Chaste: developing sustainable software for computational biology

Alexander Fletcher

University of Sheffield, UK

ZIH-Colloquium, 7th April 2022

Introduction

Motivation

- Limiting factors in computational modelling: **speed** and **reliability**

- **Speed** of progress can be limited
  - Previous models and methods are often not re-used effectively
  - Lack of formal software training can lead to ‘spaghetti code’
  - Most commercial software is closed source and difficult to extend

- The **reliability** of code is often uncertain and unprovable
  - Difficult to guarantee the reproducibility of results
  - In general, no rigorous software testing approach is taken
  - Testing whether results ‘look about right’

Introduction to Chaste

- In computational physiology and biophysics, a wide array of models are represented as continuum ODE/PDE problems, individual or agent-based discrete models, or a hybrid of these two

- Started Chaste in 2005 as a software environment for such problems
  - Most commercial software is **closed source** and difficult to extend
  - We found existing open source software **difficult to test** and extend

- Open source library of fully-tested modules for common elements

- Designed to be **easily utilised** and **readily extended** to the simulation of novel models

Introduction to Chaste

- 'Cancer, Heart, and Soft Tissue Environment'
- An open-source C++ simulation package for computational biology (https://github.com/Chaste)
- To date, development has been driven primarily by:
  - **Cardiac Chaste**: continuum modelling of cardiac electrophysiology
  - **Lung Chaste**: modelling of ventilation in lungs
  - **Cell-based Chaste**: discrete cell-based modelling of biological tissues
- Chaste is modular and extensible, providing libraries for common scientific computing infrastructure, e.g.
  - linear algebra operations
  - finite element meshes
  - ordinary and partial differential equation solvers
- Our software engineering techniques are intended to ensure **code quality**, **re-usability**, and **reliability**

Test-driven development

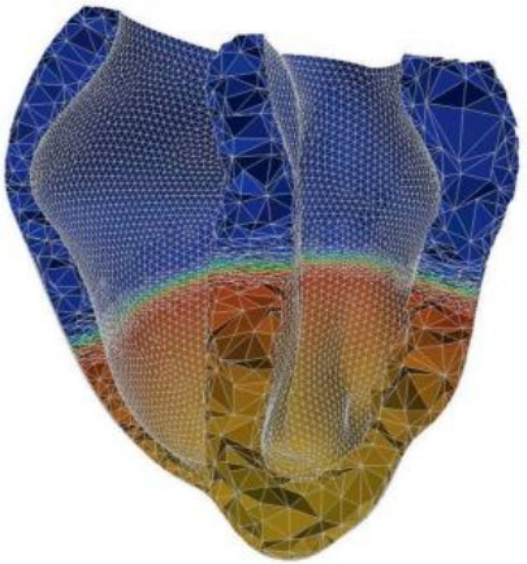
- Forces us to consider the best interface for new code, and how to test that the source code performs its function correctly
- After each 'commit', all tests are run in order to check that no functionality has been inadvertently broken
- Additional tests check for memory leaks and coverage, profile the code, ensure all code is **documented**
- Twitter: **@Chaste_Project**

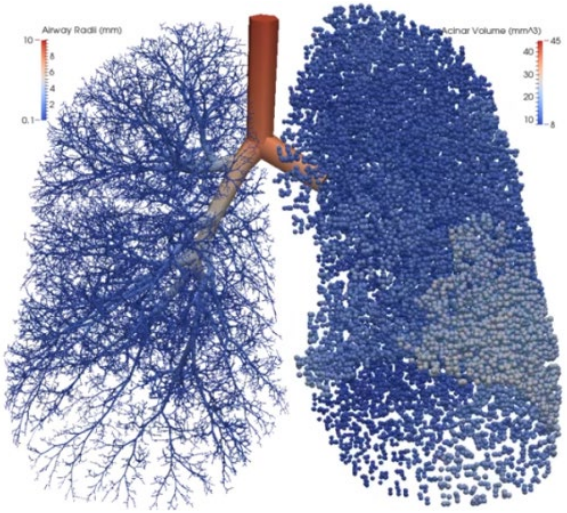
Design and implementation

- We use features of 'eXtreme Programming', allowing for fast development of working prototypes and avoiding 'paralysis through planning'
- Where possible, we use 'pair programming'
- No single person takes sole responsibility for any part of the code; simple coding standards are adhered to

Science enabled

- Since 2009, >5,000 downloads from >50 countries, >150 publications using Chaste, >5,300 citations
- Cardiac Chaste:
  - basic mechanisms of cardiac electrophysiology
  - effects of tissue structure on simulated cardiac electrical activity
  - predicting effects of drugs on cardiac activity
  - parameter inference, model selection and uncertainty quantification
  - verification and efficient numerical simulation
- Lung Chaste: patient-specific airway tree generation and flow modelling





Cell-based Chaste

Multiscale biology

Scale

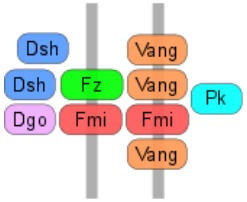
Subcellular

Cell

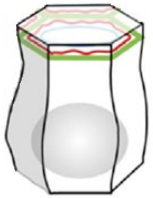
Mesoscale

Tissue

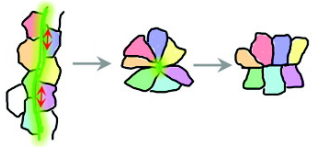
Biological examples



Protein interactions, localisation, turnover, signal transduction



Shape changes, polarisation, movement, division, apoptosis



Actomyosin cables, multicellular rosettes

Convergent extension, size control, morphogen gradients

Modelling approaches

Differential equations, master equations

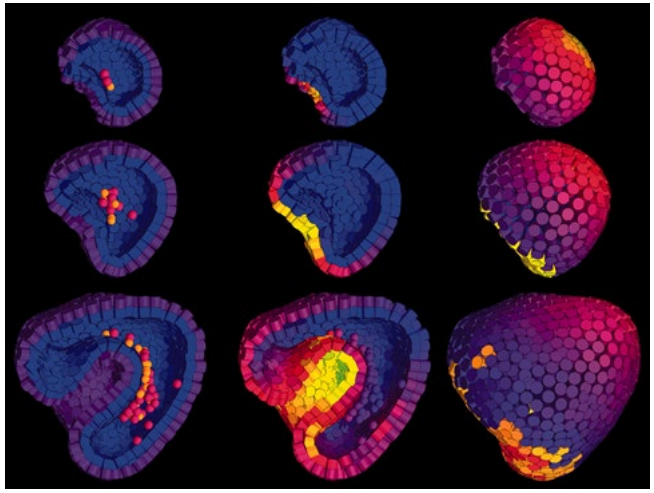
Cellular Potts models, off-lattice models

Hybrid / effective continuum models

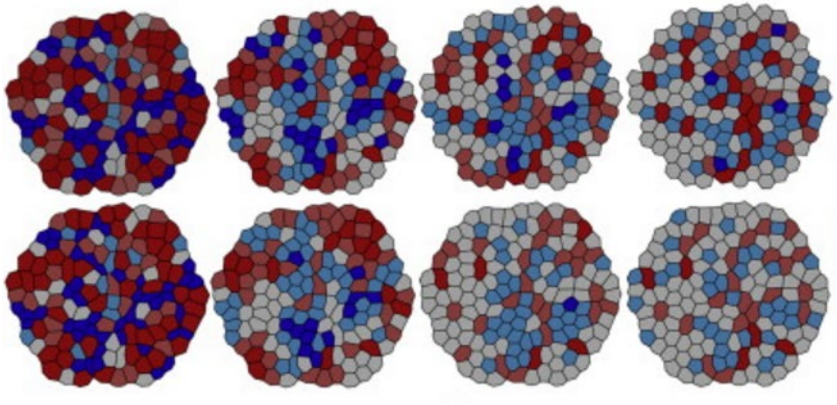
Reaction-diffusion, continuum mechanics

Fletcher & Osborne, *WIREs Mech Dis* (2021)

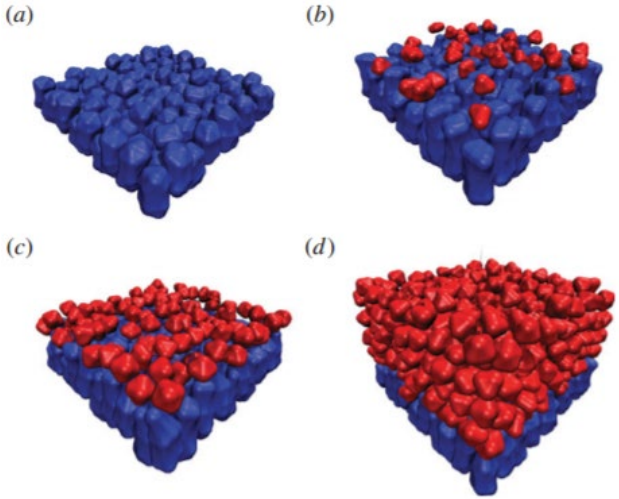
A menagerie of models



Marin-Riera et al (2016)



Xiong et al (2014)



Gord et al (2014)

Mao et al (2013)

A menagerie of models

Approach	Example(s)	Strengths	Limitations
Continuum	Morphogen gradients	Mathematical foundation	Hard to capture heterogeneity
Lattice-based	Epiboly	Fast to simulate	Hard to relate parameters
Cell centre	C. elegans germ line	Physically motivated	No cell shape
Vertex	Germ-band extension	Neighbour exchanges	No cytoskeletal remodelling
Subcellular element	Primitive streak formation	Emergent cell shape	Computationally intensive

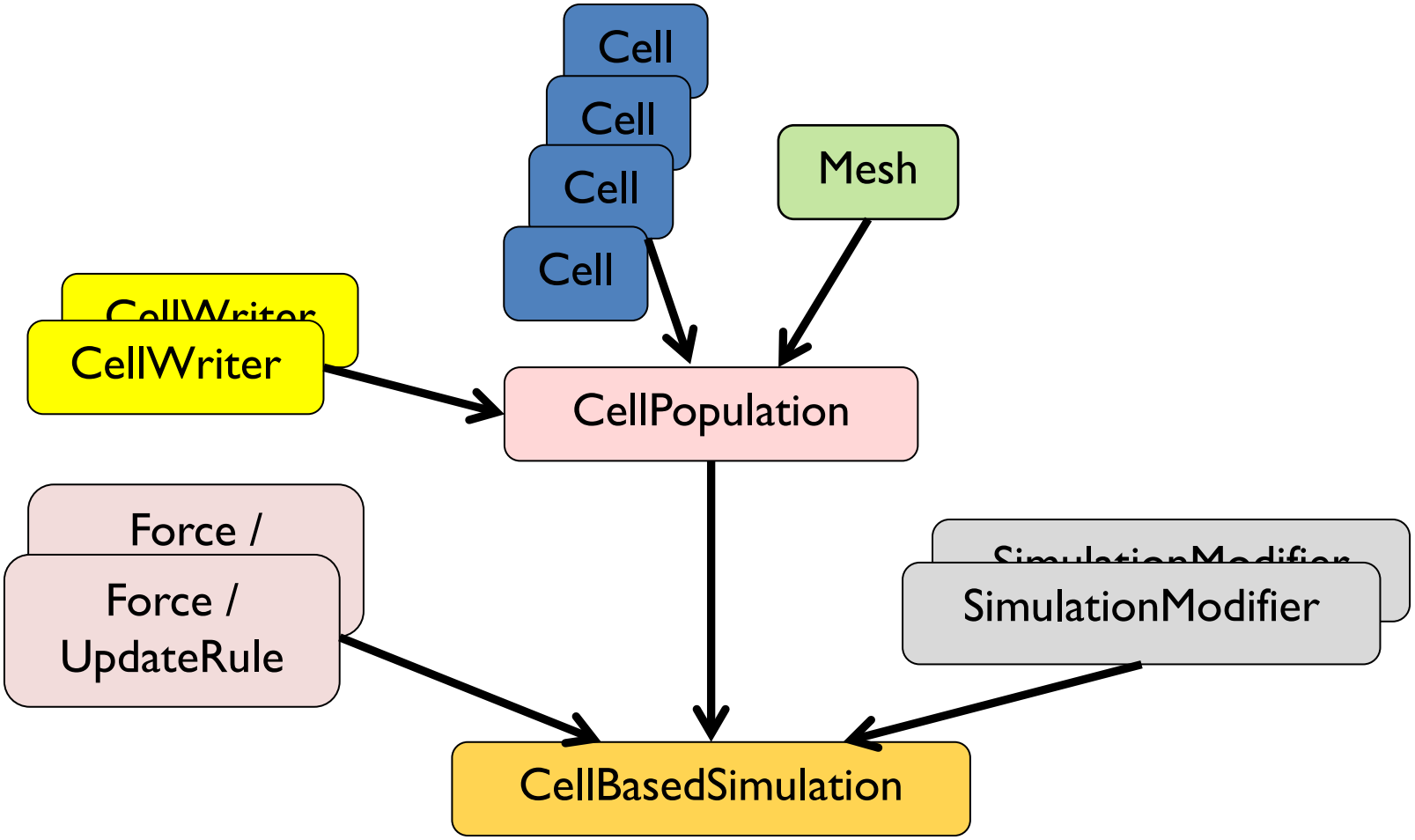
Cell-based Chaste

- **Cell-based Chaste** addresses the need for efficient and verified implementations of cell-based modelling frameworks, providing a set of extensible tools for simulating biological tissues
- A wide range of cell-based modelling frameworks have been developed that have each been successfully applied in a range of biological applications
- Cell-based Chaste includes implementations of:
  - cellular automata
  - cellular Potts models
  - cell-centre models
  - vertex models
  - immersed boundary models

Simulation structure

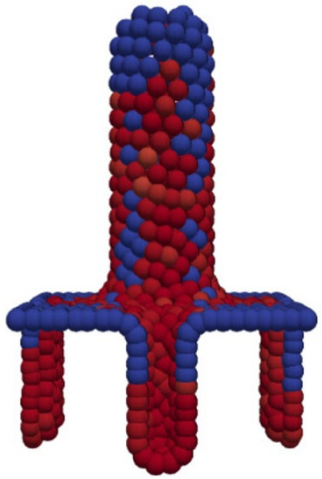
Up to three coupled model 'classes':

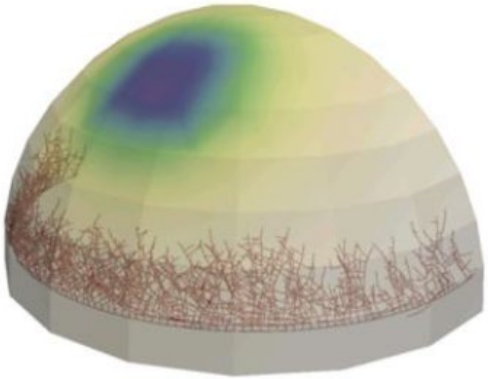
- **(sub)cellular behaviour**, e.g. progress through the cell cycle
- **cell movement** and interactions
- **transport of key signalling molecules,** e.g. morphogens

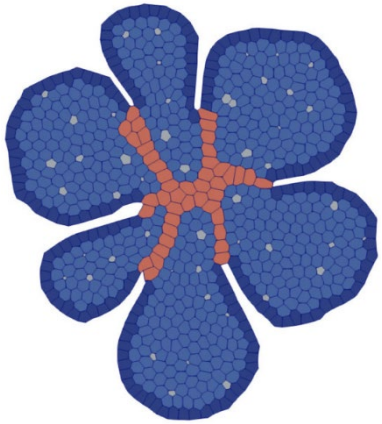


Science enabled

- Intestinal homeostasis and carcinogenesis
- Vascular tumour growth and response to therapy
- Biomechanical characterization of skin lesions
- Stem and pluripotent cells in development
- Dynamics of epithelial tissues
- Spread of sexually-transmitted infections
- Comparison of cell-based model approaches
- Calibration and parameterisation of cell-based models
- Numerical solution of cell-based models







Example application

Modelling in developmental biology

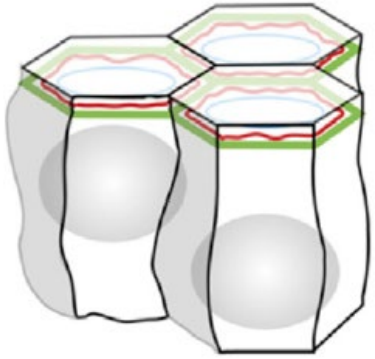
- Modelling biochemistry, mechanics, and couplings
- Symmetry breaking, pattern formation on evolving domains
- Robust emergence of ‘phenotype’ across multiple spatio-temporal scales
- State-of-the-art genetic tools and live imaging, lots of data

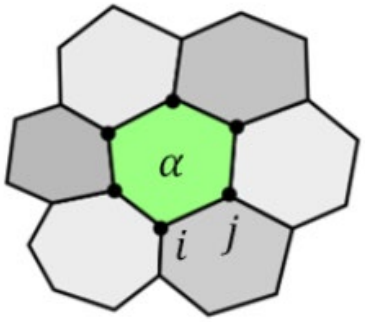
The embryo: a testbed for multiscale modelling

- Complex dynamics, coupling biochemistry and mechanics
- Symmetry breaking and pattern formation on evolving domains
- State-of-the-art genetic tools and live imaging: data at each scale

Video of fly embryonic development taken from Tomer et al, *Nat Meth* (2012)

Vertex models of epithelial tissues

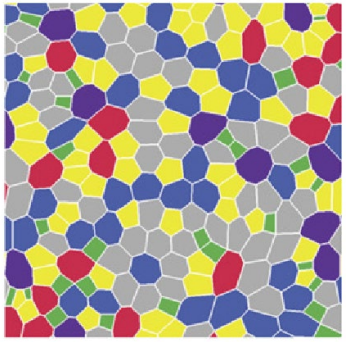


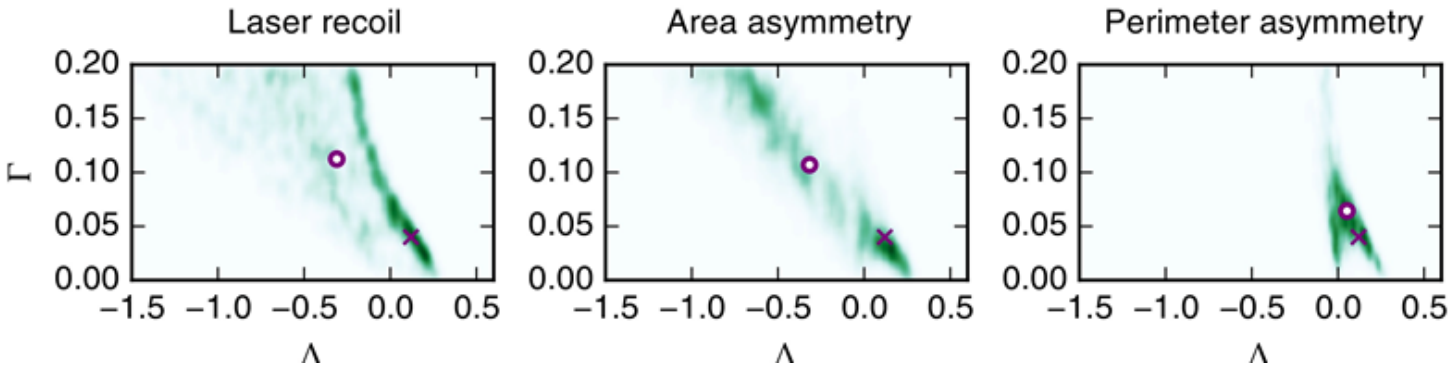


Simulation: Evolve vertices down energy gradient
+ Topological changes

$$U = \sum_{\alpha=1}^{N_c} \left[(A_\alpha - 1)^2 + \Gamma L_\alpha^2 \right] + \Lambda \sum_{ij}^{N_e} f(\{S_\alpha, S_\beta\}_{ij})$$

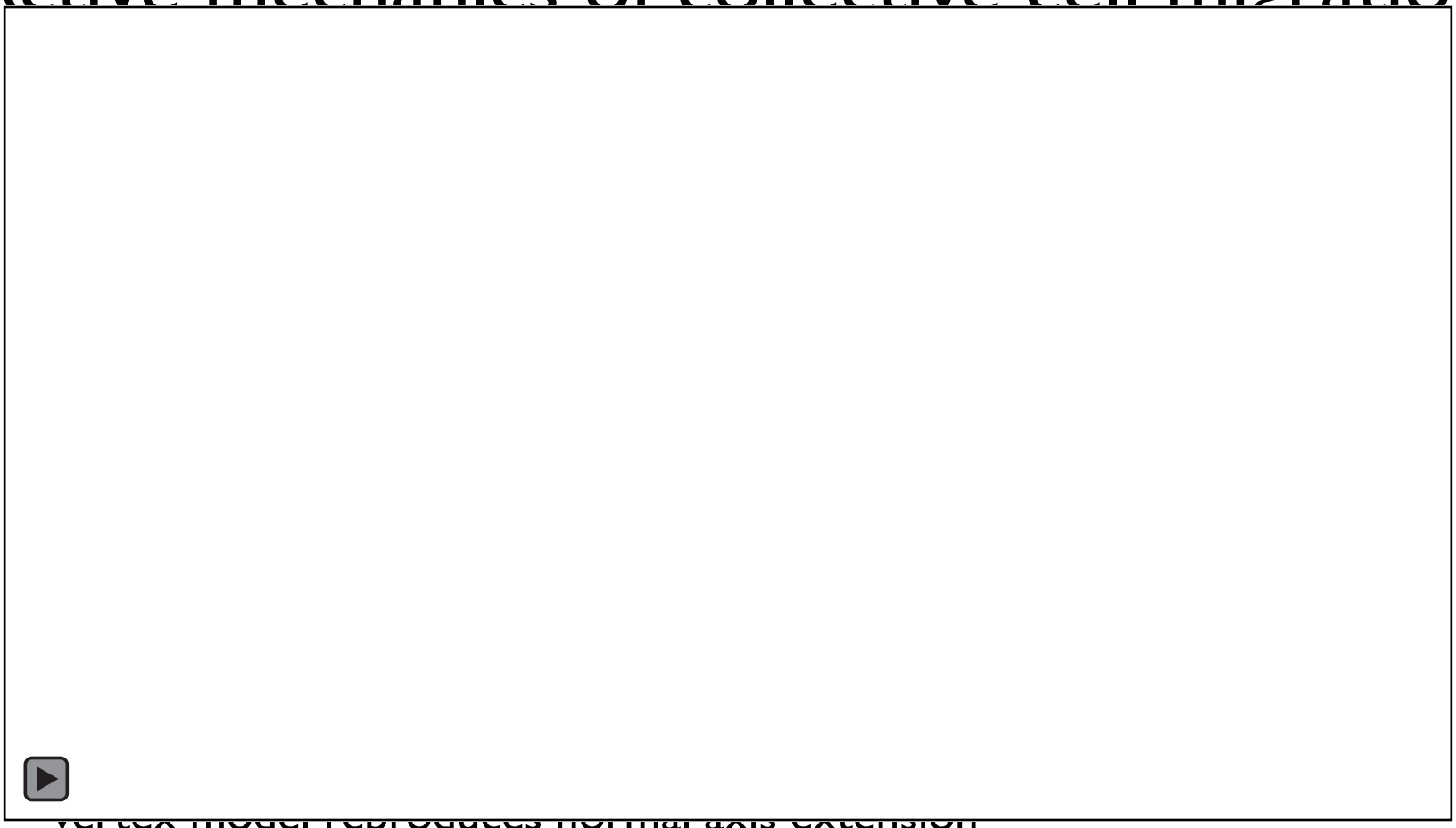
- Inspired by foam models
- Explicitly incorporate cell neighbour rearrangements
- Straightforward to generate experimentally testable summary statistics





Fletcher et al, *Prog Biophys Mol Biol* (2013)
Kursawe et al, *J Theor Biol* (2016)

Active mechanics of collective cell migration

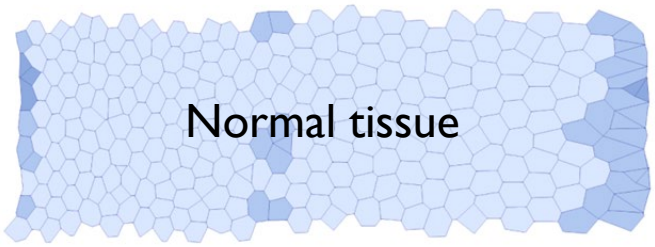


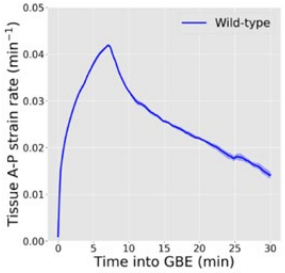
Vertex model reproduces normal axis extension

- Locally patterned intrinsic cell-cell interaction forces
- Incorporates multicellular 'rosettes' and extrinsic pulling force

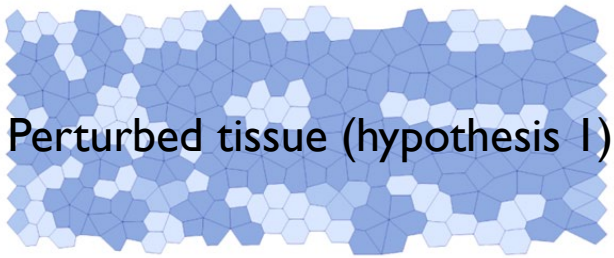
Tetley et al, *eLife* (2016)
Finegan et al, *PLOS Biol* (2019)

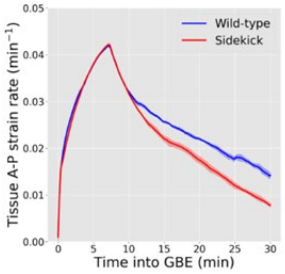
Active mechanics of collective cell migration

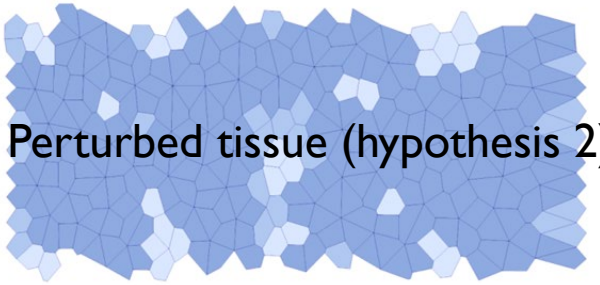


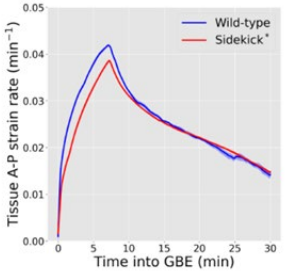


Observed normal tissue

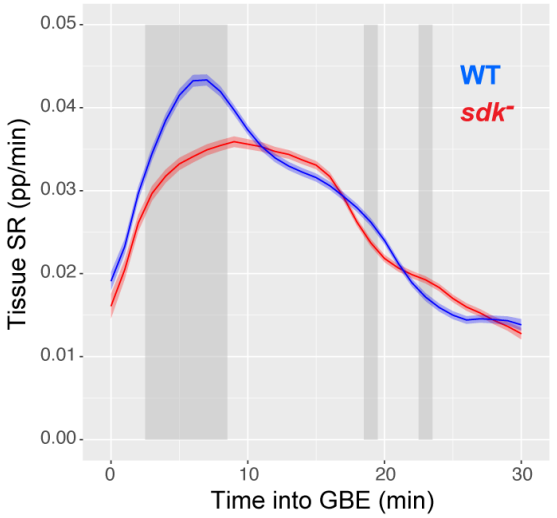








Experimental perturbation



- Use vertex model to understand behaviour when tissue is disrupted genetically
- Compare predicted tissue strain rate under different hypotheses
- Observations are explained by a delay in cell rearrangement during axis extension

Finegan et al, *PLOS Biol* (2019)

Ongoing challenges and next steps

Ongoing challenges

1. Model construction

   Fletcher et al, *Biophys J* (2014); Fletcher et al, *Phil Trans R Soc B* (2017)

2. Model calibration

   Kursawe et al, *J Theor Biol* (2016); Ruske et al, *Phys Biol* (2020)

3. Numerical solution

   Cooper et al, *SIAM J Sci Comput* (2017); Kursawe et al, *J Comput Phys* (2017)

4. Software and hardware implementation

   Harvey et al, *Comput Phys Commun* (2014); Cooper et al, *J Open Source Softw* (2020)

5. Model validation

   Fisher et al, iScience (2019); Gonay et al, *Front Endocrinol* (2021)

6. Data/code standards and benchmarks

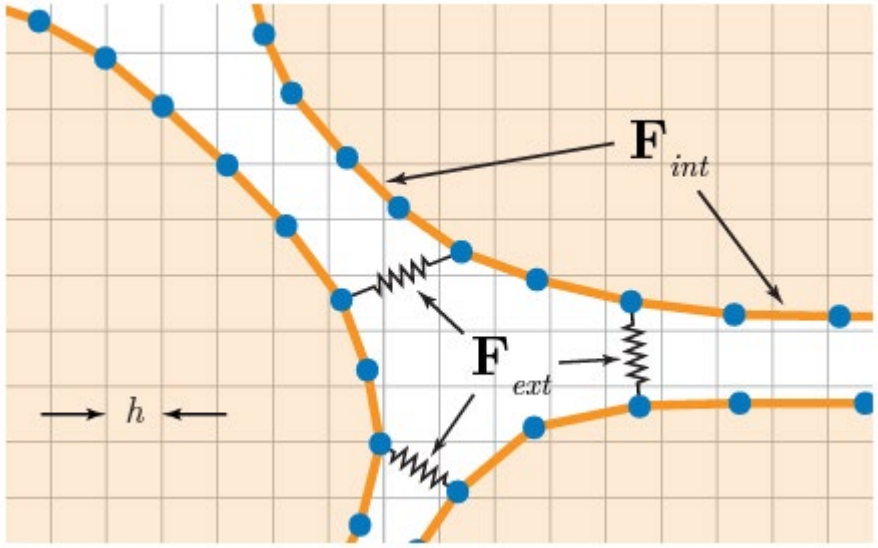
   Friedman et al, *bioRxiv*; Fletcher et al, *Prog Biophys Mol Biol* (2013)

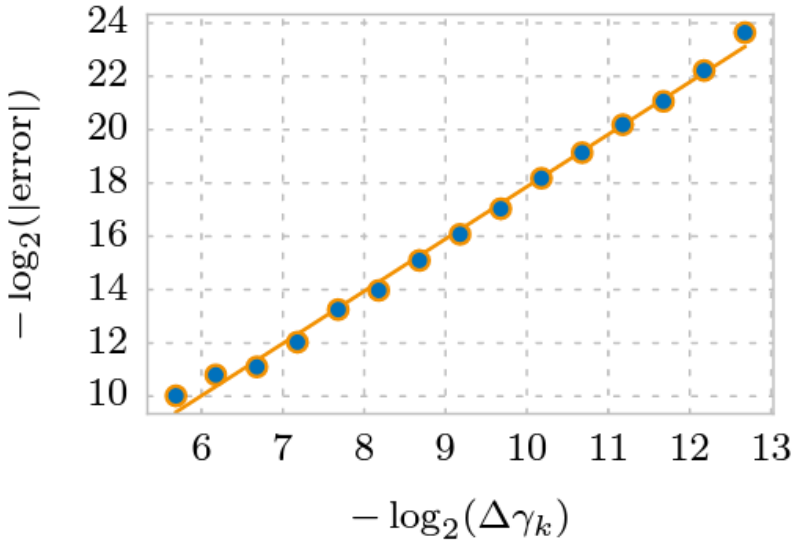
7. Comparing modelling approaches

   Pathmanathan et al, *Phys Biol* (2009); Osborne et al, *PLOS Comput Biol* (2017)

Fletcher & Osborne, *WIREs Mech Dis* (2021)

(3) Numerical analysis of cell-based models





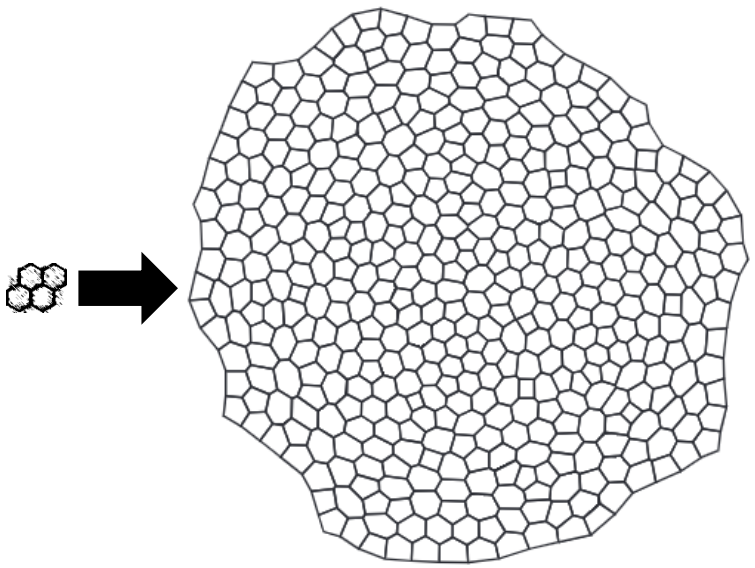
$$\frac{\mathbf{u}^{n+1}-\mathbf{u}^n}{\Delta t}+\sum_d u_d^n D_d^{\pm}\mathbf{u}^n+\frac{1}{Re}\left(\mathbf{D}^0 p^{n+1}-\sum_d D_d^+ D_d^- \mathbf{u}^{n+1}-\frac{1}{3}\mathbf{D}^0 s^n\right)-\mathbf{f}^n=0,$$

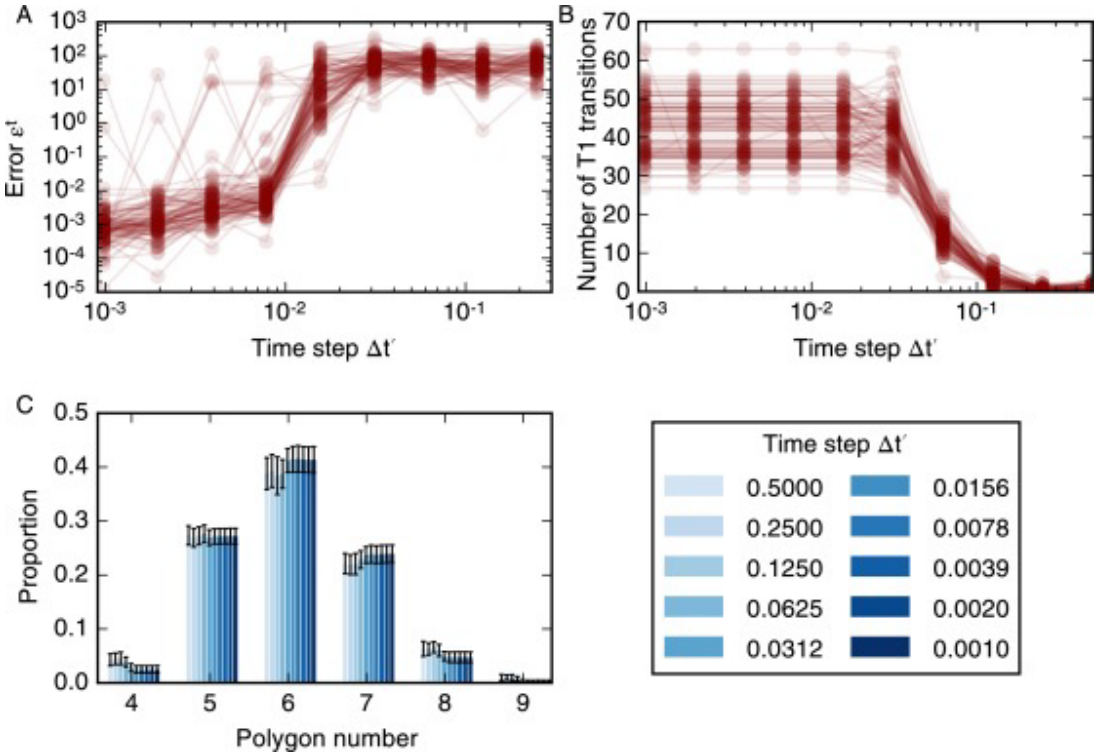
$$\mathbf{D}^0\cdot\mathbf{u}^{n+1}=s^n$$

- FFT-based numerical solution of immersed boundary model of interacting cells
- Identified constraints on relative IB node spacing and fluid mesh spacing
- Derived required scaling of membrane stiffness and cell-cell interaction strength

Cooper et al, *SIAM J Sci Comput* (2017)

(2, 3, 5) Model convergence with discrete events





- Investigate convergence of experimentally measurable summary statistics
- Vertex model of a growing tissue
- Large time steps suppress cell rearrangement
- Convergence with time step is not significantly improved by higher-order methods

Kursawe et al, *J Comput Phys* (2017)

(7) Comparing modelling approaches

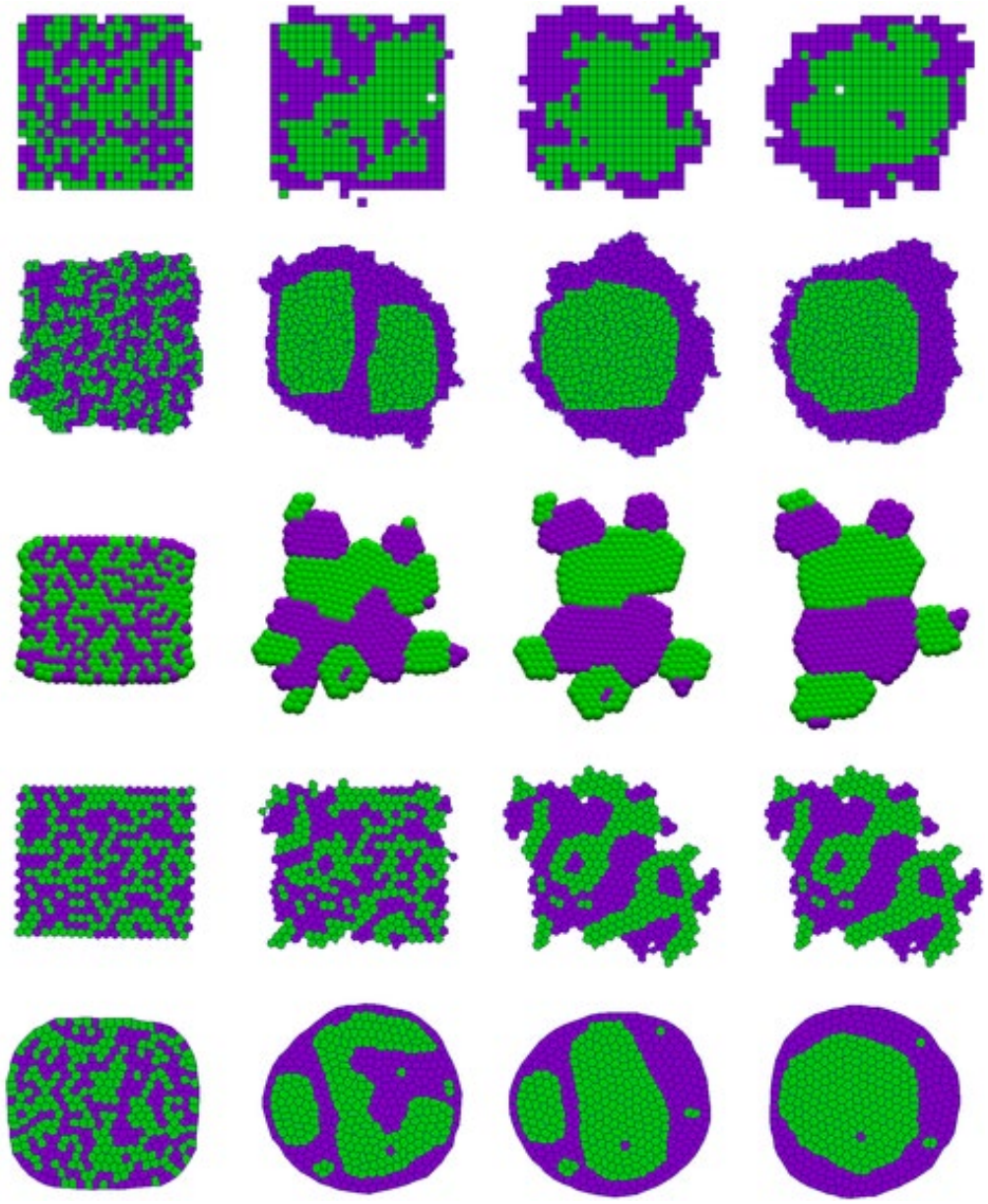
Cellular automaton

Cellular Potts model

Overlapping spheres

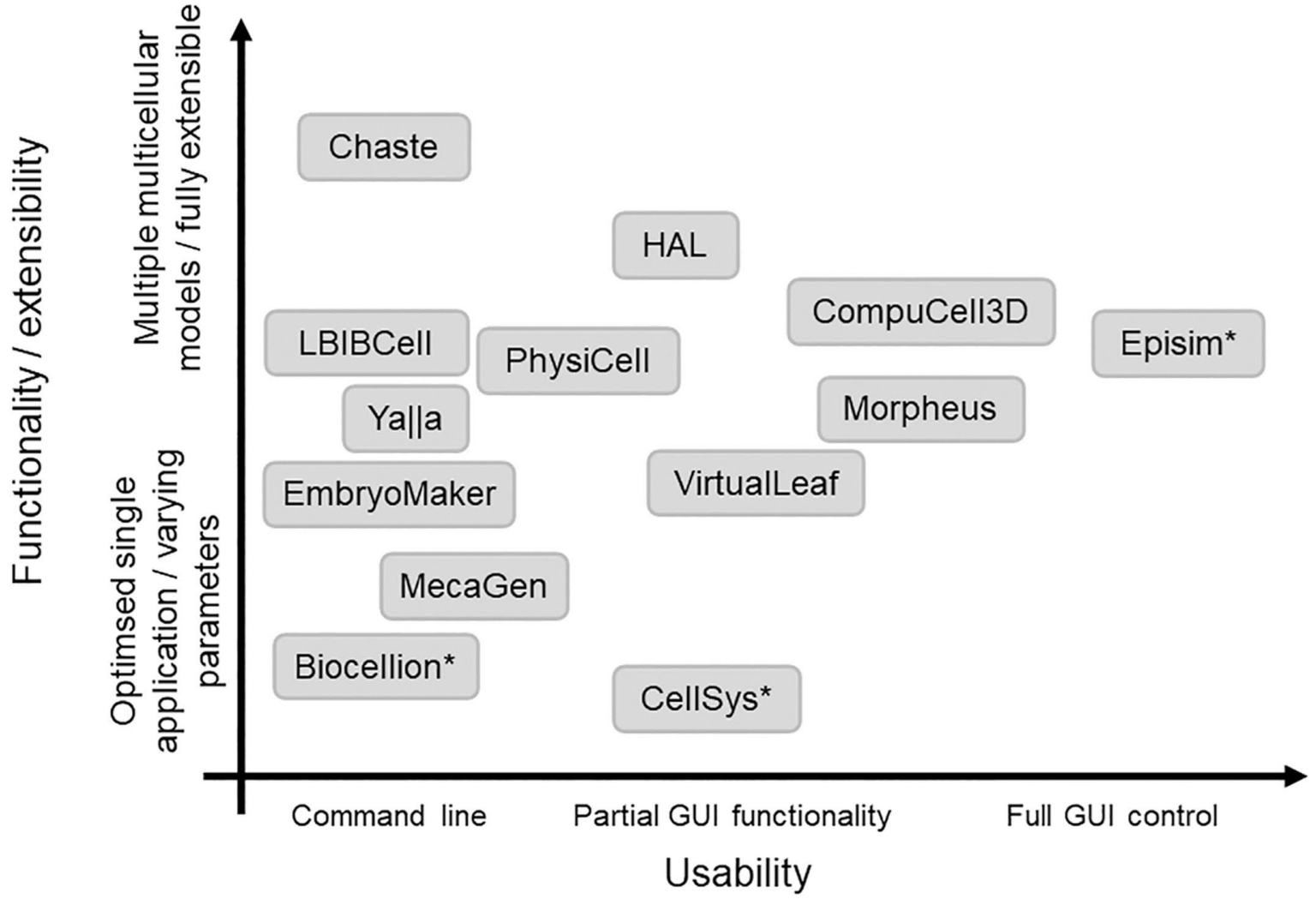
Voronoi tessellation

Vertex model



Osborne et al, *PLOS Comput Biol* (2017)

(4, 6) A software ecosystem?



BBSRC BBR grant (2021-2026)

- **Overarching aim** is to develop an efficient and flexible resource for computational biology

- We will do this by addressing each of the following objectives in a specific work package (WP):
  1. extend existing functionality for multiscale modelling of multicellular populations
  2. upgrade and future-proof our software development infrastructure
  3. improve the interoperability of our software platform
  4. lower barriers to usage and increase community engagement

WP1: Extend cell-based simulation functionality

We will add **new cell-based models** that incorporate finer-grained resolution of cell shape in 2D/3D, and allow the simulation of i**ntracellular processes**...

Milestone 1.1: Incorporate additional biophysical detail in 2D

Milestone 1.2: Implement **3D off-lattice models -** with **Jochen Kursawe**

Milestone 1.3: Incorporate discrete-continuum coupling in 3D

Milestone 1.4: Integrate state-of-the-art biological data

WP2: Utilise new software and hardware

We will substantially **modernise the codebase** to exploit modern libraries and infrastructure…

Milestone 2.1: Reduce reliance on existing physical infrastructure

Milestone 2.2: Modernise and optimise Chaste by adopting modern libraries and best practices

Milestone 2.3: **Parallelise** cell-based simulations

Milestone 2.4: Enhance **GPGPU functionality** using FLAME GPU

WP3: Facilitate interfacing with other software platforms

We will **adopt open standards and model description languages** as an important step towards future-proofing Chaste…

Milestone 3.1: Implement Python wrappers for cell-based simulations

Milestone 3.2: Enable importing of SBML models into Chaste - with **James Osborne**

Milestone 3.3: Enable conversion of MorpheusML models into Chaste - with **Morpheus developers**

Milestone 3.4: Evaluate interfacing Chaste with inference tools in Python and R

WP4: Lower barriers to usage and increase community engagement

We will modernise our **online training material** and prepare and deliver 3 Chaste **training workshops**…

Milestone 4.1: Prepare for and deliver first workshop

Milestone 4.2: Prepare for and deliver second workshop

Milestone 4.3: Prepare for and deliver third workshop

We will also organize hackathons and mini-symposia as part of larger conferences and workshops.

Mam Tor and the Great Ridge, Peak District National Park (http://tiny.cc/7hsxsz)

Acknowledgements

- Chaste development team, in particular BBR grant members:
  - Paul Richmond, Matthew Leach (Sheffield)
  - David Gavaghan, Joe Pitt-Francis, Fergus Cooper, Martin Robinson (Oxford)
  - Gary Mirams, Maurice Hendrix, Kwabena Amponsah (Nottingham)
  - James Osborne (Melbourne)
  - Jochen Kursawe (St Andrews)

- Funding:
  - UKRI-BBSRC