truth lists and determination of ERK status based on thresholding. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists of the 7th generation : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm17 8th generation

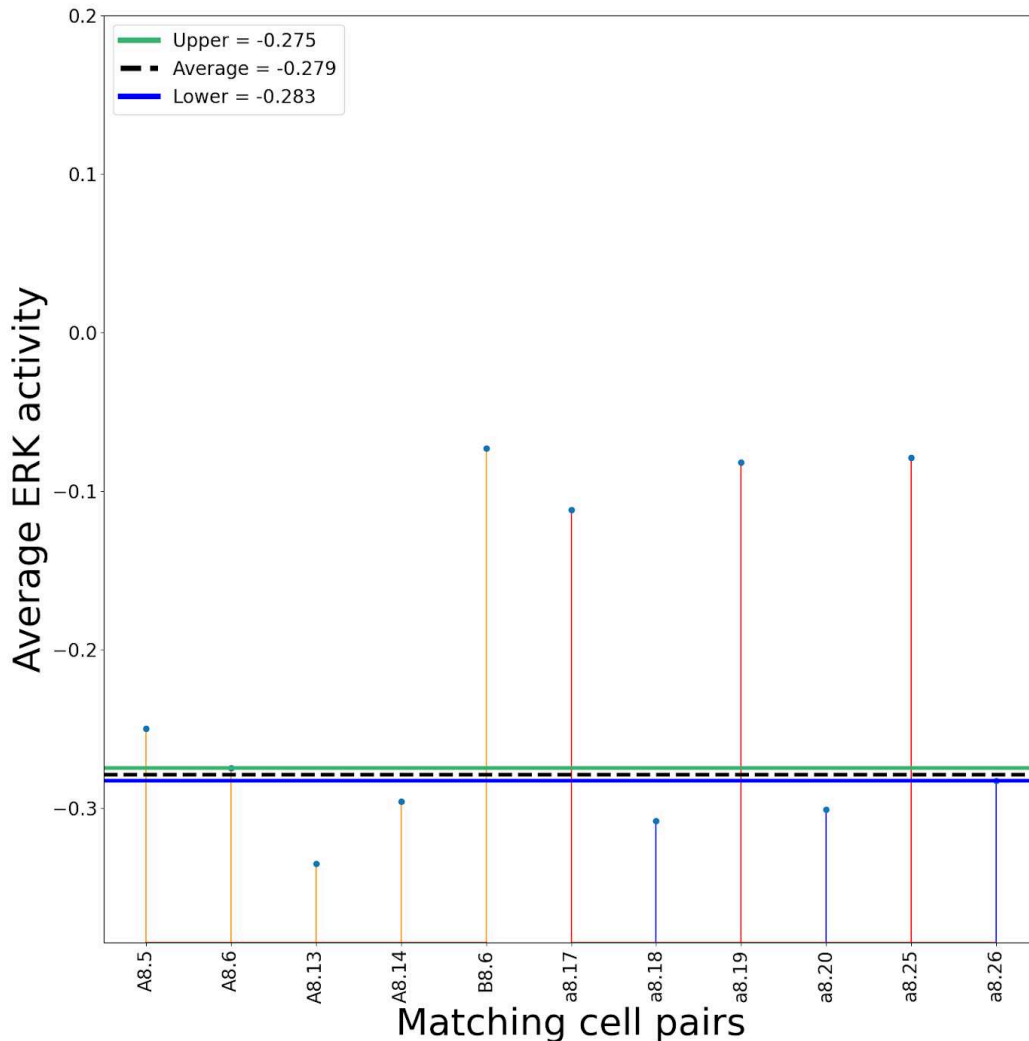


Figure S11 : Average ERK activity measured in Astec-Pm17 only for cells of the 8th generation belonging to the ground truth lists and determination of ERK status based on thresholding.

Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists of the 8th generation : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm18 all generations

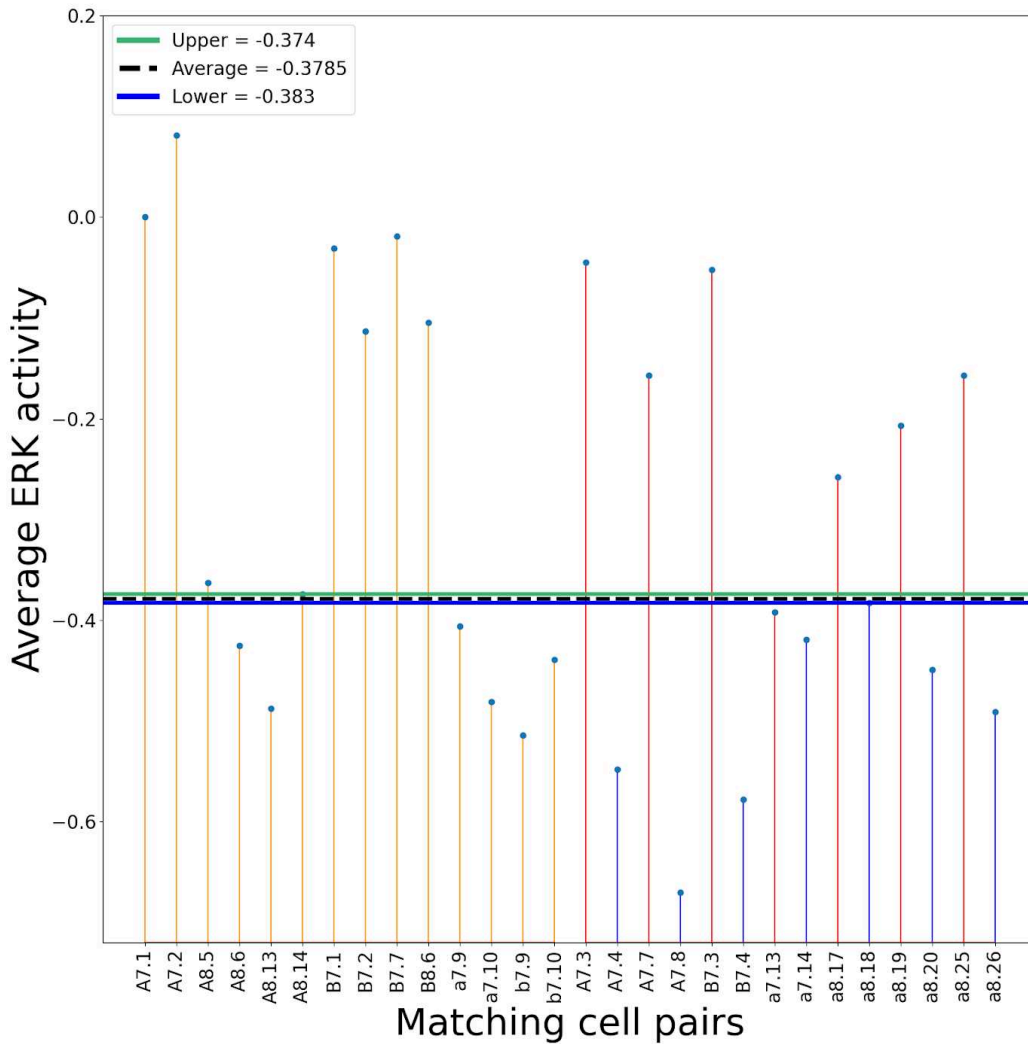


Figure S12 : Average ERK activity measured in Astec-Pm18 for cells belonging to the ground truth lists and determination of ERK status based on thresholding, all generations. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists all cell generations combined : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm18 7th generation

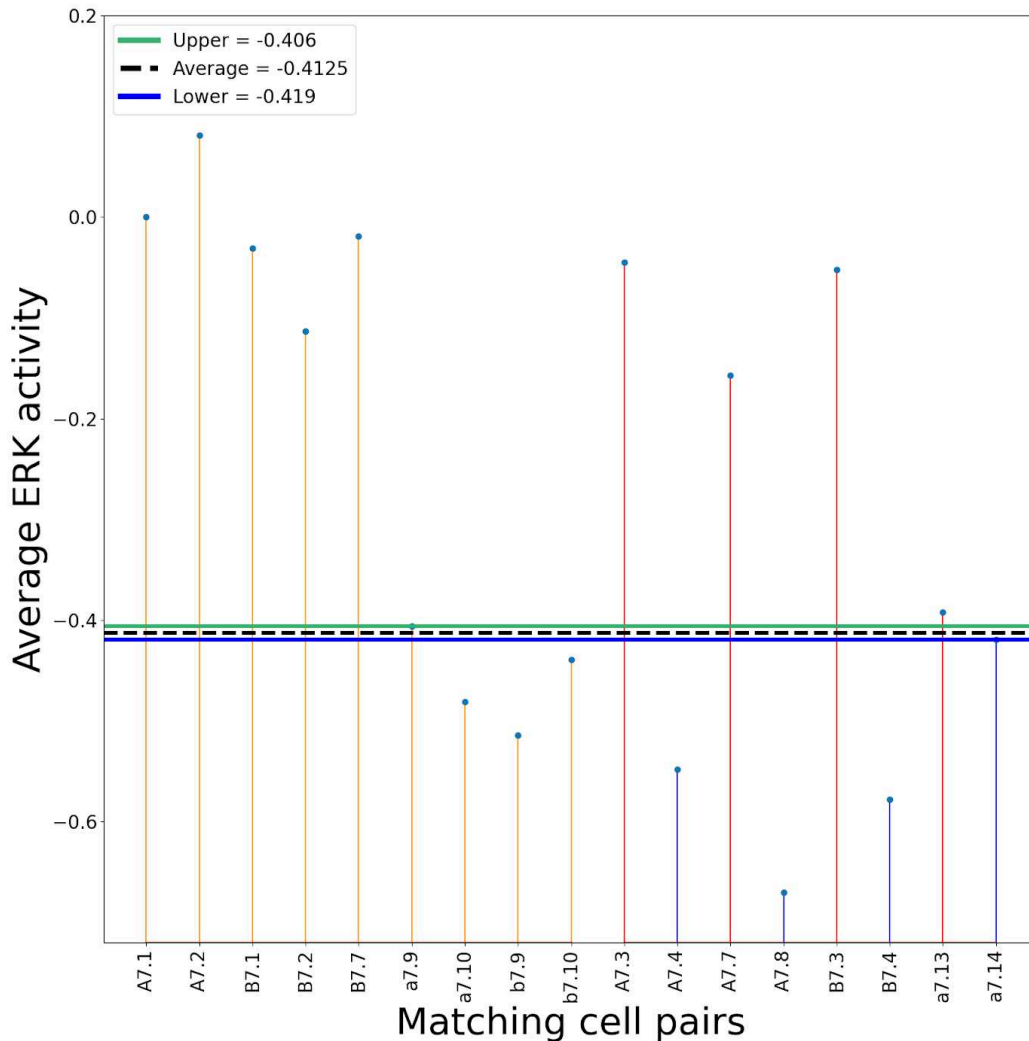


Figure S13 : Average ERK activity measured in Astec-Pm18 only for cells of the 7th generation belonging to the ground truth lists and determination of ERK status based on thresholding. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists of the 7th generation : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm18 8th generation

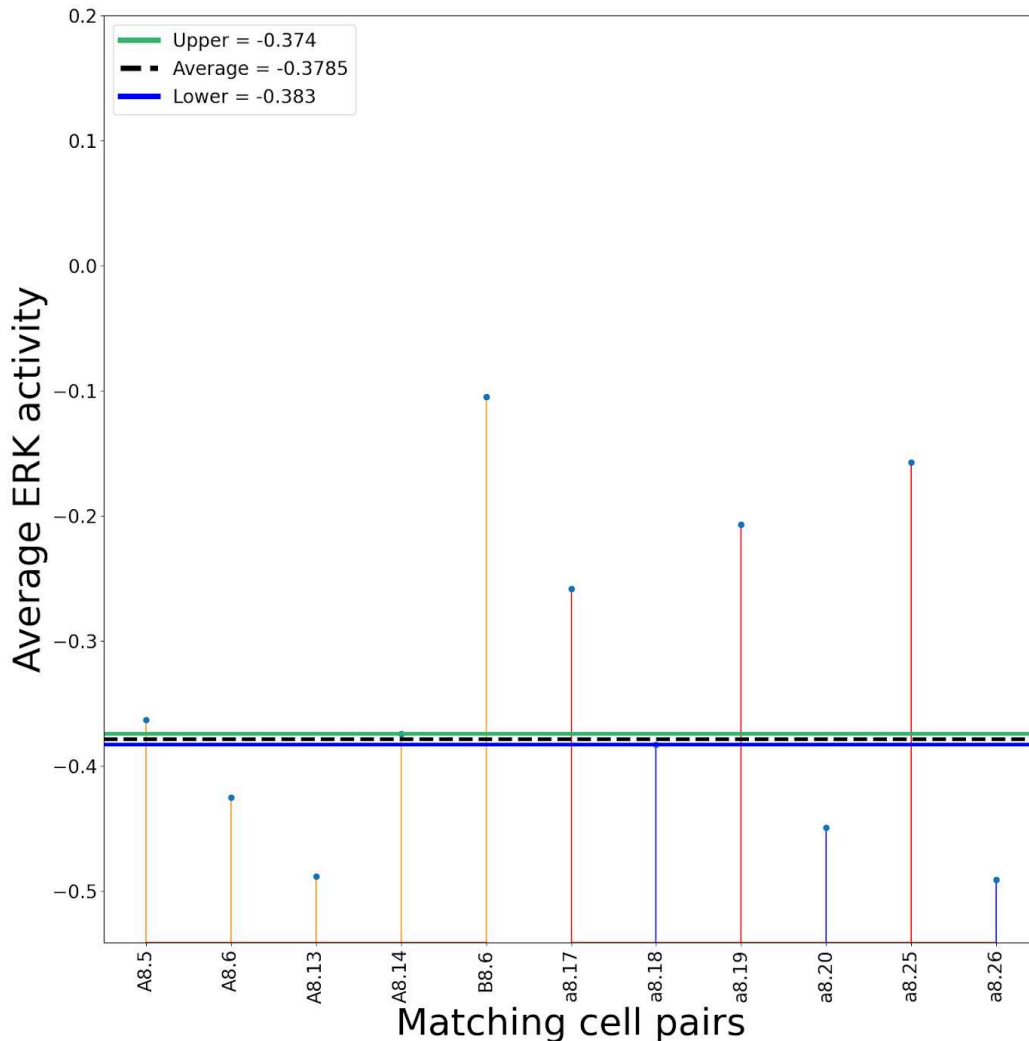


Figure S14 : Average ERK activity measured in Astec-Pm18 only for cells of the 8th generation belonging to the ground truth lists and determination of ERK status based on thresholding. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists of the 8th generation : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Astec-Pm19 8th generation

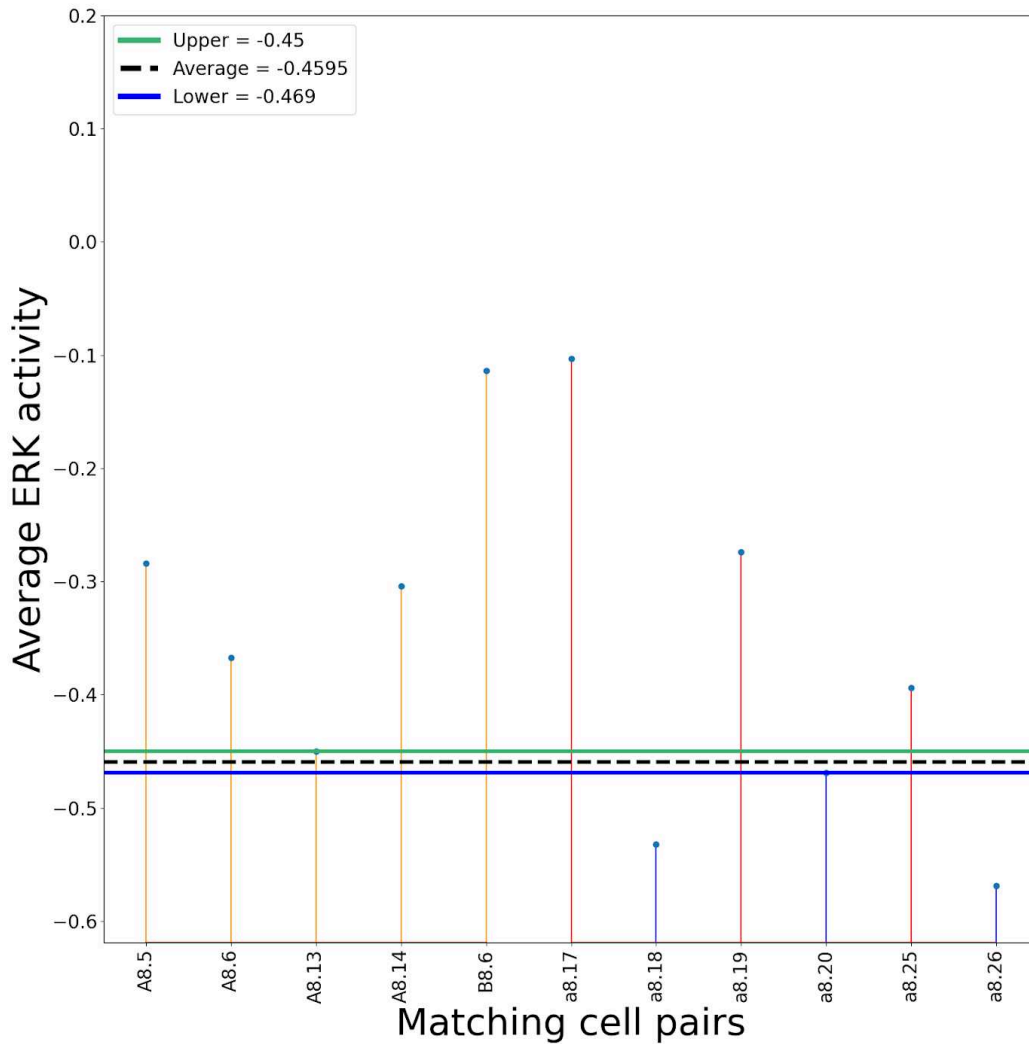


Figure S15 : Average ERK activity measured in Astec-Pm19 only for cells of the 8th generation belonging to the ground truth lists and determination of ERK status based on thresholding. Graph showing the average ERK activity measured in each left/right cell pair belonging to our ground truth cell lists of the 8th generation : positive (orange), positive (red) while their sister cells are negative (blue) with the corresponding thresholds calculated. The lower threshold is set at the level of the average ERK activity measurement of the known dp-ERK negative cell which has the highest value. The upper threshold is set at the level of the known dp-ERK positive cell which has the lowest value while remaining above the lower threshold. The average threshold value is obtained from the average of the two thresholds.

Movie S1	Astec-Pm16	ERK activity
Movie S2	Astec-Pm16	Average ERK activity
Movie S3	Astec-Pm16	Average L/R ERK activity
Movie S4	Astec-Pm17	ERK activity
Movie S5	Astec-Pm17	Average ERK activity
Movie S6	Astec-Pm17	Average L/R ERK activity
Movie S7	Astec-Pm18	ERK activity
Movie S8	Astec-Pm18	Average ERK activity
Movie S9	Astec-Pm18	Average L/R ERK activity
Movie S10	Astec-Pm19	ERK activity
Movie S11	Astec-Pm19	Average ERK activity
Movie S12	Astec-Pm19	Average L/R ERK activity

Movies S1-S12

All the movies can be downloaded here :

<https://drive.google.com/drive/folders/1tbE707MegbcQJbglKZLX3xfbA9u7Pdnk?usp=sharing>.

Discussion

Bien que de nombreuses inductions ERK-dépendantes aient été décrites chez les ascidies au cours des dernières décennies, notamment grâce à l'utilisation d'anticorps anti-dp-ERK (Nishida, 2003 ; Hudson et al., 2007 ; Yasuo et Hudson, 2007 ; Shi et Levine, 2008 ; Haupaix et al., 2013, 2014 ; Razy-Krajka et al., 2018 ; Williaume et al., 2021), cette étude suggère que nos connaissances ne sont pas exhaustives, un objectif qui pourrait être atteint par une analyse détaillée de notre atlas spatio-temporel systématique de l'activité ERK.

En effet, nous avons montré ici que la mesure de l'activité ERK moyenne par rapport au seuil fixé pour chaque génération est un bon prédicteur des inductions ERK-dépendantes. La question se pose maintenant de savoir si d'autres paramètres quantitatifs de la dynamique, tels que l'activité maximale de la signalisation ERK, son amplitude ou son activité intégrée, comme observé dans d'autres contextes développementaux (Johnson et Toettcher, 2019 ; Pokrass et al., 2020), pourraient également expliquer les inductions et lesquels sont les meilleurs prédicteurs et les plus reproductibles entre embryons. Cependant, afin de mieux caractériser la dynamique spatio-temporelle de la voie ERK dans les inductions d'ascidies et sa variabilité, une des principales limites jusqu'à présent est le manque d'un nombre suffisant d'embryons pour les divisions au-delà de la 8^{ème} génération cellulaire. Cela peut s'expliquer par la difficulté d'imager des embryons de *Phallusia mammillata* se développant correctement en raison de la phototoxicité inhérente à l'acquisition de 3 canaux d'imagerie. Pour pallier à ce problème, nous avons commencé une collaboration avec Emmanuel Faure (LIRMM, Montpellier) spécialiste en réseaux d'apprentissages profond, afin d'imager les membranes et les noyaux sur un même canal, en libérant un, et de séparer ensuite ces deux sources d'information à l'aide d'outils de déconvolution dédiés basés sur des réseaux d'apprentissage profond (Hallou et al., 2021). Pour cela j'ai généré un large dataset de 9 embryons (1135 pas de temps imagés entre les stades 16 cellules et neurula tardive) marqués au niveau des membranes (PH-GFP) et pour les noyaux (H2B-mCherry) acquis dans des canaux séparés. Ces canaux sont ensuite fusionnés sur une seule image pour chaque pas de temps et l'ensemble de ces données permet l'apprentissage de la tâche de séparation noyau/membrane par des réseaux développés par Emmanuel Faure (Figure 27). Les résultats de la déconvolution noyau/membrane sont très bons et feront l'objet d'une publication et les prochains embryons pourront être imagés avec membrane et noyau sur un même canal, rouge préférentiellement pour éviter la phototoxicité accrue dans le bleu/vert, et le ERK-KTR en infrarouge.

Image d'entrée : membranes (PH-tdTomato) et noyaux (H2B-mCherry)

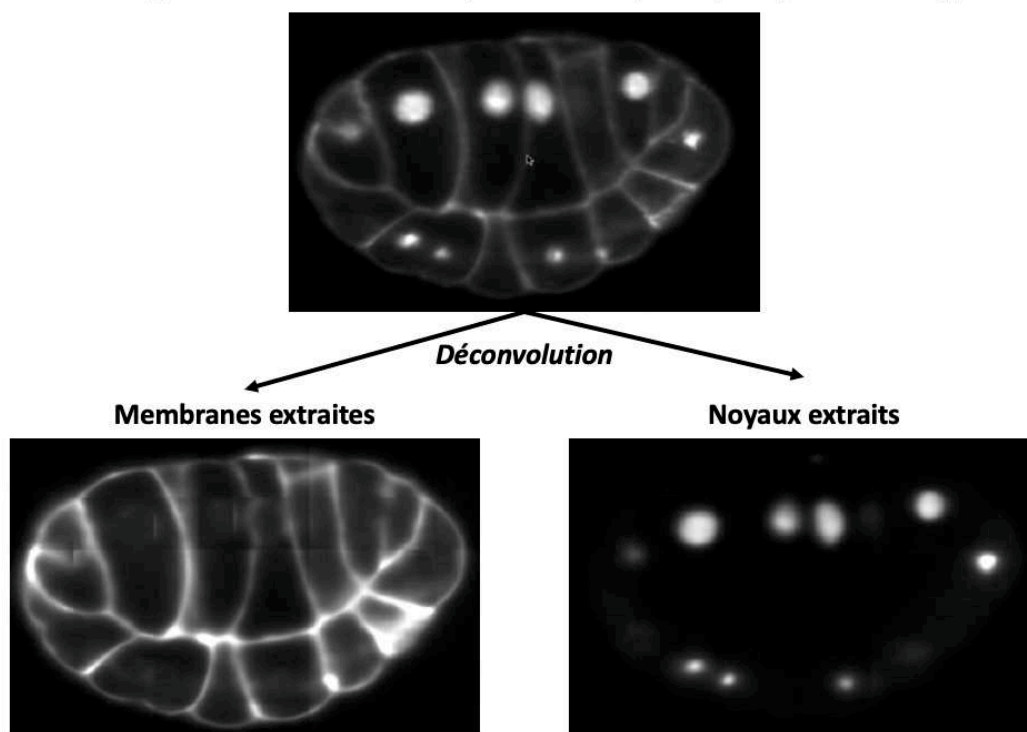


Figure 27 : Déconvolution des noyaux et des membranes basée un réseau d'apprentissage profond.

Source : Emmanuel Faure, LIRMM, Montpellier.

En plus de permettre une description complète des inductions ERK dépendantes et la caractérisation du rôle instructif de la dynamique spatio-temporelle de la voie ERK dans ces événements, l'automatisation des mesures de l'activité ERK ouvre la voie à une meilleure compréhension du comportement dépendant de la surface de contact de la voie ERK. En effet, si ce mode de signalisation ERK dépendant des surfaces de contact commence à être bien documenté pour les premiers stades du développement de l'ascidie avec des mesures quantitatives de l'activité ERK *in vivo* (Guignard et al., 2020 ; Guillaume et al., 2021), le manque de mesures pour les stades ultérieurs n'a pas permis de tester directement la généralité de ce modèle. Les données que nous avons générées nous permettent de tester avec précision ce modèle jusqu'aux cellules de 10^{ème} génération cependant, le manque de données d'expression du ligand FGF pour les stades tardifs rend cette tâche compliquée. Une alternative serait d'étudier la covariance entre les mesures précises des zones de contact et les mesures de l'activité ERK pour ces stades pour différents embryons. En outre, l'automatisation du suivi de l'activité de la voie ERK permet de la combiner avec des perturbations de la voie à différents niveaux de la cascade, ce qui est essentiel pour construire un modèle mécanistique de l'activité ERK dépendante des surfaces de contact. Par exemple, des mesures quantitatives récentes de l'activité ERK dans les cellules de la lignée "a" au stade 32 cellules, couplées à l'inhibition de la voie Ephrine/Eph, ont révélé le rôle de cette voie dans l'atténuation de la signalisation ERK, régulant ainsi la précision spatiale des inductions ERK-dépendantes de la surface de contact avec les cellules exprimant le FGF (Guillaume et al., 2021).

Enfin, il est important de mentionner que les méthodes qui ont été développées pour automatiser les mesures nucléaires et cytoplasmiques de ERK-KTR peuvent être déployées pour mesurer tout

autre processus basé sur un événement de relocalisation nucléaire renforçant l'attrait du modèle ascidie pour la construction d'atlas fonctionnels multimodaux à résolution cellulaire.

DISCUSSION ET PERSPECTIVES

RÉSUMÉ

Nos connaissances actuelles sur la voie de signalisation ERK et sur le rôle de sa dynamique spatio-temporelle au cours du développement embryonnaire sont essentiellement qualitatives et statiques et restent jusqu'à présent limitées (Biasuz et al., 2018 ; Patel et Shvartsman, 2018 ; Li et Elowitz, 2019 ; Nakamura et al., 2021). Cela est principalement dû à la difficulté de suivre l'activité en temps réel de la voie de signalisation ERK dans des embryons qui se développent. Qui plus est, dans la plupart des embryons animaux, la reproductibilité se situe à l'échelle du tissu, les comportements des cellules individuelles étant stochastiques au-delà des premières divisions et l'hétérogénéité des réponses dans les populations cellulaires rend difficile l'étude de la dynamique spatio-temporelle de ERK dans des contextes cellulaires précis.

Au cours de ma thèse, j'ai étudié la voie FGF/ERK au cours de l'embryogenèse de l'ascidie *Phallusia mammillata*, un membre des tuniciers, le groupe frère des vertébrés. Bien que les larves de têtards d'ascidies présentent le plan corporel de base des chordés, ces embryons sont beaucoup plus simples que ceux de leurs cousins vertébrés. De plus, de manière unique parmi les modèles de chordés, l'embryogenèse des ascidies se déroule selon un plan de division cellulaire invariant, de sorte que les configurations cellulaires et la progression du cycle cellulaire sont quasi invariantes (Dumollard et al., 2013; Guignard et al., 2020). Ce contexte unique permet une identification précise des cellules en fonction de leur position dans l'embryon, fournissant un cadre rigoureux pour les analyses à la résolution cellulaire. Enfin, les ascidies font un usage intensif de la signalisation ERK pour la spécification de nombreux destins (Lemaire, 2009; Hudson 2016) et il a été proposé que l'activité de la signalisation ERK est dépendante des surfaces de contact au cours de l'induction neurale au stade 32-cellules (Tassy et al., 2006).

Mon projet de thèse a donc combiné imagerie quantitative, perturbations expérimentales, modélisation théorique et développement d'outils informatiques, dans un organisme modèle animal original, afin d'étudier quantitativement la dynamique spatio-temporelle de la signalisation ERK au cours de l'embryogenèse de l'ascidie *Phallusia mammillata* et de tester la généralité du mécanisme proposé d'activité de ERK dépendante des surfaces de contacts.

Ce travail tire profit de la capacité unique qui a été développée par notre équipe au cours des années précédant mon arrivée à identifier/segmenter/suivre les cellules sur de longues périodes de leur développement avec une haute résolution temporelle, une condition essentielle pour suivre la dynamique de l'activité de la voie ERK avec une résolution cellulaire au cours de l'embryogenèse. Ce travail, interdisciplinaire, a également bénéficié du soutien de nombreux collaborateurs de long terme.

Afin de décrire le pattern spatio-temporel d'activation de la voie FGF/ERK, je me suis donc tout d'abord concentré sur le développement des méthodologies expérimentales nécessaires à la quantification manuelle *in vivo* en temps réel de l'activité ERK dans les embryons de *Phallusia mammillata* en utilisant le biosenseur ERK-KTR (Regot et al., 2014). Comme preuve de principe, cet outil a été utilisé dans le cadre de la révision d'un article afin de tester l'hypothèse selon laquelle les surfaces de contact entre les cellules qui émettent un signal et celle qui répondent fournissent les informations quantitatives qui déterminent le résultat des inductions au cours de l'embryogenèse

chez les ascidies, un modèle alternatif au modèle classique de « Morphogène » où la concentration du ligand fournit cette information (chapitre 1). Pour cela, l'activité de ERK a été mesurée grâce au ERK-KTR pour chaque cellule au stade 64-cellule dans des embryons entiers ou coupés en deux afin d'en modifier les contacts cellulaires. Ces mesures *in vivo* de la signalisation ERK ont ensuite permis la comparaison directe avec les prédictions d'un modèle computationnel de la voie ERK, qui avait déjà été développé, visant à prédire l'activité de la voie dans chaque cellule en fonction de la surface de la cellule exposée à des cellules qui émettent les ligands FGF et Ephrine. Dans tous les cas, les mesures du ERK-KTR corrélaient fortement aux prédictions du modèle, supportant que le niveau d'activité des voies de signalisation chez les ascidies est dépendent des surfaces de contact, ce qui a depuis été également montré pour la signalisation ERK au cours de l'induction neurale pour les cellules de la lignée "a" au stade 32-cellules (Williaume et al., 2021).

Cependant, la méthodologie développée dans le cadre de cette étude présentait certaines limitations, notamment le fait que l'identification des cellules et les mesures du ERK-KTR ont été effectuées manuellement et sont très chronophages. Il faudrait l'équivalent d'un an de travail à la main pour mesurer le ERK-KTR dans les noyaux de toutes les cellules nommées d'un seul embryon imagé toutes les 2 minutes entre le stade 32-cellules et le milieu de neurulation. Les mesures manuelles n'ont donc été effectuées que sur un stade donné, le stade 64-cellules et de façon discontinue, ne permettant pas d'accéder à la dynamique fine de l'activité ERK.

J'ai donc contribué au développement d'outils visant à automatiser certaines étapes de la méthodologie afin de produire un atlas spatio-temporel de l'activité ERK au cours de l'embryogenèse chez *Phallusia* et de tester le modèle de communication surface de contact dépendant sur des stades plus tardifs.

L'identification systématique de cellules équivalentes dans des embryons distincts est une condition préalable pour la réalisation d'un atlas de l'activité de ERK chez *Phallusia*. Nous avons précédemment choisi de nommer manuellement les embryons reconstruits de *Phallusia* (Guignard et al., 2020), une tâche laborieuse et sujette aux erreurs d'appréciation. Cela à tout de même permis de constituer un premier atlas morphologiques de 10 embryons de *Phallusia* entièrement nommés jusqu'à la 10^{ème} génération. En nous appuyant sur cet ensemble de données, en étroite collaboration avec Grégoire Malandain (INRIA, Sophia Antipolis), spécialiste en analyses d'images biologiques, analyses topologiques et construction d'atlas, un cadre théorique et les méthodes algorithmiques permettant l'identification automatisée et non biaisé des cellules dans les embryons stéréotypés de *Phallusia mammillata* ont été développées (chapitre 2). Après avoir montré que la mesure précise du voisinage cellulaire est suffisamment invariante dans les embryons de *Phallusia* pour représenter chaque cellule sans ambiguïté par le vecteur de ses surfaces de contact avec les cellules voisines directes, nous avons établi une méthode automatisée de nommage qui repose sur la mesure de distance du voisinage division-à-division et qui permet le transfert d'information depuis un embryon déjà nommé vers de nouveaux embryons reconstruits. Afin de rendre notre procédure de nommage automatique plus robuste et de tenir compte de la variabilité biologique du développement, nous avons établi une collection d'embryons références pour lesquels nous avons d'abord développé des méthodes informatiques afin de détecter et de corriger systématiquement les erreurs résiduelles issues du nommage manuel initial. Enfin, nous avons montré que cette approche originale, qui utilise l'information compacte et locale de la mesure du voisinage pour chaque cellule, permet de nommer

en quelques minutes très précisément toutes les cellules des embryons reconstruits de *Phallusia* jusqu'à la 10^{ème} génération, ce qui, à ce jour, n'a jamais encore été rapporté pour aucun autre modèle développemental.

En raison du niveau extrême de reproductibilité du développement embryonnaire précoce des ascidies, seules de légères différences dans la géométrie des embryons de différentes espèces sont observées (Lemaire, 2011). Nous nous attendons donc à ce que notre méthode de nommage automatique soit largement applicable pour nommer avec précision d'autres espèces d'ascidies solitaires directement à partir des embryons de *Phallusia* déjà nommés, ce qu'il faudra tester et qui pourrait permettre d'étudier la variabilité morphologique développementale plus largement au sein des d'ascidies. En raison des caractéristiques développementales des nématodes, dont les embryons se développent également avec un lignage cellulaire invariant et une grande reproductibilité de la position relative des cellules au sein de chaque espèce (Li et al., 2019), une approche similaire à la nôtre pourrait être également implémentée pour nommer automatiquement les cellules dans ces embryons, la limitation principale actuellement étant le manque de données d'embryons de nématodes imagés avec les membranes plasmiques marquées par fluorescence.

A l'avenir, il sera également intéressant d'étendre le nommage de référence pour couvrir l'ensemble du développement des ascidies jusqu'à la formation de la larve. En effet, la plupart des cellules internes ne se divisant plus (sauf les cellules du mésenchyme), il serait possible de s'inspirer des règles de Conklin pour finir de nommer à la main les cellules jusqu'à la 13^{ème} génération, génération la plus tardive observée dans la larve, à partir d'une distance géodésique par rapport au pôle animal, par exemple et de s'inspirer de cette logique pour les quelques cellules internes. Une fois les nommages établis jusqu'à la 13^{ème} génération dans plusieurs embryons de références, ceux-ci pourraient être à leur tour utilisés pour nommer automatiquement de nouveaux embryons reconstruits.

Dans l'étude présentée dans le chapitre 2, la question de savoir s'il est pertinent de nommer les cellules lorsque l'orientation de division est variable n'est pas directement adressée, et nous suggérons actuellement d'associer le nom de chaque cellule avec un indice de confiance dans le nommage. Cet indice traduit à quel point la configuration topologique du couple de cellules soeurs, qui a été automatiquement nommé, est représentée dans la population de référence. Cependant cet indice de confiance demande d'être encore raffiné car en plus de mesurer la variabilité de l'orientation de la division à nommer, il mesure également la variabilité de la division des cellules voisines à la génération précédente.

Enfin, le développement du nommage automatique pour les embryons de *Phallusia mammillata*, en facilitant l'intégration de nouveaux jeux de données dans un référentiel commun, constitue une étape majeure vers la construction d'atlas multi-échelles et l'analyse de la variabilité au cours de l'embryogénèse jusqu'en milieu de neurulation, dont une première application a été pour construire un atlas fonctionnel de l'activité de ERK.

Afin de construire cet atlas, j'ai donc ensuite travaillé, en collaboration avec Benjamin Gallean, ingénieur d'études informaticien dans notre équipe, sur le développement des méthodes expérimentales et informatiques pour automatiser les mesures du ERK-KTR dans les noyaux et le

cytoplasme de chaque cellule d'un embryon en développement (chapitre 3). Ce travail a permis de générer un premier atlas fonctionnel de la dynamique spatio-temporelle de l'activité ERK avec une résolution cellulaire et une résolution temporelle de 2 minutes sur une large période du développement embryonnaire. Ces mesures de l'activité de ERK à partir du ERK-KTR ont été mises à disposition sous la forme d'un atlas interactif en ligne à l'aide du navigateur morphodynamique en ligne, MorphoNet, pour chacun des 4 embryons analysés. Nous avons montré que la mesure de l'activité ERK moyenne récapitule ce qui est connu des inductions ERK-dépendantes chez les ascidies, ce qui soutient la pertinence biologique de notre atlas. Nous fournissons enfin des exemples montrant que ces mesures permettent d'identifier de nouveaux candidats d'inductions ERK-dépendantes. Ces premiers résultats suggèrent que nos connaissances actuelles des inductions ERK-dépendantes ne sont pas exhaustives. Cependant le nombre d'embryons analysés reste trop limité et le manque de confirmation expérimentale ne permettent pas de conclure quant à ces nouveaux candidats proposés. La question se pose également maintenant de savoir si d'autres paramètres quantitatifs de la dynamique, tels que l'activité maximale de la signalisation ERK, son amplitude ou son activité intégrée, comme observé dans d'autres contextes de développement (Johnson et Toettcher, 2019 ; Pokrass et al., 2020), pourraient également expliquer les inductions et lesquels sont les meilleurs prédicteurs et les plus reproductibles entre embryons.

Enfin, un des objectifs de l'atlas de la signalisation ERK était de pouvoir tester directement la généralité du mode de communication surface-de-contact dépendant chez les ascidies. En effet, dans l'étude publiée en 2020 (chapitre 1) (Guignard et al., 2020), les mesures du ERK-KTR ont été réalisées manuellement uniquement à un stade précoce, le stade 64-cellules. L'automatisation des mesures du ERK-KTR (chapitre 3) a alors permis la mesure de l'activité de ERK pour un embryon jusqu'à un stade tardif (Astec-Pm16, neurula tardive).

Cependant, afin de tester ce modèle du niveau d'activité de signalisation surface-de-contact dépendant, il est impératif d'intégrer les données spatiales d'expression des ligands FGF et Ephrine. Or, si de nombreuses données d'hybridation *in situ* ou de séquençage des ARNm au niveau de la cellule unique permettent d'avoir ces données pour chaque cellule jusqu'en début de gastrulation (stade 112-cellules) (Cao et al., 2019; Sladitschek et al., 2020), au-delà de ce stade les données sont plus parcellaires voire inexistantes pour de nombreuses cellules encore à ce jour, notamment chez *Phallusia*. Afin d'obtenir ces données d'expression, Ingrid Mormin, ingénieure d'études dans notre équipe est actuellement en train d'établir le protocole expérimental permettant d'imager le profil d'expression des ligands FGF et Ephrine sur des stades plus tardifs par hybridation *in situ* fluorescente quantitative (qHCR) chez *Phallusia* (Figure 28). Afin de pallier la difficulté principale qui est de pouvoir attribuer précisément le profil d'expression à des cellules clairement identifiées sur des stades tardifs et pour pouvoir quantifier précisément l'expression des ligands pour chaque cellule, Ingrid Mormin développe une méthode visant à marquer également les membranes des embryons qui sont fixés et imagés au confocal après hybridation *in situ* (Figure 28). Grâce à ce marquage, le signal pourra être quantifié au sein de chaque cellule et les embryons pourront être recalés sur les embryons reconstruits avec ASTEC déjà nommés, ce qui permettra de les nommer par transfert à partir de ces derniers. Ainsi les profils d'expression obtenus pourront être intégrés pour chaque embryon de l'atlas de l'activité ERK et, en combinaison avec les mesures précises des contacts cellulaires, ils permettront de tester la généralité du mode de communication surface de contact dépendant chez les ascidies. Enfin, les ascidies sont connues pour l'évolution rapide de leur génome

entre espèces et leur haut degré de polymorphisme génétique intraspécifique (Lemaire, 2011). Ceci suggère que le niveau d'expression des ligands peut varier et que le mode de signalisation surface de contact dépendant les tamponne, une hypothèse qu'il sera également possible de tester directement à partir de la quantification de l'expression des ligands FGF et Ephrine qui sera obtenue après analyse des données qHRC. Une prédiction de ce modèle est l'observation d'une plus grande variabilité dans la mesure quantitative de l'expression des ligands pour chaque cellule que dans les mesures des contacts cellulaires et de la signalisation ERK.

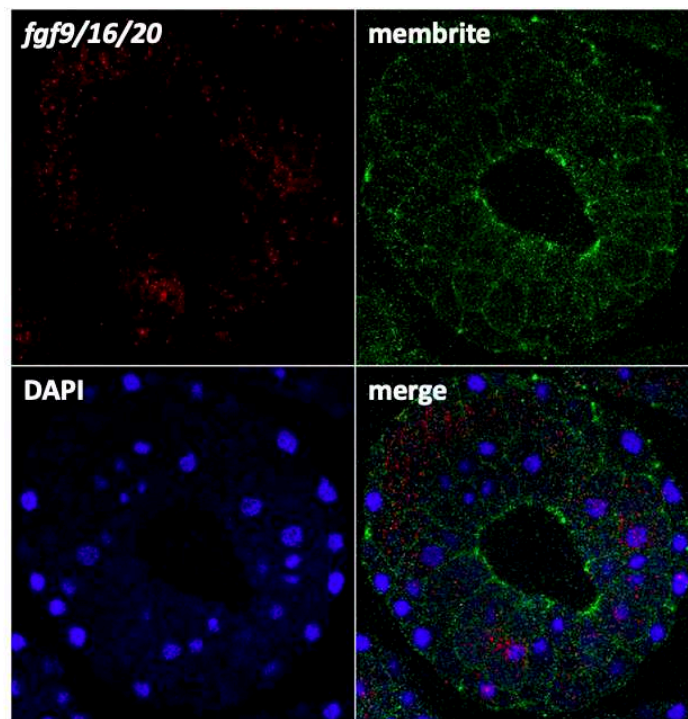


Figure 28 : Données préliminaires de l'imagerie *in situ* qHCR du *fgf9/16/20* et des marquages membranaires et nucléaires. Les sondes qHCR (quantitative Heavy Chain Reaction) ont été commandées chez molecular instruments. Une première a été désignée pour se lier spécifiquement au *fgf9/16/20* à partir du modèle de gène : phamm.g00003805 (ANISEED). Une seconde sonde couplée à un fluorochrome 546 se lie elle à la première sonde. Les membranes sont marquées avec MemBrite Fix Cell Surface Staining Kits 488/515. Les noyaux sont marqués au DAPI. Cette image d'un embryon en milieu de gastrulation a été réalisée sur un LSM980 Zeiss en mode confocal, 40X. Source : Ingrid Mormin, équipe Lemaire.

DIFFICULTÉS RENCONTRÉES

Bien que l'objectif initial ait été d'étudier et de comprendre la dynamique spatio-temporelle de la signalisation ERK au cours du développement précoce chez l'ascidie *Phallusia mammillata* et de tester la généralité du mode de signalisation surface de contact dépendant dans le cadre de mon doctorat, je me suis principalement concentré au développement et la validation d'une pipeline de méthodes expérimentales et informatiques permettant la réalisation de cet atlas de la signalisation ERK.

En effet, tout d'abord, ce projet, qui requiert la maîtrise d'outils expérimentaux et informatiques aussi divers que sophistiqués et la collaboration avec des scientifiques de formations variées (mathématiciens, physiciens, informaticiens) et, qu'il a fallu initier dans l'équipe, à nécessité un

investissement technologique lourd. Ceci notamment afin de permettre le passage à l'échelle de la pipeline de reconstruction des embryons de *Phallusia mammillata* marqués au niveau des membranes plasmiques, qui, lors de mon arrivée comportait encore de nombreuses étapes manuelles et dont le besoin et l'investissement nécessaire pour les améliorer a été initialement sous-estimé. Ce fut par exemple le cas pour l'automatisation du nommage qui a été présentée dans le chapitre 2.

D'autres, efforts visant à faciliter la reconstruction des embryons de *Phallusia*, et qui n'ont pas été présentés dans ce manuscrit, ont également été réalisés afin de faciliter la correction des segmentations en sortie d'ASTEC. En effet, en sortie d'ASTEC, les segmentations de chaque cellule pour chaque pas de temps ne sont pas parfaites. On retrouve trois types d'erreurs potentielles : deux cellules sont fusionnées (sous-segmentation), une cellule est coupée en plusieurs morceaux (sur-segmentation) ou une cellule est manquante; ASTEC étant paramétré pour induire principalement des sur-segmentations, qui sont les plus faciles à corriger. Cependant, à mon arrivée, les sur-segmentations n'étaient que difficilement corrigeables à partir de l'édition de fichiers textes visant à instruire l'exécution de programmes de traitement d'image codés en python. Ainsi, en collaboration avec Emmanuel Faure (LIRMM, Montpellier), informaticien et Benjamin Gallean, informaticien dans notre équipe, nous avons travaillé à l'élaboration d'un nouvel outil appelé "MorphoPlot"¹ adapté pour la correction de large dataset 3D+t (Figure 29 A). Ce nouvel outil repose sur le navigateur morphodynamique de MorphoNet (Leggio et al., 2019) et permet de lancer l'exécution de programmes de traitement d'image en réponse à l'interaction avec les données de segmentations directement depuis une interface utilisateur (Figure 29 A). L'utilisation de cet outil permet désormais de corriger tous les types d'erreurs et utilise l'information du lignage ce qui réduit considérablement le temps nécessaire pour la correction d'un large ensemble de données et la production d'un nouvel embryon reconstruit, de plusieurs semaines à une semaine (Figure 29 B). Ensembles, avec l'automatisation du nommage, ces efforts ont permis de faciliter considérablement la reconstruction des embryons de *Phallusia*, en passant de quelques mois de travail à quelques semaines.

¹ <https://morphonet.org/helpfiles/API/Notebooks/MorphoNetPlot.html>

A



B

Avant correction



Après correction



Figure 29 : MorphoPlot, un navigateur morphodynamique interactif pour la correction de données 3D+t complexes.

A) Illustration des segmentations d'un embryon chargées sur MorphoPlot avec le menu correction (curate) ouvert montrant différents algorithmes python de traitement d'image pouvant être exécutés directement depuis l'interface. B) Exemple d'une correction d'une sur-segmentation depuis MorphoPlot. Haut : les morceaux d'une cellule sur-segmentée sont directement manuellement identifiés à l'aide d'une même coloration. Bas : cette information de coloration est ensuite utilisée pour lancer un algorithme de fusion depuis l'interface, qui édite directement l'image de segmentation, qui est ensuite rechargée automatiquement sur MorphoNet pour continuer les corrections.

Ensuite, bien que *Phallusia mammillata* produise des gamètes tout au long de l'année, la qualité des développements embryonnaires est quant à elle ordinairement optimale au cours des périodes mars/juin et septembre/novembre. Cependant en raison du changement climatique et de difficultés à maintenir une qualité optimale d'eau de mer artificielle pour la culture des embryons, les périodes au cours desquelles le développement embryonnaire a été bon ont souvent été limitées. De plus, la fermeture du laboratoire en raison de la crise Covid en mars 2020 nous a fait perdre une saison entière d'expérimentation.

Lorsque les développements étaient bons, j'ai également été confronté à de nombreuses pannes du microscope à feuillet lumineux MuViSPIM de Luxendo, notamment juste après la reprise progressive des travaux de recherches en mai 2020 suite à la crise Covid, panne qui n'a pas pu être corrigée pendant plusieurs mois par les techniciens de Luxendo en raison des interdictions de circulations entre la France et l'Allemagne liées au Covid. Cet instrument très sophistiqué et hautement sollicité sur de longues périodes de plusieurs heures d'acquisition s'est donc avéré peu fiable pour ce qui est de la répétition des expérimentations quotidiennement, ce qui est nécessaire quand on sait qu'on ne peut imaginer qu'un seul embryon à la fois.

Enfin, lorsque les développements étaient corrects et le microscope opérationnel, une des difficultés a été l'acquisition dans le microscope d'un développement correct des embryons en raison de la

phototoxicité inhérente à l'imagerie de 3 canaux en continu sur plusieurs heures. Actuellement, les 4 embryons sélectionnés car ayant le meilleur développement présentent tout de même des défauts de développement à partir de la fin de la gastrulation ou le début de neurulation par exemple. Pour limiter la phototoxicité, nous avons mis en place, en collaboration avec Emmanuel Faure (LIRMM, Montpellier), de nouveaux outils basés sur des réseaux d'apprentissage profond permettant de séparer les membranes et les noyaux acquis sur un même canal (Figure 27). Cette stratégie pourra être appliquée pour imager de nouveaux embryons pour l'atlas de l'activité de ERK à partir de deux canaux uniquement, en évitant le vert, afin de limiter la phototoxicité et d'étudier la dynamique spatio-temporelle de l'activité ERK dans plusieurs embryons se développant correctement jusqu'à des stades tardifs. Une alternative ou une combinaison à cette première solution pour diminuer l'exposition en photon au cours de l'acquisition multi-canaux serait d'imager les membranes à une plus faible intensité et de réhausser à posteriori la qualité de ces images à partir de l'outil de débruitage basé sur des réseaux d'apprentissage profond CARE, qui permet de restaurer des images même si 60 fois moins de photons sont utilisés pendant l'acquisition (Weigert et al., 2018).

Ainsi, pour l'ensemble de ces raisons, il a été très compliqué de générer de nouveaux embryons pour l'atlas de la signalisation ERK exprimants à la fois un marquage membranaire vert (PH-GFP), un marquage des noyaux en rouge (H2B-mCherry) et le ERK-KTR en rouge lointain (ERK-KTR-iRFP) et c'est pourquoi le nombre d'embryons actuels analysés, pour les stades tardifs notamment, reste actuellement limité et limitant. Cela s'applique également pour expliquer le fait que nous n'ayons pas encore obtenu les données d'expression des ligands FGF et Ephrine par hybridation *in situ* quantitative pour les stades tardifs chez *Phallusia*, ne permettant pas de tester la généralité du modèle de signalisation ERK surface de contact dépendant chez les ascidies.

DIRECTIONS FUTURES

Vers une exploration multi-échelle de la variabilité de l'embryogénèse des ascidies

Dans l'ensemble, la combinaison i) d'ASTECC, qui permet de générer des atlas morphologiques à la résolution cellulaire pour chaque embryon à partir d'un marquage membranaire, avec les outils développés au cours de mon doctorat, qui permettent ii) le nommage automatique et l'intégration de tous les embryons reconstruits dans un référentiel commun non biaisé et iii) qui permettent la mesure automatique des signaux nucléaires et membranaires afin de suivre l'activité de n'importe quel biosenseur fluorescent dont le principe est basé sur des événements de transport nucléo-cytoplasmiques, établissent pleinement les embryons d'ascidies comme un modèle de choix pour construire des atlas multi-modaux à la résolution cellulaire et étudier la variabilité à cette différentes échelles (Figure 30).

Reconstruction automatique du développement des ascidies (ASTECC)

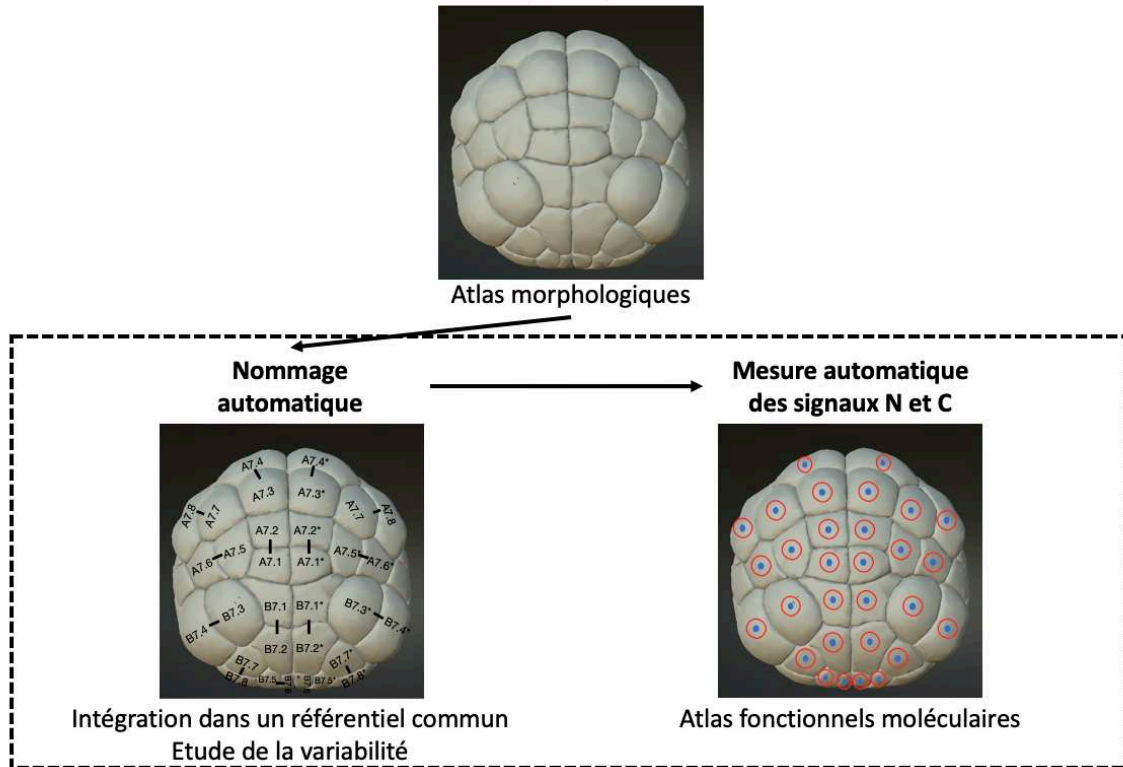


Figure 30 : Construction d'atlas multi-modaux et exploration de la variabilité développementale multi-échelles dans les embryons d'ascidie.

Images montrant un embryon de *Phallusia mammillata* au stade 64-cellules reconstruit grâce à ASTECC. Les outils développés au cours de mon doctorat (pointillés) permettent le nommage automatique des cellules et la mesure automatique des signaux de rapporteurs fluorescents dans les noyaux (N, ronds bleus) et le cytoplasme (C, disques rouges). Vue végétative. Embryon depuis MorphoNet (Leggio et al., 2019).

Tout d'abord, le nommage automatique et non biaisé des embryons reconstruits de *Phallusia mammillata*, en facilitant l'intégration de nouveaux jeux de données dans un référentiel commun, constitue une étape majeure vers l'analyse de la variabilité morphologique du développement embryonnaire chez les ascidies. En effet, par exemple, alors que le développement des ascidies est connu pour être hautement stéréotypé au cours des stades les plus précoces (jusqu'en début de gastrulation, stade 112-cellules), les analyses réalisées pour le développement des outils de nommage automatique (chapitre 2) nous ont permis d'identifier des exemples intéressants de divisions présentant une orientation de division variables. Cela suggère qu'il pourrait y avoir plus de variabilité morphologique entre embryons que ce qui est classiquement admis chez les ascidies. Ainsi, une caractérisation approfondie de la structure spatiale et temporelle de la variabilité morphologique embryonnaire permettrait de répondre à des questions importantes sur le développement et l'évolution. Tout d'abord, l'identification des cellules, tissus ou arrangements cellulaires les plus contraints pourrait éclairer la logique fondamentale des processus de développement. Par exemple, nous avons montré que les inductions cellulaires reposent sur des contacts précis entre cellules (chapitre 1), rendus fiables par l'existence d'un lignage cellulaire quasi-invariant. On pourrait donc s'attendre à ce que les interfaces entre cellules communicantes soient plus contraintes que celles entre cellules non communicantes, ce qui permettrait d'identifier géométriquement les inductions cellulaires. Ensuite, il a également été proposé qu'il existe un

modèle temporel universel de variabilité embryonnaire partagé par tous les métazoaires, appelé le modèle du sablier (Duboule, 1994; Richardson, 2012). Selon ce modèle, il existe une période centrale au cours de l'embryogenèse pendant laquelle les morphologies embryonnaires et les voies moléculaires sont plus contraintes qu'aux étapes de développement antérieures ou postérieures. Grâce au nommage, il est désormais possible de tester la validité de ce modèle de sablier chez les ascidies, en comparant par exemple la variation de la composition cellulaire entre embryons au cours du temps ou en comparant la reproductibilité de la géométrie du développement embryonnaire entre embryons à partir des outils développés pour comparer précisément le voisinage cellulaire. Cela pourra être réalisé chez *Phallusia mammillata* mais aussi pour plusieurs autres espèces d'ascidies qui ont des embryons transparents et qui pourraient être imagés puis segmentés avec ASTEC et nommés automatiquement à partir des embryons référence de *Phallusia*. Dans cette direction, nous avons commencé à imagier et reconstruire les embryons transparents de l'ascidie *Ascidella aspersa*.

Enfin, le nommage automatique combiné aux mesures automatiques de protéines fluorescentes dont le principe est basé sur une relocalisation nucléo-cytoplasmique ouvre quant à lui la voie à l'étude de la variabilité du développement embryonnaire à l'échelle moléculaire. Une première démonstration a été rapportée ici (chapitre 3) avec l'utilisation des mesures automatiques pour suivre l'activité de ERK grâce aux mesures du ERK-KTR, ayant ouvert la voie à l'étude ultérieure la variabilité de la signalisation ERK au sein d'un embryon (symétrie gauche droite, même génome) et entre embryons (polymorphisme), en lien avec la variabilité des contacts cellulaires. Ces outils pourraient être désormais appliqués pour suivre la relocalisation nucléaire de la β -caténine, par exemple. En effet, comme pour de nombreux embryons animaux, les ascidies utilisent la β -caténine nucléaire lors de la ségrégation de l'endomésoderme (ou endoderme) de l'ectoderme. Chez l'embryon d'ascidie, on sait notamment que la β -caténine nucléaire est impliquée dans la prise des décisions de destin binaire entre l'ectoderme et l'endomésoderme aux stades 16 et 32 cellules (Imai et al., 2016). Les méthodes développées dans le chapitre 3 sont donc directement applicables afin de mesurer automatiquement le signal nucléaire de la β -caténine nucléaire et générer un atlas de l'activité de la β -caténine nucléaire au cours du développement précoce chez *Phallusia*.

Vers une compréhension mécanistique du mécanisme original de signalisation surface-de-contact dépendant dans les embryons d'ascidies

Bien qu'il apparaisse maintenant clairement que le niveau d'activité de la signalisation ERK est dépendant des surfaces de contact exposées aux ligands FGF et Ephrine (Guignard et al., 2020; Guillaume et al., 2021), nous ne comprenons toujours pas comment ce mécanisme original est contrôlé au niveau moléculaire. Cette compréhension quantitative du comportement de la voie FGF/ERK ne peut être obtenue que par la modélisation mathématique du flux de l'information. La capacité de mesurer automatiquement la dynamique de l'activité de ERK dans des embryons vivants grâce au ERK-KTR permet maintenant de pouvoir construire un modèle qui récapitule la dynamique observée *in vivo* au cours du développement normal.

Dans cette direction, au cours de mon projet de thèse, en collaboration avec Bruno Leggio, physicien, et Panajot Kristofori, étudiant en master 1 biophysique que nous avons co-encadré en stage, nous avons étendu notre premier modèle mathématique très simplifié de la voie FGF/ERK (Guignard et al.,

2020) pour mieux prendre en compte la structure de la voie. En effet, alors que notre premier modèle ne tenait pas compte des composants en aval de RAS, ce nouveau modèle, toujours basé sur un système d'équations différentielles et de loi de conservation de masse, inclut désormais la cascade de signalisation depuis RAS jusqu'à ERK (Figure 31). Il inclut également une boucle d'hyperactivation entre RAS et SOS (Nakamura et al., 2016) et plusieurs boucles de rétroactions négatives entre ERK et les composants plus en amont (Figure 31) (Shin et al., 2009).

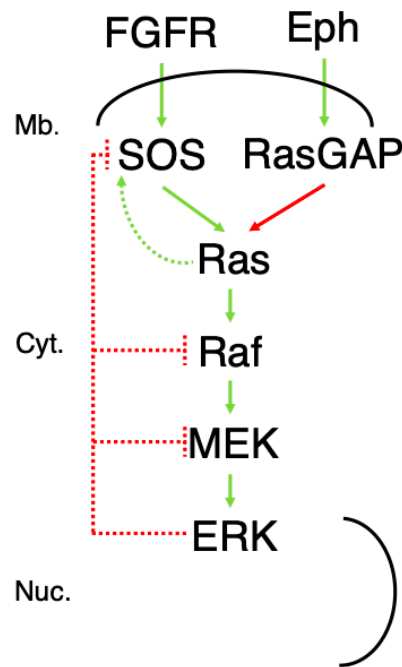


Figure 31 : Modèle étendu de la signalisation ERK.

Afin de pouvoir utiliser ce modèle plus complet pour récapituler la dynamique spatio-temporelle de la signalisation ERK dans des embryons normaux et prédire l'effet de perturbations expérimentales, il est important de contraindre les valeurs possibles des paramètres libres. Les paramètres clés du modèle comprennent notamment les constantes cinétiques des réactions et les concentrations de protéines. Les constantes cinétiques ont été déterminées biochimiquement pour la majorité des enzymes considérées chez les vertébrés. Nous pourrions donc supposer qu'en raison de la proximité phylogénétique des ascidies et des vertébrés, ces constantes s'appliquent. La concentration des composantes de la voie peut quant à elle varier fortement d'une espèce à l'autre. C'est pourquoi au cours de mon doctorat j'ai également mesuré les quantités relatives des différents composants de la voie FGFR/ERK expérimentalement chez les ascidies par spectrométrie de masse. Enfin, tous les autres paramètres libres que nous ne pourrions pas extraire de la littérature ou mesurer expérimentalement pourront être inférés par optimisation afin de récapituler au mieux les données expérimentales issues des mesures du ERK-KTR.

Ainsi, sur la base de ces travaux, il sera possible de construire un premier modèle récapitulant le flux d'informations et l'activité de la voie ERK au cours du développement normal de *Phallusia*. Celui-ci pourra être ensuite utilisé et raffiné afin de mieux comprendre les déterminants de l'activité de ERK

surface-de-contact dépendante et comment elle contribue à la robustesse du développement des ascidies face aux variations génétiques, géométriques et stochastiques. Il pourra être intéressant par exemple d'inclure dans notre modèle deux phosphatases majeures connues pour être impliquées dans le contrôle de la voie ERK : DUSP et PP1/PP2A et d'étudier l'effet de l'inhibition de ces phosphatases sur l'activité de ERK dans les prédictions du modèles et *in vivo* en combinant l'utilisation d'inhibiteurs pharmacologiques pour ces phosphatases et les mesures du ERK-KTR. L'interférence avec le fonctionnement de ces phosphatases pouvant conduire le système à une variance accrue de l'activité ERK, et donc montrer l'importance de celles-ci pour la précision de son activité plutôt que pour son niveau moyen. Enfin, les ascidies sont connues pour l'évolution rapide de leur génome et leur haut degré de polymorphisme génétique intraspécifique (Lemaire, 2011). Ceci suggère que les concentrations de protéines peuvent varier entre individus et que la signalisation ERK les tamponne, une hypothèse que nous pourrions également tester. La robustesse des prédictions du modèle face aux variations individuelles des concentrations des composants de la voie pourra être confrontée aux effets des manipulations expérimentales de ces concentrations mesurées grâce au ERK-KTR. Si les résultats diffèrent des prédictions du modèle, cela entraînera des modifications du modèle ou des paramètres.

Dans l'ensemble ce travail permettra de mieux comprendre les déterminants l'activité de ERK surface-de-contact dépendante et de la robustesse de la voie FGFR/ERK et de mieux comprendre comment émergent des dynamiques qualitativement et quantitativement différentes à partir de cette voie pourtant relativement simple. Ce travail permettra également de mieux prédire les effets non-intuitifs des perturbations génétiques pour cette voie impliquée dans la plupart des cancers.

PUBLICATIONS ANNEXES

Dans cette section sont incluses deux publications : une revue dont je suis co-auteur “The “computable egg” : Myth or useful concept?” (Biasuz et al., 2018) et un chapitre de livre dont je suis auteur “When developmental biology becomes quantitative : a big step toward “computable embryos”” (Boutet and Schierwater, 2021).



ELSEVIER

The “computable egg”: Myth or useful concept?

Kilian Biasuz¹, Bruno Leggio^{1,2}, Emmanuel Faure^{1,2,3} and Patrick Lemaire¹

Abstract

In 1994, Lewis Wolpert wondered “Will the egg be computable? That is given a total description of the fertilized egg – the total DNA sequence and the location of all proteins and RNA – could one predict how the embryo will develop?”. In this review, we discuss what computing the embryo entails and explore how advanced imaging, single-cell genomics and physical measurements are changing our view of the embryo and enable the development of multi-scale computational models that are starting to explain in physical terms how the genome controls the dynamic behavior of each cell in a developing embryo.

Addresses

¹ Centre de Recherche de Biologie cellulaire de Montpellier, CRBM, CNRS, Université de Montpellier, France

² Laboratoire Reproduction et Développement des Plantes, Univ Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA, Inria, F-69342, Lyon, France

³ Informatics Research Institute of Toulouse (IRIT), CNRS, INPT, ENSEEIHT, University of Toulouse I and III, France

Corresponding author: Lemaire, Patrick (patrick.lemaire@crbm.cnrs.fr)

Current Opinion in Systems Biology 2018, 11:91–97

This review comes from a themed issue on **Development and differentiation**

Edited by **Robert Zinzen** and **Stas Shvartsman**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 12 September 2018

<https://doi.org/10.1016/j.coisb.2018.09.003>

2452-3100/© 2018 Elsevier Ltd. All rights reserved.

Introduction

At the time it was formulated, Wolpert’s proposition [1] was purely theoretical. Even its premises, “a total description of the fertilized egg”, were inaccessible: no animal genome had yet been sequenced and GFP had not yet been fused to proteins to track the dynamic location of proteins in live embryos [2]. Over the past 25 years, technological breakthroughs in live imaging, sequencing, proteomics, and physical measurements have spurred a period of intense discovery of the parts that constitute the living matter, of their concentrations and localization in cells, and of their interactions. With such a massive increase in the volume, diversity and complexity of collected data, and the parallel

development of novel computing tools, Wolpert’s proposal is worth reexamining.

What is meant by computing the embryo? Is it worth the effort?

Embryonic development is an irreversible march forward in time. Computing the embryo rests on the idea that a mathematical description of the system can predict the future state of the embryo from the knowledge of its current state. With adequate training sets, this may be achieved by artificial intelligence methods, in particular deep learning. Indeed deep learning has been successfully used to model the hierarchical structure of the cell and make useful predictions [3]. Yet, such predictive approaches generally provide little mechanistic insight into the modeled process: past experience allows us to predict that the sun will rise every morning, but only physics can provide an explanation of the phenomenon. Likewise, models rooted in the physics or chemistry of the living matter are necessary to capture core developmental principles.

Although physical modeling of life has a long history [4], it has long remained a theoretical exercise. Insufficient measurements of physical parameters to constrain models, combined with a mostly qualitative and static description of phenotypes, made it difficult to discriminate plausible from correct hypotheses. The recent development of methods to measure kinematic and physical parameters lessened this difficulty, while making a mesoscopic physical description of biological processes more necessary than ever [5,6]. First, our brains simply cannot cope with the sheer volume of the data produced, many of which are irrelevant to the process studied. Second, biology involves multiple layers of feedback, at the origin of unintuitive non-linear behaviors. Third, biology is a multiscale process in which the macroscopic properties of cells and tissues emerge from the mesoscopic properties of molecules or subcellular structures. Finally, a precise mechanistically-rooted physical understanding of embryonic development, bridging genomic information and dynamic phenotypes, would be invaluable to guide synthetic biology approaches and increase the currently low yield of organoid production.

A major conceptual challenge, and an attraction for physicists, is that development involves morphogenetic

movements and active flows of matter and cells. As they develop, embryos thrive towards order rather than disorder. Developing systems are thus actively kept out of the thermodynamic equilibrium physicists are used to. In embryonic cells and tissues, for example, experimental measurements of parameters such as rigidity or viscosity strongly depend on both the measurement method, the spatial scale of observation and the developmental stage [7]. Fortunately, non-equilibrium features of biological tissues can often be treated under a quasi-static approximation within certain limits on both timescale and spatial extension of observation. This greatly simplifies the theoretical analysis of biological processes but restricts the applicability of such analyses to small temporal and spatial scales, unless specific theories adapted to living matter are developed, as illustrated below.

Formation of a complex organism from a fertilized cell requires the formation of the correct cell types, in the right number and in the appropriate position to ensure organ function. We will first review recent development in the systematic description of the geometric and genetic programs of individual embryonic cells. We will then discuss emerging mechanisms controlling transcription of genes, how cell fates and behaviors are choreographed by cell communication, and finally how this translates into morphogenesis.

Digital embryos with single-cell resolution

Developmental mechanisms have traditionally been studied qualitatively at the tissue-level. The breakthrough development of light sheet microscopy [8], combined with fluorescent markers for key subcellular structures ushered the field of vertebrate and invertebrate embryogenesis into an era where the behavior of each living embryonic cell and its progeny can be imaged and tracked over hours with sub-minute temporal resolution and sub-micron spatial accuracy [9].

The terabyte-size of light-sheet datasets precludes manual analysis. Dedicated computational tools have thus been developed to identify, digitalize and track each individual cell nucleus and its progeny over time [10,11]. The automated segmentation and long-time tracking of cell shapes over several cell generations, a requirement for mechanical modelling of morphogenesis, has been more challenging [12]. Using the transparent embryos of the ascidian *Phallusia mammillata*, we recently reported the accurate and automated whole-cell segmentation and tracking over 5 cell generation of membrane-labeled cells with 2-min temporal resolution [13]. This study opens the way to the systematic kinematic description of embryogenesis with cellular resolution. Current segmentation solutions, based on established segmentation algorithms [9], may still be difficult to efficiently tune to other datasets than those

they were developed for. Deep learning methods based on convolutional neural networks are revolutionizing medical imaging and start to be applied in developmental biology [14]. It will be interesting to see whether these novel approaches, alone or in combination with classical segmentation and tracking tools, will help systematic long-time cell tracking and phenotyping become a standard analysis in developmental biology. This evolution will be facilitated by the organization of contests to compare cell segmentation and tracking methods on raw and segmented/tracked reference 4D light-sheet datasets [15] deposited into public bioimage repositories [16].

Finally, frameworks to visualize, quantify, classify and compare within and between embryos the geometric and cell lineage properties at different spatial scales, from individual cells to organs, are being developed [17,18], opening the way to automated fine-grained phenotyping [13,19] of wild-type or experimentally-manipulated embryos and to an estimation of the variability of developmental processes [13,20].

These digital single-cell resolution atlases of cell properties and lineages can now be integrated with the underlying genetic program by spatially mapping the transcriptome [21–24] or proteome [25] of each individual cell in an embryo onto digital embryo atlases. Although limited sequencing depth currently only gives access to the expression profile of a minority of genes, it is sufficient to reveal the structure and dynamics of cell populations. These approaches can also suggest functional regulatory modules, whose impact on tracked cell lineages and differentiation can be assessed in small-[22,24] or large-scale [26] individual knock-downs of embryogenesis regulators.

Given the pace of technological progress in sequencing, mass spectrometry [25] and their computational analysis, we will soon have access to the RNA and partial protein content of both the egg and each individual embryonic cell, alongside its spatial position, shape and lineage, fulfilling Wolpert's premises. In the following sections, we will address how this information emerges from the genome sequence and how it generates the forces that drive morphogenesis.

From genome to transcriptomes: the physics of transcription

Precise spatiotemporal expression patterns in animal embryos results from the coordinated action of several classes of genomic regulatory elements, collectively referred to as *cis*-regulatory sequences [27]. Promoters are located around the transcription start site (TSS) of the gene and are the site of recruitment of RNA polymerase II. Enhancers are non-coding sequences located hundreds to millions of bases away from the TSS, which

control the recruitment or release of the polymerase at the promoter. Enhancers act as integration platforms for transcription factors (TFs), a class of proteins combining sequence-specific DNA-binding domain(s) with protein–protein interaction or enzymatic domains. A third class of elements, insulators, are important to control the specificity of enhancer/promoter association by creating or stabilizing DNA loops. The combination of the TFs and *cis*-regulatory elements controlling the genes involved in a given biological process defines the Gene Regulatory Network (GRN) of the process.

Several genomics methods were recently designed to unravel the repertoire of *cis*-regulatory sequences active during development, based on their open chromatin profile [28] or on their activity [29], some of which can be applied to single cells [30,31]. In parallel, the DNA-binding specificity of about 2 thirds of human TFs is known with high resolution [32]. Taken together, the combination of single-cell RNA-seq and ATAC-seq data with TF binding specificities will significantly improve current single-cell-based GRN inference algorithms (e.g., Ref [33]). The identification of an enhancer and the combination of transcription factors responsible for its activity is sufficient to mathematically predict its spatial activity [34].

The imaging of single transcripts in fixed or live cells revealed that transcription is a noisy stochastic process, causing cell-to-cell expression variability [35]. It occurs in bursts, during which tens of mRNA molecules are synthesized consecutively, interspersed with periods of transcription inactivity. Promoters have been proposed to control the number of RNA molecules transcribed per burst [36], while enhancers set the frequency of bursts, possibly through a regulation of the frequency of promoter/enhancer contacts [37].

Unexpected insight into the mechanistic origin of bursting came from the recent discovery of the prevalence in the cytoplasm or nucleus of membrane-less organelles able to undergo fusion, fission and dissipation as a result of liquid–liquid phase separation of cooperatively interacting multivalent protein assemblies [38]. Many proteins involved in transcription are indeed multivalent, including the highly repetitive C-terminal domain of RNA polymerase II. A simple model based on the law of mass-action suggested that the formation of liquid-like condensates of transcriptional regulators could explain previously poorly understood enhancer properties [39]. Consistently transcriptional coactivators, found in nuclear bodies with properties of liquid condensates, were able to drive phase separation *in vivo* and *in vitro* [40]. The presence of large condensates could also provide an explanation for the surprisingly large distance (>100 nm) separating enhancers and promoters during active transcription in *Drosophila* embryos [41,42].

In summary, the past 5 years have seen rapid progress in the identification of the *cis*-regulatory sequence repertoire for development and a paradigm shift in our physical understanding of the transcription process. These novel types of data constitute a strong incentive to further develop physical models of the different steps of transcriptional control during development.

Coordinating cell behaviors: cell communication

Both the fate and morphogenetic behavior of embryonic cells are influenced by chemical, electrical or mechanical cues emitted by neighboring cells or extracellular matrix. Here, we will restrict our discussion to the interpretation of classical chemical signaling during fate specification. Considering the diversity of fate specification events occurring during embryogenesis, surprisingly few signaling pathways exist [43], raising the issue of the specificity of response. The absolute concentration and duration of exposure to the factor and the past history of the cell have long been considered to be the major determinants of cell communication events [44]. The combination of experimental and quantitative modeling approaches now suggests that the repertoire of modalities of signal interpretation is much broader than initially thought.

Different combinations of ligands and receptors for the same pathway can lead to distinct outcomes. In chick neural crest, for example, Dll1 and Dll4 ligands individually trigger pulsatile and sustained Notch1 receptor activation dynamics, respectively, thereby inducing different cell fates [45]. The situation gets more complex when cells are exposed to multiple promiscuous ligands and receptors for the same pathway. In the BMP pathway, for example, ligands elicit non-intuitive combinatorial responses (e.g. additive, ratiometric ...) depending on the repertoire of receptors expressed by the cell [46]. The geometry of cell contacts can also impact cell communication: in both chick inner ear and ascidian embryos, the area of contact shared between emitting and receiving cells can determine signaling outcome [13,47]. Things do not get simpler inside the cell. In the Wnt and TGF- β pathways, vertebrate cells read concentration fold-change - rather than absolute level - of activated intracellular effectors, a regime more robust to noise [48,49]. Finally, recent studies revealed that the structure of the transcriptional Gene Regulatory Networks acting downstream of signaling pathways can be important to buffer cellular variability and genetic noise. In the vertebrate neural tube, mutual repression of primary sonic hedgehog transcriptional targets determines the precise spatial boundaries between motoneuron domains [50].

Novel imaging tools have also emerged to quantify the dynamics of signal transduction in single living cells of developing embryos, rather than in cell populations.

Kinases play a fundamental role in many signaling pathways. To follow their activity in real time, Regot and colleagues developed Kinase Translocation Reporters (KTRs), genetically-encoded fluorescent activity sensors that convert kinase activity into a nucleocytoplasmic shuttling equilibrium [51]. Using an KTR sensor to monitor ERK signaling in *C. elegans* vulva precursors revealed unexpected oscillatory dynamics of ERK activity [52]. Such tools will be particularly powerful to dissect how information flows through signaling pathways if the spatial range, intensity, and duration of intracellular signaling can be fine-tuned either through microfluidics devices [53] or optogenetic systems [54]. Optogenetics was recently adapted to the *Drosophila* embryo to activate with high spatiotemporal precision the Erk kinase pathway in response to light, demonstrating that development is sensitive to the spatial range and total duration but not to the intensity of Erk signaling above a certain threshold [55,56].

Signal transduction studies at the single-cell level are in their early days and the methods discussed above have only been used to monitor a single pathway component at a time. Yet, they suggest that multiplexing sensors at different level of the cascade will reveal the dynamics of information flow through the cell. Such quantitative measurements are necessary to more realistically model the growing repertoire of signaling modalities between cells.

From effector genes to phenotypes: biomechanics of morphogenesis

Morphogenesis, the dynamic process that shapes the embryo and its organs, is driven by the coordinated action of biomechanical forces acting across spatial and temporal scales, [57]. To exemplify the challenges of the physical modeling of morphogenesis, let us consider the first division of the large non-growing *C. elegans* zygote, one of the simplest and best understood animal morphogenetic processes. This division produces a large anterior cell and a smaller posterior cell, which will give rise to the germline and other tissues [58]. The process, controlled by a few tens of maternally-inherited proteins, is initiated by the establishment of a stable antero-posterior (AP) polarity in the cell cortex. This polarization leads to the posterior displacement of the spindle in response to stronger pulling forces on posterior astral microtubules. It also controls the asymmetric inheritance of multiple fate determinants, including the posterior segregation of germline P granules.

Over the years, physical models have been developed to account for actomyosin-driven cortical contractility and flow, advection and mutual antagonism of PAR polarity proteins, spindle positioning by differential pulling forces on astral microtubules and posterior localization of germline-harboring organelles called P granules

[59,60]. Modeling of these individual processes involved a diversity of soft-matter physics and fluid dynamics theories, including one which was specifically built to account for the specific out of equilibrium properties of actomyosin gels [61].

If modeling the asymmetric division of a single cell is already challenging - each of the models above only provides a partial view of the whole process -, the chances of successfully modeling complex multicellular processes in later stage embryos appear slim. Fortunately, the repertoire of embryonic cell behaviors is limited and most morphogenetic processes can be reduced to a combination of, often interdependent, oriented cell divisions, cell growth, cell division/apoptosis, cell shape changes and cell neighbor exchanges [62]. For each of these processes, a limited number of overarching/general morphogenetic principles have emerged over the past two decades that apply to multiple morphogenetic processes in vertebrate and invertebrate embryos, spurring the development of generic modelling frameworks adaptable to numerous situations [63,64]. Reassuringly, each of these morphogenetic building blocks is controlled by a limited number of proteins, about a hundred in the case of the actomyosin network [65], suggesting that their genetic control is tractable.

In summary, the field of mechanobiology has been very fruitful over the past 15 years. Two major challenges for the future will be to combine partial models accounting for short-term behaviors into a global view of biological processes, and to connect mechanical and genetic information. To this end, multi-agent models have recently been developed [66,67] to provide a more integrative view of multicellular mechanics.

Conclusions

Wolpert thought that it was unlikely that any new general principles of development would be discovered within the following 20 years [1]. There are, however, numerous examples of phenomena whose importance was largely underappreciated at the time, including the role of mechanoperception during development or the emergence of order from stochastic processes. Most strikingly, the past quarter century has brought a fully novel perspective to developmental biology, which is now seen through the combined lenses of genetics, biochemistry and physics. Physical approaches have been at the heart of many of the most important advances in the field. In some cases, they have helped formulate elegant unifying principles, such as the fundamental role of cortical actomyosin networks in morphogenesis or the importance of phase separation. The latter being by essence an all or none process provides, for example, an elegant explanation for the step-wise interpretation of linear gradients pioneered by Wolpert [68]. In other cases, physics applied to single

cells has revealed that mechanisms initially thought to be reasonably simple, such as the interpretation of chemical signals, turn out to show unexpected dynamic complexity when analyzed in single living cells.

Much of the successful modelling efforts reported here were designed to shed light on a particular process, over sufficiently short time to respect the quasi-stationary hypothesis. As a result, our physical understanding of embryogenesis is restricted to a few non-connected kernels of insight. To “compute the embryo”, increasing the number of kernels, which should be easy, will not suffice. Kernels will need to be integrated into a bigger picture. How this should be done is currently unclear, though agent-based models could be interesting tools. Inspiration could also come from whole-cell computational models bridging genotype and phenotype in *Mycoplasma genitalium* [69].

We should also consider which system is most adequate to try and build a global computational model of embryogenesis. *Drosophila melanogaster*, with its powerful genetic and cell biology tools, will remain a major driving force for the identification of core physical principles. The fly embryo may, however, be too complex to integrate these principles into a global model of embryogenesis. Simpler, more transparent embryos with few cells and invariant cell lineages, such as nematodes and ascidians, may be good alternative candidates.

Finally, the examples reported here illustrate the power of combining mathematical modeling with experimental biology. Biologists and physicists have long considered each other with skepticism, and still often struggle to find a common language to translate concepts from one discipline into the other [70]. These cultural problems would be eased by a better integration of biology and physics in university curricula [71].

Conflict of interest statement

Nothing declared.

Acknowledgements

K.B. is supported by a doctoral contract from the Epigenmed Labex. B.L. by a post-doctoral contract from the Dig-Em contract from the Agence Nationale de la Recherche (ANR-14-CE11-0013). P.L. and E.F. are members of CNRS. Work in our laboratories is supported by the Dig-Em contract.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- * of special interest
- ** of outstanding interest

1. Wolpert L: **Do we understand development?** *Science* 1994, **266**:571–572.
2. Mavrakīs M, Pourquié O, Lecuit T: **Lighting up developmental mechanisms: how fluorescence imaging heralded a new era.** *Development* 2010, **137**:373–387.
3. Ma J, Yu MK, Fong S, Ono K, Sage E, Demchak B, Sharan R, Ideker T: **Using deep learning to model the hierarchical structure and function of a cell.** *Nat Methods* 2018, <https://doi.org/10.1038/nmeth.4627>.
An interesting article that successfully uses Visible Neural Networks, based on prior knowledge of the structure of the cell, to simulate a basic eukaryotic cell
4. Thompson DW: *On growth and form.* Cambridge University Press; 1917.
5. Sharpe J: **Computer modeling in developmental biology: growing today, essential tomorrow.** *Development* 2017, **144**:4214–4225.
6. Trepāt X, Sahai E: **Mesoscale physical principles of collective cell organization.** *Nat Phys* 2018, **14**:671–682.
7. Wu P-H, Aroush DR-B, Asnacios A, Chen W-C, Dokukin ME, Doss BL, Durand-Smet P, Ekpenyong A, Guck J, Guz NV, *et al.*: **A comparison of methods to assess cell mechanical properties.** *Nat Methods* 2018, **15**:491–498.
This article compares the results of cell stiffness and viscosity measurements carried out using different measurement and analysis methods and by different groups. They report that elastic and visquous moduli can vary by up several orders of magnitude.
8. Power RM, Huisken J: **A guide to light-sheet fluorescence microscopy for multiscale imaging.** *Nat Methods* 2017, **14**:360–373.
9. Khairy K, Keller PJ: **Reconstructing embryonic development.** *Genes N Y N* 2000, **2011**:488–513.
10. Faure E, Savy T, Rizzi B, Melani C, Stašová O, Fabrèges D, Špir R, Hammons M, Cúnderlík R, Recher G, *et al.*: **A workflow to process 3D+time microscopy images of developing organisms and reconstruct their cell lineage.** *Nat Commun* 2016, **7**:8674.
11. Amat F, Lemon W, Mossing DP, McDole K, Wan Y, Branson K, Myers EW, Keller PJ: **Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data.** *Nat Methods* 2014, **11**:951–958.
12. Stegmaier J, Amat F, Lemon WC, McDole K, Wan Y, Teodoro G, Mikut R, Keller PJ: **Real-time three-dimensional cell segmentation in large-scale microscopy data of developing embryos.** *Dev Cell* 2016, **36**:225–240.
This article presents an open source software allowing accelerated segmentation of very large light-sheet datasets.
13. Guignard L, Fiuza U-M, Leggio B, Faure E, Laussu J, Hufnagel L, Malandain G, Godin C, Lemaire P: **Contact-dependent cell communications drive morphological invariance during ascidian embryogenesis.** *bioRxiv* 2018, <https://doi.org/10.1101/238741>.
This article presents the first successful segmentation and tracking of all cells in a developing animal embryo. Using modeling, it also reveals that cell inductions are controlled by the area of contact between inducing and responding cells.
14. Tokuoka Y, Yamada TG, Hiroi N, Kobayashi TJ, Yamagata K, Funahashi A: **Convolutional neural network-based instance segmentation algorithm to acquire quantitative criteria of early mouse development.** *bioRxiv* 2018, <https://doi.org/10.1101/324186>.
15. Ulman V, Maška M, Magnusson KEG, Ronneberger O, Haubold C, Harder N, Matula P, Matula P, Svoboda D, Radojević M, *et al.*: **An objective comparison of cell-tracking algorithms.** *Nat Methods* 2017, **14**:1141–1152.
This article presents the results of the Cell tracking Challenge and compares the results of several major segmentation and tracking algorithm.
16. Williams E, Moore J, Li SW, Rustici G, Tarkowska A, Chessel A, Leo S, Antal B, Ferguson RK, Sarkans U, *et al.*: **Image data resource: a bioimage data integration and publication platform.** *Nat Methods* 2017, **14**:775–781.
17. Schott B, Traub M, Schlagenhauf C, Takamiya M, Anritter T, Bartschat A, Löffler K, Blessing D, Otte JC, Kobitski AY, *et al.*: **EmbryoMiner: a new framework for interactive knowledge discovery in large-scale cell tracking data of developing embryos.** *PLoS Comput Biol* 2018, **14**, e1006128.

18. Etournay R, Merkel M, Popović M, Brandl H, Dye NA, Aigouy B, Salbreux G, Eaton S, Jülicher F, TissueMiner: **A multiscale analysis toolkit to quantify how cellular processes create tissue dynamics.** *eLife* 2016, **5**.
19. Guirao B, Rigaud SU, Bosveld F, Bailles A, López-Gay J, Ishihara S, Sugimura K, Graner F, Bellaïche Y: **Unified quantitative characterization of epithelial tissue development.** *eLife* 2015, **4**, e08519.
20. Strnad P, Gunther S, Reichmann J, Krzic U, Balazs B, Medeiros G de, Norlin N, Hiiragi T, Hufnagel L, Ellenberg J: **Inverted light-sheet microscope for imaging mouse pre-implantation development.** *Nat Methods* 2016, **13**:139–142.
21. Karaiskos N, Wahle P, Alles J, Boltengagen A, Ayoub S, Kipar C, Kocks C, Rajewsky N, Zinzen RP: **The Drosophila embryo at single-cell transcriptome resolution.** *Science* 2017, **358**:194–199.
- These 4 articles describe the first whole embryo transcriptomes at single-cell resolution in Drosophila, Zebrafish and Xenopus.
22. Farrell JA, Wang Y, Riesenfeld SJ, Shekhar K, Regev A, Schier AF: **Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis.** *Science* 2018, <https://doi.org/10.1126/science.aar3131>.
- These 4 articles describe the first whole embryo transcriptomes at single-cell resolution in Drosophila, Zebrafish and Xenopus.
23. Briggs JA, Weinreb C, Wagner DE, Megason S, Peshkin L, Kirschner MW, Klein AM: **The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution.** *Science* 2018, <https://doi.org/10.1126/science.aar5780>.
- These 4 articles describe the first whole embryo transcriptomes at single-cell resolution in Drosophila, Zebrafish and Xenopus.
24. Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM: **Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo.** *Science* 2018, <https://doi.org/10.1126/science.aar4362>.
- These 4 articles describe the first whole embryo transcriptomes at single-cell resolution in Drosophila, Zebrafish and Xenopus.
25. Budnik B, Levy E, Harmange G, Slavov N: **Mass-spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation.** *bioRxiv* 2018, <https://doi.org/10.1101/102681>.
26. Du Z, Santella A, He F, Shah PK, Kamikawa Y, Bao Z: **The regulatory landscape of lineage differentiation in a metazoan embryo.** *Dev Cell* 2015, **34**:592–607.
- This article provides a road map for the systematic exploration of the fates reached by founder cells following genetic perturbations in *C. elegans*. It highlights systemic canalization into both normal fates and fates not used during normal development.
27. Levine M, Tjian R: **Transcription regulation and animal diversity.** *Nature* 2003, **424**:147–151.
28. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ: **Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position.** *Nat Methods* 2013, **10**:1213–1218.
29. Arnold CD, Gerlach D, Stelzer C, Boryń ŁM, Rath M, Stark A: **Genome-wide quantitative enhancer activity maps identified by STARR-seq.** *Science* 2013, **339**:1074–1077.
- This article presents STARR-seq, a method to functionally test thousands of candidate enhancers *in vivo*
30. Cusanovich DA, Hill AJ, Aghamirzaie D, Daza RM, Pliner HA, Bertetch JB, Philippova GN, Huang X, Christiansen L, DeWitt WS, et al.: **A single-cell atlas of *in vivo* mammalian chromatin accessibility.** *Cell* 2018, **0**.
31. Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferrerres R, Pliner HA, Christiansen L, Qiu X, Steemers FJ, et al.: **The cis-regulatory dynamics of embryonic development at single-cell resolution.** *Nature* 2018, **555**:538–542.
32. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT: **The human transcription factors.** *Cell* 2018, **172**:650–665.
33. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine J-C, Geurts P, Aerts J, et al.: **SCENIC: single-cell regulatory network inference and clustering.** *Nat Methods* 2017, **14**:1083–1086.
34. Crocker J, Ilesley GR, Stern DL: **Quantitatively predictable control of Drosophila transcriptional enhancers *in vivo* with engineered transcription factors.** *Nat Genet* 2016, **48**:292–298.
35. Lenstra TL, Rodriguez J, Chen H, Larson DR: **Transcription dynamics in living cells.** *Annu Rev Biophys* 2016, **45**:25–47.
36. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F: **Mammalian genes are transcribed with widely different bursting kinetics.** *Science* 2011, **332**:472–474.
37. Fukaya T, Lim B, Levine M: **Enhancer control of transcriptional bursting.** *Cell* 2016, **166**:358–368.
38. Banani SF, Lee HO, Hyman AA, Rosen MK: **Biomolecular condensates: organizers of cellular biochemistry.** *Nat Rev Mol Cell Biol* 2017, **18**:285–298.
39. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA: **A phase separation model for transcriptional control.** *Cell* 2017, **169**:13–23.
- The authors present the hypothesis that phase separation of transcriptional regulators may explain some puzzling behaviors of (super-)enhancers.
40. Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, et al.: **Coactivator condensation at super-enhancers links phase separation and gene control.** *Science* 2018, **361**. eaar3958.
- This article shows that coactivators can form liquid-like condensates *in vitro* and *in vivo*.
41. Chen H, Levo M, Barinov L, Fujioka M, Jaynes JB, Gregor T: **Dynamic interplay between enhancer–promoter topology and gene activity.** *Nat Genet* 2018, <https://doi.org/10.1038/s41588-018-0175-z>.
42. Lim B, Heist T, Levine M, Fukaya T: **Visualization of transcription in living Drosophila embryos.** *Mol Cell* 2018, **70**:287–296.e6.
43. Perrimon N, Pitsouli C, Shilo BZ: **Signaling mechanisms controlling cell fate and embryonic patterning.** *Cold Spring Harb Perspect Biol* 2012, **4**. a005975–a005975.
44. Sagner A, Briscoe J: **Morphogen interpretation: concentration, time, competence, and signaling dynamics.** *Wiley Interdiscip Rev Dev Biol* 2017, **6**. n/a-n/a.
45. Nandagopal N, Santat LA, LeBon L, Sprinzak D, Bronner ME, Elowitz MB: **Dynamic ligand discrimination in the Notch signaling pathway.** *Cell* 2018, **172**:869–880.e19.
- This article demonstrates that the area of cell contacts between Notch and delta expressing cells impact sNotch signalling, thus coupling cell geometry to signaling output
46. Antebi YE, Linton JM, Klumpe H, Bintu B, Gong M, Su C, McCardell R, Elowitz MB: **Combinatorial signal perception in the BMP pathway.** *Cell* 2017, **170**:1184–1196.e24.
47. Shaya O, Binshtok U, Hersch M, Rivkin D, Weinreb S, Amir-Zilberstein L, Khamaisi B, Oppenheim O, Desai RA, Goodyear RJ, et al.: **Cell-cell contact area affects Notch signaling and notch-dependent patterning.** *Dev Cell* 2017, **40**:505–511.e6.
48. Frick CL, Yarka C, Nunns H, Goentoro L: **Sensing relative signal in the Tgf-β/Smad pathway.** *Proc Natl Acad Sci* 2017, **114**:E2975–E2982.
- The authors monitored the dynamics of Smad3 activation in living cells. They report that TGF-β signaling is encoded in fold change of Smad3 nuclear localization and that this fold change shows less cell-to-cell variability than absolute levels.
49. Goentoro L, Kirschner MW: **Evidence that fold-change, and not absolute level, of β-catenin dictates Wnt signaling.** *Mol Cell* 2009, **36**:872–884.
50. Balaskas N, Ribeiro A, Panovska J, Dessaud E, Sasai N, Page KM, Briscoe J, Ribes V: **Gene regulatory logic for reading**

- the sonic hedgehog signaling gradient in the vertebrate neural tube. *Cell* 2012, **148**:273–284.
51. Regot S, Hughey JJ, Bajar BT, Carrasco S, Covert MW: **High-sensitivity measurements of multiple kinase activities in live single cells.** *Cell* 2014, **157**:1724–1734.
 52. de la Cova C, Townley R, Regot S, Greenwald I: **A real-time biosensor for ERK activity reveals signaling dynamics during *C. elegans* cell fate specification.** *Dev Cell* 2017, **42**:542–553.e4.
- The article reported the first observation of oscillatory dynamics of the ERK activity in developing embryos. It led researcher to speculate that *in vivo* different levels of EGF produce different frequencies of pulsatile ERK activity.
53. Song J, Ryu H, Chung M, Kim Y, Blum Y, Lee SS, Pertz O, Jeon NL: **Microfluidic platform for single cell analysis under dynamic spatial and temporal stimulation.** *Biosens Bioelectron* 2018, **104**:58–64.
 54. Toettcher JE, Weiner OD, Lim WA: **Using optogenetics to interrogate the dynamic control of signal transmission by the ras/erk module.** *Cell* 2013, **155**:1422–1434.
 55. Johnson HE, Goyal Y, Pannucci NL, Schüpbach T, Shvartsman SY, Toettcher JE: **The spatiotemporal limits of developmental Erk signaling.** *Dev Cell* 2017, **40**:185–192.
- This work reports the first use of optogenetics in *Drosophila* embryos to manipulate Erk signaling *in vivo*, showing that development is sensitive to spatiotemporal change but not intensity of Erk signaling.
56. Johnson HE, Shvartsman SY, Toettcher JE: **Signaling dynamics control cell fate in the early *Drosophila* embryo.** *bioRxiv* 2018, <https://doi.org/10.1101/342998>.
 57. Heisenberg C-P, Bellaïche Y: **Forces in tissue morphogenesis and patterning.** *Cell* 2013, **153**:948–962.
 58. Rose L, Gonczy P: **Polarity establishment, asymmetric division and segregation of fate determinants in early *C. elegans* embryos.** *WormBook* 2014, <https://doi.org/10.1895/wormbook.1.30.2>.
 59. Gross P, Kumar KV, Grill SW: **How active mechanics and regulatory biochemistry combine to form patterns in development.** *Annu Rev Biophys* 2017, **46**:337–356.
 60. Lang CF, Munro E: **The PAR proteins: from molecular circuits to dynamic self-stabilizing cell polarity.** *Development* 2017, **144**:3405–3416.
 61. Kruse K, Joanny JF, Jülicher F, Prost J, Sekimoto K: **Generic theory of active polar gels: a paradigm for cytoskeletal dynamics.** *Eur Phys J E* 2005, **16**:5–16.
- In this landmark article, the authors present a first general theory for active viscoelastic materials made of polar filaments such as actin.
62. Etournay R, Popović M, Merkel M, Nandi A, Blasse C, Aigouy B, Brandl H, Myers G, Salbreux G, Jülicher F, *et al.*: **Interplay of cell dynamics and epithelial tension during morphogenesis of the *Drosophila* pupal wing.** *eLife* 2015, **4**, e07090.
 63. Pierre A, Sallé J, Wühr M, Minc N: **Generic theoretical models to predict division patterns of cleaving embryos.** *Dev Cell* 2016, **39**:667–682.
 64. Nedelec F, Foethke D: **Collective Langevin dynamics of flexible cytoskeletal fibers.** *New J Phys* 2007, **9**:427.
 65. Zaidel-Bar R, Zhenhuan G, Luxenburg C: **The contractome – a systems view of actomyosin contractility in non-muscle cells.** *J Cell Sci* 2015, **128**:2209–2217.
 66. Delile J, Herrmann M, Peyriéras N, Doursat R: **A cell-based computational model of early embryogenesis coupling mechanical behaviour and gene regulation.** *Nat Commun* 2017, **8**:13929.
 67. Nissen SB, Perera M, Gonzalez JM, Morgani SM, Jensen MH, Sneppen K, Brickman JM, Trusina A: **Four simple rules that are sufficient to generate the mammalian blastocyst.** *PLoS Biol* 2017, **15**, e2000737.
 68. Dine E, Gil AA, Uribe G, Brangwynne CP, Toettcher JE: **Protein phase separation provides long-term memory of transient spatial stimuli.** *Cell Syst* 2018, **6**:655–663.e5.
 69. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW: **A whole-cell computational model predicts phenotype from genotype.** *Cell* 2012, **150**:389–401.
- This article reports the first successful whole-cell model of a bacterial parasite, integrating 28 functional modules.
70. Bentovim L, Harden TT, DePace AH: **Transcriptional precision and accuracy in development: from measurements to models and mechanisms.** *Development* 2017, **144**:3855–3866.
 71. Bauer G, Fakhri N, Kicheva A, Kondev J, Kruse K, Noji H, Riveline D, Saunders TE, Thattai M, Wieschaus E: **The science of living matter for tomorrow.** *Cell Syst* 2018, **6**:400–402.

to produce the diversity of fate specification events occurring during embryogenesis (Perrimon et al. 2012). Despite this, many simple questions remain unanswered, such as: “What is the mechanism regulating the dose-response to increasing concentrations of ligands or receptors?” “How are ligand concentration and time of exposure integrated by cells?” To deepen the understanding of the principles which govern embryonic development, it is important to combine quantitative experimental approaches at the cellular scale with dynamic mathematical models including mechanistic details. For example, the recent development in quantitative imaging, sequencing, proteomics and physical measurements have allowed us to refine the historical morphogen concept, in which diffusible signaling molecules are proposed to coordinate cell fate specification and tissue formation using concentration-dependent mechanisms (a static readout), because it was insufficient to describe or model the complexities of patterning observed with these techniques in developing embryos (Garcia et al. 2020; Huang and Saunders 2020; Jaeger and Verd 2020; Rogers and Müller 2020; Schloop et al. 2020).

While physical modeling of life has a long history (Thompson 1917), it has remained a theoretical exercise for a long time: insufficient measurements of physical parameters for constraining models coupled with a largely qualitative and static description of phenotypes have rendered it difficult to apply physics to developing embryos and even to single cells. The recent technological breakthrough mentioned previously, however, reduced this difficulty while making “computable embryos” through a precise physical description of embryonic development more necessary than ever to capture key developmental concepts and bridge genomic information and dynamic phenotypes (Biasuz et al. 2018). First, our brains are simply unable to cope with the large amount of data generated, much of which are unrelated to the mechanism being studied. Second, biology involves several layers of feedback, resulting in unintuitive non-linear behaviors. Third, biology is a multiscale process in which macroscopic properties of cells and tissues arise from the mesoscopic properties of molecules or subcellular structures.

Ascidians definitively constitute a model of choice to build a global computational model of embryogenesis. Embryonic development is a continuous progression in time. The “computable embryo” is based on the idea that a mathematical description of the system can predict the future state of the embryo from the knowledge of its current state. This global computational model of embryogenesis at the single-cell, genome-wide and whole-embryo level is a challenging task and will only be achieved using the most appropriate developmental systems (Biasuz et al. 2018).

Solitary ascidian embryos seem to be good candidates for this breakthrough. At first glance, one would rather think of the *Drosophila melanogaster* or vertebrate embryos for this role. Indeed, thanks to decades of research, a deep understanding of core developmental mechanisms has been achieved, and powerful genetic and cell biology tools exist.

20.9.3 WHEN DEVELOPMENTAL BIOLOGY BECOMES QUANTITATIVE: A BIG STEP TOWARD “COMPUTABLE EMBRYOS”

The transition from a single fertilized cell to a complex organism, with various cell types that compose its tissues in the correct numbers and their fine regulation in space and time, is the question at the heart of developmental biology. Decades of research in this field have designed a broad portrait of the fundamental processes involved during embryogenesis: from the description of the genetic programs of embryonic cells and the mechanisms regulating gene transcription to how cell fates and behaviors are coordinated by cell communication and the way this translates into morphogenesis.

Developmental mechanisms have traditionally been studied at the tissue level in a qualitative manner. For example, consider the current view of the classical chemical signaling during fate specification. A surprisingly small number of signaling pathways involving cell surface receptor and activating ligands act in widely different cellular contexts

These embryos, while remaining a significant motor for defining new concepts, may, however, be too complicated to incorporate these concepts into a global model of embryogenesis. In contrast, ascidian embryos, as nematodes, are simpler and develop stereotypically with few cells and invariant cell lineages, so that each cell can be named and found at the same position in all embryos (Lemaire 2009). Unlike those of nematodes (Goldstein 2001), ascidian embryo geometries have even remained essentially unchanged since the emergence of the group, around 400 million years ago, despite extensive genomic divergence (Delsuc et al. 2018; Lemaire 2011). The development of ascidians is also characterized by earlier fate restriction than most animal embryos: 94 of the 112 early gastrula cells in the ascidian *Ciona* are fate restricted, each contributing to a single larval tissue type (Nishida 1987). Moreover, ascidians are closely related to vertebrates, as they belong to the vertebrate sister-group, but ascidians kept their genomic simplicity. Indeed, they diverged before the two rounds of whole-genome duplication events which occurred in the vertebrate lineage leading to the apparition of multiple paralogues for each gene (Dehal and Boore 2005), with potentially slightly divergent activities. Finally, ascidian embryos are small (~130 μm) and transparent, and they develop rapidly externally in sea water up to the larval stage (~12 h), making them very easy to image. Thus, ascidian embryos provide a rigid framework that allows combination of analyses at cellular resolution with mathematical modeling.

These advantageous properties of ascidian embryos, especially *Phallusia mammillata* embryos, which are fully transparent, combined with the breakthrough development of light sheet microscopy (Power and Huisken 2017), have enabled the production of the first digitized version of a metazoan embryo (Figure 20.7a) (Guignard et al. 2020). Based on automatic whole-cell segmentation and tracking over five cell generations of membrane-labeled cells with two-minute temporal resolution, this research offers a complete description of early ascidian embryo development, accounting for each cell in the ten embryos analyzed. Moreover, this quantitative and dynamic atlas of cell positions and geometries can be associated with the known cell fates and interactively explored through the MorphoNet online morphological browser (Leggio et al. 2019). These “digital embryos” show that ascidian development is reproducible down to the scale of cell–cell contacts and, combined with modeling and experimental manipulations, it allows us to establish contact area-dependent inductions as an alternative to classical morphogen gradients. This work opens the door to quantitative single-cell morphology and mechanical morphogenesis modeling.

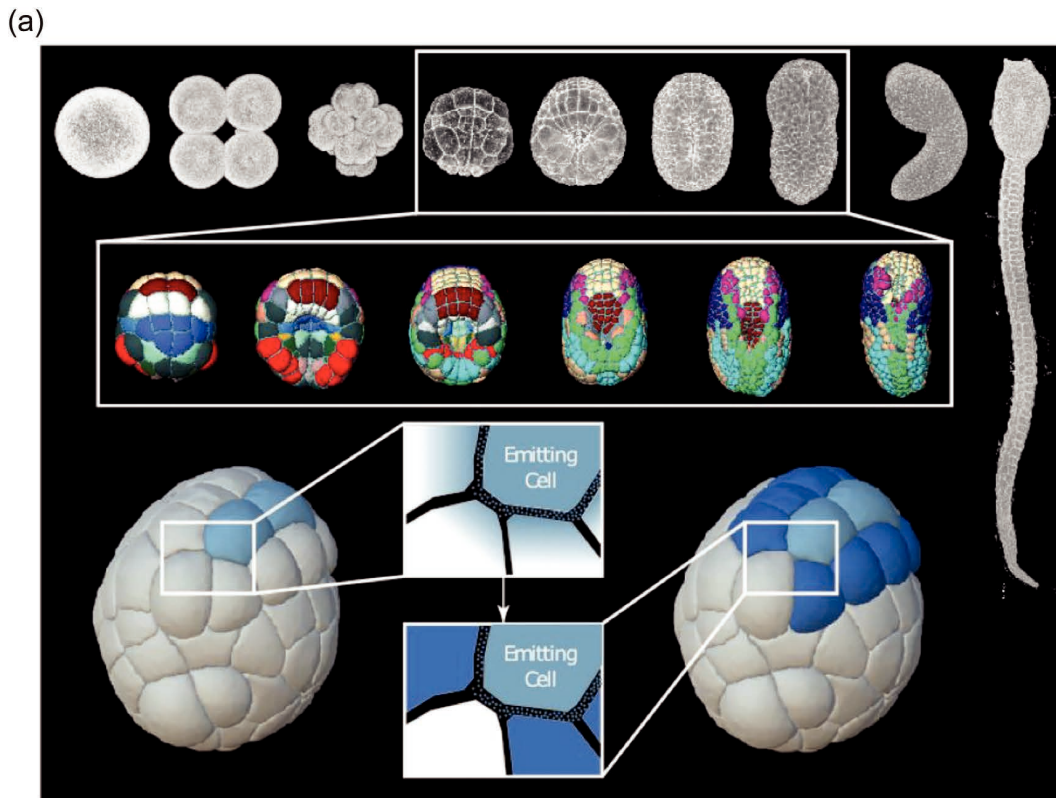
In parallel with this work, another group combined high-resolution single-cell transcriptomics (single-cell RNA sequencing) and light-sheet imaging to build the first full comprehensive atlas which describes the genome-wide gene expression of every single cell of an embryo in the early stages of development, showing the evolution from a single cell up to gastrulation in the ascidian *Phallusia mammillata* (Sladitschek et al. 2020). By providing a complete

representation of the gene expression programs, which instruct individual cells to form the different cell types necessary to build an embryo, and therefore by allowing us to know precisely cell-specific expression of transcription factors at the single-cell level, this study will significantly enhance current single-cell-based gene regulatory network inference algorithms (Aibar et al. 2017) and will help to further develop single-cell-based physical models of the different steps of transcriptional control during development. Moreover, these single-cell gene expression data will feed several layers of physical description of biological processes. For example, identification of cell-adhesion molecules will allow the refining of morphogenetic models, such as oriented cell divisions, cell shape changes or cell neighbor exchanges models (Etournay et al. 2015), thereby linking mechanical and genetic information at the cellular resolution.

In spite of the convenient properties and the recent advances that have been realized thanks to the ascidian embryo model, there is still a long way to go to be able to “compute the embryo”.

Typically, studies at the single-cell level are in their early days, as can be illustrated by signal transduction studies. The MAPK/ERK signaling pathway is one of the important embryonic signaling pathways used by vertebrates and invertebrates, controlling many physiological processes (Lavoie et al. 2020), and is the main inducing pathway in early ascidian embryos (Lemaire 2009). The signaling cascade from the activation of the transmembrane receptor to the phosphorylation of the ERK nuclear targets is well described (Figure 20.7b) (Lavoie et al. 2020). Our current knowledge of this pathway is, however, mostly static, and an integrated understanding of its spatio-temporal dynamics is lacking (Patel and Shvartsman 2018). For example, it has been shown that the ERK pathway can trigger two qualitatively different types of ERK activity: pulsatile or continuous (Aikin et al. 2019). To understand these non-intuitive results, it is important to combine quantitative experimental approaches at the cellular resolution with dynamic mathematical models including mechanistic details. Genetically encoded fluorescent activity sensors that convert kinase activity into nucleocytoplasmic events have been recently developed (Durandau et al. 2015; Regot et al. 2014), and these tools can now be coupled with optogenetic systems in order to activate the ERK pathway with high spatiotemporal accuracy at different levels (Gagliardi and Pertz 2019). However, these techniques were only used to track a single pathway component at a time. Yet they suggest that multiplexing sensors at different levels of the cascade could reveal the dynamics of information flow through the cell. Such quantitative measures are required to more realistically model the catalogue of cell-signaling modalities (Biasuz et al. 2018).

The technological breakthroughs of the last quarter of the century have brought a whole new perspective to developmental biology, which is now seen through the combined lenses of mathematical modeling and experimental biology. A major challenge for the future will now be to integrate



(b)

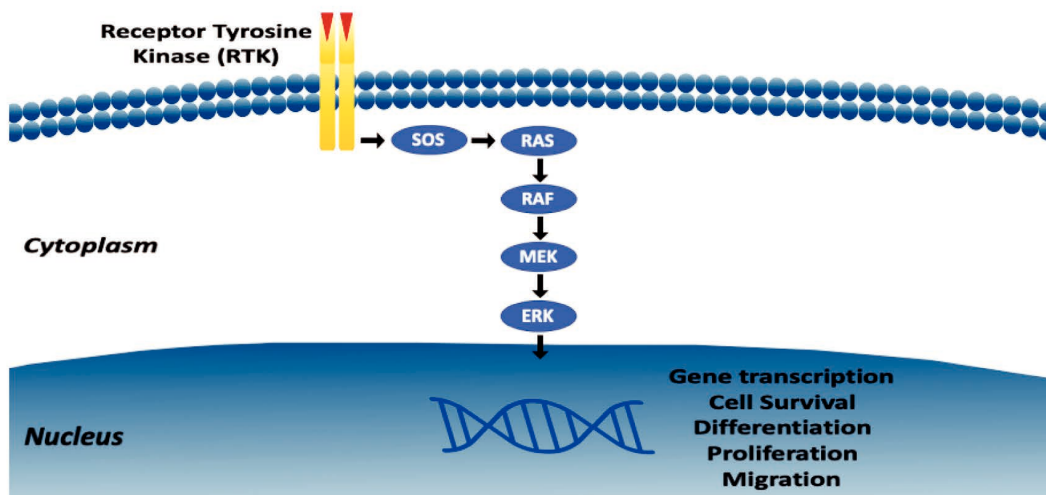


FIGURE 20.7 (a) Digitalization of *Phallusia mammillata* embryogenesis reveal contact area-dependent cell inductions. (Top) Light-sheet imaging of cell membranes (not shown) combined with automated cell segmentation and tracking allowed reconstruction of *Phallusia* embryogenesis between the 64-cell and initial tailbud stages. Digital embryos represented here are color-coded with cell fates. (Bottom) Illustration of the contact area-dependent mode of cell inductions. Light blue cells emit inducing extracellular signals (left). Among the neighbor cells which receive the signal, only the dark blue cells, which have the largest surface of contact with emitting cells, are induced (right). Digital embryos have been explored through the MorphoNet online morphological browser. (b) Simplified representation of the MAPK/ERK signaling pathway. [a] Figure courtesy of Leo Guignard & Kilian Biasuz; [b] figure courtesy of Kilian Biasuz.)

partial models accounting for short-term activities into a global view of biological processes. Indeed, most of the modeling efforts were designed to shed light on specific processes over a short period of time. As a consequence, our physical knowledge of embryogenesis is reduced to a few unconnected kernels of insight. Increasing the number of kernels is imperative to “compute the embryo” but will not suffice: kernels will need to be incorporated into a bigger picture. The solitary ascidian embryos, which are simple and transparent and contain a relatively small number of cells and invariant cell lineages are perfect candidates to integrate these principles into a global model of embryogenesis.

RÉFÉRENCES

- Acampora, D., Gulisano, M., Broccoli, V., Simeone, A., 2001. Otx genes in brain morphogenesis. *Prog. Neurobiol.* 64, 69–95. [https://doi.org/10.1016/s0301-0082\(00\)00042-3](https://doi.org/10.1016/s0301-0082(00)00042-3)
- Affolter, M., Basler, K., 2007. The Decapentaplegic morphogen gradient: from pattern formation to growth regulation. *Nat. Rev. Genet.* 8, 663–674. <https://doi.org/10.1038/nrg2166>
- Albeck, J.G., Mills, G.B., Brugge, J.S., 2013. Frequency-Modulated Pulses of ERK Activity Transmit Quantitative Proliferation Signals. *Mol. Cell* 49, 249–261. <https://doi.org/10.1016/j.molcel.2012.11.002>
- Alexandre, C., Baena-Lopez, A., Vincent, J.-P., 2014. Patterning and growth control by membrane-tethered Wingless. *Nature* 505, 180–5. <https://doi.org/10.1038/nature12879>
- Amos, W.B., White, J.G., 2003. How the confocal laser scanning microscope entered biological research. *Biol. Cell* 95, 335–342. [https://doi.org/10.1016/s0248-4900\(03\)00078-9](https://doi.org/10.1016/s0248-4900(03)00078-9)
- Antebi, Y.E., Nandagopal, N., Elowitz, M.B., 2017. An operational view of intercellular signaling pathways. *Curr. Opin. Syst. Biol.* 1, 16–24. <https://doi.org/10.1016/j.coisb.2016.12.003>
- Aoki, K., Kumagai, Y., Sakurai, A., Komatsu, N., Fujita, Y., Shionyu, C., Matsuda, M., 2013. Stochastic ERK Activation Induced by Noise and Cell-to-Cell Propagation Regulates Cell Density-Dependent Proliferation. *Mol. Cell* 52, 529–540. <https://doi.org/10.1016/j.molcel.2013.09.015>
- Balaskas, N., Ribeiro, A., Panovska, J., Dessaud, E., Sasai, N., Page, K.M., Briscoe, J., Ribes, V., 2012. Gene Regulatory Logic for Reading the Sonic Hedgehog Signaling Gradient in the Vertebrate Neural Tube. *Cell* 148, 273–284. <https://doi.org/10.1016/j.cell.2011.10.047>
- Bertrand, V., Hudson, C., Caillol, D., Popovici, C., Lemaire, P., 2003. Neural tissue in ascidian embryos is induced by FGF9/16/20, acting via a combination of maternal GATA and Ets transcription factors. *Cell* 115, 615–627. [https://doi.org/10.1016/s0092-8674\(03\)00928-0](https://doi.org/10.1016/s0092-8674(03)00928-0)
- Biasuz, K., Leggio, B., Faure, E., Lemaire, P., 2018. The “computable egg”: Myth or useful concept? *Curr. Opin. Syst. Biol.* 11, 91–97. <https://doi.org/10.1016/j.coisb.2018.09.003>
- Blum, Y., Mikelson, J., Dobrzyński, M., Ryu, H., Jacques, M.-A., Jeon, N.L., Khammash, M., Pertz, O., 2019. Temporal perturbation of ERK dynamics reveals network architecture of FGF2/MAPK signaling. *Mol. Syst. Biol.* 15, e8947. <https://doi.org/10.15252/msb.20198947>
- Bouchemousse, S., Liautard-Haag, C., Bierne, N., Viard, F., 2016. Distinguishing contemporary hybridization from past introgression with postgenomic ancestry-informative SNPs in strongly differentiated *Ciona* species. *Mol. Ecol.* 25, 5527–5542. <https://doi.org/10.1111/mec.13854>
- Boutet, A., Schierwater, B. (Eds.), 2021. *Handbook of Marine Model Organisms in Experimental Biology: Established and Emerging*. CRC Press, Boca Raton. <https://doi.org/10.1201/9781003217503>
- Bray, S.J., 2006. Notch signalling: a simple pathway becomes complex. *Nat. Rev. Mol. Cell Biol.* 7, 678–689. <https://doi.org/10.1038/nrm2009>
- Brozovic, M., Dantec, C., Dardaillon, J., Dauga, D., Faure, E., Gineste, M., Louis, A., Naville, M., Nitta, K.R., Piette, J., Reeves, W., Scornavacca, C., Simion, P., Vincentelli, R., Bellec, M., Aicha, S.B., Fagotto, M., Guérault-Bellone, M., Haeussler, M., Jacox, E., Lowe, E.K., Mendez, M., Roberge, A., Stolfi, A., Yokomori, R., Brown, C.T., Cambillau, C., Christiaen, L., Delsuc, F., Douzery, E., Dumollard, R., Kusakabe, T., Nakai, K., Nishida, H., Satou, Y., Swalla, B., Veeman, M., Volff, J.-N., Lemaire, P., 2018. ANISEED 2017: extending the integrated ascidian database to the exploration and evolutionary comparison of genome-scale datasets. *Nucleic Acids Res.* 46, D718–D725. <https://doi.org/10.1093/nar/gkx1108>
- Brunetti, R., Gissi, C., Pennati, R., Caicci, F., Gasparini, F., Manni, L., 2015. Morphological evidence that the molecularly determined *Ciona intestinalis* type A and type B are different species: *Ciona robusta* and *Ciona intestinalis*. *J. Zool. Syst. Evol. Res.* 53, 186–193. <https://doi.org/10.1111/jzs.12101>
- Butler, M.T., Wallingford, J.B., 2017. Planar cell polarity in development and disease. *Nat. Rev. Mol. Cell Biol.* 18, 375–388. <https://doi.org/10.1038/nrm.2017.11>
- Cao, C., Lemaire, L.A., Wang, W., Yoon, P.H., Choi, Y.A., Parsons, L.R., Matese, J.C., Wang, W., Levine,

- M., Chen, K., 2019. Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* 571, 349–354. <https://doi.org/10.1038/s41586-019-1385-y>
- Čapek, D., Müller, P., 2019. Positional information and tissue scaling during development and regeneration. *Development* 146, dev177709. <https://doi.org/10.1242/dev.177709>
- Castro-González, C., Ledesma-Carbayo, M.J., Peyriéras, N., Santos, A., 2012. Assembling models of embryo development: Image analysis and the construction of digital atlases. *Birth Defects Res. Part C Embryo Today Rev.* 96, 109–120. <https://doi.org/10.1002/bdrc.21012>
- Chabry, L., 1887. *Chabry: Embryologie normale et tératologique des Ascidies*. Felix Alcan Ed. Paris.
- Chazaud, C., Yamanaka, Y., Pawson, T., Rossant, J., 2006. Early Lineage Segregation between Epiblast and Primitive Endoderm in Mouse Blastocysts through the Grb2-MAPK Pathway. *Dev. Cell* 10, 615–624. <https://doi.org/10.1016/j.devcel.2006.02.020>
- Cole, A.G., Meinertzhagen, I.A., 2004. The central nervous system of the ascidian larva: mitotic history of cells forming the neural tube in late embryonic *Ciona intestinalis*. *Dev. Biol.* 271, 239–262. <https://doi.org/10.1016/j.ydbio.2004.04.001>
- Collier, J.R., Monk, N.A.M., Maini, P.K., Lewis, J.H., 1996. Pattern Formation by Lateral Inhibition with Feedback: a Mathematical Model of Delta-Notch Intercellular Signalling. *J. Theor. Biol.* 183, 429–446. <https://doi.org/10.1006/jtbi.1996.0233>
- Conklin, E.G., 1905. Mosaic development in ascidian eggs. *J. Exp. Zool.* 2, 145–223. <https://doi.org/10.1002/jez.1400020202>
- Couso, J.P., Bishop, S.A., Martinez Arias, A., 1994. The wingless signalling pathway and the patterning of the wing margin in *Drosophila*. *Dev. Camb. Engl.* 120, 621–636. <https://doi.org/10.1242/dev.120.3.621>
- Crick, F., 1970. Diffusion in Embryogenesis. *Nature* 225, 420–422. <https://doi.org/10.1038/225420a0>
- Danesh, S.M., Villasenor, A., Chong, D., Soukup, C., Cleaver, O., 2009. BMP and BMP receptor expression during murine organogenesis. *Gene Expr. Patterns* 9, 255–265. <https://doi.org/10.1016/j.gexp.2009.04.002>
- De Jossineau, C., Soulé, J., Martin, M., Anguille, C., Montcourrier, P., Alexandre, D., 2004. Delta-promoted filopodia mediate long-range lateral inhibition in *Drosophila*. *Nature*. <https://doi.org/10.1038/nature02157>
- de la Cova, C., Townley, R., Regot, S., Greenwald, I., 2017. A Real-Time Biosensor for ERK Activity Reveals Signaling Dynamics during *C. elegans* Cell Fate Specification. *Dev. Cell* 42, 542–553.e4. <https://doi.org/10.1016/j.devcel.2017.07.014>
- de Medeiros, G., Norlin, N., Gunther, S., Albert, M., Panavaite, L., Fiuza, U.-M., Peri, F., Hiiragi, T., Krzic, U., Hufnagel, L., 2015. Confocal multiview light-sheet microscopy. *Nat. Commun.* 6, 8881. <https://doi.org/10.1038/ncomms9881>
- De Robertis, E.M., 2009. Spemann's organizer and the self-regulation of embryonic fields. *Mech. Dev.* 126, 925–941. <https://doi.org/10.1016/j.mod.2009.08.004>
- Débarre, D., Olivier, N., Supatto, W., Beaurepaire, E., 2014. Mitigating Phototoxicity during Multiphoton Microscopy of Live *Drosophila* Embryos in the 1.0–1.2 μm Wavelength Range. *PLoS ONE* 9, e104250. <https://doi.org/10.1371/journal.pone.0104250>
- Dehal, P., Boore, J.L., 2005. Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. *PLOS Biol.* 3, e314. <https://doi.org/10.1371/journal.pbio.0030314>
- Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D.M., Harafuji, N., Hastings, K.E.M., Ho, I., Hotta, K., Huang, W., Kawashima, T., Lemaire, P., Martinez, D., Meinertzhagen, I.A., Necula, S., Nonaka, M., Putnam, N., Rash, S., Saiga, H., Satake, M., Terry, A., Yamada, L., Wang, H.-G., Awazu, S., Azumi, K., Boore, J., Branno, M., Chin-Bow, S., DeSantis, R., Doyle, S., Francino, P., Keys, D.N., Haga, S., Hayashi, H., Hino, K., Imai, K.S., Inaba, K., Kano, S., Kobayashi, K., Kobayashi, M., Lee, B.-I., Makabe, K.W., Manohar, C., Matassi, G., Medina, M., Mochizuki, Y., Mount, S., Morishita, T., Miura, S., Nakayama, A., Nishizaka, S., Nomoto, H., Ohta, F., Oishi, K., Rigoutsos, I., Sano, M., Sasaki, A., Sasakura, Y., Shoguchi, E., Shin-i, T., Spagnuolo, A., Stainier, D., Suzuki,

- M.M., Tassy, O., Takatori, N., Tokuoka, M., Yagi, K., Yoshizaki, F., Wada, S., Zhang, C., Hyatt, P.D., Larimer, F., Detter, C., Doggett, N., Glavina, T., Hawkins, T., Richardson, P., Lucas, S., Kohara, Y., Levine, M., Satoh, N., Rokhsar, D.S., 2002. The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* 298, 2157–2167. <https://doi.org/10.1126/science.1080049>
- Delsuc, F., Brinkmann, H., Chourrout, D., Philippe, H., 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439, 965–968. <https://doi.org/10.1038/nature04336>
- Delsuc, F., Philippe, H., Tsagkogeorga, G., Simion, P., Tilak, M.-K., Turon, X., López-Legentil, S., Piette, J., Lemaire, P., Douzery, E.J.P., 2018. A phylogenomic framework and timescale for comparative studies of tunicates. *BMC Biol.* 16, 39. <https://doi.org/10.1186/s12915-018-0499-2>
- Delsuc, F., Tsagkogeorga, G., Lartillot, N., Philippe, H., 2008. Additional molecular support for the new chordate phylogeny. *Genesis* 46, 592–604. <https://doi.org/10.1002/dvg.20450>
- Deng, C.X., Wynshaw-Boris, A., Shen, M.M., Daugherty, C., Ornitz, D.M., Leder, P., 1994. Murine FGFR-1 is required for early postimplantation growth and axial organization. *Genes Dev.* 8, 3045–3057. <https://doi.org/10.1101/gad.8.24.3045>
- Denk, W., Strickler, J.H., Webb, W.W., 1990. Two-photon laser scanning fluorescence microscopy. *Science* 248, 73–76. <https://doi.org/10.1126/science.2321027>
- Denoëud, F., Henriët, S., Mungpakdee, S., Aury, J.-M., Da Silva, C., Brinkmann, H., Mikhaleva, J., Olsen, L.C., Jubin, C., Cañestro, C., Bouquet, J.-M., Danks, G., Poulain, J., Campsteijn, C., Adamski, M., Cross, I., Yadetie, F., Muffato, M., Louis, A., Butcher, S., Tsagkogeorga, G., Konrad, A., Singh, S., Jensen, M.F., Cong, E.H., Eikeseth-Otteraa, H., Noel, B., Anthouard, V., Porcel, B.M., Kachouri-Lafond, R., Nishino, A., Ugolini, M., Chourrout, P., Nishida, H., Aasland, R., Huzurbazar, S., Westhof, E., Delsuc, F., Lehrach, H., Reinhardt, R., Weissenbach, J., Roy, S.W., Artiguenave, F., Postlethwait, J.H., Manak, J.R., Thompson, E.M., Jaillon, O., Pasquier, L.D., Boudinot, P., Liberles, D.A., Volff, J.-N., Philippe, H., Lenhard, B., Crollius, H.R., Wincker, P., Chourrout, D., 2010. Plasticity of Animal Genome Architecture Unmasked by Rapid Evolution of a Pelagic Tunicate. *Science* 330, 1381–1385. <https://doi.org/10.1126/science.1194167>
- Dessaud, E., Yang, L.L., Hill, K., Cox, B., Ulloa, F., Ribeiro, A., Mynett, A., Novitch, B.G., Briscoe, J., 2007. Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. *Nature* 450, 717–720. <https://doi.org/10.1038/nature06347>
- Duboule, D., 1994. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Dev. Camb. Engl. Suppl.* 135–142.
- Dumollard, R., Hebras, C., Besnardeau, L., McDougall, A., 2013. Beta-catenin patterns the cell cycle during maternal-to-zygotic transition in urochordate embryos. *Dev. Biol.* 384, 331–342. <https://doi.org/10.1016/j.ydbio.2013.10.007>
- Durandau, E., Aymoz, D., Pelet, S., 2015. Dynamic single cell measurements of kinase activity by synthetic kinase activity relocation sensors. *BMC Biol.* 13, 55. <https://doi.org/10.1186/s12915-015-0163-z>
- Ericson, J., Rashbass, P., Schedl, A., Brenner-Morton, S., Kawakami, A., van Heyningen, V., Jessell, T.M., Briscoe, J., 1997. Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling. *Cell* 90, 169–180. [https://doi.org/10.1016/s0092-8674\(00\)80323-2](https://doi.org/10.1016/s0092-8674(00)80323-2)
- Ferree, P.L., Deneke, V.E., Di Talia, S., 2016. Measuring time during early embryonic development. *Semin. Cell Dev. Biol.* 55, 80–88. <https://doi.org/10.1016/j.semcd.2016.03.013>
- Fey, D., Matallanas, D., Rauch, J., Rukhlenko, O.S., Kholodenko, B.N., 2016. The complexities and versatility of the RAS-to-ERK signalling system in normal and cancer cells. *Semin. Cell Dev. Biol., Cardiac Regeneration* 58, 96–107. <https://doi.org/10.1016/j.semcd.2016.06.011>
- Filonov, G.S., Piatkevich, K.D., Ting, L.-M., Zhang, J., Kim, K., Verkhusha, V.V., 2011. Bright and stable

- near infra-red fluorescent protein for in vivo imaging. *Nat. Biotechnol.* 29, 757–761. <https://doi.org/10.1038/nbt.1918>
- Förster, Th., 1948. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* 437, 55–75. <https://doi.org/10.1002/andp.19484370105>
- Freeman, M., 2000. Feedback control of intercellular signalling in development. *Nature* 408, 313–319. <https://doi.org/10.1038/35042500>
- Frick, C.L., Yarka, C., Nunns, H., Goentoro, L., 2017. Sensing relative signal in the Tgf- β /Smad pathway. *Proc. Natl. Acad. Sci. U. S. A.* 114, E2975–E2982. <https://doi.org/10.1073/pnas.1611428114>
- Fritz, R.D., Letzelter, M., Reimann, A., Martin, K., Fusco, L., Ritsma, L., Ponsioen, B., Fluri, E., Schulte-Merker, S., van Rheenen, J., Pertz, O., 2013. A Versatile Toolkit to Produce Sensitive FRET Biosensors to Visualize Signaling in Time and Space. *Sci. Signal.* 6, rs12–rs12. <https://doi.org/10.1126/scisignal.2004135>
- Gagliardi, P.A., Dobrzyński, M., Jacques, M.-A., Dessauges, C., Ender, P., Blum, Y., Hughes, R.M., Cohen, A.R., Pertz, O., 2021. Collective ERK/Akt activity waves orchestrate epithelial homeostasis by driving apoptosis-induced survival. *Dev. Cell* 56, 1712–1726.e6. <https://doi.org/10.1016/j.devcel.2021.05.007>
- Gerhart, J., 1999. 1998 warkany lecture: Signaling pathways in development. *Teratology* 60, 226–239. [https://doi.org/10.1002/\(SICI\)1096-9926\(199910\)60:4<226::AID-TERA7>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1096-9926(199910)60:4<226::AID-TERA7>3.0.CO;2-W)
- Gillies, T.E., Pargett, M., Minguet, M., Davies, A.E., Albeck, J.G., 2017. Linear integration of ERK activity predominates over persistence detection in Fra-1 regulation. *Cell Syst.* 5, 549–563.e5. <https://doi.org/10.1016/j.cels.2017.10.019>
- Goentoro, L., Kirschner, M.W., 2009. Evidence that Fold-Change, and Not Absolute Level, of β -Catenin Dictates Wnt Signaling. *Mol. Cell* 36, 872–884. <https://doi.org/10.1016/j.molcel.2009.11.017>
- Goglia, A.G., Wilson, M.Z., Jena, S.G., Silbert, J., Basta, L.P., Devenport, D., Toettcher, J.E., 2020. A Live-Cell Screen for Altered Erk Dynamics Reveals Principles of Proliferative Control. *Cell Syst.* 10, 240–253.e6. <https://doi.org/10.1016/j.cels.2020.02.005>
- Goldstein, B., 2001. On the evolution of early development in the Nematoda. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 356, 1521–1531. <https://doi.org/10.1098/rstb.2001.0977>
- Goyal, Y., Jindal, G.A., Pelliccia, J.L., Yamaya, K., Yeung, E., Futran, A.S., Burdine, R.D., Schüpbach, T., Shvartsman, S.Y., 2017. Divergent effects of intrinsically active MEK variants on developmental Ras signaling. *Nat. Genet.* 49, 465–469. <https://doi.org/10.1038/ng.3780>
- Greene, L.A., Tischler, A.S., 1976. Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor. *Proc. Natl. Acad. Sci.* 73, 2424–2428. <https://doi.org/10.1073/pnas.73.7.2424>
- Guignard, L., 2015. Quantitative analysis of animal morphogenesis: from high-throughput laser imaging to 4D virtual embryo in ascidians (phdthesis). Université Montpellier.
- Guignard, L., Fiúza, U.-M., Leggio, B., Laussu, J., Faure, E., Michelin, G., Biasuz, K., Hufnagel, L., Malandain, G., Godin, C., Lemaire, P., 2020. Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis. *Science* 369. <https://doi.org/10.1126/science.aar5663>
- Guo, G., Huss, M., Tong, G.Q., Wang, C., Sun, L.L., Clarke, N.D., Robson, P., 2010. Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Dev. Cell* 18, 675–685. <https://doi.org/10.1016/j.devcel.2010.02.012>
- Gurdon, J.B., Dyson, S., Johnston, D.S., 1998. Cells' Perception of Position in a Concentration Gradient. *Cell* 95, 159–162. [https://doi.org/10.1016/S0092-8674\(00\)81747-X](https://doi.org/10.1016/S0092-8674(00)81747-X)
- Halaban, R., Zhang, W., Bacchiocchi, A., Cheng, E., Parisi, F., Ariyan, S., Krauthammer, M., McCusker, J.P., Kluger, Y., Sznol, M., 2010. PLX4032, a selective BRAF(V600E) kinase inhibitor, activates the ERK pathway and enhances cell migration and proliferation of BRAF melanoma cells. *Pigment Cell Melanoma Res.* 23, 190–200. <https://doi.org/10.1111/j.1755-148X.2010.00685.x>
- Hannanta-anan, P., Chow, B.Y., 2016. Optogenetic Control of Calcium Oscillation Waveform Defines

- NFAT as an Integrator of Calcium Load. *Cell Syst.* 2, 283–288. <https://doi.org/10.1016/j.cels.2016.03.010>
- Harvey, C.D., Ehrhardt, A.G., Cellurale, C., Zhong, H., Yasuda, R., Davis, R.J., Svoboda, K., 2008. A genetically encoded fluorescent sensor of ERK activity. *Proc. Natl. Acad. Sci.* 105, 19264–19269. <https://doi.org/10.1073/pnas.0804598105>
- Hatano, N., Mori, Y., Oh-hora, M., Kosugi, A., Fujikawa, T., Nakai, N., Niwa, H., Miyazaki, J., Hamaoka, T., Ogata, M., 2003. Essential role for ERK2 mitogen-activated protein kinase in placental development. *Genes Cells* 8, 847–856. <https://doi.org/10.1046/j.1365-2443.2003.00680.x>
- Haupaix, N., Stolfi, A., Sirour, C., Picco, V., Levine, M., Christiaen, L., Yasuo, H., 2013. p120RasGAP mediates ephrin/Eph-dependent attenuation of FGF/ERK signals during cell fate specification in ascidian embryos. *Dev. Camb. Engl.* 140, 4347–4352. <https://doi.org/10.1242/dev.098756>
- Hayashi, S., Ogura, Y., 2020. ERK signaling dynamics in the morphogenesis and homeostasis of *Drosophila*. *Curr. Opin. Genet. Dev., Developmental Mechanisms, patterning and evolution* 63, 9–15. <https://doi.org/10.1016/j.gde.2020.01.004>
- Henrique, D., Schweisguth, F., 2019. Mechanisms of Notch signaling: a simple logic deployed in time and space. *Development* 146, dev172148. <https://doi.org/10.1242/dev.172148>
- Hill, R.J., Sternberg, P.W., 1992. The gene *lin-3* encodes an inductive signal for vulval development in *C. elegans*. *Nature* 358, 470–476. <https://doi.org/10.1038/358470a0>
- Hino, N., Rossetti, L., Marín-Llauradó, A., Aoki, K., Trepát, X., Matsuda, M., Hirashima, T., 2020. ERK-Mediated Mechanochemical Waves Direct Collective Cell Polarization. *Dev. Cell* 53, 646–660.e8. <https://doi.org/10.1016/j.devcel.2020.05.011>
- Hiratsuka, T., Fujita, Y., Naoki, H., Aoki, K., Kamioka, Y., Matsuda, M., 2015. Intercellular propagation of extracellular signal-regulated kinase activation revealed by in vivo imaging of mouse skin. *eLife* 4, e05178. <https://doi.org/10.7554/eLife.05178>
- Holland, L.Z., Gibson-Brown, J.J., 2003. The *Ciona intestinalis* genome: When the constraints are off. *BioEssays* 25, 529–532. <https://doi.org/10.1002/bies.10302>
- Hotta, K., Mitsuhashi, K., Takahashi, H., Inaba, K., Oka, K., Gojobori, T., Ikeo, K., 2007. A web-based interactive developmental table for the ascidian *Ciona intestinalis*, including 3D real-image embryo reconstructions: I. From fertilized egg to hatching larva. *Dev. Dyn.* 236, 1790–1805. <https://doi.org/10.1002/dvdy.21188>
- Hudson, C., 2016. The central nervous system of ascidian larvae. *WIREs Dev. Biol.* 5, 538–561. <https://doi.org/10.1002/wdev.239>
- Hudson, C., Darras, S., Caillol, D., Yasuo, H., Lemaire, P., 2003. A conserved role for the MEK signalling pathway in neural tissue specification and posteriorisation in the invertebrate chordate, the ascidian *Ciona intestinalis*. *Development* 130, 147–159. <https://doi.org/10.1242/dev.00200>
- Hudson, C., Lemaire, P., 2001. Induction of anterior neural fates in the ascidian *Ciona intestinalis*. *Mech. Dev.* 100, 189–203. [https://doi.org/10.1016/S0925-4773\(00\)00528-1](https://doi.org/10.1016/S0925-4773(00)00528-1)
- Hudson, C., Yasuo, H., 2005. Patterning across the ascidian neural plate by lateral Nodal signalling sources. *Development* 132, 1199–1210. <https://doi.org/10.1242/dev.01688>
- Huff, K., End, D., Guroff, G., 1981. Nerve growth factor-induced alteration in the response of PC12 pheochromocytoma cells to epidermal growth factor. *J. Cell Biol.* 88, 189–198.
- Imai, K.S., Hudson, C., Oda-Ishii, I., Yasuo, H., Satou, Y., 2016. Antagonism between β -catenin and Gata.a sequentially segregates the germ layers of ascidian embryos. *Dev. Camb. Engl.* 143, 4167–4172. <https://doi.org/10.1242/dev.141481>
- Imai, K.S., Levine, M., Satoh, N., Satou, Y., 2006. Regulatory blueprint for a chordate embryo. *Science* 312, 1183–1187. <https://doi.org/10.1126/science.1123404>
- Ingham, P.W., McMahon, A.P., 2001. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev.* 15, 3059–3087. <https://doi.org/10.1101/gad.938601>
- Jacques, M., Dobrzyński, M., Gagliardi, P.A., Sznitman, R., Pertz, O., 2021. CODEX, a neural network approach to explore signaling dynamics landscapes. *Mol. Syst. Biol.* 17. <https://doi.org/10.15252/msb.202010026>

- Jafar-Nejad, H., Tien, A.-C., Acar, M., Bellen, H.J., 2006. Senseless and Daughterless confer neuronal identity to epithelial cells in the *Drosophila* wing margin. *Dev. Camb. Engl.* 133, 1683–1692. <https://doi.org/10.1242/dev.02338>
- Johnson, H.E., Djabrayan, N.J.V., Shvartsman, S.Y., Toettcher, J.E., 2020. Optogenetic Rescue of a Patterning Mutant. *Curr. Biol.* 30, 3414–3424.e3. <https://doi.org/10.1016/j.cub.2020.06.059>
- Johnson, H.E., Goyal, Y., Pannucci, N.L., Schüpbach, T., Shvartsman, S.Y., Toettcher, J.E., 2017. The Spatiotemporal Limits of Developmental Erk Signaling. *Dev. Cell* 40, 185–192. <https://doi.org/10.1016/j.devcel.2016.12.002>
- Johnson, H.E., Toettcher, J.E., 2019. Signaling Dynamics Control Cell Fate in the Early *Drosophila* Embryo. *Dev. Cell* 48, 361–370.e3. <https://doi.org/10.1016/j.devcel.2019.01.009>
- Kao, S., Jaiswal, R.K., Kolch, W., Landreth, G.E., 2001. Identification of the Mechanisms Regulating the Differential Activation of the MAPK Cascade by Epidermal Growth Factor and Nerve Growth Factor in PC12 Cells. *J. Biol. Chem.* 276, 18169–18177. <https://doi.org/10.1074/jbc.M008870200>
- Katz, W.S., Hill, R.J., Clandinin, T.R., Sternberg, P.W., 1995. Different Levels of the *C. elegans* Growth Factor LIN-3 Promote Distinct Vulval Precursor Fates. *Cell* 11.
- Keduka, E., Kaiho, A., Hamada, M., Watanabe-Takano, H., Takano, K., Ogasawara, M., Satou, Y., Satoh, N., Endo, T., 2009. M-Ras evolved independently of R-Ras and its neural function is conserved between mammalian and ascidian, which lacks classical Ras. *Gene* 429, 49–58. <https://doi.org/10.1016/j.gene.2008.10.001>
- Khairy, K., Keller, P.J., 2011. Reconstructing embryonic development. *genesis* 49, 488–513. <https://doi.org/10.1002/dvg.20698>
- Kicheva, A., Bollenbach, T., Wartlick, O., Jülicher, F., Gonzalez-Gaitan, M., 2012. Investigating the principles of morphogen gradient formation: from tissues to cells. *Curr. Opin. Genet. Dev.* 22, 527–532. <https://doi.org/10.1016/j.gde.2012.08.004>
- Knowles, D.W., Biggin, M.D., 2013. Building quantitative, three-dimensional atlases of gene expression and morphology at cellular resolution. *Wiley Interdiscip. Rev. Dev. Biol.* 2, 767–779. <https://doi.org/10.1002/wdev.107>
- Kolch, W., 2000. Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. *Biochem. J.* 351, 289–305.
- Komatsu, N., Aoki, K., Yamada, M., Yukinaga, H., Fujita, Y., Kamioka, Y., Matsuda, M., 2011. Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol. Biol. Cell* 22, 4647–4656. <https://doi.org/10.1091/mbc.e11-01-0072>
- Kopan, R., 2012. Notch Signaling. *Cold Spring Harb. Perspect. Biol.* 4, a011213. <https://doi.org/10.1101/cshperspect.a011213>
- Kumagai, Y., Naoki, H., Nakasyo, E., Kamioka, Y., Kiyokawa, E., Matsuda, M., 2015. Heterogeneity in ERK activity as visualized by in vivo FRET imaging of mammary tumor cells developed in MMTV-Neu mice. *Oncogene* 34, 1051–1057. <https://doi.org/10.1038/onc.2014.28>
- Kuratani, S., Wada, H., Kusakabe, R., Agata, K., 2006. Evolutionary embryology resurrected in Japan with a new molecular basis: Nori Satoh and the history of ascidian studies originating in Kyoto during the 20th century. *Int. J. Dev. Biol.* 50. <https://doi.org/10.1387/ijdb.062154sk>
- Lavoie, H., Gagnon, J., Therrien, M., 2020. ERK signalling: a master regulator of cell behaviour, life and fate. *Nat. Rev. Mol. Cell Biol.* 21, 607–632. <https://doi.org/10.1038/s41580-020-0255-7>
- Lavoie, H., Therrien, M., 2015. Regulation of RAF protein kinases in ERK signalling. *Nat. Rev. Mol. Cell Biol.* 16, 281–298. <https://doi.org/10.1038/nrm3979>
- Leavesley, S.J., Rich, T.C., 2016. Overcoming Limitations of FRET Measurements. *Cytom. Part J. Int. Soc. Anal. Cytol.* 89, 325–327. <https://doi.org/10.1002/cyto.a.22851>
- Leggio, B., Laussu, J., Carlier, A., Godin, C., Lemaire, P., Faure, E., 2019. MorphoNet: an interactive online morphological browser to explore complex multi-scale data. *Nat. Commun.* 10, 2812. <https://doi.org/10.1038/s41467-019-10668-1>
- Lemaire, P., 2011. Evolutionary crossroads in developmental biology: the tunicates. *Development*

- 138, 2143–2152. <https://doi.org/10.1242/dev.048975>
- Lemaire, P., 2009. Unfolding a chordate developmental program, one cell at a time: Invariant cell lineages, short-range inductions and evolutionary plasticity in ascidians. *Dev. Biol.* 332, 48–60. <https://doi.org/10.1016/j.ydbio.2009.05.540>
- Lemaire, P., Kodjabachian, L., 1996. The vertebrate organizer: structure and molecules. *Trends Genet. TIG* 12, 525–531. [https://doi.org/10.1016/s0168-9525\(97\)81401-1](https://doi.org/10.1016/s0168-9525(97)81401-1)
- Li, P., Elowitz, M.B., 2019. Communication codes in developmental signaling pathways. *Development* 146, dev170977. <https://doi.org/10.1242/dev.170977>
- Li, P., Markson, J.S., Wang, S., Chen, S., Vachharajani, V., Elowitz, M.B., 2018. Morphogen gradient reconstitution reveals Hedgehog pathway design principles. *Science* 360, 543–548. <https://doi.org/10.1126/science.aao0645>
- Liao, B.-K., Oates, A.C., 2017. Delta-Notch signalling in segmentation. *Arthropod Struct. Dev.* 46, 429–447. <https://doi.org/10.1016/j.asd.2016.11.007>
- Malfant, M., Darras, S., Viard, F., 2018. Coupling molecular data and experimental crosses sheds light about species delineation: a case study with the genus *Ciona*. *Sci. Rep.* 8, 1480. <https://doi.org/10.1038/s41598-018-19811-2>
- Marshall, C.J., 1995. Signaling: Transient versus Sustained Extracellular Signal-Regulated Kinase Activation. *Cell* 80, 179–185.
- Maryu, G., Matsuda, M., Aoki, K., 2016. Multiplexed Fluorescence Imaging of ERK and Akt Activities and Cell-cycle Progression. *Cell Struct. Funct.* 41, 81–92. <https://doi.org/10.1247/csf.16007>
- Matise, M.P., Joyner, A.L., 1999. Gli genes in development and cancer. *Oncogene* 18, 7852–7859. <https://doi.org/10.1038/sj.onc.1203243>
- Min, H., Danilenko, D.M., Scully, S.A., Bolon, B., Ring, B.D., Tarpley, J.E., DeRose, M., Simonet, W.S., 1998. Fgf-10 is required for both limb and lung development and exhibits striking functional similarity to *Drosophila* branchless. *Genes Dev.* 12, 3156–3161. <https://doi.org/10.1101/gad.12.20.3156>
- Miyawaki, A., 2003. Visualization of the Spatial and Temporal Dynamics of Intracellular Signaling. *Dev. Cell* 4, 295–305. [https://doi.org/10.1016/S1534-5807\(03\)00060-1](https://doi.org/10.1016/S1534-5807(03)00060-1)
- Moreau, M., Leclerc, C., Néant, I., 2020. The saga of neural induction: almost a century of research. *Med. Sci. MS* 36, 1018–1026. <https://doi.org/10.1051/medsci/2020172>
- Morgan, T.H., 1901. *Regeneration*, Macmillan & Co., London.
- Morrison, D.K., 2012. MAP Kinase Pathways. *Cold Spring Harb. Perspect. Biol.* 4, a011254. <https://doi.org/10.1101/cshperspect.a011254>
- Murphy, L.O., MacKeigan, J.P., Blenis, J., 2004. A Network of Immediate Early Gene Products Propagates Subtle Differences in Mitogen-Activated Protein Kinase Signal Amplitude and Duration. *Mol. Cell. Biol.* 24, 144–153. <https://doi.org/10.1128/MCB.24.1.144-153.2004>
- Murphy, L.O., Smith, S., Chen, R.-H., Fingar, D.C., Blenis, J., 2002. Molecular interpretation of ERK signal duration by immediate early gene products. *Nat. Cell Biol.* 4, 556–564. <https://doi.org/10.1038/ncb822>
- Nakakuki, T., Birtwistle, M.R., Saeki, Y., Yumoto, N., Ide, K., Nagashima, T., Brusch, L., Ogunnaike, B.A., Okada-Hatakeyama, M., Kholodenko, B.N., 2010. Ligand-Specific c-Fos Expression Emerges from the Spatiotemporal Control of ErbB Network Dynamics. *Cell* 141, 884–896. <https://doi.org/10.1016/j.cell.2010.03.054>
- Nakamura, A., Goto, Y., Kondo, Y., Aoki, K., 2021. Shedding light on developmental ERK signaling with genetically encoded biosensors. *Development* 148, dev199767. <https://doi.org/10.1242/dev.199767>
- Nakamura, Y., Hibino, K., Yanagida, T., Sako, Y., 2016. Switching of the positive feedback for RAS activation by a concerted function of SOS membrane association domains. *Biophys. Physicobiology* 13, 1–11. https://doi.org/10.2142/biophysico.13.0_1
- Nandagopal, N., Santat, L.A., LeBon, L., Sprinzak, D., Bronner, M.E., Elowitz, M.B., 2018. Dynamic Ligand Discrimination in the Notch Signaling Pathway. *Cell* 172, 869–880.e19.

<https://doi.org/10.1016/j.cell.2018.01.002>

- Neumann, C.J., Cohen, S.M., 1997. Long-range action of Wingless organizes the dorsal-ventral axis of the *Drosophila* wing. *Dev. Camb. Engl.* 124, 871–880. <https://doi.org/10.1242/dev.124.4.871>
- Newbern, J., Zhong, J., Wickramasinghe, R.S., Li, X., Wu, Y., Samuels, I., Cherosky, N., Karlo, J.C., O’Loughlin, B., Wikenheiser, J., Garghesha, M., Doughman, Y.Q., Charron, J., Ginty, D.D., Watanabe, M., Saitta, S.C., Snider, W.D., Landreth, G.E., 2008. Mouse and human phenotypes indicate a critical conserved role for ERK2 signaling in neural crest development. *Proc. Natl. Acad. Sci.* 105, 17115–17120. <https://doi.org/10.1073/pnas.0805239105>
- Nicol, D., Meinertzhagen, I.A., 1988. Development of the Central Nervous System of the Larva of the Ascidian, *Ciona intestinalis* L. 30.
- Nishida, H., 2005. Specification of embryonic axis and mosaic development in ascidians. *Dev. Dyn.* 233, 1177–1193. <https://doi.org/10.1002/dvdy.20469>
- Nishida, H., 1987. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme: III. Up to the tissue restricted stage. *Dev. Biol.* 121, 526–541. [https://doi.org/10.1016/0012-1606\(87\)90188-6](https://doi.org/10.1016/0012-1606(87)90188-6)
- Nishida, H., Satoh, N., 1985. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme: II. The 16- and 32-cell stages. *Dev. Biol.* 110, 440–454. [https://doi.org/10.1016/0012-1606\(85\)90102-2](https://doi.org/10.1016/0012-1606(85)90102-2)
- Nishida, H., Satoh, N., 1983. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme: I. Up to the eight-cell stage. *Dev. Biol.* 99, 382–394. [https://doi.org/10.1016/0012-1606\(83\)90288-9](https://doi.org/10.1016/0012-1606(83)90288-9)
- Nishida, H., Sawada, K., 2001. *macho-1* encodes a localized mRNA in ascidian eggs that specifies muscle fate during embryogenesis. *Nature* 409, 724–729. <https://doi.org/10.1038/35055568>
- Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., Makita, R., Kurihara, H., Morin-Kensicki, E.M., Nojima, H., Rossant, J., Nakao, K., Niwa, H., Sasaki, H., 2009. The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass. *Dev. Cell* 16, 398–410. <https://doi.org/10.1016/j.devcel.2009.02.003>
- Nolo, R., Abbott, L.A., Bellen, H.J., 2000. Senseless, a Zn finger transcription factor, is necessary and sufficient for sensory organ development in *Drosophila*. *Cell* 102, 349–362. [https://doi.org/10.1016/s0092-8674\(00\)00040-4](https://doi.org/10.1016/s0092-8674(00)00040-4)
- Nusse, R., 2012. Wnt Signaling. *Cold Spring Harb. Perspect. Biol.* 4, a011163–a011163. <https://doi.org/10.1101/cshperspect.a011163>
- Ogura, Y., Wen, F.-L., Sami, M.M., Shibata, T., Hayashi, S., 2018. A Switch-like Activation Relay of EGFR-ERK Signaling Regulates a Wave of Cellular Contractility for Epithelial Invagination. *Dev. Cell* 46, 162-172.e5. <https://doi.org/10.1016/j.devcel.2018.06.004>
- Ohta, N., Satou, Y., 2013. Multiple signaling pathways coordinate to induce a threshold response in a chordate embryo. *PLoS Genet.* 9, e1003818. <https://doi.org/10.1371/journal.pgen.1003818>
- Park, H.L., Bai, C., Platt, K.A., Matisse, M.P., Beeghly, A., Hui, C.C., Nakashima, M., Joyner, A.L., 2000. Mouse *Gli1* mutants are viable but have defects in SHH signaling in combination with a *Gli2* mutation. *Development* 127, 1593–1605. <https://doi.org/10.1242/dev.127.8.1593>
- Patel, A.L., Shvartsman, S.Y., 2018. Outstanding questions in developmental ERK signaling. *Dev. Camb. Engl.* 145, dev143818. <https://doi.org/10.1242/dev.143818>
- Perrimon, N., Pitsouli, C., Shilo, B.-Z., 2012. Signaling Mechanisms Controlling Cell Fate and Embryonic Patterning. *Cold Spring Harb. Perspect. Biol.* 4, a005975–a005975. <https://doi.org/10.1101/cshperspect.a005975>
- Pokrass, M.J., Ryan, K.A., Xin, T., Pielstick, B., Timp, W., Greco, V., Regot, S., 2020. Cell-Cycle-Dependent ERK Signaling Dynamics Direct Fate Specification in the Mammalian Preimplantation Embryo. *Dev. Cell* 55, 328-340.e5. <https://doi.org/10.1016/j.devcel.2020.09.013>
- Power, R.M., Huisken, J., 2017. A guide to light-sheet fluorescence microscopy for multiscale imaging.

- Nat. Methods 14, 360–373. <https://doi.org/10.1038/nmeth.4224>
- Purvis, J.E., Lahav, G., 2013. Encoding and Decoding Cellular Information through Signaling Dynamics. *Cell* 152, 945–956. <https://doi.org/10.1016/j.cell.2013.02.005>
- Rahimi, N., Averbukh, I., Haskel-Ittah, M., Degani, N., Schejter, E.D., Barkai, N., Shilo, B.-Z., 2016. A WntD-Dependent Integral Feedback Loop Attenuates Variability in *Drosophila* Toll Signaling. *Dev. Cell* 36, 401–414. <https://doi.org/10.1016/j.devcel.2016.01.023>
- Rauen, K.A., 2013. The RASopathies. *Annu. Rev. Genomics Hum. Genet.* 14, 355–369. <https://doi.org/10.1146/annurev-genom-091212-153523>
- Razy-Krajka, F., Gravez, B., Kaplan, N., Racioppi, C., Wang, W., Christiaen, L., 2018. An FGF-driven feed-forward circuit patterns the cardiopharyngeal mesoderm in space and time. *eLife* 7, e29656. <https://doi.org/10.7554/eLife.29656>
- Razy-Krajka, F., Lam, K., Wang, W., Stolfi, A., Joly, M., Bonneau, R., Christiaen, L., 2014. Collier/OLF/EBF-Dependent Transcriptional Dynamics Control Pharyngeal Muscle Specification from Primed Cardiopharyngeal Progenitors. *Dev. Cell* 29, 263–276. <https://doi.org/10.1016/j.devcel.2014.04.001>
- Regot, S., Hughey, J.J., Bajar, B.T., Carrasco, S., Covert, M.W., 2014. High-Sensitivity Measurements of Multiple Kinase Activities in Live Single Cells. *Cell* 157, 1724–1734. <https://doi.org/10.1016/j.cell.2014.04.039>
- Ribes, V., Briscoe, J., 2009. Establishing and Interpreting Graded Sonic Hedgehog Signaling during Vertebrate Neural Tube Patterning: The Role of Negative Feedback. *Cold Spring Harb. Perspect. Biol.* 1, a002014. <https://doi.org/10.1101/cshperspect.a002014>
- Richardson, M.K., 2012. A Phylotypic Stage for All Animals? *Dev. Cell* 22, 903–904. <https://doi.org/10.1016/j.devcel.2012.05.001>
- Roelink, H., Porter, J.A., Chiang, C., Tanabe, Y., Chang, D.T., Beachy, P.A., Jessell, T.M., 1995. Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproteolysis. *Cell* 81, 445–455. [https://doi.org/10.1016/0092-8674\(95\)90397-6](https://doi.org/10.1016/0092-8674(95)90397-6)
- Ruiz i Altaba, A., 1998. Combinatorial Gli gene function in floor plate and neuronal inductions by Sonic hedgehog. *Development* 125, 2203–2212. <https://doi.org/10.1242/dev.125.12.2203>
- Sagner, A., Briscoe, J., 2017. Morphogen interpretation: concentration, time, competence, and signaling dynamics. *Wiley Interdiscip. Rev. Dev. Biol.* 6, e271. <https://doi.org/10.1002/wdev.271>
- Salazar, V.S., Gamer, L.W., Rosen, V., 2016. BMP signalling in skeletal development, disease and repair. *Nat. Rev. Endocrinol.* 12, 203–221. <https://doi.org/10.1038/nrendo.2016.12>
- Samakovlis, C., Hacohen, N., Manning, G., Sutherland, D.C., Guillemin, K., Krasnow, M.A., 1996. Development of the *Drosophila* tracheal system occurs by a series of morphologically distinct but genetically coupled branching events. *Development* 122, 13.
- Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W.K., Luna, A., La, K.C., Dimitriadoy, S., Liu, D.L., Kantheti, H.S., Saghafeina, S., Chakravarty, D., Daian, F., Gao, Q., Bailey, M.H., Liang, W.-W., Foltz, S.M., Shmulevich, I., Ding, L., Heins, Z., Ochoa, A., Gross, B., Gao, J., Zhang, H., Kundra, R., Kandath, C., Bahceci, I., Dervishi, L., Dogrusoz, U., Zhou, W., Shen, H., Laird, P.W., Way, G.P., Greene, C.S., Liang, H., Xiao, Y., Wang, C., Iavarone, A., Berger, A.H., Bivona, T.G., Lazar, A.J., Hammer, G.D., Giordano, T., Kwong, L.N., McArthur, G., Huang, C., Tward, A.D., Frederick, M.J., McCormick, F., Meyerson, M., Van Allen, E.M., Cherniack, A.D., Ciriello, G., Sander, C., Schultz, N., 2018. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* 173, 321–337. <https://doi.org/10.1016/j.cell.2018.03.035>
- Santos, S.D.M., Verveer, P.J., Bastiaens, P.I.H., 2007. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nat. Cell Biol.* 9, 324–330. <https://doi.org/10.1038/ncb1543>
- Sardet, C., McDougall, A., Yasuo, H., Chenevert, J., Pruliere, G., Dumollard, R., Hudson, C., Hebras, C., Le Nguyen, N., Paix, A., 2011. Embryological Methods in Ascidians: The Villefranche-sur-Mer

- Protocols, in: Pelegri, F.J. (Ed.), *Vertebrate Embryogenesis, Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp. 365–400. https://doi.org/10.1007/978-1-61779-210-6_14
- Sasagawa, S., Ozaki, Y., Fujita, K., Kuroda, S., 2005. Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nat. Cell Biol.* 7, 365–373. <https://doi.org/10.1038/ncb1233>
- Schweisguth, F., 2000. Fonctions et régulation de l'activité de signalisation du récepteur Notch. *médecine/sciences* 16, 186. <https://doi.org/10.4267/10608/1620>
- Seo, H.-C., Edvardsen, R.B., Maeland, A.D., Bjordal, M., Jensen, M.F., Hansen, A., Flaate, M., Weissenbach, J., Lehrach, H., Wincker, P., Reinhardt, R., Chourrout, D., 2004. Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. *Nature* 431, 67–71. <https://doi.org/10.1038/nature02709>
- Shilo, B.-Z., Barkai, N., 2017. Lighting Up ERK Activity. *Dev. Cell* 40, 115–116. <https://doi.org/10.1016/j.devcel.2016.12.016>
- Shin, S.-Y., Rath, O., Choo, S.-M., Fee, F., McFerran, B., Kolch, W., Cho, K.-H., 2009. Positive- and negative-feedback regulations coordinate the dynamic behavior of the Ras-Raf-MEK-ERK signal transduction pathway. *J. Cell Sci.* 122, 425–435. <https://doi.org/10.1242/jcs.036319>
- Sladitschek, H.L., Fiuza, U.-M., Pavlinic, D., Benes, V., Hufnagel, L., Neveu, P.A., 2020. MorphoSeq: Full Single-Cell Transcriptome Dynamics Up to Gastrulation in a Chordate. *Cell* 181, 922–935.e21. <https://doi.org/10.1016/j.cell.2020.03.055>
- Small, K.S., Brudno, M., Hill, M.M., Sidow, A., 2007. A haplome alignment and reference sequence of the highly polymorphic *Ciona savignyi* genome. *Genome Biol.* 8, R41. <https://doi.org/10.1186/gb-2007-8-3-r41>
- Spemann, H., Mangold, H., 1924. Über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Arch. Für Mikrosk. Anat. Entwicklungsmechanik* 599–638.
- Spitz, F., Furlong, E.E.M., 2012. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* 13, 613–626. <https://doi.org/10.1038/nrg3207>
- Sprinzak, D., Lakhanpal, A., LeBon, L., Garcia-Ojalvo, J., Elowitz, M.B., 2011. Mutual Inactivation of Notch Receptors and Ligands Facilitates Developmental Patterning. *PLoS Comput. Biol.* 7, e1002069. <https://doi.org/10.1371/journal.pcbi.1002069>
- Srinivasan, R., Zabuawala, T., Huang, H., Zhang, J., Gulati, P., Fernandez, S., Karlo, J.C., Landreth, G.E., Leone, G., Ostrowski, M.C., 2009. Erk1 and Erk2 Regulate Endothelial Cell Proliferation and Migration during Mouse Embryonic Angiogenesis. *PLOS ONE* 4, e8283. <https://doi.org/10.1371/journal.pone.0008283>
- Stach, T., Anselmi, C., 2015. High-precision morphology: bifocal 4D-microscopy enables the comparison of detailed cell lineages of two chordate species separated for more than 525 million years. *BMC Biol.* 13, 113. <https://doi.org/10.1186/s12915-015-0218-1>
- Stapornwongkul, K.S., de Gennes, M., Cocconi, L., Salbreux, G., Vincent, J.-P., 2020. Patterning and growth control in vivo by an engineered GFP gradient. *Science* 370, 321–327. <https://doi.org/10.1126/science.abb8205>
- Sternberg, P.W., Horvitz, H.R., 1986. Pattern formation during vulval development in *C. elegans*. *Cell* 44, 761–772. [https://doi.org/10.1016/0092-8674\(86\)90842-1](https://doi.org/10.1016/0092-8674(86)90842-1)
- Sternberg, P.W., Horvitz, H.R., 1984. THE GENETIC CONTROL OF CELL LINEAGE DURING NEMATODE DEVELOPMENT 36.
- Stolfi, A., Gainous, T.B., Young, J.J., Mori, A., Levine, M., Christiaen, L., 2010. Early Chordate Origins of the Vertebrate Second Heart Field. *Science* 329, 565–568. <https://doi.org/10.1126/science.1190181>
- Stolfi, A., Levine, M., 2011. Neuronal subtype specification in the spinal cord of a protovertebrate. *Dev. Camb. Engl.* 138, 995–1004. <https://doi.org/10.1242/dev.061507>
- Stolfi, A., Lowe, E.K., Racioppi, C., Ristatore, F., Brown, C.T., Swalla, B.J., Christiaen, L., 2014. Divergent mechanisms regulate conserved cardiopharyngeal development and gene expression in distantly related ascidians. *eLife* 3, e03728.

- <https://doi.org/10.7554/eLife.03728>
- Swarup, S., Verheyen, E.M., 2012. Wnt/Wingless signaling in *Drosophila*. *Cold Spring Harb. Perspect. Biol.* 4, a007930. <https://doi.org/10.1101/cshperspect.a007930>
- Tassy, O., Daian, F., Hudson, C., Bertrand, V., Lemaire, P., 2006. A Quantitative Approach to the Study of Cell Shapes and Interactions during Early Chordate Embryogenesis. *Curr. Biol.* 16, 345–358. <https://doi.org/10.1016/j.cub.2005.12.044>
- Toettcher, J.E., Weiner, O.D., Lim, W.A., 2013. Using Optogenetics to Interrogate the Dynamic Control of Signal Transmission by the Ras/Erk Module. *Cell* 155, 1422–1434. <https://doi.org/10.1016/j.cell.2013.11.004>
- Tokuoka, Y., Yamada, T.G., Mashiko, D., Ikeda, Z., Hiroi, N.F., Kobayashi, T.J., Yamagata, K., Funahashi, A., 2020. 3D convolutional neural networks-based segmentation to acquire quantitative criteria of the nucleus during mouse embryogenesis. *Npj Syst. Biol. Appl.* 6, 32. <https://doi.org/10.1038/s41540-020-00152-8>
- Tolkin, T., Christiaen, L., 2016. Rewiring of an ancestral Tbx1/10-Ebf-Mrf network for pharyngeal muscle specification in distinct embryonic lineages. *Development* 143, 3852–3862. <https://doi.org/10.1242/dev.136267>
- Tomer, R., Denes, A.S., Tessmar-Raible, K., Arendt, D., 2010. Profiling by Image Registration Reveals Common Origin of Annelid Mushroom Bodies and Vertebrate Pallium. *Cell* 142, 800–809. <https://doi.org/10.1016/j.cell.2010.07.043>
- Traverse, S., Gomez, N., Paterson, H., Marshall, C., Cohen, P., 1992. Sustained activation of the mitogen-activated protein (MAP) kinase cascade may be required for differentiation of PC12 cells. Comparison of the effects of nerve growth factor and epidermal growth factor. *Biochem. J.* 288, 351–355.
- Tsagkogeorga, G., Turon, X., Galtier, N., Douzery, E.J.P., Delsuc, F., 2010. Accelerated evolutionary rate of housekeeping genes in tunicates. *J. Mol. Evol.* 71, 153–67. <https://doi.org/10.1007/s00239-010-9372-9>
- Turing, A.M., 1952. The chemical basis of morphogenesis. *R. Soc.* 237.
- Uhlitz, F., Sieber, A., Wyler, E., Fritsche-Guenther, R., Meisig, J., Landthaler, M., Klinger, B., Blüthgen, N., 2017. An immediate–late gene expression module decodes ERK signal duration. *Mol. Syst. Biol.* 13, 928. <https://doi.org/10.15252/msb.20177554>
- Vandame, P., Spriet, C., Riquet, F., Trinel, D., Cailliau-Maggio, K., Bodart, J.-F., 2014. Optimization of ERK Activity Biosensors for both Ratiometric and Lifetime FRET Measurements. *Sensors* 14, 1140–1154. <https://doi.org/10.3390/s140101140>
- Veeman, M., Reeves, W., 2015. Quantitative and in toto imaging in ascidians: Working toward an image-centric systems biology of chordate morphogenesis. *genesis* 53, 143–159. <https://doi.org/10.1002/dvg.22828>
- Vergassola, M., Deneke, V.E., Di Talia, S., 2018. Mitotic waves in the early embryogenesis of *Drosophila*: Bistability traded for speed. *Proc. Natl. Acad. Sci.* 115. <https://doi.org/10.1073/pnas.1714873115>
- Vieites-Prado, A., Renier, N., 2021. Tissue clearing and 3D imaging in developmental biology. *Development* 148, dev199369. <https://doi.org/10.1242/dev.199369>
- Waddington, C.H., 1940. Organizers and Genes. *J. Anat.* 75, 141–142.
- Wartlick, O., Kicheva, A., González-Gaitán, M., 2009. Morphogen gradient formation. *Cold Spring Harb. Perspect. Biol.* 1, a001255. <https://doi.org/10.1101/cshperspect.a001255>
- Watson, U., Jain, R., Asthana, S., Saini, D.K., 2018. Spatiotemporal Modulation of ERK Activation by GPCRs. *Int. Rev. Cell Mol. Biol.* 338, 111–140. <https://doi.org/10.1016/bs.ircmb.2018.02.004>
- Weigert, M., Schmidt, U., Boothe, T., Müller, A., Dibrov, A., Jain, A., Wilhelm, B., Schmidt, D., Broaddus, C., Culley, S., Rocha-Martins, M., Segovia-Miranda, F., Norden, C., Henriques, R., Zerial, M., Solimena, M., Rink, J., Tomancak, P., Royer, L., Jug, F., Myers, E.W., 2018. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* 15, 1090–1097. <https://doi.org/10.1038/s41592-018-0216-7>

- White, J.G., Amos, W.B., Fordham, M., 1987. An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy. *J. Cell Biol.* 105, 41–48. <https://doi.org/10.1083/jcb.105.1.41>
- Whitmarsh, A.J., 2007. Regulation of gene transcription by mitogen-activated protein kinase signaling pathways. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1773, 1285–1298. <https://doi.org/10.1016/j.bbamcr.2006.11.011>
- Williaume, G., de Buyl, S., Sirour, C., Haupaix, N., Bettoni, R., Imai, K.S., Satou, Y., Dupont, G., Hudson, C., Yasuo, H., 2021. Cell geometry, signal dampening, and a bimodal transcriptional response underlie the spatial precision of an ERK-mediated embryonic induction. *Dev. Cell.* <https://doi.org/10.1016/j.devcel.2021.09.025>
- Wilson, M.Z., Ravindran, P.T., Lim, W.A., Toettcher, J.E., 2017. Tracing Information Flow from Erk to Target Gene Induction Reveals Mechanisms of Dynamic and Combinatorial Control. *Mol. Cell* 67, 757-769.e5. <https://doi.org/10.1016/j.molcel.2017.07.016>
- Wolpert, L., 1994. Do We Understand Development? *Science* 266, 571–572. <https://doi.org/10.1126/science.7939707>
- Wolpert, L., 1969. Positional information and the spatial pattern of cellular differentiation. *J. Theor. Biol.* 25, 1–47. [https://doi.org/10.1016/S0022-5193\(69\)80016-0](https://doi.org/10.1016/S0022-5193(69)80016-0)
- Xiong, F., Tentner, A.R., Hiscock, T.W., Huang, P., Megason, S.G., 2018. Heterogeneity of Sonic Hedgehog Response Dynamics and Fate Specification in Single Neural Progenitors (preprint). *Developmental Biology.* <https://doi.org/10.1101/412858>
- Yamanaka, Y., Lanner, F., Rossant, J., 2010. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* 137, 715–724. <https://doi.org/10.1242/dev.043471>
- Yan, D., Lin, X., 2009. Shaping Morphogen Gradients by Proteoglycans. *Cold Spring Harb. Perspect. Biol.* 1, a002493. <https://doi.org/10.1101/cshperspect.a002493>
- Zagorski, M., Tabata, Y., Brandenberg, N., Lutolf, M.P., Tkačik, G., Bollenbach, T., Briscoe, J., Kicheva, A., 2017. Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science* 356, 1379–1383. <https://doi.org/10.1126/science.aam5887>
- Zecca, M., Basler, K., Struhl, G., 1996. Direct and long-range action of a wingless morphogen gradient. *Cell* 87, 833–844. [https://doi.org/10.1016/s0092-8674\(00\)81991-1](https://doi.org/10.1016/s0092-8674(00)81991-1)
- Zhang, J., Campbell, R.E., Ting, A.Y., Tsien, R.Y., 2002. Creating new fluorescent probes for cell biology. *Nat. Rev. Mol. Cell Biol.* 3, 906–918. <https://doi.org/10.1038/nrm976>