3. Cellular and Multicellular Dynamics of Life

Background

For centuries, scientists have been fascinated by the remarkable precision with which organisms develop and the stability yet plasticity that they can show in response to environmental signals. To achieve an accurate and predictive understanding of a complex living system, we must understand how cells respond to intrinsic and extrinsic signals to become organised in time and space. There are immense opportunities to achieve a molecular understanding of how living systems function, respond to ever-changing environments, and evolve.

Many cellular and multicellular processes only make sense in the context of their natural environment inside a cell or organism, or at the interface of these settings with the outside world. To survive and reproduce, organisms – and their underlying molecular, physical, and chemical properties – are highly attuned to these environments, and can respond rapidly and adapt to environmental cues. At the same time, mechanisms have evolved to produce a stable set of internal physical and chemical conditions that can resist environmental fluctuations and ensure the survival and propagation of a species.

Living systems emerge from a web of complex dynamic interactions across scales, and rely on integration of molecular, chemical, and physical cues of different types and origins. These include biochemical interactions to make reactions occur more rapidly or build higher-order structures from the molecular building blocks of life (Chapter 2: Molecular Building Blocks in Context); metabolic or nutritional cues that change the internal and external environment; mechanical interactions to push or pull; and geometric properties such as size or shape. While the impact of the environment on phenotypic outcome, referred to as phenotypic plasticity, is well described at the organism and population levels, the underlying molecular mechanisms remain relatively uncharacterised. The key challenge now is to **reveal the molecular and cellular mechanisms that underlie developmental and phenotypic plasticity.** For a given genotype, the questions are: how much variation is possible, what is due to intrinsic variation or response to extrinsic signals, and how is this achieved?

A major goal is therefore to **unravel the genetic and environmental sources of variability in living systems, to understand responsiveness at the cellular level and in a multicellular context**. This will require novel experimental strategies that allow us to measure, model, and perturb these complex multiscale networks of interactions to reach a fundamental understanding of living systems: how they form, how they respond, and how they evolve.

EMBL is in a unique position to address these challenges, with its foundations of truly interdisciplinary and fundamental research, its strength in technology development and innovation, and its unique collaborative culture. The study of the cellular and multicellular dynamics of life in context will require development of theoretical concepts for living systems, alongside experimental strategies with the right level of granularity to answer specific questions. Developing these experimental and theoretical underpinnings will enable scientists to address key conceptual issues in modern biology, and will have a fundamental impact on the life sciences in general.

The Opportunity

Living systems are extremely complex, not only due to the large number of components they contain, but especially due to their highly interconnected, dynamic feedback regulation, which integrates different modalities (e.g. chemical, physical) and bridges many scales in space and time. Of particular importance is the way in which the different levels of organisation in living systems are interlinked, influencing and feeding back

on each other in a reciprocal manner. In other words, organisms both **shape their environment** and, at the same time, **respond to it**. For example, the molecular structure of cells influences the mechanical properties of the tissues they form. In turn, tissue mechanics impact on cellular signalling and cellular structure. This creates a self-organising and self-referential system of extreme complexity, which raises the question: **what are the logical principles that underlie the emergence of life?** Key considerations include:

- **Robustness**. A fundamental property of living systems is their ability to maintain their function and structure despite internal and external fluctuations in a large number of components and environmental parameters. For instance, how do cells scale their structures to different sizes to maintain order and function in time and space? How do embryos develop normally despite changing metabolic and nutritional cues? How do environmental cues influence the ability of tissues to regenerate? Therefore, a key question is: what mechanisms underlie robustness at the cellular, tissue, and organismal scale? Conceptually, researchers need to understand what determines the limits of robustness, beyond which living systems deteriorate into disease states.
- **Plasticity**. Living systems also have the ability to integrate their internal and external conditions to dynamically adjust their appearance and function; in other words, to generate phenotypic plasticity. Multiple questions remain unanswered in this context. For example, how do dynamically fluctuating molecular networks allow cells to permanently explore different functions? How do environmental changes lead to a functional adaptation and how are these adaptations locked into cellular programmes and genetic or epigenetic memory? What is the contribution of heritable genetic and non-genetic components to this phenotypic plasticity? How are multiple cues integrated across different modalities and scales? How does tissue geometry impact biochemical intracellular signalling?

The dynamic responsiveness of living systems to their environment, linked to **robustness** and **plasticity**, can be analysed across different temporal and spatial scales of life – from the organismal to the cellular and subcellular levels. For the first time, an array of technologies exist to obtain high-quality, quantitative, and dynamic molecular data across scales. Nevertheless, novel experimental strategies and technology development will be required to obtain direct readouts and means of perturbing in a spatio-temporally controlled manner. These perturbations must be made not only at the molecular level, but also by manipulating environmental cues including physical and chemical parameters. Such a combined approach, which integrates new experimental model systems, quantitative measurement, precisely controlled perturbations, and theoretical modelling, is now possible across scales. To test theoretical predictions effectively, this new Programme also includes Data Sciences (Chapter 8: Data Sciences) and Theory (Chapter 9: Theory at EMBL) to explore the vast amounts of quantitative and dynamic data that will be generated. EMBL will be able to provide conceptually new levels of understanding of the logical principles that underlie cellular and multicellular life, such as robustness and plasticity, in the context of the environment.

Research Aims

In the next scientific programme, EMBL aims to develop:

I. **New experimental strategies to reveal the mechanisms underlying the responsiveness of living systems to their environment**. EMBL's goal is to develop strategies to address questions about robustness and plasticity in response to environmental cues in a wide range of contexts,

including embryonic development, regeneration, and disease states. Of particular importance is the use of suitable, genetically tractable experimental model systems, both *in vivo* and *in vitro*, which will enable a systematic and functional interrogation of the interplay between environmental cues and their effects on living systems, at the mechanistic level. Current research and pilot projects (see below), which span the cellular to the organismal scale, will be the bases to address these questions. EMBL will apply a highly interdisciplinary approach, combining new technology development for quantification (ii), perturbation (iii), and theory and modelling (iv).

II. **New technologies to generate, integrate, and share quantitative dynamic data on relevant molecular, physical, and chemical parameters**. Living systems depend not only on molecular cues, but also integrate mechanical, geometrical, metabolic, nutritional, and other cues across scales in time and space. One key challenge that must therefore be tackled is the ability to generate **quantitative** and **dynamic** measurements of these **multimodal** parameters in living systems. EMBL will continue to pioneer technologies to allow integrated data capture, spanning multiple types of measurement in space and time. Simultaneously, it is critical to further develop methods that allow the integration, visualisation, and sharing of these multimodal and multidimensional datasets. This will enable them to be used, in combination with functional experimentation, to answer fundamental biological questions.

III. **New technologies for precise perturbations of biological systems and environments**. Probing, disrupting, and perturbing biological systems (e.g. through classical genetics, or chemical and physical means) have always been at the heart of experimental biology. A vast array of sophisticated and powerful perturbation technologies now exist with increasingly high-throughput approaches enabling multiple components of a modality to be systematically addressed. These range from the panoply of CRISPR/Cas9 or small molecule screens, to highly specific optogenetic and degron approaches that can modulate gene expression and protein function in space and time, especially in *in vitro* models. However, perturbing the physical or chemical parameters inside and outside biological systems in a precisely controlled manner is technically more challenging. In addition, for physiological, cellular, and multicellular systems like primary cells, organoids, or the developing embryo, perturbations are often only low-throughput. A key objective is therefore to develop the next generation of automated technologies to perform systematic molecular, physical, and chemical perturbations of physiological biological systems. Moreover, as key regulatory processes usually occur at precise moments in time and/or positions in space, a new generation of perturbation techniques that can be spatio-temporally controlled in living organisms will be developed at EMBL.

IV. **Extracting correlations and carrying out predictive computer modelling**. Mathematical models of biological processes, along with data-driven computer simulations, make it possible to create testable predictions about mechanisms. The study of biology has always been driven by a continuous interplay between hypotheses and experimental observations, but is classically based on intuitive logical reasoning. Today, the level of complexity often requires non-intuitive approaches, largely due to non-linearities, complex feedback mechanisms, and the sheer number of interacting components. The formalisation of hypotheses into mathematical models is thus an essential way forward to develop a deeper understanding of biological mechanisms. Furthermore, creating data-driven computer simulations makes it possible to test specific predictions about the underlying mechanisms, explore large sets of possible solutions, and determine the key parameters that distinguish between these solutions. When based on quantitative measurements of the real biological system, sets of alternative predictions can

thus drive the next round of experiments, which can verify or refute the predictions of the model. This important activity will rely on an increased number of theoretical biologists, and close links with the new Theory at EMBL theme (Chapter 9: Theory at EMBL).

The combined experimental and theoretical strategy outlined above forms a continuous scientific cycle where hypotheses, mathematical models, and computer simulations continuously improve and integrate knowledge, driven by new quantitative measurements and targeted perturbations of key parameters. Conversely, the experimental design is continuously refined to test more specific and precise questions driven by model predictions. In this way, EMBL will move towards a fully quantitative, dynamic, and predictive understanding of the cellular and multicellular dynamics of life.

EMBL's Approach

New Experimental Strategies to Study the Mechanisms Underlying Responsiveness to Environmental Cues

EMBL's approach is to develop a cutting-edge experimental and theoretical strategy that combines measurements and controlled environmental parameters, while performing rigorous multimodal quantification of dynamic cellular responses across scales. The development of new model systems and interdisciplinary experimental strategies will have a central role. EMBL has ongoing research in this area and has initiated several pilot projects that aim to tackle this challenge in different contexts, such as embryonic development, regeneration, and disease. One ongoing pilot project involves a marine model system, the sea anemone *Nematostella vectensis*, which is being developed at EMBL as an excellent model to study the effects of environmental and nutritional cues, for example in the process of tentacle regeneration. Importantly, genome editing possibilities in *Nematostella* have rendered it a powerful genetic model system. Combined with the ability to generate quantitative data across very different modalities, such as spatially resolved transcriptomics and **spatial metabolomics** (Figure MD1), cellular dynamics, and even the quantification of mechanical properties (Tech Dev Box TD2_MD, and Tech Dev Box TD3_MO), this new model system provides an outstanding opportunity to address the mechanisms underlying the integration of environmental cues, such as nutrition, and the way these mechanisms link to cellular programs of regeneration.

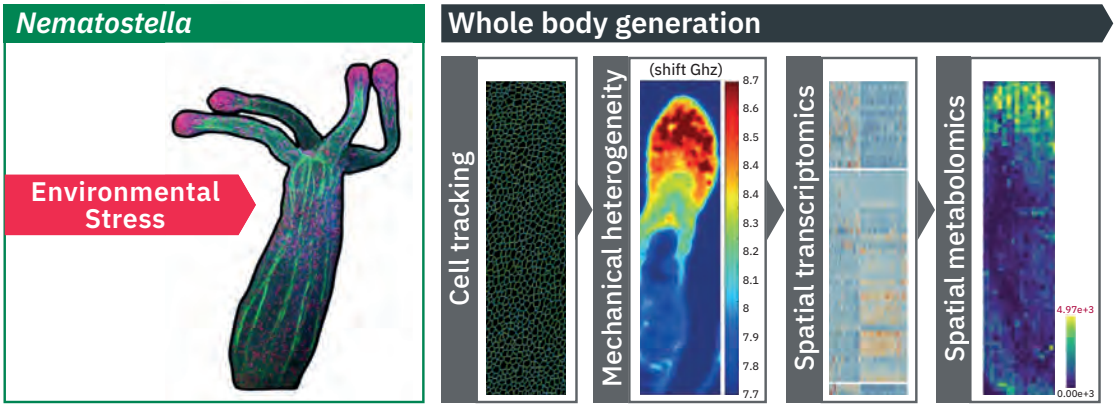


Figure MD1 | The sea anemone *Nematostella vectensis* is a novel genetic model for studying regeneration in the context of organism–environment interactions across multiple scales.
Spatially resolved omics data and imaging-based approaches to probe tissue properties *in vivo* are used to define the multicellular dynamics of regeneration in wild-type and mutant animals when environmentally stressed by injury, heat shock, or chemicals.

At the organismal and population scales, EMBL researchers will study communities of the fruit fly ***Drosophila melanogaster*** in controlled laboratory ecosystems to address the effects of changes in environmental conditions (Chapter 7: Planetary Biology). Following defined perturbations at the environmental and metabolic levels, it will be possible to rigorously quantify the effect of these perturbations on gene regulatory networks, signalling, metabolism, and even behavioural responses, both at the level of embryonic development in individual organisms, and at the population level. Again, it is the combination of a genetically tractable model system with the ability to experimentally control environmental cues and to obtain multimodal quantifications of the dynamic responses *in vivo*, that will provide novel avenues and insight into these fundamental questions.

At the cellular level, the influence of the cellular environment is being investigated with a particular focus on the reciprocal interactions between cellular programs and environmental cues such as physical forces (e.g. plasma membrane tension). Altered physical forces are increasingly recognised as playing an important role in allowing cells to sense their external context, leading to dynamic cellular responses including effects on cell fate and cell motility. At the same time, it is necessary to investigate how cells cope with changes in internal or external conditions and yet show **robust** cellular function. By combining newly developed optogenetic tools (Tech Dev Box TD1_MD) and Brillouin microscopy (Tech Dev Box TD2_MD), EMBL scientists aim to address **how physical cues translate into cellular functionality**. Such biophysical analyses will be complemented by integrated informatics strategies that aim to disentangle genetic and environmental factors impacting cellular metrics. For example, a major EMBL study has exploited extensive proteomics datasets to identify the significant impact of altered diet on individual proteotype, particularly on the nuclear pore complex stoichiometry, which in turn has broad cellular influence.

As well as being impacted by physical and chemical environmental variables, humans can be affected by human behaviour and **social interactions** (Chapter 6: Human Ecosystems). EMBL researchers are developing novel experimental strategies that enable quantification of the effect of social defeat experiences in mice at multiple scales, from neural activity (using calcium imaging) to the detailed extraction of synaptic ultrastructure using high-resolution synchrotron X-ray holographic nanotomography. The pilot project will establish the feasibility of several critical aspects of linking large-scale single unit neural activity recording data to neural ultrastructure and gene expression. It is hoped that, once established, the pipeline could be offered as an EMBL service.

Along these lines, and to investigate how environmental cues impact **disease states**, researchers at EMBL also revert to highly controllable *in vitro* models and drive the development of novel 3D vascularised *in vitro* models of the blood–brain barrier, to study, for instance, malaria pathogenesis and the interaction with brain vasculature (Chapter 5: Infection Biology, and Chapter 6: Human Ecosystems). The *in vitro* strategy complements the *in vivo* organismal studies mentioned above, and could reveal fundamental mechanisms by which environmental and genetic information are integrated and underlie the plasticity and robustness of phenotypic outcome.

New Technologies to Generate Quantitative Data

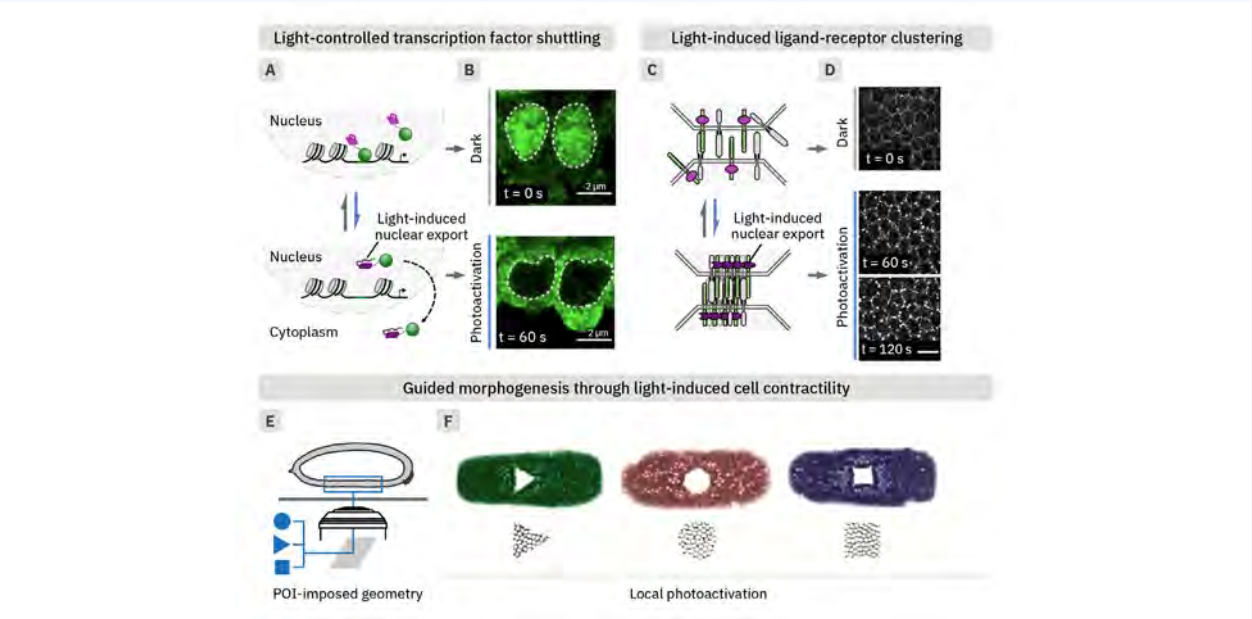
Not all the techniques and tools needed to measure the multiscale nature of life in context, with the required molecular, physical, and chemical modalities, are currently available. For example, techniques to measure changes in physical properties during cellular and developmental morphogenesis and how they are dynamically interacting with, and driven by, the underlying molecular networks, or tools to ask how morphogenesis is affected by changes in the chemical environment are highly needed. To address these challenges, EMBL will develop new technologies that allow dynamic and quantitative measurements of molecular, chemical, and physical properties of cells and multicellular systems. A key goal will be to **develop non-invasive methods**

that perturb the living biological system as little as possible, and seamlessly integrate them with invasive or destructive techniques where no alternative exists to measure the required parameters.

A key set of parameters that currently cannot be measured sufficiently are the physical properties of biological systems. For example, measuring the tension, viscoelasticity, and force anisotropy of the surface and interior of cells and tissues is essential to construct maps of biophysical properties across space and time and connect them with their molecular networks. EMBL is already very strong in biophysical technologies to determine a range of parameters, and has leading expertise in fluorescence correlation spectroscopy (viscosity), laser ablation (tissue tension), atomic force microscopy (AFM), micropipette aspiration (low-frequency surface mechanics, membrane tension), and magnetic droplets (force anisotropy). In addition, EMBL groups have recently developed Brillouin microscopy for non-invasively probing high-frequency mechanics inside cells and organisms (Tech Dev Box TD2_MD). Additional promising methods for future developments include probing single-cell mechanics by acoustic scattering, magnetic twisting cytometry, or optical tweezers.

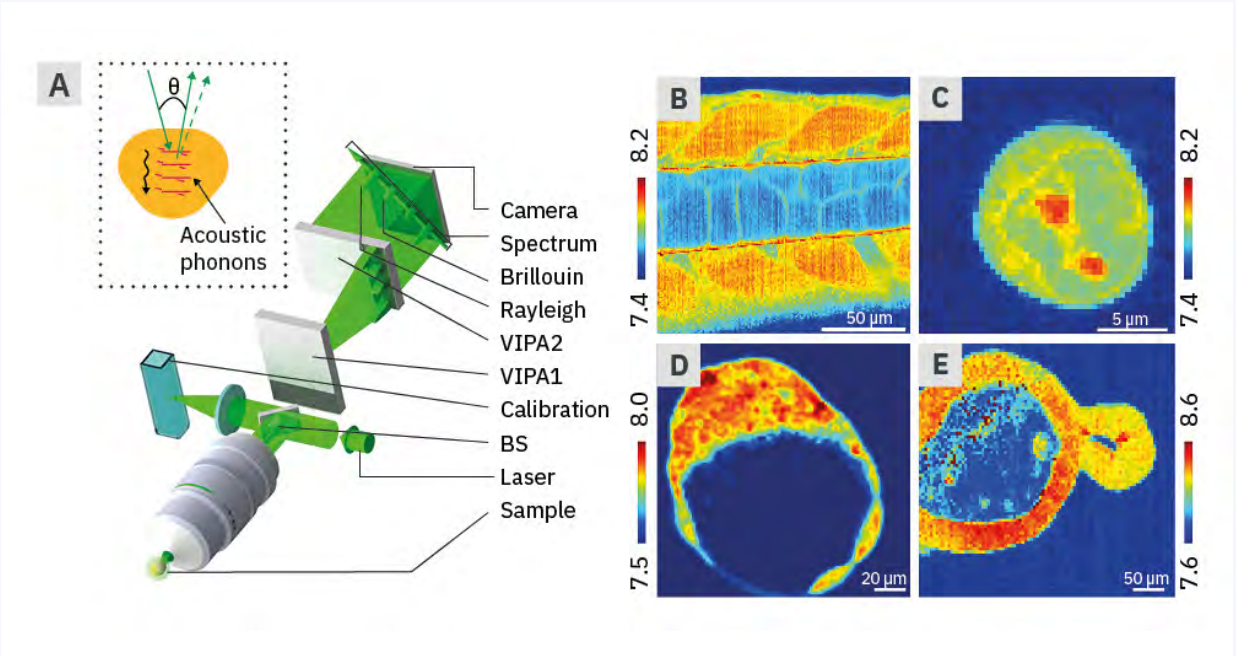
Technology Development Box TD1_MD | Optogenetics.

Optogenetics is a powerful technique that was initially developed to control neuronal activity in living animals. EMBL scientists have pioneered this technology for use in living embryos, with the aim of dynamically controlling protein activity with high spatio-temporal precision as development proceeds. By tagging genes of interest at their endogenous locus with photosensitive protein domains, the activity of endogenous proteins can be controlled at will during animal development. In this figure, a few examples of cellular processes that have been put under optogenetic control are illustrated. **(A–B)** Light-mediated control of the nuclear localisation of the key mesodermal transcription factor Twist during *Drosophila* embryonic development. In less than a minute, the endogenous pool of Twist can be moved into and out of the nuclei. **(C–D)** Light-mediated plasma membrane clustering of the endogenous Delta signalling protein allows fast and reversible inhibition of Notch signalling during *Drosophila* development. **(E–F)** Optogenetic activation of Rho signalling using two-photon illumination allows precise subcellular activation of apical constriction and morphogenesis (tissue invagination) in live *Drosophila* embryos, following the spatial pattern of photoactivation (ROI = region of interest). In the future, the combination of optogenetics with advanced microscopy techniques will allow scientists to elucidate the causal roles of many different parameters during development, giving new mechanistic insights into embryonic patterning and providing the essential data to build predictive models of development.



Technology Development Box TD2_MD | Brillouin microscopy.

Mechanical properties of cells and tissues, such as elasticity and viscosity, are important in determining biological function. However, current biophysical techniques used in the field to assess these biological functions exhibit intrinsic limitations. To enable 3D measurements of viscoelastic properties at high spatial and temporal resolution in biological samples, the Prevedel Group is developing methodologies based on Brillouin light scattering. This new approach, coined 'Brillouin microscopy', offers a conceptually novel way to probe elastic and viscous properties of biological materials with subcellular spatial resolution and in a non-contact and label-free fashion. Together with the Diz-Muñoz Group, researchers are applying this technique to questions in cell biology, such as cytoskeletal mechanics during cell division, and are working towards establishing this technology as a more widely used, groundbreaking tool in mechanobiology. EMBL researchers are systematically investigating the relationship between the spectra measured by Brillouin microscopy and common mechanical parameters used in the field, such as the ones derived from atomic force microscopy. Furthermore, techniques to decipher the role of mechanics in morphogenesis and tissue self-organisation are applied by studying early mouse embryogenesis and generating mechanical 'atlases' of organisms such as *Platynereis*, which can be linked to spatial gene expression and ultrastructural maps in this animal. This would allow researchers, for the first time, to move biomechanical studies to the molecular regime.



TD1_MD | Brillouin microscopy and its applications in biology. (A) The principle of the Brillouin light scattering interaction with intrinsic acoustic phonons (inset) and a schematic illustration of the imaging system (BS: beamsplitter; VIPA: virtually imaged phase array). Brillouin frequency shift image of **(B)** a live zebrafish tail tissue, **(C)** a mouse embryonic stem cell, **(D)** a preimplantation mouse embryo, and **(E)** an intestinal organoid. Colour bars denote Brillouin frequency shift in GHz, with higher shift indicating higher elasticity or 'stiffness'.

Another key challenge in the future is to **correlate multiple methodologies** so that different parameters and data obtained at different resolutions and frequency regimes can be integrated. An example would be a combined Brillouin and AFM instrument, enabling the integration of internal and surface mechanics, as well as high- and low-frequency mechanics. These biophysical methods also need to be combined with the enormous power of fluorescence imaging technologies to probe multiple molecular properties of living systems. Combining Brillouin and light-sheet microscopy, for example, could be used to probe dynamic changes in cytoskeletal structure and changes in stiffness inside an embryo simultaneously. The aim is to make these tools universally applicable and combinable to allow comprehensive measurement of the biophysical state and dynamics of cells and tissues. This would enable EMBL to create quantitative maps of biophysical properties across space and time, and would provide key new parameters to establish accurate physical models of living systems.

Just as important as pushing the biophysical tools, EMBL researchers need to extend the ability to **comprehensively sample molecular properties of living systems in real time and in 3D**. At the comprehensive scale (e.g. using genomics, transcriptomics, proteomics, and metabolomics), this is currently only possible by using snapshots at discrete time points, and thereby precludes the dynamic analysis of live systems. Although some multimodal single-cell methods have been developed, the data are sparse, with complex interdependencies between the quality of the different data types. These single-cell technologies are very recent and constantly improving. EMBL scientists have been at the forefront of their development and use, including the development of computational models to interpret these new types of data. EMBL is also a leader in the development of real-time imaging technologies, ranging from super-resolution, light-sheet, and multiphoton imaging, to alternative approaches such as photoacoustic microscopy, for which molecular reporters are becoming available. EMBL researchers will develop methods to directly correlate and integrate single-cell omics readouts with single-cell imaging. A pioneering step in this direction was recently made at EMBL in another marine model system, *Phallusia,* where single-cell transcriptomics could be integrated with dynamic 4D light microscopy data of the early development from fertilised egg to gastrulation (Figure MD2).

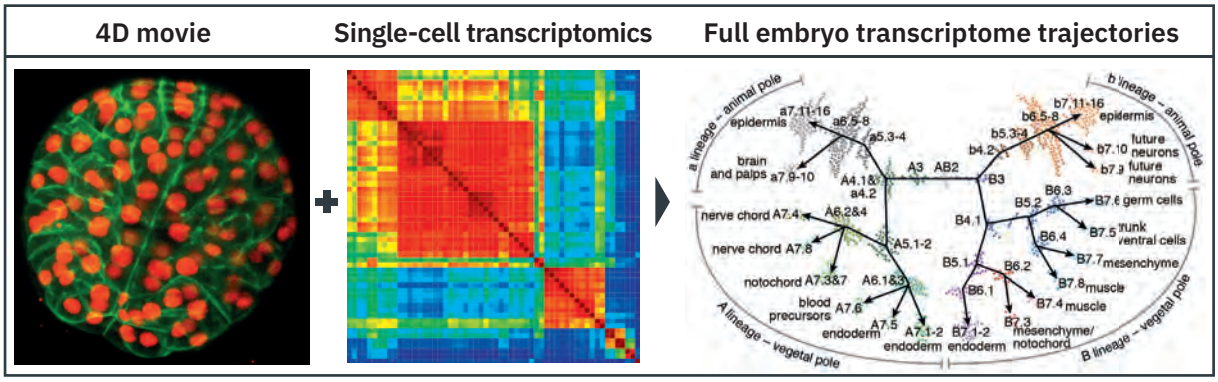


Figure MD2 | Combined 4D real-time imaging and correlative single-cell transcriptome analysis of an embryo of the marine ascidian *Phallusia mammillata*.
3D light-sheet microscopy live-embryo movies of cell lineages were combined with correlative complete single-cell dissection and transcriptome analysis to create complete transcriptome trajectories of all cells in the embryo during gastrulation (http://digitalembryo.org).

One key direction for the future includes the development of multimodal spatial omics technologies to measure multiple molecular parameters from the same single cell in developing embryos *in situ*. A second goal is to develop better and more multiplexable labelling technologies for different super-resolution and real-time imaging technologies to non-invasively read out many molecular entities with high subcellular

precision, ideally to the single-molecule level. These crucial labelling and multiplexing tools will largely come from breakthroughs in chemical biology, where EMBL has recently established a new group that focuses on fluorescent dye and novel reporter development. The development of this next generation of new technologies will be essential to generate quantitative data for multiscale biology in context. These technologies can then be made available as new services to member states and beyond, via the new EMBL Imaging Centre (Chapter 10: Scientific Services).

New Methods of Data Integration, Visualisation, and Sharing

A major challenge for multiscale biology is to successfully integrate models and data relating to cellular and multicellular biology from multiple modalities, as well as multiple spatial and temporal scales. Key aspects to overcome include: (i) how to go beyond data aggregation and integration and use the datasets to create useful mechanistic inferences and testable dynamic predictions, and (ii) how to make data openly accessible and reusable so the community can extract new knowledge and create the next generation of models. Interoperability of heterogeneous experimental databases with different modes of information is already possible, and EMBL plans to be at the forefront of this revolution. To bridge the gap between single-cell spatial omics and dynamic, imaging-based molecular and biophysical properties of cells and tissues, data from both domains need to be integrated into unified frameworks, in the form of time-resolved 3D maps (or '4D atlases'), which can be interactively browsed and annotated by researchers. This will require multimodal data platforms and interactive data visualisation interfaces for virtual reality to be developed.

At the cellular scale, an initial project was accomplished at EMBL with the dynamic protein atlas of human cell division, which integrates dynamic concentration and subcellular localisation data of the proteins driving cell division into an interactive 4D model that can be browsed in virtual reality (Figure MD3). At the multicellular or organism level, a pilot project is currently coming to fruition at EMBL on the marine worm *Platynereis*. In this project, a cellular gene expression atlas is correlated with serial block-face electron microscopy data for a specific developmental state of the entire organism. This enables molecular markers to be registered with segmented nuclei and cell types for every single cell within the organism (Tech Dev Box TD2_SS). The ultimate goal for such multiscale organismal atlases is to have longitudinal data over developmental time – or the lifespan of the organism – to capture dynamic changes in cellular transitions.

The development of multiscale and multimodal organismal atlases provides in-depth curation, annotation, standardisation, re-analysis, and integration of independent datasets: for example, molecular data with morphological, metabolic, and mechanical measurements. These types of knowledge bases can provide broad user communities with access to high-quality biological information and analytical tools for data discovery. Here, the unique expertise of EMBL in data warehousing and data sharing will be leveraged to develop the future 4D tools to visualise multimodal biology in context, across spatial and temporal scales. The recent launch of the BioImage Archive at EMBL-EBI (www.ebi.ac.uk/bioimage-archive), together with the unified metadata concept in the BioStudies database, already provides a foundation for linking molecular and spatial data. Equally important is the strong expertise of EMBL research groups in computational image analysis, using artificial intelligence approaches as well as morphometric shape models. The biological atlases they develop will help scientists to achieve one of the ultimate goals of the field: formulating predictive mechanistic models of the dynamics of these systems. Such computational models will link causal relationships at different scales to the emergent dynamics that explain the growth, movement, morphology, and behaviour of cellular and multicellular systems.

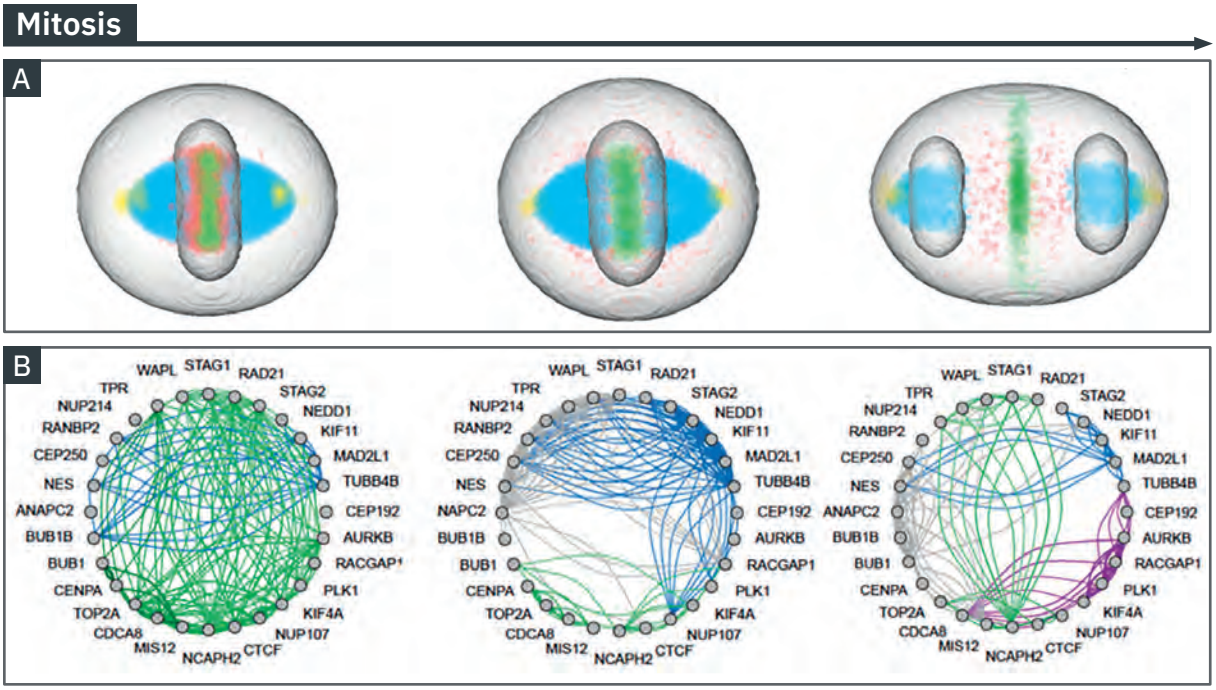


Figure MD3 | Dynamic protein atlas of human cell division.
By integrating experimental and computational approaches, EMBL researchers identified approximately 600 proteins that are needed for a human cell to divide. In an ongoing effort, the timing and subcellular location of those proteins and their interactions are currently identified, enabling the formulation of a data-driven computational model that could predict this dynamic molecular network and relate it to changing cellular boundaries. **(A)** Image analysis and mathematical modelling created a computational mitotic cell model that integrated and visualised protein locations. **(B)** Machine learning was employed to compare the dynamic subcellular fluxes between proteins, allowing the prediction of protein complexes, the temporal order of their (dis)assembly, and the abundance of their subunits. The established approach is generic and can be conceptually transferred to other cellular functions (www.mitocheck.org).

New Technologies for Precise Perturbations of Biological Systems and Environments

Creating precisely controlled perturbations at large scale is essential, both in constructing models and testing their predictions, and to acquire a deep understanding of biological functions to address longstanding mechanistic questions. Single-cell atlas projects are currently generating a wealth of information about which genes are expressed in different cell types. These expression states result from highly interconnected and combinatorial transcriptional networks. Dissecting their regulatory input and functional output in terms of cell function and tissue formation is a pressing current challenge within multiscale biology, which requires rigorous and often systematic experimental testing, combining systematic perturbations with single-cell genomics and other methods.

Recent advances in CRISPR technology have enabled high-throughput screens in cell culture models, either by deleting genetic sequences or by silencing or activating them using dead Cas9 (dCas9) linked to specific domains. Both these techniques are currently impractical to perform at large scale in tissues or embryos. Fundamentally, these ‘genetic’ perturbations are often too crude. This is because a mechanism typically unfolds only at a specific time, such as a critical stage of the cell cycle or development, and/or in a specific place, such as a subcellular compartment or cell type. Beyond qualitative on/off (present/absent) types of manipulations, more tunable and dynamic control needs to occur in order to engineer quantitative changes (e.g. reducing the gene dose or protein affinity) to test quantitative predictions in cellular and multicellular systems.

There is therefore an urgent need for **new technologies that allow systematic perturbations that can be precisely controlled in time and space**. Such new types of perturbation technologies will allow completely new types of questions to be addressed, such as: how are cell-state transitions modulated by perturbing the dynamics of regulator action, or how do perturbations at different timescales affect tissue morphogenesis?

To address this, EMBL researchers are developing new approaches to perturb cellular and multicellular systems with higher throughput, more quantitative precision, and temporal control. For example, most gene regulatory networks do not follow a simple on/off regulatory logic. To alter the levels of upstream regulators in a quantitative manner, perturbations will need to be made over a broad dynamic range, using synthetic programmable transcription factors (such as TALE activators and repressors) and epigenetic modifiers (such as dCas9 activators and repressors). Both technologies have been validated rigorously and applied in cell culture models (Perturb-seq), but currently only at single loci in multicellular systems. Developing methods to enable the systematic use of such systems in primary cells, embryos, and tissues will provide a comprehensive framework for the quantitative control of gene expression, and allow tunable gene regulatory outputs to be created.

Another major effort of EMBL researchers will be to develop systems to control regulators in time. Complex systems are not static atlases, but are rather highly dynamic multimodal networks that constantly change in both their molecular components, such as their levels, interactions, and regulation, and their biophysical properties. In essence, living systems operate as 'open' systems that are out of equilibrium, continuously exploring different states and rapidly transitioning from one state to the next. Although inherent in all living systems, the precise manipulation of dynamic networks in time has remained a huge challenge, given the pleiotropic nature of most essential regulators, and the irreversible nature of most genetic perturbations. To perturb a system in real time, researchers at EMBL are therefore developing multiple methods to control or modulate the timing of molecular components in cellular and multicellular systems. This includes the development of microfluidics-based systems to synchronise and control the timing of the cell cycle or organism development, by modulating, for example, the timing of oscillating systems. Opto- and chemogenetics are another very exciting pair of technologies that EMBL scientists are developing in multiple directions to induce different types of perturbations in live cells and developing embryos. Here, the rapid mislocalisation of key effector proteins within the cell can be induced by a laser beam or the addition of a small molecule that allows proteins or RNAs to switch from an active to an inactive state (Tech Dev Box TD1_MD). The power of such opto- and chemogenetic methods is their highly dynamic and reversible nature. Such precisely controlled temporal perturbations will yield new types of data, which will be complemented by the development of new computational methods to model dynamic trajectories combined with inference-driven experimental designs to select the most informative time points to perturb.

Extracting Correlations and Carrying Out Predictive Computer Modelling

In order to reveal the logic of dynamic living systems and their dynamic response to changing environments, EMBL will combine novel experimental model systems and the rich four-dimensional and multimodal data with theoretical and modeling approaches. Beyond the integrated multimodal databases and portals described above, EMBL will focus on the theoretical foundations of the field, to harness the rich four-dimensional molecular and physical data and discern the underlying logic of dynamic living systems.

A first level of prediction stems from quantitative correlations within the data. If both physical aspects of a cell (e.g. membrane tension) and certain molecular measurements (e.g. protein abundance) are found to be correlated in space across many samples, this can suggest a functional interaction. New sophisticated statistical methods and machine learning techniques have become powerful tools for finding such similarities and patterns within large datasets, and can span both omics and imaging data. Developing these important

approaches further, for multiscale biology, will be pursued together with EMBL's Data Sciences programme (Chapter 8: Data Sciences).

A second level of prediction can be used to define mechanisms. Understanding the dynamic mechanisms behind the behaviour of multiscale biological systems requires formalising hypotheses into mathematical models that can be used to simulate the system computationally. This is because behaviours emerge at higher levels through the nonlinear interactions of numerous components that are structured at lower levels, such as the emergent cellular behaviours arising from the dynamic interactions of regulatory and structural molecules. Mathematical models need to provide non-intuitive but concrete predictions, which can be tested with further experiments. For example, a model should ideally predict the change of a cell's stiffness after receiving an external signal, or the altered morphology of an embryo after knocking down a gene.

EMBL is developing, and will continue to pursue, a variety of mathematical modelling and computational simulation approaches to create **quantitative models** of dynamic 3D cellular and multicellular systems. Useful models will be those that produce alternative hypotheses and distinct concrete predictions that can be tested experimentally. One example of an ongoing multicellular modelling project that EMBL is tackling is mammalian organogenesis (Figure MD4). This data-driven project exemplifies the complete modelling cycle. Such a modelling framework can search through very large parameter spaces by running the slightly altered simulations millions of times; however, simulations are only useful when compared to real data, such as when the outputs of the simulations can be automatically and quantitatively compared to the atlas of real gene expression patterns. Computer simulations can, in principle, take into account different modalities of the properties of multiscale biological systems. These include molecular properties that cross different layers; physical properties such as linear forces, pressures, or tensions; mechanical properties such as elasticity, viscosity, plasticity, or anisotropy; and geometric properties such as morphometry, vectorial orientations of processes, and intercalation.

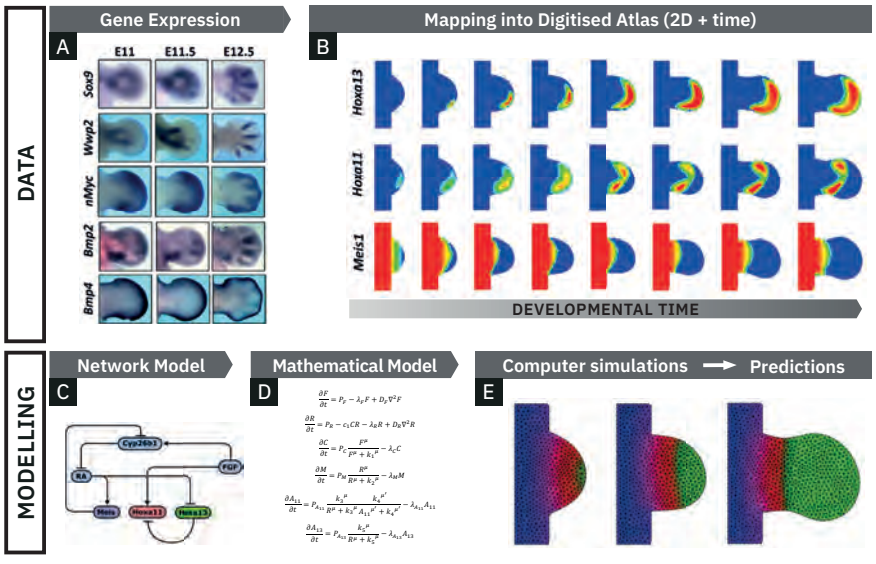


Figure MD4 | Computational modelling of limb development can create quantitative and testable predictions about dynamic gene expression patterns.
(A) The molecular expression patterns for hundreds of genes can be assayed within the whole tissue and imaged. **(B)** The spatial patterns of genes at many time points during development can then be digitised to create a 2D atlas over time and space. **(C)** Hypotheses that may explain the gene expression dynamics can be represented as network models, and **(D)** formalised into mathematical models. **(E)** These models can then be explored in a computer simulation that makes concrete predictions about the gene expression patterns over all time points during development. Importantly, the quantitative predictions can then be compared automatically with the digitised data (shown in B), to determine how accurate the model's predictions are, and ultimately to guide the researcher towards better hypotheses and more informative perturbation experiments.

Modelling complex systems is non-trivial. In practice researchers have to make choices about how to simplify biological systems to make them amenable to systematic experimental measurement and to computer simulations. A key decision to answer any multiscale biology question is also to find the right level of granularity or level of abstraction, which may range from understanding how every known molecular entity fits into the overall picture to capturing just enough quantitative data to make correct predictions about the functional cell or tissue-level behaviours of interest, such as the overall morphology of an organ. These alternatives are not just a matter of choice, but in fact represent a genuine deep scientific question, at the heart of studying complex multiscale biology. The close integration of **theoretical coarse-graining approaches** that are applicable to multiscale biology is key to guide these decisions. Such an integrated approach will also build on and feed into the new Theory at EMBL theme (Chapter 9: Theory at EMBL). Combining experimental, modelling, and theoretical research is urgently needed to understand how life responds and adapts to its ever-changing environment.

Impact

Understanding the biological principles that underlie cellular and multicellular life and its robustness and plasticity in response to changing environmental context, will not only provide fundamental new insight into what underlies **normal development and healthy life**, but will also be essential to revealing how living systems can deteriorate in ageing, disease, or through disruptive environmental changes. It is important to understand what determines the limits of robustness, beyond which living systems deteriorate into disease states. Thus the novel insights from this research area will also provide a basis for the rational design of cells and tissues, ranging from patient-derived induced pluripotent stem cells to powerful *in vitro* models of disease. EMBL research groups are already moving into tissue engineering; for example, integrating human perfusable vascular networks with other cell types to create *in vitro* tissues that enable the study of placental dysfunction, cardiac tissue regeneration, drug delivery to tumour models, and even the pathogenic processes of malarial infection in a model of the human blood–brain barrier (Figure IB2). The deeper our understanding of multicellular dynamics, the greater our capacity to study human diseases *in vitro*.

The study of cellular and multicellular life in context will address these fundamental questions by providing **the next generation of experimental strategies and technologies** to generate quantitative and dynamic molecular and physical data to perturb biological systems with exquisite control in space and time. EMBL develops technology with the goal of transferring it as efficiently as possible to the wider scientific community. Therefore these new tools will be designed to be generally applicable to many biological systems, and will be shared with EMBL's member state community via new scientific services and training activities as soon as possible (Chapter 10: Scientific Services, and Chapter 11: Training).

Exploring cellular and multicellular dynamics will also provide high-value **data portals for cellular and multicellular models**. EMBL will use its strong foundations in open data provision to share large and multimodal datasets effectively, to set up new core EMBL data services, and in an easily accessible manner to maximally enable community impact (Chapter 8: Data Sciences). In addition, EMBL will develop new mathematical models and computer simulations of dynamic living systems, which will allow the exploration of much larger parameter spaces than is experimentally possible, and again share them with the community to use, build on, and further contribute to. EMBL will abstract from such models general principles and best practices for moving biology from the big data, quantitative discipline it is today, into a predictive science for the future.