

Application-driven development of modules for an image analysis pipeline in three- dimensional fluorescence microscopy

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich Biowissenschaften

der Johann Wolfgang Goethe-Universität

in Frankfurt am Main

von

Alexander Schmitz

aus Daun (Eifel)

Frankfurt am Main, 2017

(D30)

vom Fachbereich Biowissenschaften der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekanin: Prof. Dr. Meike Piepenbring

Gutachter:

Prof. Dr. Ernst H. K. Stelzer (Goethe-Universität Frankfurt am Main)

Prof. Dr. Ina Koch (Goethe-Universität Frankfurt am Main)

Datum der Disputation: _____

Zusammenfassung

In der Zell- und Entwicklungsbiologie werden Modellsysteme eingesetzt, in denen natürliche Gegebenheiten repräsentativ abgebildet werden. Als Modelle eignen sich v.a. Systeme, die leicht kultivierbar und genetisch modifizierbar sind. So werden entwicklungsbiologische Fragestellungen an Insekten, wie der Fruchtfliege, *Drosophila melanogaster*, dem roten Reismehlkäfer, *Tribolium castaneum*, oder der Feldgrille, *Gryllus bimaculatus*, erforscht. Auch Pflanzen eignen sich für die Analyse von Entwicklungsprozessen. Beispielsweise dient die Entwicklung von Seitenwurzeln in *Arabidopsis thaliana* (Ackerschmalwand) als Modell für die post-embryonale Organogenese. Um die Physiologie und Pathologie von Zellen, Geweben und Organen zu analysieren, werden in der Zellbiologie dreidimensionale Zellkulturen (z.B. Sphäroide oder Organoide) verwendet. Sphäroide sind im Labor einfach herzustellen und erlauben eine reproduzierbare Charakterisierung. In allen Systemen werden Nachbarschaft und Interaktionen von Zellen sowie physikochemische Einflüsse repräsentativ für ein echtes Gewebe abgebildet. Um diese Zusammenhänge zu analysieren, bedarf es geeigneter bildgebender Beobachtungsmethoden. Mit Hilfe der Fluoreszenzmikroskopie können Aufnahmen von spezifisch markierten Molekülen generiert werden.

Die konfokale Fluoreszenzmikroskopie und die Lichtscheiben-Fluoreszenzmikroskopie (engl. *light sheet-based fluorescence microscopy*, LSFM) bieten die Möglichkeit, Bildstapel von dreidimensionalen Proben unter Wahrung ihrer Integrität zu generieren. Die hohe Aufnahmegeschwindigkeit und der geringe Energieeintrag eines LSFM ermöglichen Aufnahmen lebender Proben über lange Zeiträume mit hoher räumlicher und zeitlicher Auflösung. Eine bedeutende Eigenschaft von LSFM ist die Rotation der Proben zwischen den Aufnahmen. So können mehrere Blickrichtungen aufgenommen werden, die zusammengenommen ein sehr detailliertes Abbild der Probe liefern. Die Kombination von LSFM mit optischen Aufhellverfahren (engl. *optical clearing*) ermöglicht eine Steigerung der Eindringtiefe im Vergleich zu nicht-aufgehellten Proben. Die dreidimensionale Struktur bleibt dabei zum Großteil erhalten und kann in ihrer Gesamtheit analysiert werden.

Allgemein ermöglicht die Qualität der mit Fluoreszenzmikroskopie erstellten Aufnahmen die Erhebung quantitativer Daten zu Morphologie, Funktionalität und Dynamik auf verschiedenen Detailstufen. Kompartimente wie Zellkerne, aber auch gesamte Zellen können im Kontext des Gesamtsystems, bzw. als Funktion der Zeit, analysiert werden. Der Rückschluss auf Struktur

und Funktion eines biologischen Systems aus den quantitativen Daten gilt als *bildbasierte Systembiologie*. Diese birgt jedoch einige Herausforderungen. Beispielsweise fallen für eine typische Langzeitaufnahme von *Tribolium castaneum* über einen Zeitraum von 150 Stunden bereits Daten im Umfang von ca. 1,5 Terabyte an, die aus mehreren hunderttausend Einzelbildern bestehen. Um das Potential der Daten zu nutzen, muss die dreidimensionale Fluoreszenzmikroskopie mit einem Bildanalysestrang (engl. *image analysis pipeline*) verknüpft werden.

Ziel dieser Arbeit war es, Module eines Bildanalysestrangs zu identifizieren, zu entwickeln und zu optimieren. Die Module umfassen die Organisation, Reduktion und Visualisierung von Bilddaten, die Rekonstruktion von Bilddaten mehrerer Blickrichtungen, die automatische Erkennung (Segmentierung) und zeitliche Verfolgung (engl. *tracking*) von Zellkernen und die Extraktion von Objektmerkmalen. Die Module wurden anwendungsorientiert entwickelt um die Anwendbarkeit auf Daten aus der dreidimensionalen Fluoreszenzmikroskopie zu gewährleisten. Die zu Grunde liegenden Bilddaten stammten aus aktuellen Forschungsprojekten in der Entwicklungsbiologie von Insekten und Pflanzen, sowie aus der Zellbiologie.

Das erste Modul führt eine Strukturierung der Datensätze durch, reduziert die Datenmenge durch Zuschneiden und Bildkompression und berechnet Projektionen entlang verschiedener Raumrichtungen. Die Einstellungen des Modules werden in einer Benutzeroberfläche kontrolliert. Für die Bildkompression wurden die Kompressionsverfahren ZIP und JPEG2000 auf die Kompressionsleistung und die resultierende Bildqualität evaluiert. Bei gleichbleibender Bildqualität komprimierte ZIP die Bilddaten auf etwa 40% des Rohdatenvolumens. Mit JPEG2000 konnten die Daten auf unter 1% des Rohdatenvolumens reduziert werden. Um dabei die Qualität der Daten bestmöglich zu erhalten, wurde vorab eine Intensitätsanpassung durchgeführt. Relevante Strukturen, wie z.B. Zellkerne, blieben erhalten und konnten durch eine Segmentierung identifiziert werden. Die Evaluierung zeigte, dass die ZIP Kompression im täglichen Umgang mit den Daten als Standard verwendet werden kann. JPEG2000 eignet sich aufgrund der notwendigen Anpassung der Intensitätswerte für eine Langzeitspeicherung der Daten.

Das zweite Modul erlaubt die Aufbereitung der im ersten Modul generierten Projektionen und korrigiert in diesen Bildern Intensitätsschwankungen, die als Funktion der Zeit auftreten. Zudem wird der Bildhintergrund automatisch entfernt. Das Programm kann über eine Benutzeroberfläche gesteuert werden und wurde als Teil einer Publikation in *Nature Protocols* veröffentlicht. Für den Austausch und die einheitliche Visualisierung der Datensätze wurde die Webapplikation *BugCube* als drittes Modul des Bildanalysestrangs entwickelt. Die Datensätze können dadurch teilweise oder vollständig heruntergeladen, sowie selektiv in einem

Webbrowser dargestellt werden. BugCube ermöglicht dadurch eine einheitliche Präsentation von Daten und erleichtert beispielsweise den Datenaustausch mit Kooperationspartnern.

Für die Rekonstruktion von LSFM Aufnahmen aus verschiedenen Blickrichtungen wurde die Erweiterung *Multiview Reconstruction* der Software *FIJI* evaluiert. Die Erweiterung führt eine automatische Registrierung, Fusionierung und Entfaltung der Bilddaten durch. Im Vorfeld meiner Arbeit konnten Frederic Strobl und Alexander Ross die gesamte Embryogenese von *Tribolium castaneum* und *Gryllus bimaculatus* mit dem LSFM aufnehmen. Die Erweiterung wurde hinsichtlich Automatisierung und Effizienz optimiert. In den Zeitserien aller *Tribolium* Datensätze lag der Registrierungsfehler zwischen 0.8 und 1.4 μm . Durch Nutzung der parallelisierten Implementierung auf Grafikprozessoren konnte die Berechnungszeit der Rekonstruktion um den Faktor 15 gesenkt werden. In meiner Arbeit konnte ich zeigen, dass die *FIJI* Erweiterung als Modul für die Rekonstruktion von LSFM Aufnahmen aus verschiedenen Blickrichtungen geeignet ist. Somit konnten wir die ersten qualitativ hochwertigen Rekonstruktionen während der Embryogenese beider Insekten erstellen.

Im Weiteren wurde ein Modul für die quantitative Charakterisierung entwickelt. Die Grundlage bildet eine saatkernbasierte Wasserscheidentransformation, mit der eine Zellkernsegmentierung erreicht wird. Für die zuverlässige Erkennung von Saatkernen habe ich einen Laplace-Gauß-Filter implementiert, der über mehrere Skalen die Positionen der Zellkerne erkennt. Um eine Leistungsbewertung der Zellkernsegmentierung durchzuführen, wurde ein Programm entwickelt, mit dem aus Bilddaten Validierungsdatensätze (engl. *ground truth*) erstellt werden konnten. Das Programm wurde in der Zeitschrift *BMC Bioinformatics* veröffentlicht. In dicht gepackten Sphäroiden erreichte die implementierte Segmentierung eine Quote von 88% korrekt erkannten Zellkernen. Die Zellkernsegmentierung lieferte die Grundlage für Simulationsmodelle, in denen strukturell unterschiedliche Bereiche in Sphäroiden mit Korrelationsfunktionen identifiziert wurden. Die Ergebnisse dieser Zusammenarbeit wurden in der Zeitschrift *Royal Society Interface* veröffentlicht. Neben Position, Volumen und Form erfasst die quantitative Charakterisierung lokale Strukturparameter, die anhand eines Formmodells (engl. *alpha shape*) und von zwei Zellgraphen ermittelt werden. Das Modul liefert somit eine umfassende Charakterisierung auf der Ebene einzelner Zellen, der Zelnachbarschaft und des gesamten Zellverbandes.

In dieser Arbeit wurde das Modul in vier Anwendungen verwendet, um den Einfluss verschiedener Stressbedingungen auf die Morphologie und Anordnung von Zellen in Endpunktaufnahmen optisch aufgehellter Sphäroide zu analysieren. Mit Hilfe des Formmodells konnten Morphologie und Volumen der Sphäroide präzise bestimmt werden. Die Berechnung der Oberfläche des Formmodells erlaubte die Bestimmung der relativen Position von Zellen

innerhalb der Sphäroide. Die generierten Zellgraphen bestimmten die Nachbarschaftsparameter aller Zellen und identifizierten lokale Muster in der Zellanordnung.

In der ersten Anwendung habe ich in Zusammenarbeit mit Christian Mattheyer und Dr. Sabine Fischer den Einfluss der Sphäroidgröße auf die interne Struktur von Tumorsphäroiden analysiert. Die Analyse aus der quantitativen Charakterisierung ermöglichte die Erkennung von strukturellen Zonen innerhalb großer Sphäroide mit mehr als 30.000 Zellen. Während im äußeren Bereich der Sphäroide eine hohe Zelldichte vorherrschte, war diese im zentralen Bereich deutlich reduziert. Zudem konnten wir zeigen, dass sich die Dicke des äußeren Bereichs nicht proportional zur Sphäroidgröße verhielt. Das Modul zur quantitativen Charakterisierung wurde zusammen mit den Ergebnissen dieser Anwendung als Artikel in der Zeitschrift *Scientific Reports* eingereicht und akzeptiert.

In einer zweiten Anwendung habe ich in Zusammenarbeit mit Dr. Nariman Ansari den Einfluss der Apoptose, die vor allem im Bereich der Krebsbehandlung eine Rolle spielt, auf die Zellkern- und Sphäroidmorphologie analysiert. Wir konnten zeigen, dass das Sphäroidvolumen nach einer 72-stündigen Behandlung mit dem Apoptose induzierenden Wirkstoff Staurosporin auf bis zu 32% des Ausgangsvolumens reduziert wurde. Auch Zellkerne verloren bei 72-stündiger Behandlung bis zu 57% ihres Volumens. Anhand der umfangreichen quantitativen Charakterisierung konnte aufgezeigt werden, dass sich die Reduktion des Zellkernvolumens im gesamten Sphäroid, unabhängig von der Größe gleich verhielt.

Die Induktion von Autophagie stellt einen vielversprechenden Therapieansatz, v.a. in der Krebstherapie dar. In Zusammenarbeit mit Katharina Hötte, Michael Koch und Dr. Sabine Fischer habe ich den Einfluss von Langzeitbehandlungen mit Autophagie induzierenden und Autophagie inhibierenden Wirkstoffen auf die Zellkern- und Sphäroidmorphologie analysiert. Die Induktion der Autophagie führte zu einer signifikanten Reduktion der finalen Zellzahl um 34% und des Sphäroidvolumens um 48%. Gleichzeitig konnten wir zeigen, dass sich die Zelldichte bei Induzierung der Autophagie um 15% erhöhte.

Auch mechanische Kräfte können erheblichen Einfluss auf die Struktur und Funktion von Zellverbänden haben. In Zusammenarbeit mit Isabell Smyrek habe ich den Einfluss von mechanischem Stress auf Sphäroide und deren interne Morphologie studiert. Mit Hilfe der umfangreichen quantitativen Charakterisierung konnten wir zeigen, dass sich die Behandlung deutlich auf das Zellkernvolumen auswirkte. Nach 24-stündiger Behandlung waren die Zellkernvolumina um 21% kleiner im Vergleich zu unbehandelten Sphäroiden.

Ein weiteres Modul des Bildanalysestrangs wurde für die Extraktion von Zellkernpositionen und der Abstammungshistorie entwickelt. In Zusammenarbeit mit Dr. Daniel von Wangenheim, Dr. Alexis Maizel und Dr. Jens Fangerau wurde das Programm auf Langzeitaufnahmen der

Seitenwurzel von *Arabidopsis thaliana* angewendet. Mit Hilfe des Programms konnten Zellen aus fünf Datensätzen bis zu 30 Stunden verfolgt und eine Rekonstruktion von insgesamt 62 Zellstammbäumen erstellt und umfassend analysiert werden. Unsere Arbeit ermöglichte die Erstellung eines vierdimensionalen Atlas der Seitenwurzelentwicklung in *Arabidopsis thaliana*. Anhand der Daten konnten wir eine dominante Rolle der zentral gelegenen Zellen identifizieren. Um Muster in der Orientierung von Zellteilungen zu erkennen, wurden diese anhand der Ausrichtung ihrer Zellteilungsachse automatisch klassifiziert. Auf Basis dieser Einteilung entdeckten wir, ausgenommen vom ersten Teilungszyklus der Zellen, ein alternierendes Muster in der Orientierung der Zellteilungen in allen Teilungszyklen. Weiterhin konnten wir feststellen, dass sich in der Seitenwurzel neue Zellschichten zu festen Zeitpunkten in der Entwicklung bildeten. Das entwickelte Modul und die Ergebnisse der Anwendung wurden in der Zeitschrift *Current Biology* veröffentlicht.

In der vorliegenden Arbeit konnte ich die Module in interdisziplinären Forschungsprojekten der Zell- und Entwicklungsbiologie auf Datensätze aus der dreidimensionalen Fluoreszenzmikroskopie anwenden. Die Module ermöglichten die Organisation, Verarbeitung, Visualisierung und Analyse der Datensätze. Die entwickelten Module eines automatisierten Bildanalysestrangs haben ein hohes Potential für die Anwendung in weiteren Forschungsprojekten.

In zukünftigen Arbeiten können standardisierte Verarbeitungsschritte, wie die Erstellung von Projektionen, die Kompression oder die Entfaltung von Datensätzen bereits während der Aufnahme abgearbeitet werden. Das Modul für die Rekonstruktion von LSFM Aufnahmen kann für die Erzeugung von Datensätzen eingesetzt werden, in denen Analysen zu Bewegungsmustern und Stammbäumen von Zellen durchgeführt werden können. Dies setzt die Etablierung einer geeigneten, automatischen Methode zur Verfolgung von Zellen als Funktion der Zeit auf Basis der bestehenden Lösungen und meiner Arbeit voraus. Das entwickelte Modul zur quantitativen Charakterisierung kann hinsichtlich der Interpretation anderer Strukturen (z.B. Membranen) oder funktionaler Färbungen (z.B. apoptotische Zellen) erweitert werden. Der komplette Ansatz lässt sich ebenso auf Lebendaufnahmen und ähnliche Systeme (z.B. Organoide) anwenden. Mit Hilfe einer umfangreichen quantitativen Charakterisierung lässt sich der Einfluss von physikalischen und chemischen Faktoren auf die Physiologie von Zellen, bzw. Geweben analysieren.

Die Perspektive eines Bildanalysestrangs liegt in einer gemeinschaftlichen Herangehensweise mit regem Austausch von Wissen und spezialisierten Lösungen für die Probleme der Bildanalyse. Dieser Ansatz kann nicht nur in der bildbasierten Systembiologie, sondern auch in der medizinischen Diagnostik und in industriellen Hochdurchsatzverfahren Anwendung finden.

Summary

Research in cell and developmental biology requires the application of three-dimensional model systems that reproduce the natural environment of cells. Processes in developmental biology are therefore studied in entire systems like insects or plants. In cell biology, three-dimensional cell cultures (e.g. spheroids or organoids) model the physiology and pathology of cells, tissues or organs. In all systems, the cellular neighborhood and interactions, but also physicochemical influences, are realistically presented. The production and handling of these model systems is rather simple and allows for reproducible characterization.

Confocal and light sheet-based fluorescence microscopy (LSFM) enable the observation of these systems while maintaining their three-dimensional integrity. LSFM is applicable to imaging live samples at high spatio-temporal resolution over long periods of time. The quality of the acquired datasets enables the extraction of quantitative features about morphology, functionality and dynamics in the context of the complete system. This approach is referred to as *image-based systems biology*. Exploiting the potential of the generated datasets requires an image analysis pipeline for data management, visualization and the retrieval of biologically meaningful values.

The goal of this thesis was to identify, develop and optimize modules of the image analysis pipeline. The modules cover data management and reduction, visualization, reconstruction of multiview image datasets, the segmentation and tracking of cell nuclei and the extraction of quantitative features. The modules were developed in an application-driven manner to test and ensure their applicability to real datasets from three-dimensional fluorescence microscopy. The underlying datasets were taken from research projects in developmental biology in insects and plants, as well as from cell biology.

The datasets acquired in fluorescence microscopy are typically complex and require common image processing steps in order to manage, visualize, and analyze the datasets. The first module accomplishes automatic structuring of large image datasets, reduces the data amount by image cropping and compression and computes maximum projection images along different spatial directions. The second module corrects for intensity variations in the generated maximum projection images that occur as a function of time. The program was published as a part of an article in *Nature Protocols*. Another developed module named *BugCube* provides a web-based platform to visualize and share the processed image datasets.

In LSFM, samples can be rotated in-between two acquisitions enabling the generation of multiview image datasets. Prior to my work, Frederic Strobl and Alexander Ross acquired the complete embryogenesis of the red flour beetle, *Tribolium castaneum*, and the field cricket, *Gryllus bimaculatus*, with LSFM. I evaluated a plugin for the software *FIJI* as a module for the reconstruction of such datasets. The plugin was optimized for automation and efficiency. We obtained the first high quality three-dimensional reconstructions of *Tribolium* and *Gryllus* datasets.

Optical clearing increases the penetration depth into samples, thus providing endpoint images of entire three-dimensional objects with cellular detail. This work contributes a quantitative characterization module that was applied to endpoint images of optically cleared spheroids. A program for the generation of ground truth datasets was developed in order to evaluate the cell nuclei segmentation performance. The program was part of a paper that was published in *BMC Bioinformatics*. Using the program, I could show that the cell nuclei segmentation is robust and accurate. Approaches from computational topology and graph theory complete the segmentation of cell nuclei. Thus, the developed module provides a comprehensive quantitative characterization of spheroids on the level of the individual cell, the cell neighborhood and the whole cell aggregate. The module was employed in four applications to analyze the influence of different stress conditions on the morphology and cellular arrangement of cells in spheroids. The module was accepted for publication in *Scientific Reports* along with the results for one application. The cell nuclei segmentation further provided a data source for simulation models that used correlation functions to identify structural zones in spheroids. These results were published in *Royal Society Interface*.

The final part of this work presents a module for cell tracking and lineage reconstruction. In collaboration with Dr. Alexis Maizel, Dr. Jens Fangerau and Dr. Daniel von Wangenheim, I developed a module to track the positions of all cells involved in lateral root formation in *Arabidopsis thaliana* and used the extracted positions for extensive data analysis. We reconstructed the cell lineages and established the first atlas of all founder cells that contribute to the formation. The analysis of the retrieved data allowed us to study conserved and individual patterns in lateral root formation. The atlas and parts of the analysis presented in this thesis were published in *Current Biology*.

In this thesis, I developed modules for an image analysis pipeline in three-dimensional fluorescence microscopy and applied them in interdisciplinary research projects. The modules enabled the organization, processing, visualization and analysis of the datasets. The perspective of the image analysis pipeline is not restricted to image-based systems biology. With ongoing development of the image analysis pipeline, it can also be a valuable tool for medical diagnostics or industrial high-throughput approaches.

Acknowledgements

Ich bedanke mich herzlich bei allen Personen, die mich während meiner Promotion begleitet haben: Meinem Doktorvater, Professor Ernst H. K. Stelzer, für die herzliche Aufnahme in seine Arbeitsgruppe, den Freiraum und die Unterstützung in allen Situationen. Professor Ina Koch, die mich seit meiner Diplomarbeit in ihrer Arbeitsgruppe unterstützend begleitet hat. Ein besonderer Dank gilt Dr. Sabine Fischer, die meine Arbeit betreut, sich immer Zeit genommen und meine analytische Denkweise geprägt hat.

Ich danke insbesondere allen Kollegen und externe Kollaborationspartnern, mit denen ich an interessanten Projekten zusammen arbeiten konnte: Dr. Daniel von Wangenheim, Dr. Alexis Maizel und Dr. Jens Fangerau für die Zusammenarbeit am Seitenwurzelprojekt. Frederic Strobl für die Zusammenarbeit in vielen erfolgreichen Projekten. Ich danke Dr. Nariman Ansari für die Zusammenarbeit im ProMEBS Projekt und allen daraus entstandenen Folgeprojekten, Katharina Hötte für die Zusammenarbeit im Autophagie-Projekt und Isabell Smyrek für die Zusammenarbeit am Projekt der mechanischen Stressinduzierung in Sphäroiden.

Ich möchte weiterhin den Mitgliedern des „Brainlab“, Biena Mathew, Dr. Sabine Fischer und Sven Plath, für die stets tolle Atmosphäre während der letzten Jahre danken. Mein Dank gilt ebenso allen weiteren Kollegen der Arbeitsgruppe für die gemeinsame Zeit: Dr. Francesco Pampaloni, Berit Reinhardt, Sigrun Becker, Michaela Koller, Victor Perez Meza, Heinz Schewe, Till Moreth und Lotta Hof. Allen Alumnis und Studenten mit denen ich während meiner Promotion zusammen arbeiten konnte: Christian Mattheyer, Dr. Bo-Jui Chang, Jenia Schlegel, Alexander Ross, Helena Domínguez Moreno, Manuel Sommer, Konstantin Kabat vel Job, Dr. Sebastian Pfeiffer, Roli Richa Jannis Lehmann und Alexander Atzberger. Ein besonderer Dank gilt den tatkräftigen und jederzeit hilfsbereiten Jungs des HRZ, Kai Richter und Sascha Rodzies.

Ich danke meinen Eltern, Sonja und Bruno Schmitz, für die Unterstützung während meines gesamten Studiums und der Promotion. Ein besonderer Dank gilt meiner Freundin, Isabell Smyrek, für die Unterstützung in allen Lebenslagen.

Table of Contents

Zusammenfassung	i
Summary	vi
Acknowledgements	viii
Table of Contents.....	ix
List of Figures	xiii
List of Tables.....	xv
List of Abbreviations.....	xvi
1 Introduction	1
1.1 Workflow of image-based systems biology	1
1.1.1 Image acquisition with advanced light microscopy	2
1.1.2 Components of an image analysis pipeline.....	4
1.2 Biological model systems and applications.....	8
1.2.1 Tribolium castaneum and Gryllus bimaculatus embryos.....	8
1.2.2 Three-dimensional cell cultures as representative <i>in vitro</i> models for tissues	9
1.2.3 Lateral root development in <i>Arabidopsis thaliana</i>	13
1.3 Scope of this thesis.....	15
2 Material	16
2.1 Tribolium datasets.....	16
2.1.1 Preparation of Tribolium embryos	16
2.1.2 Long-term fluorescence live imaging of Tribolium embryos with LSFM	16
2.2 Gryllus datasets	17
2.2.1 Preparation of Gryllus embryos.....	17
2.2.2 Long-term fluorescence live imaging of Gryllus embryos with LSFM.....	17
2.3 Spheroid datasets.....	18
2.3.1 Cell culture and spheroid preparation.....	18
2.3.2 Drug Treatment of T47D spheroids.....	18
2.3.3 Cylindrical constriction of HC11 spheroids	18
2.3.4 Staining and immunofluorescence.....	19
2.3.5 Optical clearing and sample preparation	19
2.3.6 Image acquisition with CLSM and LSFM.....	20
2.4 Arabidopsis datasets.....	20
2.4.1 Plant growth conditions	20
2.4.2 Image acquisition with LSFM	20
2.5 Hardware and implementation details.....	21
2.5.1 Hardware	21

2.5.2	Implementation.....	21
3	Methods	22
3.1	Automated data management	22
3.1.1	Data structure and file name convention	22
3.1.2	Data reduction by three-dimensional cropping.....	23
3.1.3	Data reduction by ZIP and JPEG2000 compression.....	23
3.1.4	Generation of maximum projections	24
3.2	Intensity normalization of time-lapse maximum projections	24
3.2.1	Background removal	24
3.2.2	Intensity normalization	24
3.3	Web-based data visualization and sharing with <i>BugCube</i>	24
3.4	Multiview reconstruction	25
3.4.1	Multiview reconstruction of <i>Tribolium</i> datasets	25
3.4.2	Multiview reconstruction of a <i>Gryllus</i> dataset.....	26
3.5	Multiscale quantitative characterization.....	27
3.5.1	Pre-processing	27
3.5.2	Initial cell nuclei segmentation.....	28
3.5.3	Marker point detection	28
3.5.4	Watershed-based decomposition of cell nuclei clusters	29
3.5.5	Evaluation of segmentation performance	31
3.5.6	Geometrical modelling with alpha shapes	31
3.5.7	Cell graph construction.....	32
3.5.8	Guidelines for setting up parameter values.....	33
3.5.9	Random cell position model	37
3.6	Computer-aided cell tracking and lineage reconstruction	37
3.6.1	<i>TrackGen</i> combines cell tracking and lineage tracing.....	37
3.6.2	Automated lineage reconstruction	38
3.6.3	Registration of the Arabidopsis datasets.....	38
3.6.4	Quantitative lateral root properties	39
3.6.5	Quantitative cell division properties.....	39
4	Results.....	41
4.1	An automated data management module	41
4.1.1	Workflow of the automated data management module	41
4.1.2	Capabilities of the automated data management module	42
4.1.3	Evaluation of ZIP and JPEG2000 image compression	43
4.2	Intensity normalization of time-lapse maximum projections	47
4.3	<i>BugCube</i> enables web-based visualization and data sharing.....	48
4.4	Reconstruction of multiview image datasets.....	49
4.4.1	Overview of the reconstruction procedure.....	49
4.4.2	Multiview reconstruction of <i>Tribolium</i> datasets.....	50
4.4.3	Multiview reconstruction of <i>Gryllus</i> datasets	54

4.5	Multiscale quantitative characterization.....	56
4.5.1	Overview of the multiscale quantitative characterization module.....	57
4.5.2	Interactive control of cell nuclei segmentation.....	58
4.5.3	Post-processing the cell nuclei segmentation results.....	59
4.6	Quantitative characterization of spheroids under stress conditions.....	60
4.6.1	Application I: nutrient and oxygen deficiency.....	60
4.6.2	Application II: apoptosis induction.....	71
4.6.3	Application III: autophagy induction.....	77
4.6.4	Application IV: effects of mechanical perturbation.....	80
4.7	Computer-aided cell tracking and lineage reconstruction.....	84
4.7.1	Five datasets capturing lateral root development.....	84
4.7.2	<i>TrackGen</i> enables cell tracking and lineage tracing.....	86
4.7.3	Complete reconstruction of cell lineages.....	87
4.7.4	Assignment of cell files.....	88
4.7.5	Classification of cell division types.....	89
4.7.6	Complete atlas of cell lineages for five lateral root datasets.....	90
4.7.7	Growth dynamics of lateral root development.....	91
4.7.8	Cellular composition of the lateral root.....	93
4.7.9	Spatio-temporal patterns of cell divisions.....	94
5	Discussion.....	100
5.1	Multidimensional image datasets require automated data management.....	100
5.2	BugCube enables web-based visualization and sharing.....	101
5.3	Reconstruction of multiview image datasets.....	102
5.3.1	The first multiview reconstruction of <i>Tribolium</i> and <i>Gryllus</i> embryogenesis.....	102
5.3.2	A general workflow for reconstruction of multiview image datasets.....	102
5.4	Multiscale quantitative characterization.....	103
5.4.1	Densely packed cell nuclei in diverse spheroid datasets are accurately identified.....	103
5.4.2	Alpha shapes provide a geometrical model for arbitrarily shaped cell aggregates.....	104
5.4.3	Cell graphs investigate the spatial topology of cell aggregates.....	104
5.5	Quantitative characterization of spheroids under stress conditions.....	104
5.5.1	Application I: nutrient and oxygen deficiency.....	105
5.5.2	Application II: apoptosis induction.....	105
5.5.3	Application III: autophagy induction.....	107
5.5.4	Application IV: acute mechanical constriction.....	108
5.6	Computer-aided cell tracking and lineage reconstruction.....	109
5.6.1	Simplified cell tracking and lineage reconstruction.....	109
5.6.2	The first four-dimensional atlas of lateral root development.....	110
5.6.3	Lateral root development is conserved on the organ level.....	110
5.6.4	Cellular contribution of cell lineages and cell files is variable.....	110
5.6.5	Organ-level coupling in the occurrence of formative cell divisions.....	111
5.6.6	Developmental instability and self-organization in lateral root development.....	111

6	Outlook and Conclusion	113
6.1	Future directions of the image analysis pipeline	113
6.1.1	Multiview reconstruction as a basis for cell tracking and lineage tracing	113
6.1.2	Generalized usage of the multiview reconstruction procedure	114
6.1.3	Extension and application of the multiscale quantitative characterization	115
6.1.4	Extension of the pipeline towards automated high-throughput processing	116
6.2	Concluding remarks	117
7	References	119
8	Supplementary Material	129

List of Figures

Figure 1 Workflow of an image-based systems biology approach.	1
Figure 2 Illumination and detection principle in fluorescence microscopes.	3
Figure 3 Comparative analysis of developmental processes during insect embryogenesis.	9
Figure 4 Lateral root development in <i>Arabidopsis thaliana</i>	15
Figure 5 Mounting principle for <i>Tribolium</i> embryos.	16
Figure 6 Mounting principle for <i>Gryllus</i> embryos.	17
Figure 7 Overview of cylindrical constriction of HC11 spheroids.	19
Figure 8 Schematic illustration of the immersion watershed transform.	30
Figure 9 Framework of the automated data management module.	42
Figure 10 <i>JobCreator</i> user interface.	43
Figure 11 Evaluation of JPEG2000 compression for different scaling factors.	45
Figure 12 Qualitative evaluation of JPEG2000 compression for <i>Tribolium</i> dataset 1.	46
Figure 13 Intensity normalization of time-lapse maximum projections.	48
Figure 14 User interface of the program <i>BugCube</i>	49
Figure 15 Multiview reconstruction of three <i>Tribolium</i> image datasets.	51
Figure 16 Renderings of a fused <i>Tribolium</i> image dataset.	52
Figure 17 Image quality of images after WAF and MVD.	53
Figure 18 Qualitative comparison of raw images, WAF and MVD.	54
Figure 19 Multiview reconstruction of a <i>Gryllus</i> image dataset.	55
Figure 20 Three-dimensional renderings of a reconstructed <i>Gryllus</i> embryo.	56
Figure 21 Illustration of the main steps of the quantitative characterization module.	57
Figure 22 User interface for cell nuclei segmentation.	58
Figure 23 User interface for cell nuclei post-processing.	59
Figure 24 Image quality of three-dimensional datasets.	60
Figure 25 Efficiency of local thresholding in the initial cell nuclei segmentation step.	62
Figure 26 Marker point detection with the Laplacian of Gaussian filter.	63
Figure 27 Marker-controlled watershed segmentation improves the separation of cell nuclei.	64
Figure 28 Evaluation of parameter values for initial segmentation and marker point detection.	65
Figure 29 Quality of the cell nuclei segmentation in different regions of a large T47D spheroid.	66
Figure 30 Surface approximation using alpha shapes.	68
Figure 31 Surface approximation and characterization of T47D spheroids.	69
Figure 32 Medium and large spheroids show differences in internal structure.	71
Figure 33 Examples of a Staurosporine-treated and untreated T47D spheroids.	72
Figure 34 Cell nuclei segmentation performance in T47D spheroids.	73
Figure 35 Comparison of T47D spheroid volume, cell numbers and shape for all conditions.	74

Figure 36 Cell nucleus volume in T47D spheroids for all conditions.	76
Figure 37 Single planes along x-y of an optically cleared and triple-stained T47D spheroid.	77
Figure 38 Cell nuclei segmentation performance in T47D spheroids.	78
Figure 39 Cellular density and nuclear volume are reduced following prolonged autophagy modulation.	79
Figure 40 Example of a constricted and a non-constricted (control) HC11 spheroid.	80
Figure 41 Cell nuclei segmentation performance in HC11 spheroids.	81
Figure 42 Quantitative results obtained for constricted and non-constricted HC11 spheroids.	83
Figure 43 Cell nucleus volume and cell density as a function of depth in HC11 spheroids.	84
Figure 44 Live recording of lateral root development from initiation to emergence.	85
Figure 45 Tracking and lineage tracing of cells in lateral root development.	87
Figure 46 Schematic illustration of the cellular disposition in lateral roots and assignment of cell files. ...	88
Figure 47 Classification of cell division types in lateral root development.	89
Figure 48 Growth and shape dynamics in lateral root development.	92
Figure 49 Contribution of cell files and cell lineages to the lateral root.	94
Figure 50 Spatio-temporal occurrence of the three cell division types anticlinal, periclinal and radial.	96
Figure 51 Alternation of cell division orientation in lateral root development.	98
Figure 52 Occurrence of cell layers in lateral root development.	99
Figure 53 Cell nuclei tracking in <i>Tribolium castaneum</i>	114
Figure 54 Future application scenarios of the quantitative characterization module.	115
Figure 55 A multiscale approach for image-based systems biology.	116
Supplementary Figure 1 Renderings of the surface approximation.	129
Supplementary Figure 2 Spatial localization of cell nuclei with a volume between 300 and 600 μm^3	129
Supplementary Figure 3 Plots of the mean cell density versus the normalized distance to the surface ...	130
Supplementary Figure 4 Local cell density in all datasets.	131
Supplementary Figure 5 Rendering of cell nuclei locations and approximated surface.	133
Supplementary Figure 6 Comparison of cell nucleus volume in T47D spheroids	133
Supplementary Figure 7 Alpha shape surface for selected T47D spheroids.	135
Supplementary Figure 8 Quantitative results for all four experiments.	136
Supplementary Figure 9 Quantitative results for all experiments.	137
Supplementary Figure 10 Lineage trees of lateral root founder cells in all datasets.	140
Supplementary Figure 11 Plot of the exponential growth profiles and fitting parameters.	141
Supplementary Figure 12 Plot of the volume as a function of cell number and fitting parameters.	142

List of Tables

Table 1 File name convention used by the automated data management module.	23
Table 2 The underlying data structure of <i>BugCube</i>	25
Table 3 Parameter values for multiview reconstruction of <i>Tribolium</i> and <i>Gryllus</i> datasets.	26
Table 4 Features extracted by the quantitative characterization module.	35
Table 5 Parameter values of the quantitative characterization module	36
Table 6 Performance of ZIP and JPEG2000 compression.	45
Table 7 Evaluation of segmentation performance for different regions within the spheroid.	67
Table 8 Excerpt of the file for dataset #120830 generated by the program <i>TrackGen</i>	86
Table 9 Summary of all analyzed lateral root datasets.	87
Table 10 Accuracy of the cell division classification for all cell division rounds.	90
Table 11 Growth rates, doubling times, height growth and volume growth for all lateral root datasets. ...	92
Supplementary Table 1 Runtime of the quantitative characterization module.	131
Supplementary Table 2 Quantitative measures for all datasets.	132
Supplementary Table 3 Number of T47D spheroids per condition and time point.	134
Supplementary Table 4 Volume of T47D spheroids	134
Supplementary Table 5 Number of cells in T47D spheroids.	134
Supplementary Table 6 Average diameter of T47D spheroids.	134
Supplementary Table 7 Average cell nucleus volume in T47D spheroids	135
Supplementary Table 8 Number of T47D spheroids per condition.	136
Supplementary Table 9 Quantitative results for pooled datasets.	137
Supplementary Table 10 Number of HC11 spheroids per condition.	137
Supplementary Table 11 Quantitative results for pooled datasets.	138
Supplementary Table 12 Contribution of cell files to the lateral root.	143
Supplementary Table 13 Number of anticlinal, periclinal and radial cell divisions.	143
Supplementary Table 14 First occurrence of anticlinal, periclinal and radial cell divisions.	143
Supplementary Table 15 Occurrence of cell layers.	144

List of Abbreviations

1)	CCD	charge-coupled device
2)	CLSM	confocal laser scanning microscopy
3)	CPU	central processing unit
4)	CUDA	compute unified device architecture
5)	DCG	Delaunay cell graph
6)	DNA	deoxyribonucleic acid
7)	DoG	difference of Gaussian
8)	DoM	difference of mean
9)	FOV	field of view
10)	FN	false negative
11)	FP	false positive
12)	GFP	green fluorescent protein
13)	GUI	general user interface
14)	GPU	graphics processing unit
15)	GT	ground truth
16)	HCS	high-content screening
17)	JPEG	joint photographic experts group
18)	LRP	lateral root primordium
19)	LSFM	light sheet-based fluorescence microscopy
20)	LZW	Lempel-Ziv-Welch algorithm
21)	mDSLIM	monolithic digital scanned laser light sheet-based fluorescence microscope
22)	mTOR	mammalian target of Rapamycin
23)	MVD	multiview deconvolution
24)	NA	numerical aperture
25)	NDS	normalized distance to the surface
26)	PCA	principal component analysis
27)	PCG	proximity cell graph
28)	PSF	point spread function
29)	RCP	random cell position
30)	sCMOS	scientific complementary metal-oxide-Semiconductor
31)	SNR	signal-to-noise ratio
32)	TIF	tagged image file
33)	TP	true positive
34)	WAF	weighted average fusion
35)	XLSX	Microsoft Excel workbook format
36)	YFP	yellow fluorescent protein

1 Introduction

Systems biology aims to understand a biological system in its entirety, its individual components and their relationship. In the last decades, classical systems biology has focused on the analysis of regulatory networks and biochemical pathways in cells, organs or organisms (Chuang et al., 2010). With advanced light microscopy, the structure and the dynamics of intact biological systems can be studied in high spatio-temporal resolution (Antony et al., 2013). Image-based systems biology exploits the information content of the acquired images, providing quantitative data and thereby establishes the connection between experimental and theoretical science (Figge and Murphy, 2015).

1.1 Workflow of image-based systems biology

Typically, image-based systems biology is composed of an experimental and an image analysis pipeline (Figure 1). In the experimental pipeline, an experiment is realized in order to address a scientific question about the biological system. The object of interest (specimen) is prepared for observation with a suitable microscope (Kherlopian et al., 2008). The second part encompasses a series of image analysis modules, which extract the relevant information from the acquired data. The following sections introduce the image acquisition systems and components of an image analysis pipeline.

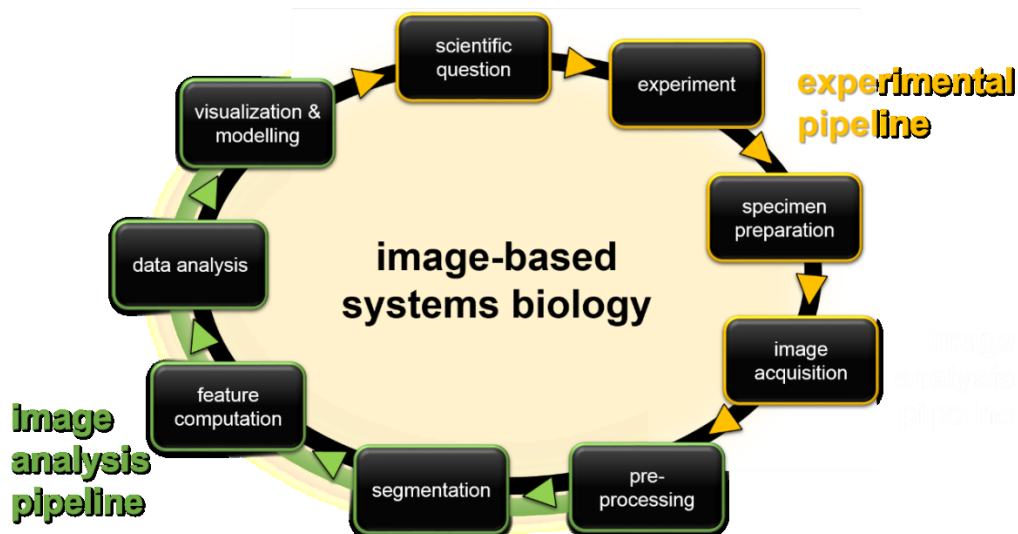


Figure 1 Workflow of an image-based systems biology approach. The workflow consists of an experimental and an image analysis pipeline. Segmentation and feature computation enable the extraction of static or dynamic object features made available for subsequent data analysis, visualization and

modelling. A reconsideration of the image analysis results is essential to optimize parameters of an experiment or plan the experiments for new scientific questions.

1.1.1 Image acquisition with advanced light microscopy

High quality images are an essential requirement to perform quantitative image analysis. Advanced systems such as confocal and light sheet-based fluorescence microscopy (LSFM) enable imaging intact specimen at the cellular level (Ansari et al., 2014; Verveer et al., 2007). The acquired images provide the spatio-temporal distribution of structures specifically labeled with fluorophores. The fundamental task of a fluorescence microscope is to illuminate a specimen at the excitation wavelength and to detect the fluorescence emitted by fluorophores. Filters block out the excitation light while passing the emission light to the detection system of the microscope.

Widefield and confocal fluorescence microscopy

In widefield fluorescence microscopy, the entire field of view (FOV) is illuminated and a camera simultaneously detects the emitted fluorescence. In confocal microscopy, the fluorescence light passes through a pinhole before it reaches the detector, which allows discriminating against out-of-focus light. Confocal laser scanning microscopy (CLSM) provides high lateral and axial resolution and have optical sectioning capability (Lichtman and Conchello, 2005). Widefield and confocal microscopes suffer from some limitations. First, the illumination path and the detection path are identical (epi-fluorescence). The same objective lens focuses the excitation light on the sample and the emitted fluorescence to the detection system. Illumination of the focal plane in the specimen therefore results in the illumination of the entire fluorophore distribution in the specimen above and below the focal plane (Figure 2a, b). Thus, fluorophores are several hundred times more often excited than actually observed by the detection system (Keller et al., 2008; Verveer et al., 2007). The number of detectable fluorophores is finite. Unless the observed specimen renews or reproduces the fluorophores, their consumption due to degradation results in the disappearance of the fluorescence signal (photobleaching). Second, the constant or repeated illumination poses a considerable energy load onto the specimen that can have severe consequences on live specimen. For example, the excitation light might lead to the degradation of endogenous organic substances (Wagner et al., 2010). Degraded substances are no longer available for the metabolism of the specimen. Moreover, the degradation often leads to the generation of reactive species that can be toxic to the specimen (phototoxicity). For example, light irradiation produces oxygen radicals inducing oxidative stress and DNA damage which impair cell viability and function (Frigault, Lacoste et al. 2009). Biological experiments rely on the observation of the specimen under physiological conditions without exerting a severe impact. Therefore, the amount of light energy that the specimen is exposed to during an experiment should be minimized (Stelzer, 2015). Third, the amount of light scattering and absorption limits the penetration depth into the specimen

(Reynaud et al., 2008). Consequently, fluorescence microscopes usually cannot penetrate more than several tens to hundreds of microns into the specimen (Ntziachristos, 2010).

Light sheet-based fluorescence microscopy (LSFM)

LSFM addresses all above-mentioned problems and has become a valuable technique for cell and developmental biology (Pampaloni et al., 2015; Strobl and Stelzer, 2016). In LSFM, the illumination and detection light paths are decoupled (Figure 2c). Two objective lenses are used in an azimuthal arrangement for illumination of the specimen and the detection of the emitted fluorescence (Stelzer and Lindek, 1994). A light sheet that overlaps with the focal plane of the detection system illuminates the specimen. Thus, only fluorophores in the focal plane get excited and the overlapping provides true optical sectioning (Greger et al., 2007).

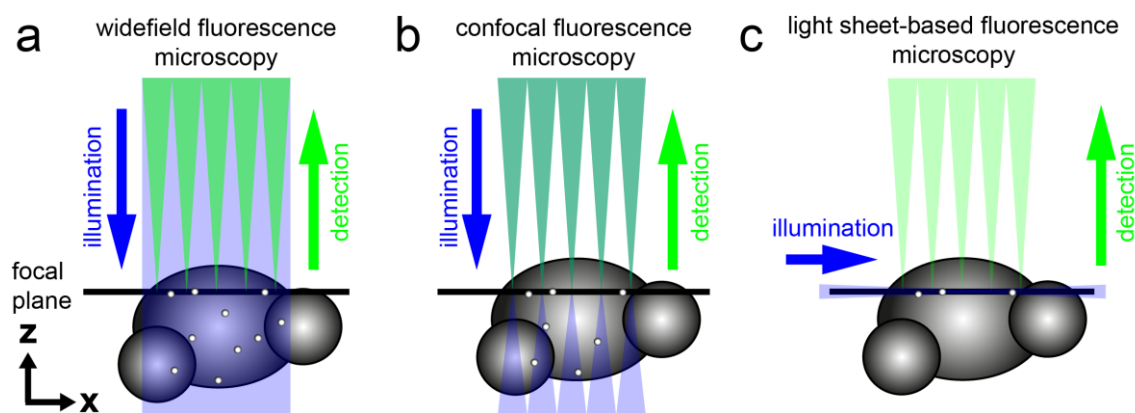


Figure 2 Illumination and detection principle in fluorescence microscopes. In widefield (a) and confocal (b) microscopes, the fluorophores in the entire specimen are illuminated. The detection of emitted fluorescence occurs along the same path as the illumination. In a light sheet-based fluorescence microscope (c), a sheet of light illuminates the specimen and overlaps with the focal plane. The emitted fluorescence is detected perpendicular to the illumination path.

The LSFM detects the emitted fluorescence in the FOV in parallel (Pampaloni et al., 2015). Modern scientific cameras such as Charge-Coupled Device (CCD) or scientific Complementary Metal-Oxide-Semiconductor (sCMOS) cameras allow for high acquisition speed and provide a 10 to 100-fold higher signal-to-noise ratio (SNR) than confocal microscopes (Keller et al., 2011). The massive parallelization provides high dynamic ranges of up to 14 bits because more photons can be detected simultaneously (Reynaud et al., 2008). An optical section of the specimen, i.e. a two-dimensional image with spatial dimension x and y , is obtained. High precision stages move the specimen along the optical axis (z axis) through the light sheet allowing for the acquisition of multiple optical sections that are combined into a stack of images (z stack or image stack). A z stack captures the three-dimensional fluorophore distribution in the entire specimen.

LSFM adopts several advantages of confocal microscopy such as the optical sectioning capability, a true axial resolution and combines them with a high SNR, minimal photobleaching and phototoxicity (Stelzer, 2015). The advantages of LSFM render it particularly useful for the

long-term observation of live specimen (Keller et al., 2008; Strobl et al., 2015). A powerful feature of LSFM is that specimen can be recorded along multiple view directions by rotating the specimen around the rotation axis (y axis). This feature provides complementary information and is particularly useful for imaging large, light scattering or light absorbing specimen (Huisken et al., 2004; Swoger et al., 2007). Optical clearing increases object transparency and thus the penetration depth into larger specimens (Dodt et al., 2007). In combination with optical clearing, LSFM enables imaging of objects even in the range of a millimeter (Keller and Dodt, 2012) and the systematic retrieval individual cell properties in their native microenvironment (Schmitz et al., 2017).

1.1.2 Components of an image analysis pipeline

Biological research relies on image datasets as an information source in order to understand molecular and cellular mechanisms. The main goal of image analysis is the conversion of image datasets into meaningful and interpretable data, ideally to simple numbers. Advanced light microscopes capture data about all three spatial dimensions, multiple time points, views, wavelengths or even biological specimens. This rapidly generates millions of images and Terabytes of data volume. The resulting big data from microscopy cannot be readily visualized and analyzed (Meijering et al., 2016). For example, a long-term multiview dataset capturing the embryogenesis of the red flour beetle *Tribolium castaneum* consists of 229 planes (1040×392 pixels) per image stack, eight view directions, 301 time points and one channel which results in 2408 image stacks (551432 single images) and a total of 1.5 Terabyte data volume (Strobl et al., 2015). This motivates the generation of automated image analysis pipelines that perform the processing steps in a high throughput manner (Antony et al., 2013). These requirements have led to the new field “bioimage informatics” in which researchers focus on the application-driven development and optimization of algorithms for processing biological image data (Eliceiri et al., 2012; Myers, 2012).

The ultimate goal is to replace human effort by automated pipelines and at the same time improving the objectivity of image analysis. An image analysis pipeline is a composition of processing steps that typically cover image pre-processing, segmentation, feature computation, data analysis, visualization and modeling (Figure 1). Although commonly illustrated as an unidirectional sequence, an image analysis pipeline needs to have a flexible and modularized architecture (Meijering et al., 2016). Depending on the application, not all modules are required and a feedback from one module guides the application of other modules.

Software packages such as *Mathematica* (Wolfram Research Inc.) or *Matlab* (MathWorks Inc.) offer powerful platforms for the implementation of image analysis pipelines. The platforms integrate algorithms from image analysis, graph theory, statistics and computational topology. Moreover, the platforms can be extended by specialized packages for image analysis such as the

Insight Segmentation and Registration Toolkit (ITK) (Yoo et al., 2002), the *Visualization Toolkit* (VTK) (Schroeder et al., 1996), *FJII* (Schindelin et al., 2012) and *R* (R Development Core Team, 2016).

Automated data management and image pre-processing

The acquired multidimensional image datasets require automated steps for data management, data reduction and image quality improvement. Cropping the image dataset to the region of interest (ROI), or image compression are essential steps to reduce the amount of data volume. Lossless compression methods such as ZIP compression allow a substantial reduction of the data amount. A promising alternative method is lossy image compression with the JPEG2000 standard (Bernas et al., 2006). In JPEG2000 compression, the image is first transformed into the wavelet frequency domain producing as many wavelet coefficients as pixels exist in the image. In a subsequent quantization step, the coefficients are compressed. Image pre-processing steps improve image quality and enhance features present in the images. Examples include the correction of inhomogeneous illumination, the reduction of image background or the recovery of decreased signal intensities, e.g. when thick specimens were acquired (Uchida, 2013). Image filtering (convolution) reduces image noise, increase image contrast, or prepares images for subsequent object detection steps (González and Woods, 2010). Additive noise and the influence of the point spread function (PSF) of the microscope degrade the resulting image quality. Deconvolution methods model the PSF as a convolution kernel and perform a reconstruction of the latent image (Wallace et al., 2001). Deconvolution reduces the amount of out-of-focus blur and background noise thereby increasing image contrast and the apparent image quality (Shaw, 2006).

Pre-processing methods for three-dimensional image reconstruction are often required when images of large specimen were acquired. For example, large specimens that do not fit in a single FOV of the microscope can be acquired with tiled image stacks. The reconstruction of tiled image stacks (stitching) involves the registration of the tiles and subsequent image fusion to obtain a single image stack capturing the complete specimen. Efficient reconstruction of tiled image stacks into a single three-dimensional image is commonly achieved using Fourier transform-based approaches (Preibisch et al., 2009).

In LSFM, large specimen can also be acquired from multiple view directions. Such multiview image datasets are reconstructed by image registration and fusion (multiview reconstruction), yielding image datasets with almost isotropic resolution (Swoger et al., 2007). Image registration and fusion methods include intensity-based (Huisken et al., 2004; Preibisch et al., 2008; Swoger et al., 2007), segmentation-based (Keller et al., 2008), wavelet-based (Rubio-Guivernau et al., 2012) or bead-based (Preibisch et al., 2010) approaches. In the bead-based approach, the specimen is prepared in conjunction with fluorescent beads. The beads are then

detected in the acquired images and serve as landmarks for subsequent image registration. Bead-based approaches are accurate, fast, independent of the sample and allow for automatic evaluation of the registration quality (Preibisch et al., 2010). Recently, the bead-based approach has been implemented in conjunction with multiview deconvolution (Preibisch et al., 2014). The software implements an iterative Richardson Lucy deconvolution procedure performing a maximum-likelihood estimation of the underlying image.

Segmentation and feature computation exemplified for cell nuclei

After pre-processing, an image analysis pipeline requires the extraction of objects from the image (segmentation) and the extraction of quantitative and biologically relevant features (feature computation). Features include intensity-related (e.g. the mean intensity and maximum intensity), texture-related (Depeursinge et al., 2014) and shape-related (Pincus and Theriot, 2007) properties of cells that allow for classification with pattern recognition techniques (Shamir et al., 2010; Uchida, 2013). Cell nuclei are the appropriate reference for the spatial position of a cell. They further provide valuable information about the health state of a cell, which is especially important in diagnostic and drug development approaches (Zink et al., 2004). Three-dimensional microscopy typically provides image stacks with fluorescently labelled cell nuclei. The image stacks thus provide a resource for the quantitative characterization of individual cells and their environment (Bilgin et al., 2013). An accurate segmentation of cell nuclei is an essential starting point for a quantitative characterization. The challenges of cell nuclei segmentation are due to the complexity and variability of nuclear appearance across images. Low image quality (i.e. low signal to noise ratio) complicates the distinction between cell nuclei and background noise. Imperfect staining or intrinsic cell characteristics lead to fluctuations of the signal intensity, which commonly causes over- or under-segmentation. Fluctuations in the background increase the number of false positive detections of cell nuclei.

The two main properties used for segmentation of cell nuclei use similarity (region-based methods) or discontinuity (boundary-based methods) of pixel intensities. The high variability of cell nuclei staining and volume in spheroids impedes accurate cell nuclei segmentation using simple methods such as global thresholding. Furthermore, imaging large samples *in toto* limits the achievable lateral and axial resolution. In combination with high cell density, this results in a considerable amount of apparently touching cell nuclei in the images that are difficult to separate (Mathew et al., 2015). A variety of powerful cell nuclei segmentation methods including iterative thresholding (Keller et al., 2008), level sets (Bergeest and Rohr, 2012), graph cut (Qi, 2014), gradient flow tracking (Li et al., 2007), lines-of-sight (Mathew et al., 2015) or watershed methods (Cheng and Rajapakse, 2009) address this problem. The regions of interest that result from the watershed segmentation can be constrained by marker points that guide the

watershed segmentation procedure. Exemplary marker point detection algorithms are Euclidean distance-maps (Cheng and Rajapakse, 2009), the shrinkage of level sets, partial differential equations (Mikula et al., 2011) or Laplacian of Gaussian (LoG) filtering (Lindeberg, 1998; Lowe, 2004).

Towards a multiscale quantitative characterization

Features of cell nuclei need to be integrated with computed features from other fields including computational topology, graph theory and spatial statistics to obtain a multiscale quantitative characterization a biological system (Schmitz et al., 2017). In computational topology, efficient algorithms exist for the approximation of surface, volume or shape of objects (Edelsbrunner and Harer, 2010; Edelsbrunner and Mücke, 1994). Alpha shapes are a generalization of the convex hull that can be used to obtain geometrical models of cell networks (Edelsbrunner et al., 1983). Cell graphs consist of vertices representing the cell nuclei and edges representing the neighborhood relationship between any two cells. They have been used to analyze the two-dimensional spatial cell network in breast tissue (Bilgin et al., 2007), malignant glioma (Demir et al., 2005) or Hodgkin's lymphoma (Schäfer et al., 2015). This concept can also be translated to analyze and describe three-dimensional cell networks.

Tracking and lineage reconstruction of cells

The acquisition of biological systems with high spatial and temporal resolution allows studying dynamics in a spatio-temporal context. Methods for the identification of objects over time (tracking) and relating different objects with each other (lineage reconstruction) are required (Meijering et al., 2009). Previous studies employed the advantages of LSFM to study cell migration during embryogenesis of zebrafish (Keller et al., 2008), chicken (Rozbicki et al., 2015), mouse (Ichikawa et al., 2013), or *Drosophila melanogaster* (Amat et al., 2014). The approach enables the construction of digital cell atlases (i.e. virtual embryos) of embryogenesis (Khairy and Keller, 2011). Virtual embryos generated for different model organisms allow a comparative analysis of developmental processes (Kalinka and Tomancak, 2012).

Comprehensive data analysis, visualization and mathematical modelling

A substantial portion of the image analysis pipeline resides in the analysis, visualization and mathematical modelling of the retrieved quantitative information. Data analysis includes method for data mining, statistical testing, classification of identified objects or machine learning in order to identify patterns of the underlying system (Chessel, 2017). Many experiments require programs that allow for inspection, correction and annotation of computed quantitative results. For example, results from automated segmentation or tracking require tools for interactive visualization (Wait et al., 2014; Walter et al., 2010). The ultimate goal of image analysis is to provide Mathematical models that describe and generalize the underlying

biological system (Buck et al., 2012), predicts the outcome of alterations made to the system and therefore helps to develop new scientific questions.

1.2 Biological model systems and applications

In this thesis, I developed several modules for an image analysis pipeline. The development was driven by applications with different biological model systems. The following sections provide biological and motivational background information on the model systems and the applications in which the image analysis modules were used.

1.2.1 *Tribolium castaneum* and *Gryllus bimaculatus* embryos

Insects represent excellent model organisms for developmental biology. Larger and more complex insect model organisms such as the red flour beetle *Tribolium castaneum* (Strobl and Stelzer, 2014), the Mediterranean fruit fly *Ceratitidis capitata*, or the field cricket *Gryllus bimaculatus* (Nakamura et al., 2010) have emerged as valuable model systems to study embryonic events of insect development. *Tribolium* is a holometabolous species that passes through the developmental stages of egg, larva, pupa and adult. The embryogenesis of *Tribolium* shows distinct characteristics to that of other insect models such as *Drosophila*. For example, the embryogenesis exhibits short-germ development, embryonic leg development, extensive extra-embryonic membrane formation and non-involuted head development (summarized in Strobl and Stelzer, 2014). *Gryllus* is a hemimetabolous organism that passes through the developmental stages egg, nymph and adult.

Tribolium and *Gryllus* are several times larger than *Drosophila* or *Ceratitidis* (Figure 3). This requires the adaption of sample preparation, image acquisition and subsequent image processing. Image acquisition suffers from three main limitations. First, biological tissue scatters and absorbs light and thus the penetration depth into the specimen is limited. Second, only a fraction of the embryo is observable. Large specimen require the acquisition of image stacks along multiple views in order to obtain a full three-dimensional image. Third, in order to achieve high spatial resolution, the specimen is acquired with high magnification objectives. *In toto* imaging of *Tribolium*, *Ceratitidis* and *Drosophila* embryos within a single camera field of view is feasible. However, the *Gryllus* embryo has approximately 80 times the volume of a *Drosophila* embryo and requires the acquisition of tiled image stacks that in conjunction capture the entire embryo.

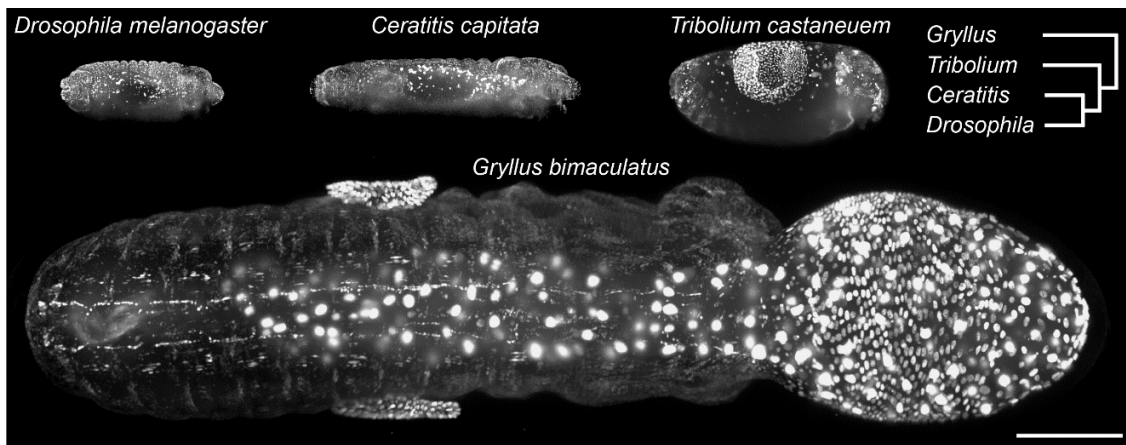


Figure 3 Comparative analysis of developmental processes during insect embryogenesis. Maximum projections along z of four different insect embryos shown in dorsal view during dorsal closure. All embryos express a nuclear localized or Histone-linked green fluorescent protein (GFP) subtype under control of a ubiquitous promoter. The box in the upper right corner provides the phylogenetic relationship. Microscope: images of *Drosophila*, *Ceratitis* and *Tribolium* were acquired with the mDSLM, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 10x NA 0.3 N-Achroplan, camera: Andor Clara. The *Gryllus* dataset was acquired with tiled image stacks (overlap 40%) with the LightSheet Z.1, illumination objective: 10x NA0.2, detection objective: 20x NA 1.0, camera: Sony ICX 285. Stitching was performed with the *Image Stitching* software. Scale bar: 300 μ m. Detailed information about the datasets and the software for stitching can be found in the Material and Methods sections. Figure adapted from Strobl *et al.*, 2017.

The long-term acquisition of such specimens produces Terabytes of complex image datasets that cannot be readily analyzed. Automated processing modules are required to manage, reduce and visualize the datasets. Furthermore, the tiled and multiview image datasets need to be fused by computational methods to obtain a single and complete image of the specimens at each time point.

1.2.2 Three-dimensional cell cultures as representative *in vitro* models for tissues

Experiments in cell biology rely on the assumption that cell cultures reflect the physiology of the tissue. Two-dimensional *in vitro* cell cultures fail to represent many tissue-specific architecture aspects such as mechanical strain, cell-cell interactions, or nutrient flow. Three-dimensional cell cultures reflect the cellular arrangement and interactions of real tissues (Pampaloni *et al.*, 2007). Spheroids are model systems of low complexity that allow for adequate and reproducible characterization (Sutherland, 1988). The aggregation and proliferation of cells leads to a ball of cells. A concentric cell layering consisting of a necrotic core surrounded by layers of quiescent and proliferative cells is observed for spheroids with a diameter of more than 400-500 μ m (Kunz-Schughart *et al.*, 2004). In addition, cells in spheroids develop characteristic properties of their ancestral tissue. Examples include aggregates of mouse embryonic stem cells (van den Brink *et al.*, 2014) or beating cardiomyocyte spheroids (Bartholomä *et al.*, 2005). Spheroids have high potential for fundamental research questions on cell differentiation, cancer biology and drug response (Ravi *et al.*, 2015).

Studies in cell biology rely on the retrieval of the characteristics of individual cells in spheroids. Image-based characterization of spheroids and spheroid growth has high potential for three-dimensional cell biology research (Jagiella et al., 2016). Most attempts to retrieve these properties have relied on spheroid disintegration, histological sectioning, or the use of rather small spheroids that for example do not show a concentric layering. A histological section provides insight into the internal morphology of a spheroid (Jagiella et al., 2016). A drawback is that histological sectioning destroys the integrity of the spheroid. The combination of LSFM and optical clearing allows to produce endpoint measurements of entire spheroids and has been used to study drug induced cell death in different regions (Wenzel et al., 2014).

Investigating the impact of different stress conditions on spheroids

In the context of three-dimensional cell cultures, the focus in this thesis was set on the investigation of different stress conditions. These were realized in four applications that analyze the morphological impact on spheroids:

- Application I: nutrient and oxygen deficiency
- Application II: apoptosis induction
- Application III: autophagy induction
- Application IV: acute mechanical constriction

Application I: nutrient and oxygen deficiency

Distinct microenvironmental conditions (e.g. nutrient and oxygen supply conditions) in solid tumors lead to the formation of different zones comprising proliferating, quiescent and necrotic cells (Thoma et al., 2014). Similarly, spheroids have a diffusion limit of about 150-200 μm to many molecules including oxygen (Lin and Chang, 2008). Thus, spheroids with diameters of more than 400-500 μm display a concentric cell layering, in which a necrotic core is surrounded by a layer of quiescent cells and an outer rim of proliferating cells (Kunz-Schughart et al., 2004). This zonation is critical for anti-cancer therapeutics. For example, the zonation has a huge impact on therapeutic response (Andre et al., 2010; Kim et al., 2011; Shekhar, 2011). Wenzel *et al.* showed that the therapy of T47D spheroids with cytostatic drugs (e.g. Paclitaxel, Cisplatin) mainly affects outer regions of proliferating cells, whereas the inner core region of dormant cells shows no effect (Wenzel et al., 2014). Drug penetration models that incorporate the effects of local drug, oxygen and nutrient concentration have confirmed these results. For example, it has been shown that a cellular barrier effect alone can lead to poor response to drugs (Frieboes et al., 2009).

Small spheroids of less than 200 μm in diameter lack the prominent concentric layering and central necrosis. The information gain from small spheroids, e.g. for evaluating drug effects on specific sub-populations of cells within spheroids, is thus limited. This is underpinned by

mathematical models of spheroids suggesting that changes in the state of each individual cell (Powathil et al., 2014) and the three-dimensional architecture (Kempf et al., 2013) must be monitored when studying therapeutic efficacy. Thus, the retrieval of morphometric cell nuclei features in intact, differently-sized spheroids is an essential step in understanding tumor development (Bredel-Geissler et al., 1992).

Application II: apoptosis induction

High content screening (HCS) combines high-throughput imaging and automated quantitative characterization (Antony et al., 2013). HCS has become an integral part of drug discovery and drug evaluation (Zanella et al., 2010). The main goal of HCS assays is to collect quantitative data to evaluate the properties and effectiveness of different drugs (e.g. in cancer therapy) in one approach. Modern HCS are performed with three-dimensional cell cultures (Friedrich et al., 2009). In contrast to biochemical assays, image analysis is not limited to population averages and provides information at the single cell level (Antony et al., 2013).

Traditional HCS approaches have relied on image analysis of individual sectional planes or maximum projection images (Reid et al., 2014; Vinci et al., 2012). These approaches only provide an average of the biological response of all analyzed cells and do not represent the variance of cellular responses (Levsky and Singer, 2003). However, the exposure of cells to different microenvironments and stresses leads to heterogeneous cell populations in tissues. Proliferation occurs mainly in well-nourished outer regions, while cells in nutrient deprived and hypoxic regions remain in a dormant state (Kyle et al., 2012). Consequently, cytostatic therapy with compounds strongly affect cells in the outer spheroid regions whereas the effect is lower on dormant cells in the quiescent and core regions (Thoma et al., 2014). It is therefore important that HCS approaches comprise quantitative methods to consider drug responses on a cellular level.

Caspases are central components that act as intracellular proteases and mediate apoptosis in cells. Different signaling pathways control the activation of caspases. However, the molecular mechanisms of drug-mediated activation of caspases are not well understood (McIlwain et al., 2013). Apoptosis-inducing drugs have been evaluated for their efficiency against breast cancer in previous studies (Imamura et al., 2015; Liedtke et al., 2010; Reid et al., 2014; Wang et al., 2013). Staurosporine acts as a competitive inhibitor of protein kinases and prevents ATP from binding to the kinase (Tanramluk et al., 2009). Staurosporine has also been shown to induce apoptosis in two-dimensional cell cultures of T47D breast cancer cells (Mooney et al., 2002).

The morphology of cell nuclei has been shown to correlate with cell proliferation activity (Smolle et al., 1989) and have shown diagnostic value in oncology (Saito et al., 2016). Similarly, cells that undergo apoptosis present features that are readily recognizable using

fluorescence microscopy and quantitative analysis. During apoptosis, organelles and membranes are conserved for some time, while the nucleus undergoes early degeneration and shrinks due to DNA loss (Bacus et al., 1989). Consequently, morphological features of the cell nucleus are indicators of apoptosis. For example, Mooney and coworkers have determined the apoptosis induction in T47D cells using a DNA fragmentation assay and have confirmed the effect by the changes in nuclear morphology using fluorescent microscopy (Mooney et al., 2002). Three-dimensional datasets with cellular resolution of *in toto* acquired cell cultures had not been available in these studies and thus, the effect of apoptosis induction on the single cell level could not have been investigated.

Application III: autophagy induction

Autophagy is a catabolic process involved in the maintenance of cellular homeostasis by the degradation and recycling of cellular components. Under nutrient-rich conditions, autophagy represents a way to recycle damaged organelles, particularly in quiescent and terminally differentiated cells (Rabinowitz and White, 2010). During stress conditions (e.g. starvation), the autophagy-mediated breakdown of cellular components promotes cell survival. For example, the ability of lymphocytes to consume environmental nutrients depends on the availability of growth factors. In the absence of growth factors, nutrient supply is maintained by autophagy and cells shrink to about 50% of the size (Lum et al., 2005).

The mammalian target of Rapamycin (mTOR) signaling is a key regulator of autophagy. It controls cell growth, proliferation, survival and metabolism in response to extracellular signals such as nutrient availability and growth factors (Efeyan and Sabatini, 2010). Under nutrient-rich conditions, autophagy is inhibited by the activity of the mTOR signaling pathway (Laplante and Sabatini, 2012). Rapamycin is an approved drug that interacts with and inhibits the activity of the mTOR signaling pathway, thus inducing autophagy (Ravikumar et al., 2004).

Defects in autophagy regulation (e.g. deregulation of the mTOR signaling pathway) are involved in the development of various diseases, including cancers (Aredia et al., 2012). However, the role of autophagy in cancer progression is controversial. On one hand, it acts tumor suppressive due to its function as a recycling mechanism of damaged cellular components. In contrast, autophagy can be tumor promoting in established tumors, where autophagy promotes the survival of tumor cells under metabolic and therapeutic stress (Degenhardt et al., 2006; Guo et al., 2013).

Breast cancer is one of the most common malignancies in women worldwide. Although mortality rates are decreasing due to improved medication and early detection, many aspects of breast cancer progression are still unknown. Previous studies indicate an upregulation of the mTOR signaling pathway in breast cancers (Kerekatte et al., 1995) and thereby decreases levels

of autophagy. Therefore, mTOR inhibitors such as Rapamycin are investigated as alternative cell death pathways for anti-tumor therapy (Schleicher et al., 2010). The development of successful autophagy-modulating therapies (e.g. using Rapamycin) against cancer requires a solid understanding of the dual role of autophagy in normal tissues and tumors (Seto, 2012).

Application IV: acute mechanical constriction

The behavior of cells in a cellular network such as healthy and tumor tissue are affected by cell density (Böttger et al., 2015), contact inhibition or mechanical pressure (Cheng et al., 2009; Drasdo and Höhme, 2005; Jagiella et al., 2016; Montel et al., 2011). For example, mechanical compression of tumor tissue drives cancer cells towards to develop an invasive phenotype (Tse et al., 2012). Spheroids provide a suitable model to study the effects of different stress conditions on tissue. Several studies investigated the role of mechanical stress on spheroid morphology, cell proliferation and apoptosis in the context of tumor biology. Compressive stress during long-term growth of spheroids led to reduced cell proliferation (Montel et al., 2011) and induction of apoptosis (Cheng et al., 2009). The impact of compressive stress was found to be dependent on the duration of the stress (Delarue et al., 2014). Reversible inhibition of proliferation was observed on the scale of several hours, whereas several days of applied compressive stress induced mitotic arrest of cells (Desmaison et al., 2013). Compressive stress applied to spheroids generated from mouse colon carcinoma cells increased the cell density two-fold, affected the cellular organization and restricted cell proliferation to the outer rim of the spheroid. More dead cells were observed in the center of the spheroid (Alessandri et al., 2013). These results are supported by mathematical models of compressed tumor growth (Ciarletta et al., 2013).

Mechanical stress also plays a fundamental role healthy tissue. The mammary gland is a highly dynamic organ, which undergoes massive modifications during development and postnatal phases. For example, the tissue completely remodels in the involution phase. During involution, the mammary gland morphology is remodeled into the pre-pregnant state. Mechanical cues are expected to play an important role during involution. For example, it is assumed that milk stasis induces mechanical stress on alveolar cells thus resulting in stress sensing mechanisms that trigger involution (Quaglino et al., 2009). Studying the effects of induced constriction on individual cells in spheroids provides valuable knowledge to understand complex processes such as involution. However, existing work has focused on confined growth of spheroids in the context of tumor biology and little is known on induced constriction of tissue.

1.2.3 Lateral root development in *Arabidopsis thaliana*

Shape and function of organs is controlled by the organization and behavior of cells. Plant cells are turgid and surrounded by a cell wall that restricts their mobility. Hence, cell growth and cell division predominantly determine shape and function of plant organs. Stereotypic and

deterministic processes control the early plant embryogenesis, in which almost invariant patterns of cell divisions lead to the specification of the apical-basal axis as well as shoot and root meristems (De Smet et al., 2010; Yoshida et al., 2014).

Plants constantly adapt their root systems to varying environmental conditions such as the availability of water, nutrients or light. This ability is commonly referred to as phenotypic plasticity (Bradshaw, 2006) and encompasses environmentally induced changes of morphology, physiology or behavior. Consequently, the architecture of the plant root and shoot system is determined by post-embryonic organ formation. A well-established model system to study post-embryonic organ formation is the root system of *Arabidopsis thaliana*.

Meristematic cells in the apical meristem of the primary root continuously generate new cells that are arranged in cell files. These cells elongate and form concentric tissue layers that give rise to the root vasculature, pericycle, cortex, endodermis and epidermis (Dolan et al., 1993) (Figure 4a). Some cells get primed to form lateral organs. In response to local accumulation of auxin and oscillating gene expression, small founder cells situated in the pericycle cell layer adjacent to the vasculature in the center of the primary root initiate the formation of a lateral root primordium (LRP) (Van Norman et al., 2013). The lateral root founder cells re-enter the cell cycle and divide anticlinal (parallel to the shoot-root axis) in an asymmetric fashion and later periclinal (normal to the shoot-root axis) (De Smet, 2012). These formative cell divisions create a dome-shaped LRP that grows through the overlaying cell layers to emerge from the primary root (Figure 4b). The overlaying tissues accommodate the passage (Vermeer et al., 2014) and have a major influence on the shape of the LRP (Lucas et al., 2013; Vermeer et al., 2014).

Traditionally, lateral root development has been subdivided into stages according to the number of generated cell layers (Malamy and Benfey, 1997) (Figure 4c). This staging system relied on two-dimensional images of fixed and optically cleared primary roots. Expression markers, specific for a certain type of tissue, display the type and lineage of each cell. Such attempts at cell lineage reconstruction lacked the spatio-temporal resolution to infer cell lineages throughout the complete development of LRPs (Kurup et al., 2005; Lucas et al., 2013). Moreover, it has been impossible to investigate whether stereotypic and deterministic patterns similar to those identified in plant embryogenesis exist in the post-embryonic formation of organs. To investigate these patterns, tools for computer-aided cell tracking and lineage reconstruction are required.

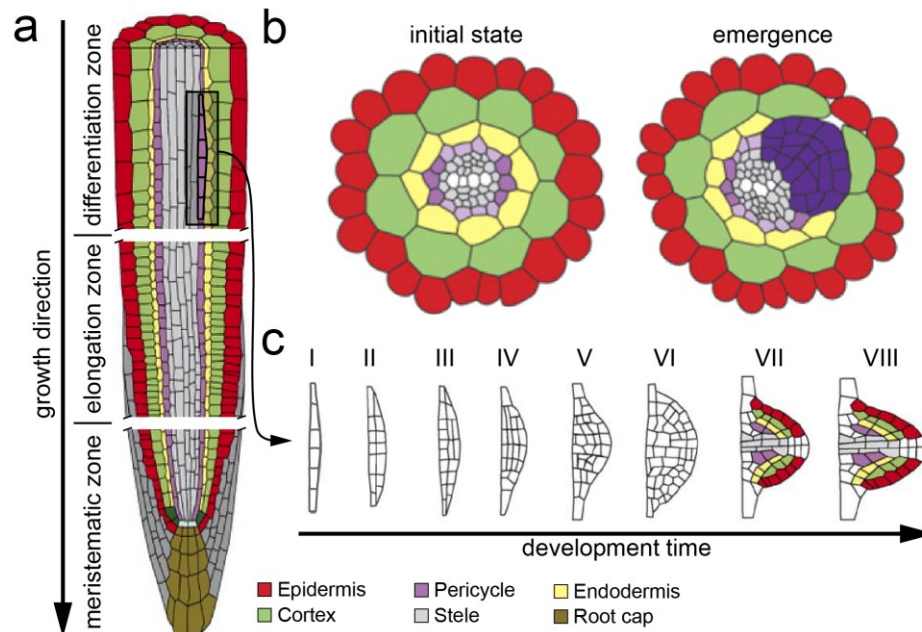


Figure 4 Lateral root development in *Arabidopsis thaliana*. (a) Longitudinal cut through the primary root. Lateral roots originate from cells in the pericycle cell layer. The priming of founder cells occurs in the elongation zone, whereas lateral root development and emergence occur in the differentiation zone of the primary root. (b) Cross sections of the primary root in the initial state (left) and during emergence (right) of a lateral root. (c) Lateral root formation is classically subdivided into eight stages (I-VIII). Panels a and b were adapted from Péret, De Rybel, *et al.*, 2009, panel c was adapted from Péret, Larrieu, *et al.*, 2009.

1.3 Scope of this thesis

The aim of this thesis was the identification, development and optimization of modules for an image analysis pipeline in three-dimensional fluorescence microscopy. The modules cover the following aspects:

- Automated data management as a prerequisite to structure, reduce and visualize multidimensional image datasets
- Normalization of intensity variations in time series maximum projections
- Web-based visualization and sharing of image datasets
- Reconstruction of multiview image datasets including bead-based registration, fusion and deconvolution
- Multiscale quantitative characterization including automated cell nuclei segmentation, feature computation, computational topology and cell graphs
- Computer-aided cell tracking and lineage reconstruction

In order to ensure the applicability to real image datasets from fluorescence microscopy, the modules were developed in an application-driven approach. The image datasets were obtained from research projects in developmental biology of insects and plants, as well as from cell biology.

2 Material

2.1 Tribolium datasets

2.1.1 Preparation of Tribolium embryos

Tribolium embryos were prepared for imaging as described previously (Strobl et al., 2015). Here, Tribolium embryos of the EFA-nGFP transgenic line expressing nuclear-localized GFP (Sarrazin et al., 2012) were fixed on an agarose hemisphere (1% (wt/vol) low-melt agarose). This ensured that the embryo remained in the same location throughout the entire long-term acquisition. To allow for bead-based registration of the datasets, fluorescent beads of 1.0 μm diameter were embedded into the agarose hemisphere (Figure 5).

2.1.2 Long-term fluorescence live imaging of Tribolium embryos with LSFM

Three time-lapse datasets of Tribolium embryos were acquired (dataset 1, 2 and 3) with the monolithic digital scanned laser light sheet-based fluorescence microscope (mDSLm), using an axial pitch of 2.58 μm were acquired at an interval of 30 minutes for 150 hours from eight view directions at each time step. The mDSLm was equipped with a 2.5x NA 0.06 EC Epiplan-Neofluar illumination objective, a 10x NA 0.3 W N-Achroplan detection objective, a CCD camera (Andor Clara, 1394×1040 pixels, pixel pitch: 0.645 μm). Each of the acquired datasets consisted of 2,408 image stacks with a total raw data volume of up to 1.5 Terabyte. The datasets further served as test datasets to evaluate the performance of ZIP and JPEG2000 compression. Embryo preparation and imaging were performed by Frederic Strobl.

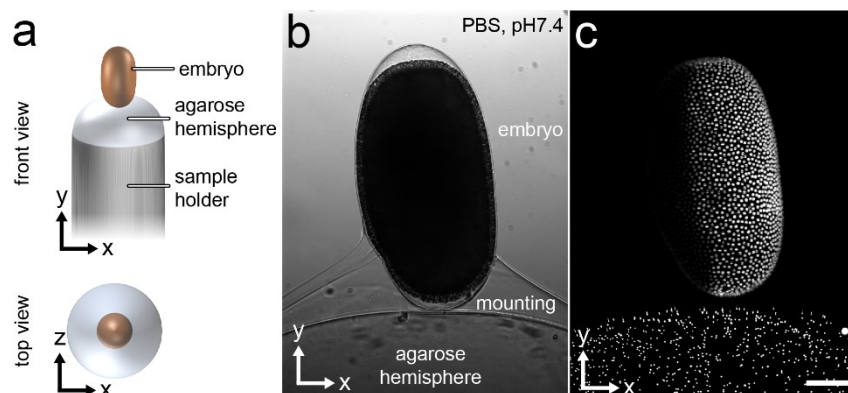


Figure 5 Mounting principle for Tribolium embryos. (a) Illustration of the Tribolium embryo mounting principle in front and top view. A drop of liquid agarose is placed on top of the tip of the sample holder. The embryo is then fixed on top of the created agarose half sphere. Not that for illustration purposes, the size of the embryo was increased. (b) Transmission light image of a Tribolium embryo (dataset 1) fixed on the agarose half sphere (dataset DS0001). (c) Fluorescence light image of the embryo shown in panel b. Fluorescent beads in the agarose hemisphere are visible. Microscope: mDSLm,

illumination objective: CZ 5x/NA 0.16, detection objective: CZ 10x/NA 0.50, camera: Andor Clara. Scale bar: 100 μm . Images in panels b and c were provided by Frederic Strobl.

2.2 Gryllus datasets

2.2.1 Preparation of Gryllus embryos

Culturing, preparation and imaging of Gryllus embryos was performed by the former Master student Alexander Ross. Gryllus embryos of the H2B-eGFP transgenic line were cultured and prepared according to the Master thesis of Alexander Ross. Briefly, a metal holder with an opening at the top (“Cobweb holder”) was filled with 1% (wt/vol) low-melt agarose and soaked up until a thin film of agarose remained in the opening. The Gryllus embryo was mounted onto the agarose film and covered with a second film of agarose (Figure 6). This ensured that the embryo remained in the same location throughout image acquisition. To allow for bead-based registration of the datasets, fluorescent beads of 1.0 μm diameter were embedded into the agarose film.

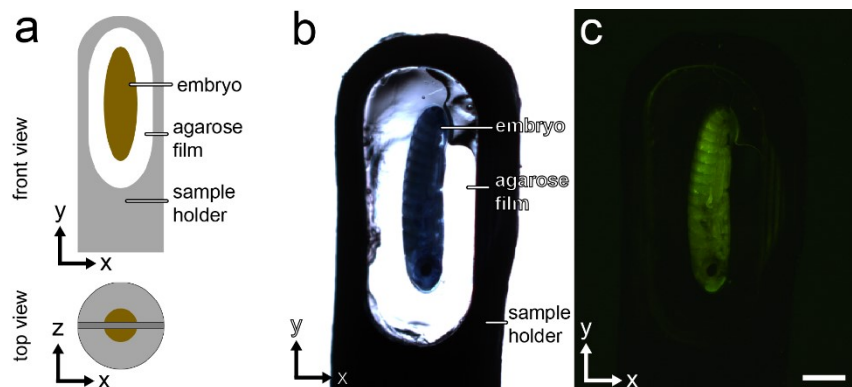


Figure 6 Mounting principle for Gryllus embryos. (a) Schematic of the Gryllus embryo mounting principle in front and top view. A drop of liquid agarose is placed on the opening of the sample holder generating a thin agarose film. The embryo is placed on the agarose film and fixed with another drop of agarose. Note that for illustration purposes, the size of the embryo was increased. (b) Transmission light (b) and fluorescence (c) image (GFP filter set) of sample holder with a mounted Gryllus embryo under the stereo microscope (Zeiss Discovery.V8). Scale bar: 600 μm . Images in panels b and c were provided by Frederic Strobl.

2.2.2 Long-term fluorescence live imaging of Gryllus embryos with LSM

Image stacks were acquired with the Zeiss LightSheet Z.1 fluorescence microscope, using an axial pitch of 4 μm were acquired at an interval of 30 minutes for up to 143 hours 30 minutes from four view directions (0° , 90° , 180° and 270°) at each time step. The Z.1 was equipped with a 10 \times NA 0.2 objective for excitation and a 20 \times NA 1.0 detection objective, a Sony ICX 285 CCD camera (1036 \times 1388 pixels, pixel pitch: 1 μm). To capture the complete embryo at 20 \times magnification, four tiles were acquired along the y axis with an overlap of 40%. The dataset used in this thesis consists 4592 image stacks, with 254 single planes per stack. In total, 287 time points, four view directions and four tiles per time point and view direction were acquired resulting in about 3.2 Terabyte data volume.

2.3 Spheroid datasets

2.3.1 Cell culture and spheroid preparation

The T47D (human ductal breast epithelial tumor) cell line were purchased from ATCC (ATCC-HTB-133) cultured in RPMI1640 (Gibco) supplemented with 10% FBS (Gibco, Lot: 10270 and Lot: 41Q6640K) and 2 mM L-glutamine (Roth) (Gibco) at 37 °C, with 5% CO₂.

HC11 cells (murine non-tumorigenic mammary epithelial cell line) were a gift from B. Groner at the Georg Speyer Haus in Frankfurt am Main. Cells were cultured in RPMI1640 (Gibco) supplemented with 10% FBS (Gibco, Lot: 41Q6640K), 2 mM L-glutamine (Roth) (Gibco), 5 µg/ml human recombinant insulin (Sigma) and 10 ng/ml murine EGF (Peprotech) at 37°C, with 5% CO₂.

Spheroids were formed by liquid overlay. Cell suspensions containing 500, 1,000, 2,000, 5,000 or 10,000 were placed in U-well plates coated with a non-adhesive layer (Thermo Scientific, Nunc HydroCell 96 Well, 174908) or in agarose-coated 96 well plates. To obtain compact spheroids, each plate was centrifuged and cultured for the indicated time at 37°C und 5% CO₂ in the incubator.

2.3.2 Drug Treatment of T47D spheroids

Application II. After spheroid formation, the cells were treated with 6.7 µM Staurosporine (Sigma Aldrich) diluted in 0.1% DMSO or 0.1% DMSO for further 3 days.

Application III. After spheroid formation for 3 days, the spheroids were treated with Rapamycin (Sigma Aldrich) or DMSO for further 7 days. The medium containing Rapamycin was refreshed every three days.

2.3.3 Cylindrical constriction of HC11 spheroids

Mechanical perturbation of HC11 spheroids was achieved by placing spheroids into agarose capillaries with a diameter of 200 µm (Figure 7a). In a control experiment, HC11 spheroids were placed into agarose capillaries with a diameter of 300 µm. To generate agarose capillaries, 2% low-melt agarose in PBS was injected to fill a red color-coded capillary (Blaubrand – intraMark, Brand). Subsequently, a glass rod (Hilgenberg) with the diameter of either 200 µm or 300 µm was inserted into the agarose (Figure 7b). After solidifying, the glass rod was removed. The agarose column was then placed into a petri dish and cut into approximately 1 cm long pieces. Spheroids were pipetted on the opening of the agarose tubes (Figure 7c, panels 1 and 2) and were sucked up by capillary forces (Figure 7c, panel 3). The spheroid-containing agarose tube (Figure 7d) was placed into fresh growth medium and incubated for either 24 or 96 hours.

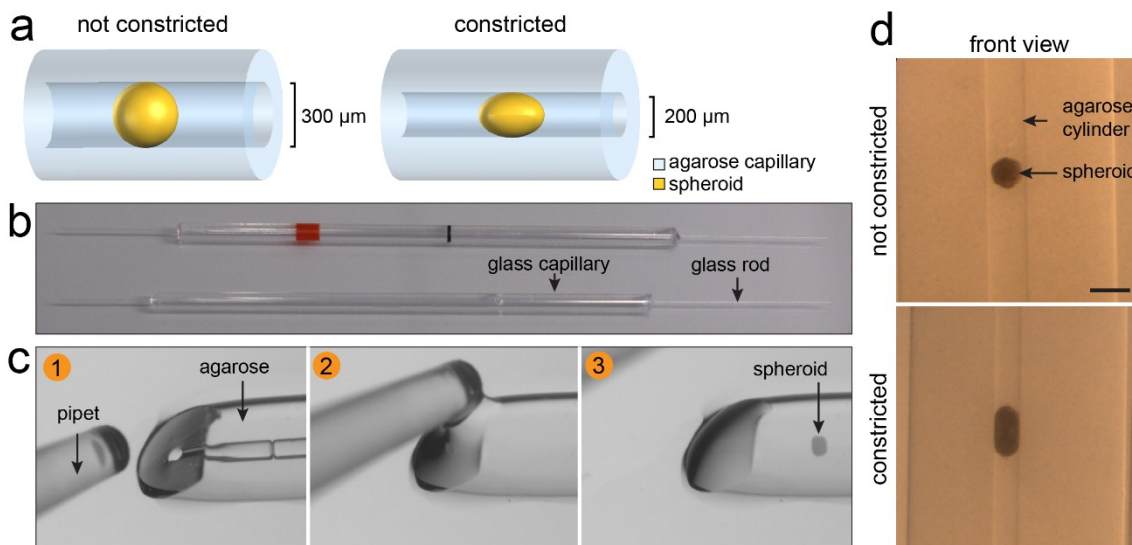


Figure 7 Overview of cylindrical constriction of HC11 spheroids. (a) Spheroids are placed in agarose capillaries with an inner diameter of either 300 μm or 200 μm. (b) Glass material used to generate agarose capillaries. (c) The spheroid is released from the pipet on the opening of the agarose capillary and sucked up by capillary forces. (d) Front view of spheroid in a 300 μm (upper panel) and a 200 μm (lower panel) agarose capillary. Scale bar: 300 μm. Isabell Smyrek acquired and provided the images in panels b, c and d.

2.3.4 Staining and immunofluorescence

Application I and II. Spheroids were fixed for 1 hour in 4% paraformaldehyde (Sigma Aldrich) and afterwards washed twice for 5 minutes in 1xPBS. The nuclei were stained for 24 hours in 5 μM Draq5 (Biostatus) according to the instruction manual.

Application III and IV. Spheroids inside the agarose capillaries were fixed with 4% PFA for 20 minutes at RT. After fixation, spheroids were carefully removed from the agarose tubes. After washing with PBS (3x10 minutes), the spheroids were permeabilized with 0.3% Triton X-100 for 30 minutes. Spheroids were placed in blocking solution for one hour at room temperature (RT). Primary cleaved-caspase 3 antibody (Cell Signaling) was diluted in blocking solution and spheroids were incubated for 24 hours at 37°C and 600 rpm shaking in a thermomixer (Eppendorf). Spheroids were rinsed in PBS (3x5 minutes, 3x20 minutes), before incubation with 1 μg/ml DAPI (Merck), Phalloidin-AF546 (Life Technologies) and the secondary anti rabbit AF488 antibody diluted in blocking solution for 4 hours at 37°C and 600 rpm shaking in a thermomixer. Spheroids were rinsed in PBS (3x5 minutes, 3x20 minutes).

2.3.5 Optical clearing and sample preparation

Spheroids were dehydrated with increasing concentrations of ethanol in deionized water at room temperature. The dehydrated spheroids were transferred into BABB solution (1:2 v/v, benzyl alcohol: benzyl benzoate) with a refractive index of 1.56 and incubated until transparency. Samples were transferred into rectangular glass capillaries (Hilgenberg), which were previously cleaned with 1% Hellmanex (Hellma Analytics) in deionized water for 2 hours shaking on a thermomixer at 70°C. The capillary was mounted onto a stainless steel sample holder and then

placed inside the microscope incubation chamber, which was filled with thiodiethanol (TDE, Sigma).

2.3.6 Image acquisition with CLSM and LSFM

Application I, III and IV. Images were acquired with a monolithic Digital Scanned Laser Light Sheet-based Fluorescence Microscope (mDSLML), which is a subtype of LSFM (Keller and Stelzer, 2008). We define the illumination axis as x , the detection axis as z and the axis orthogonal to x and z as y .

A Carl Zeiss Epiplan-Neofluar 2.5x/NA 0.06 or a Plan-Neofluar 5x/0.16 objective lens for excitation, a Carl Zeiss N-Achroplan 20x/NA 0.5 objective lens for emission and a high-resolution Andor Neo 5.5 sCMOS or Andor Clara CCD camera for image recording were used. Image stacks were acquired with an axial pitch of 1.29 μm .

Application II. Samples were transferred into Cyclo-Olefin Copolymer 96-Well plates (Greiner). Images were acquired with the Carl Zeiss LSM 780 laser scanning confocal microscope. The microscope was equipped with a LD LCI Plan-Apochromat 25x 0.8 NA objective lens (Carl Zeiss). The image stack dimensions were set to 1024 x 1024 pixels per image plane with a pixel pitch of 0.55 μm and a z spacing of 2.81 μm .

2.4 Arabidopsis datasets

2.4.1 Plant growth conditions

Arabidopsis thaliana plants were grown under standard conditions (Lucas et al., 2013), prepared for imaging and acquired using light sheet-based fluorescence microscopy (LSFM) by Dr. Daniel von Wangenheim. In brief, lateral root formation was recorded *in vivo* in transgenic *Arabidopsis thaliana* plants expressing a pan-nuclear marker (pUBQ10::H2B-RFP) and a pan-plasma membrane marker (pUBQ10::YFP-PIP1;4). Further, the plants expressed a nuclear reporter (pGATA23::nls-GUS-GFP) that specifically marked pericycle cells primed to become lateral root founder cells (De Rybel et al., 2010). This event marks the location of lateral root formation in a plant root. The rotation by 90° (gravistimulation) for six hours of seven-day-old plants induced the formation of lateral roots. Imaging lateral root formation was performed for up to three days at an interval of five minutes with LSFM. Due to the intrinsic properties of LSFM, the imaging process did not affect the viability of the plants. All plants survived the imaging process without any morphological abnormalities and continued their life cycle.

2.4.2 Image acquisition with LSFM

Image stacks consisting of 233 planes with an axial pitch of 0.645 μm were acquired at an interval of 5 minutes for up to 30 hours. Details about plant culture, preparation and time-lapse imaging can be found in the recent publication (von Wangenheim et al., 2016).

2.5 Hardware and implementation details

2.5.1 Hardware

Computations were conducted on a high performance workstation comprising two six-core CPUs (X5650, Intel Corporation), 96 Gigabyte DDR3 memory, seven Tesla M2070 (NVIDIA Corporation) graphics cards for GPU processing, running Windows Server 2012 R2.

2.5.2 Implementation

Modules of the image analysis pipeline were developed and implemented in *Mathematica* (Wolfram Research, Inc., version 9 or higher) and *FIJI* (version 1.51d). For further implementation details of the modules, see section Methods. Three-dimensional visualizations were generated in *Mathematica* or *Arivis* (<http://www.arivis.com>).

3 Methods

3.1 Automated data management

The module for automated data management comprises a front end user interface application (*JobCreator*) to create job files. Both programs were developed in *Mathematica*. The *JobCreator* writes a job description file in XLSX format that it sends to the back end (*JobProcessor*) that performs the bulk of image processing.

3.1.1 Data structure and file name convention

The module for automated data management requires the input data in TIF format. The first step is the conversion of the multidimensional input dataset into a generalized file structure and file names that follow a generalized naming convention. The actual sorting of the input dataset is not important (i.e. the program works with input datasets that consist of image stacks in TIF format or datasets where single planes are given as individual TIF files). The conversion step ensures that after processing, each TIF file contains an image stack and the file name directly indicates the file content. The module automatically generates file names that are comprised of a set of unique identifiers followed by a variable number of digits.

The user first specifies a name for the job and the input directory of the dataset. Each dataset is expected to belong to a certain experiment for which the user specifies a name in the *JobCreator* user interface. The experiment name will later be the leading string in the file name (e.g. “Tribolium2016-tif”). For a collection of datasets of an experiment, a unique identifier number followed by four digits labels each dataset (e.g. Tribolium2016-DS0001tif”). After the dataset, two characters compose the unique identifier of each dimension followed by a variable number of digits containing the actual value (Table 1).

The input structure of the underlying image dataset needs to be specified by entering the equivalent for each of the identifiers in the *JobCreator* user interface. At any time, the user can check the input directory for files matching the specified identifier equivalents. The *JobCreator* interface will display the size of each identified dimension (e.g. the number of time points or channels). For example, a dataset that comprises three channels, where the file names contain the identifier “chn” and the user needs to enter “chn” below the identifier “CH”. Afterwards, the user needs to specify the output directory.

Table 1 File name convention used by the automated data management module.

identifier short	identifier name	description
LC	location	The location identifier is a unique representation of either a region of interest (ROI) in a specimen (e.g. head, thorax, abdomen of a Tribolium embryo) or a complete specimen in a multiwell experiment, where each well contains one embryo (e.g. “Tribolium2016-DS0001 LC1tif”).
SL	sub-location	The sub-location identifier is a unique representation for a ROI at a defined location (e.g. “Tribolium2016-DS0001LC1 SL1tif”). For example, the sub-location could indicate head, thorax and abdomen of a Tribolium embryo in a multiwell experiment, where each well corresponds to one location.
TP	time point	The time point identifier represents the time point in a time-lapse recording (e.g. “Tribolium2016-DS0001LC1SL1 TP001tif”).
DR	direction	The direction identifier represents information from which direction (view angle) the specimen was recorded. The identifier is followed by a variable number of digits (e.g. “Tribolium2016-DS0001LC1SL1TP001 DR01tif”).
CH	channel	The channel identifier represents the information of different signals that were recorded in separate image channels followed by a variable number of digits, (e.g. “Tribolium2016-DS0001LC1SL1TP001DR01 CH1tif”).
PL	image plane	This plane identifier represents the image plane followed by a variable number of digits, (e.g. “Tribolium2016_DS0001LC1SL1TP001DR01CH1 PL001tif”).

Identifiers for which no data is available are per default omitted in the file name. The number of digits that follow an identifier is restricted by the size of the dimension (e.g. if 301 time points were acquired, the number of digits after the identifier TP is restricted to three digits).

3.1.2 Data reduction by three-dimensional cropping

After the basic setup of the job, the user specifies parameter values for three-dimensional cropping of the image dataset. A separate user interface allows to interactively specify cropping dimensions in x, y and z and additionally perform rotation around the z axis. The specified cropping settings apply to all image stacks of the complete dataset.

3.1.3 Data reduction by ZIP and JPEG2000 compression

The automated data management module further compresses all generated stacks using the built-in, lossless ZIP compression of the TIF format. The result is one compressed TIF file per image stack. A copy of the image dataset is generated and compressed with the JPEG2000 standard (part 1 of the ISO 15444 standard). JPEG2000 compression works only for single

planes. Thus, the result is a ZIP archive per image stack containing a JPEG2000 file (JP2) per image plane.

3.1.4 Generation of maximum projections

The module uses a macro for the open source software *FIJI* (version 1.51d) to compute maximum projections. For efficiency, the generation of z maximum projections is parallelized on the level of the image stacks, i.e. multiple projections are computed in parallel.

3.2 Intensity normalization of time-lapse maximum projections

The program allows the removal of background information and the adjustment of intensity variations in maximum projections along z as a function of time. A user-friendly interface allows the setup of parameter values and running the procedure in batch mode. The program is available for download (Strobl et al., 2015).

3.2.1 Background removal

Maximum projections are pre-processed using Gaussian filtering with a kernel size of 7×7 pixels. An intensity threshold is computed for each pre-processed image using Kittler-Illingworth's minimum error threshold method (Cho et al., 1989). If necessary, the computed threshold values can be adjusted by multiplication with a constant suitable factor. Small background regions are filled to obtain a binary image. Optionally, morphological opening is performed to smooth object boundaries. The program selects the largest object in the binary image and computes the mean background intensity of the remaining image region. The binary image is multiplied with the raw maximum projection and the mean background intensity is subtracted.

3.2.2 Intensity normalization

In a second step, the intensity of the time-lapse is adjusted. The program determines the two background-corrected maximum projections with minimum and maximum mean intensity and computes the mean intensity of these two values. The program chooses a reference maximum projection, whose mean intensity is closest to the computed value and multiplies all other maximum projection images such that they match the reference mean intensity.

3.3 Web-based data visualization and sharing with *BugCube*

The *BugCube* program was developed in *Mathematica* and is provided as a file in computable document format (CDF). The CDF file is available via the web-browser CDF player plugin (<https://www.wolfram.com/cdf-player/>). *BugCube* provides a web-interface to browse, visualize and download multidimensional image datasets. Data access is handled via an FTP server (<ftp://stelzer-ftp.physbio.uni-frankfurt.de>). The program was originally developed for *Tribolium*

image datasets acquired from four view angles consisting of one channel. It can easily be extended towards more views or more dimensions (e.g. multiple channels) or for usage with other datasets.

For generation of the required files (e.g. maximum projections along z) the above-described data management module is suitable. *BugCube* further requires downsampled versions of the image stacks compressed in a ZIP archive. A custom program that completes the *BugCube* program generates these files. Given a directory containing a multidimensional dataset (projections represented in image stacks as TIF files, and raw image stacks as TIF files), the program splits the data into six subdirectories (Table 2). The data is additionally compressed by adding the files to ZIP archives.

Table 2 The underlying data structure of *BugCube*. Each TIF file is compressed into a ZIP archive.

directory	content
zStacks	image stacks (one TIF file per image stack).
tStacks	time series of maximum projections along z (one TIF file per time series).
zStacksSinglePlanes	image planes of the dataset (one TIF file per image plane).
tStacksSinglePlanes	containing all raw projection images (one TIF file per projection image).
zStacksPreview	downsampled image stacks for preview (one TIF file per z stack).
tStacksPreview	containing the down-sampled t-stacks for preview purposes (one TIF file per tStack).

3.4 Multiview reconstruction

3.4.1 Multiview reconstruction of *Tribolium* datasets

Registration and fusion of the *Tribolium* datasets was performed using the *Multiview Reconstruction* software (Preibisch et al., 2010, 2014) available as a plugin for *FIJI* (<http://imagej.net/Multiview-Reconstruction>). The plugin comprises four modules (dataset specification, bead detection, bead registration and image fusion). After specification of the dataset, the plugin first identifies the subpixel location of fluorescent beads in each view with a difference of Gaussian (DoG) or difference of mean (DoM) filter.

Second, the plugin identifies corresponding beads in all views based on a geometric local descriptor that represents each bead by its location, the location of the three closest beads and their relative orientation towards the bead. The best alignment of two descriptors is identified by pairwise rotation invariant matching of descriptors. Locally similar constellations are identified and each matching yields a transformation model (correspondence candidates). The largest subset of true corresponding beads (i.e. that have a similar descriptor and suggest a similar

transformation model) is identified using the random sample consensus (RANSAC) algorithm. True corresponding beads are registered by a global minimization of the displacement yielding the final affine transformation model. The displacement error summarized the registration performance.

For image fusion, the plugin provides two methods: weighted average fusion (WAF) and multiview deconvolution (MVD). The MVD additionally applies iterative Bayesian-based derivation of multiview deconvolution (Preibisch et al., 2014).

We took advantage of the built-in Compute Unified Device Architecture (CUDA) implementation to speed up the steps of bead detection and multiview deconvolution in the software. To enable GPU usage, CUDA version 6.5 was installed and the precompiled dynamic link library (DLL) files for CUDA (*SeparableConvolutionCUDA.dll* for bead detection and *FourierConvolutionCUDALib.dll* for multiview deconvolution) were placed in the *FIJI* directory. The relevant settings for reconstruction of Tribolium datasets are given in Table 3.

3.4.2 Multiview reconstruction of a *Gryllus* dataset

Stitching of tiled image stacks was performed using the *Image Stitching* software (Preibisch et al., 2009) available as a plugin for *FIJI* (http://imagej.net/Image_Stitching). In brief, the plugin computes all possible translations in x, y and z using the Fourier shift theorem. The translation with the highest cross correlation measure yields the best overlap. For more than two tiles, a global optimization selects the most appropriate translation of all input tiles. We chose an overlap parameter value of 30% for stitching the tiled stacks. All remaining parameters were set to default values.

Registration and fusion of the *Gryllus* datasets was performed with the *Multiview Reconstruction* software (Preibisch et al., 2010, 2014) available as a plugin for *FIJI* (<http://imagej.net/Multiview-Reconstruction>). The settings used for multiview reconstruction are given in Table 3.

Table 3 Parameter values for multiview reconstruction of Tribolium and Gryllus datasets.

Parameter	Tribolium datasets	Gryllus dataset
Bead detection		
Bead detection method	difference of Gaussian (DoG)	difference of Gaussian (DoG)
sigma (pixels)	3.3405	1.8585
Threshold	0.0061	0.0034
Bead registration		
Allowed error for RANSAC	2	5
Significance required for	10	10

a descriptor match		
Regularize model	yes	no
Lambda	0.10	0.10
Transformation model	rigid	rigid
Weighted average fusion (WAF)		
Process views in parallel	All	All
Blend images smoothly	yes	yes
Interpolation	Linear Interpolation	Linear Interpolation
Multiview deconvolution (MVD)		
Image dimensions cropped (x×y×z)	600×1000×600	
ImgLib2 container	ArrayImg	
ImgLib2 container FFTs	ArrayImg	
Save memory	no	
Type of iteration	Efficient Bayesian – Optimization II	
Image weights	Virtual weights / Precompute weights for all views	
OSEM acceleration	1	
Number of iterations	8	
Adjust blending parameters	no	
Use Tikonov regularization	yes	
Tikonov parameter	0.0060	
Compute	In 256×256×256 blocks	
Compute on	GPU	
PSF estimation	Extract from beads	

3.5 Multiscale quantitative characterization

The quantitative characterization pipeline consists of pre-processing, cell nuclei segmentation, alpha shape generation, cell graph computation and subsequent feature extraction. The pipeline was developed in an application-driven manner using image datasets of cellular spheroids and is controlled by several parameters. Please note that parameter settings for each application are given in Table 5.

3.5.1 Pre-processing

Raw image stacks were cropped to the region of interest in *FIJI* (version 1.51d). Background subtraction was performed using the function “Subtract Background” with a radius of 20 pixels

(Application III) or 25 pixels (Application II, III). If image stacks were acquired with anisotropic voxel sizes, they were scaled along the Z dimension and missing information was interpolated (parameter: *ImageZScalingFactor*). For Application II, the appropriate value for the parameter *ImageZScalingFactor* was 5.1. However, for a better segmentation performance, the parameter was set to 3.5 such that cell nuclei appeared less elongated. After cell nuclei segmentation, the obtained results were corrected accordingly. For computational efficiency, the resulting three-dimensional image was resized by a suitable factor (parameter: *ImageScalingFactor*). In the remainder of this section, the obtained three-dimensional image is termed $f_r(x, y, z)$.

3.5.2 Initial cell nuclei segmentation

Noise is reduced by Gaussian filtering with a three-dimensional convolution kernel (parameter: *NucleiFilterRange*). The mean background intensity t_{global} of the convolved image is determined by Otsu's method (Otsu, 1979) or minimum error thresholding (Kittler and Illingworth, 1986) (parameter: *NucleiThresholdMethod*). Local thresholding is applied per sectional plane along the dimensions x, y and z. The resulting binary images are multiplied to obtain the initial segmentation. For each pixel the local threshold t_{local} is

$$t_{local} = m_{local} + \gamma t_{global},$$

where m_{local} corresponds to the mean intensity measured in a specified range (parameter: *NucleiThresholdRange*) given in pixels and t_{global} is the mean background intensity. The factor γ controls the impact of the mean background intensity (parameter: *NucleiBackgroundFactor*). Each pixel with an intensity greater than t_{local} is set to 1, all others are set to 0. Holes, i.e. regions that because of minor segmentation errors are falsely detected as background, are removed. Therefore, the initial segmentation is inverted to obtain the complement image. Holes are identified as small foreground regions and those smaller than a predefined threshold are removed (parameter: *HoleFillingRange*). The image is inverted again to obtain the hole-corrected initial binary image $f_b(x, y, z)$.

3.5.3 Marker point detection

For the decomposition of connected components in $f_b(x, y, z)$, we use a three-dimensional marker-controlled immersion watershed algorithm. The marker positions are obtained by a multiscale Laplacian of Gaussian (LoG) filter algorithm. Thereby, a blob refers to a region of a three-dimensional image, in which the intensities vary within a sufficiently small range of values. First, the pre-processed image $f_r(x, y, z)$ is inverted to obtain $f_i(x, y, z)$. The LoG corresponds to first convolving the image $f_i(x, y, z)$ by a Gaussian kernel $g(x, y, z; \sigma)$. Thus, we get

$$f_g(x, y, z; \sigma) = g(x, y, z; \sigma) * f_i(x, y, z),$$

where σ is the standard deviation (scale) of the Gaussian kernel. Then the Laplacian operator ∇^2 is applied to the convolved image to obtain the LoG response. In our approach, the image $f_i(x, y, z)$ is processed at multiple scales $\sigma_i \in \{\sigma_{min}, \dots, \sigma_{max}\}$. The minimal and maximal scales σ_{min} and σ_{max} are determined using the relationship

$$r = \sqrt{2}\sigma,$$

between radius r of a blob-like object and the scale σ of the LoG. We measured the minimal (r_{min}) and maximal (r_{max}) radius of cell nuclei *a priori* in the images (*NucleiSeedDetectionMinRadius* and *NucleiSeedDetectionMaxRadius*) and computed σ_{min} and σ_{max} . To achieve scale-invariance, the LoG response at scale σ is normalized by multiplication with σ^3 such that the LoG response at scale σ is given by

$$LoG(x, y, z; \sigma) = \sigma^3 \nabla^2 f_g(x, y, z; \sigma).$$

For computational efficiency, we iteratively compute the maximum response $LoG_{max}(x, y, z; \sigma_{min}, \dots, \sigma_{max})$ over scales given by

$$LoG_{max}(x, y, z; \sigma_{min}, \dots, \sigma_{max}) = \max_{\sigma_{min} \leq \sigma \leq \sigma_{max}} LoG(x, y, z; \sigma).$$

A maximum transform detects extended maxima in the obtained maximum response $LoG_{max}(x, y, z; \sigma_{min}, \dots, \sigma_{max})$. A voxel is considered an extended maximum if no voxel in its direct neighborhood has a higher local intensity. The obtained binary image contains the locations of extended maxima (marker points) and is multiplied with the initial segmentation $f_b(x, y, z)$ to discard extended maxima detected in the background. Extended maxima that are in close proximity to each other are merged by increasing the size of the marker points using a morphological dilation operator with a round structuring element (parameter: *NucleiSeedDilation*). The resulting binary image $f_m(x, y, z)$ specifies the marker points that are used to initialize the subsequent watershed algorithm.

3.5.4 Watershed-based decomposition of cell nuclei clusters

With the initial segmentation $f_b(x, y, z)$, the inverted pre-processed image $f_i(x, y, z)$ and the markers in $f_m(x, y, z)$, we use the immersion watershed algorithm (Vincent and Soille, 1991) to achieve a decomposition of connected cell nuclei clusters. The immersion watershed algorithm is a region-based image segmentation method that relies on the image intensities. The algorithm interprets the image as a topographic intensity relief that is gradually immersed resulting in catchment basins. The marker points defined in $f_m(x, y, z)$ represent the initial set of catchment basins and have different labels, (e.g. unique numbers). The algorithm then iteratively increases

the intensity threshold from the minimum to the maximum intensity observed in $f_i(x, y, z)$. In each iteration, all pixels that fall below the current intensity threshold are labelled with the label of the closest catchment basin. The closest catchment basin is determined using the geodesic distance function. It corresponds to the shortest path length between two pixels among all paths between the two pixels. Please note that in this context a path can only use pixels that already have a label. The catchment basins are sequentially expanded and watersheds (label 0) are formed where different catchment basins meet during the immersion process. The final set of uniquely labelled catchment basins and watersheds is a partitioning $f_w(x, y, z)$ of the image $f_i(x, y, z)$ where each catchment basin contains one cell nucleus (Figure 8).

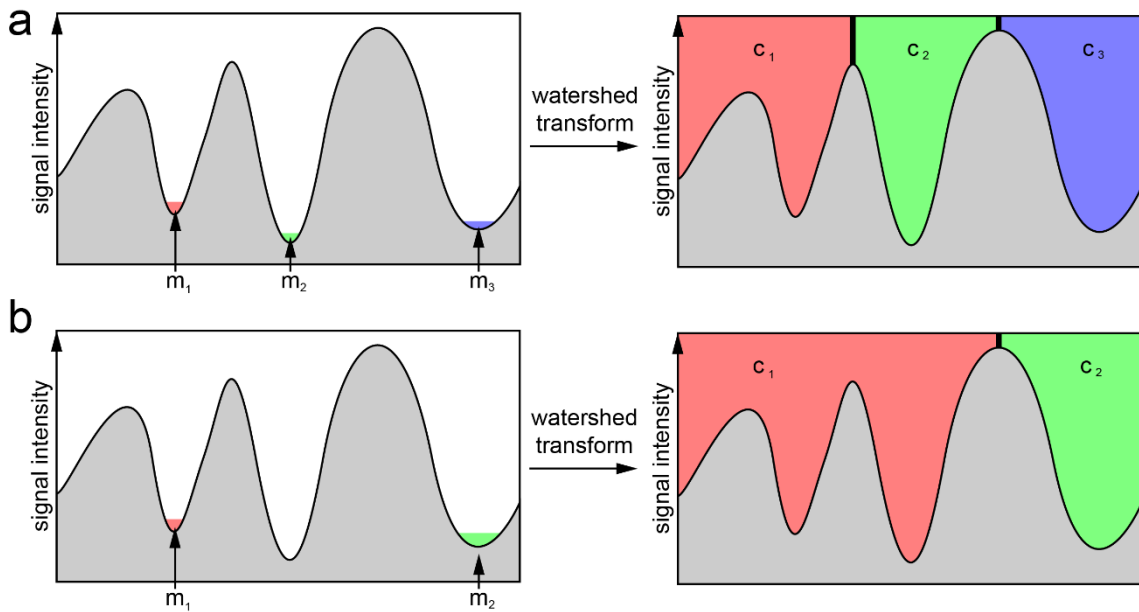


Figure 8 Schematic illustration of the immersion watershed transform. Two examples of the watershed transform applied to a one-dimensional signal. (a) Three marker points m_1 , m_2 and m_3 result in a separation into three catchment basins c_1 , c_2 and c_3 , with boundaries (watershed lines) at the local maxima (indicated in black). (b) For two marker point m_1 and m_2 , catchment basin c_1 floods over a peak during the immersion process until a boundary is formed with catchment basin c_2 .

The immersion process does not necessarily stop the catchment basins at the borders of cell nucleus but extends until other catchment basins are met. Thus, we multiply the resulting matrix $f_w(x, y, z)$ with the initial binary image $f_b(x, y, z)$ resulting in the final segmentation of cell nuclei $f_s(x, y, z)$.

Cell nucleus-like objects in $f_s(x, y, z)$ are selected by incorporating lower and upper volume thresholds. Based on the measurements of minimal and maximal radii of cell nuclei, the mean volume is approximated as a sphere with equivalent radius and we obtain an approximate lower and upper threshold in voxels (parameters: *NucleiMinCount* and *NucleiMaxCount*). For the remaining objects, we extract intensity-related and morphological features of each identified component. The morphological features include the volume in number of voxels, centroid, intensity-weighted centroid, mean, minimum and maximum distance to the centroid, number of voxels on the surface and the bounding box. Based on the voxels of each component, we use

principal component analysis to determine the principal directions and the extensions along these directions as a measure of shape and orientation of the component. Intensity-related features include the mean, minimum and maximum intensity and the intensity standard deviation. All intensity-related features are rescaled to the range $[0, 1]$. Please refer to Table 4 for a complete list of extracted cell nucleus features.

3.5.5 Evaluation of segmentation performance

Ground truth (GT) datasets of cell nuclei centroids were created by manually marking the approximate centroids of cell nuclei in dataset L3 of Application I. We generated a total of three ground truth (GT) datasets by cropping three sub-regions of $100 \times 100 \times 100$ voxels out of the raw image. The centroids of all cell nuclei within the sub-regions were manually identified with a custom program implemented in *Mathematica* version 10.2. Based on the generated GT, we computed the number of correctly detected cell nuclei (true positives, TP), the number of cell nuclei that were falsely detected by the segmentation (false positives, FP) and the number of cell nuclei that were not detected by the segmentation (false negatives, FN). To compute these numbers, we used the following algorithm: if exactly one centroid of the segmentation is found within a spherical neighborhood of twelve voxels of a centroid in the GT, we count it as TP and delete it from the list. If multiple centroids are found within this neighborhood range, the closest centroid is considered as TP. The FP and FN were obtained using $FP = N_{SC} - TP$ and $FN = N_{GT} - TP$, where N_{SC} is the number of centroids determined by the segmentation and N_{GT} is the number of centroids in the GT. We determined the metrics recall, precision and F score with values ranging from zero (worst performance) to one (best performance):

$$recall = \frac{TP}{TP + FN}$$

$$precision = \frac{TP}{TP + FP}$$

$$F\ score = 2 \times \frac{precision \times recall}{precision + recall}$$

3.5.6 Geometrical modelling with alpha shapes

Alpha shapes are a generalization of the convex hull, i.e., every convex hull is an alpha shape, but not every alpha shape is a convex hull (Edelsbrunner et al., 1983). For a detailed description of the underlying concepts, please refer to the book “*Computational Topology – An Introduction*” (Edelsbrunner and Harer, 2010). Given the set of n cell nuclei centroids, $P = \{p_0, \dots, p_n\}$, we used the alpha shapes approach to obtain a geometrical model of the cell aggregate.

Requirements for alpha shapes

Several requirements have to be met for the alpha shapes approach to be applicable as a geometrical model of cell aggregates. First, the cell nuclei centroids P are required to be in general position. Thus, in three-dimensional space, no more than two centroids lie on a two-dimensional hyperplane. The detected cell nuclei centroids have subpixel accuracy, thus the general position requirement is usually fulfilled. Alpha shapes are known to work well for high density point sets. Low density point sets will require large values of alpha in order to connect. Thus, alpha shapes are not appropriate for hollow cell aggregates or when the cell deviates from a uniform distribution.

Edelsbrunner algorithm for alpha shapes

An efficient algorithm to obtain the alpha shape for a given set of points P with real number α is the Edelsbrunner algorithm (Edelsbrunner and Mücke, 1994; Edelsbrunner et al., 1983). The alpha shape is constructed as a subgraph of the Delaunay triangulation. Each triangle in the Delaunay triangulation is associated with the radius of the smallest circle containing the triangle (circumcircle). For a real number α , all triangles for which the radius of the circumcircle is larger than α are removed. The alpha shape corresponds to the union of the remaining triangles, i.e. the alpha shape is a subset of the Delaunay triangulation.

Alpha shape of the cell aggregate

To discriminate against false positive detection of cell nuclei, all points in P whose pairwise distance is below a certain threshold (parameter: *OutlierDistanceThreshold*) are connected. The largest connected component is regarded as the cell aggregate and remaining components are removed. The subset S of P contains the points that belong to the largest component. Subsequently, the alpha shape is approximated for S using the above-described Edelsbrunner algorithm for alpha shapes (parameter: *Alpha*). The alpha shape represents a geometrical model and is further used to compute volume and compute the surface area of the cell aggregate. The relative location of each cell nucleus is determined by computing its distance to the alpha shape surface.

3.5.7 Cell graph construction

A graph G is a pair (V, E) consisting of a set of vertices V and a set of edges E , each a pair (u, w) of vertices. A graph is simple, if no two edges connect the same two vertices and no edge joins a vertex to itself (loop). The graph is undirected, thus, edges have no orientation and the edge (u, w) is identical to the edge (w, u) . In a simple graph, a path between two vertices u and v is a sequence of vertices $u = u_0, u_1, \dots, u_k = v$ with an edge between every pair of vertices u_i and u_{i+1} for each $0 \leq i \leq k - 1$. A simple graph is connected if there is a path between every pair of vertices.

A cell graph is a simple graph without loops, with vertices being the set of cell nuclei centroids and edges being the neighborhood relationship between any two cells. We derive two graph representations representing the spatial distribution of cells within cell aggregates. The proximity cell graph is given by $PCG(V_S, E_{PCG})$ where V_S is the vertex set and E_{PCG} is the edge set of the graph. In the PCG, cells are neighbors if they are closer than a certain distance. Thus, we obtain an edge (u, w) between two vertices u and w if the Euclidean distance between u and w is less than a predefined threshold (parameter: *EdgeDistanceThreshold*). The Delaunay cell graph is given by $DCG(V_S, E_{DCG})$ where V_S is the vertex set and E_{DCG} is the edge set of the graph. The DCG graph is constructed based on a Delaunay triangulation to estimate, which cells are neighbors and in physical contact. An edge (u, w) is created between two vertices u and w if the corresponding points are connected in the Delaunay triangulation and the Euclidean distance between u and w is less than a predefined threshold (parameter: *EdgeDistanceThreshold*). Edge weights in both graphs are assigned according to the Euclidean distance.

We extract the degree of each vertex as the number of neighbors. Further, the minimum, maximum, mean and standard deviation of the distance to neighbors for each vertex v is given by the weights of all edges incident to v . Please refer to Table 4 for a complete list of extracted features for the cell aggregate and individual cell nuclei.

3.5.8 Guidelines for setting up parameter values

The following paragraphs describe suitable initial values for the most important parameters used in cell nuclei segmentation, alpha shape computation and cell graph generation. The effect of different values for all parameters can be investigated using the provided user interface.

The value of the parameter *ImageZScalingFactor* is a calibration factor that specifies the scaling of the z dimension in order to obtain an isotropic voxel size in x, y and z dimension. For example, a factor of four is appropriate for an image stack that was recorded with a pixel pitch of 0.645 μm in x and y and a z spacing of 2.58 μm . The parameter *ImageScalingFactor* allows resizing the three-dimensional image by a multiple of the original dimensions. For example, a value of 0.5 will result in a three-dimensional image with half the number of pixels in x, y, and z. The computation time for downsampled images is drastically reduced.

The parameter *NucleiFilterRange* specifies the range of the Gaussian filter used to suppress background noise before the mean background intensity is determined. The range in which the local threshold is computed is given by the parameter *NucleiThresholdRange*. A suitable starting value is the mean diameter of the cell nuclei to be detected in the image. The parameter *NucleiBackgroundFactor* controls the influence of the mean background intensity on the locally computed threshold value. For a value of 0 the threshold value is solely based on the locally

computed mean intensity. Higher values will increase the influence of the background intensity and thus increase the locally computed threshold value. Depending on the quality of the input images this value should be slightly increased (e.g. to 0.25) in order to ensure that no cell nuclei are detected in the background region. If background subtraction has been performed in the images (e.g. the background intensity is effectively 0) the value for *NucleiBackgroundFactor* can be set to 0.

The parameters *NucleiSeedDetectionMinRadius* and *NucleiSeedDetectionMaxRadius* specify the minimal and maximal radius of the Laplacian of Gaussian filter algorithm. These two parameters allow fine-tuning the cell nuclei segmentation such that it tolerates cell nuclei of different diameters to a certain extent. Ideally, the minimal radius is set to the radius of the smallest cell nuclei, whereas the maximal radius is set to the radius largest cell nuclei to be detected.

Objects that lie far outside the main aggregate of cells (e.g. single, separated cells) are eliminated by setting the value of the parameter *OutlierDistanceThreshold*. A suitable value is the mean distance of cells from each other. All objects that are not within this mean distance will be separated from the main aggregate of cells and discarded. The alpha shape approximation is solely influenced by the parameter *Alpha*. A suitable value is the expected maximal distance between two neighboring cells in the cell aggregate. For higher values, the alpha shape will correspond to the convex hull, whereas for low values the alpha shape will be reduced to the cell nuclei centroids. The parameter *EdgeDistanceThreshold* specifies the maximal allowed distance between two cell nuclei in order to be connected by an edge in the PCG and DCG. The parameter value drastically affects the number of edges generated in the PCG.

Table 4 Features extracted by the quantitative characterization module.

	Feature name	Description
Cell aggregate features	ProximityCellGraph (PCG)	Proximity cell graph $PCG(V_S, E_{PCG})$ with vertices V_S and edges E_{PCG} , describing the whole cell aggregate. Vertices V_S have positions $\{x, y, z\}$ and represent cell nuclei centroids. An edge $(u, w) \in E_{PCG}$ exists if the Euclidean distance between the two vertices u and w is less than a threshold. The edge weight is determined by the Euclidean distance between the two vertices u and w .
	DelaunayCellGraph (DCG)	Delaunay cell graph $DCG(V_S, E_{DCG})$ with vertices V_S representing and edges E_{DCG} , describing the whole cell aggregate. Vertices V_S have positions $\{x, y, z\}$ and represent cell nuclei centroids. An edge $(u, w) \in E_{DCG}$ exists if it is part of the Delaunay triangulation of V_S and the Euclidean distance between the two vertices u and w is less than a threshold. The edge weight is determined by the Euclidean distance between the two vertices u and w .
	Surface	Surface of the cell aggregate constructed from all cell nuclei locations S using the alpha shapes approach.
	Centroid	Geometric center $\{x, y, z\}$ of the cell aggregate given by the arithmetic mean of all cell nuclei locations.
	MinDistanceSurface	Euclidean distance of <i>Centroid</i> to its closest point on <i>Surface</i> .
	Volume	Total number of voxels of the volume enclosed by and including <i>Surface</i> .
	SurfaceArea	The number of voxels of <i>Surface</i> .

	Feature name	Description
Cell nuclei features (segmentation, surface)	Label	Unique integer that identifies a cell nucleus.
	Mask	A binary three-dimensional array in which 1s refer to the voxel that are part of a cell nucleus.
	BoundingBox	Axes-oriented minimal bounding box $\{\{x_{min}, y_{min}, z_{max}\}, \{x_{max}, y_{max}, z_{max}\}\}$ that contains a cell nucleus.
	Count	Total number of voxels of a cell nucleus.
	PerimeterCount	Total number of voxels on the surface of a cell nucleus.
	MinCentroidDistance, MaxCentroidDistance, MeanCentroidDistance	Minimum, maximum and mean distance of all voxels to the geometric center of a cell nucleus.
	Centroid	Geometric center $\{x, y, z\}$ of a cell nucleus.
	MainAxes	Main axes of orientation $\{p_1, p_2, p_3\}$ of a cell nucleus measured by principal components analysis. In particular, p_1 is the vector along the first, p_2 along the second and p_3 along the third principal component, respectively.
	Extension	Extension $\{e_1, e_2, e_3\}$ of a cell nucleus along its three main axes of orientation $\{p_1, p_2, p_3\}$ in number of voxels, where e_1 is the extension along p_1 , e_2 along p_2 and e_3 along p_3 , respectively.
	TotalIntensity, MeanIntensity, StandardDeviationIntensity, MinIntensity, MaxIntensity	Total, mean, standard deviation, minimum and maximum value of the intensity distribution of a cell nucleus measured in the raw volume. Intensity features are normalized to the range $[0, 1]$.
	IntensityCentroid	Intensity weighted geometric center $\{x, y, z\}$ of a cell nucleus.
	SurfaceDistance, SurfaceNearest	Minimal distance of the cell nucleus centroid to the aggregate surface in number of voxels and the corresponding surface point $\{x, y, z\}$.
SurfaceOrientation	Angle in degree between major axis vector p_1 of the cell nucleus and the aggregate surface. An angle of 0° indicates that p_1 is parallel to <i>Surface</i> , whereas for an angle of 90° , p_1 is orthogonal <i>Surface</i> .	

Table 5 Parameter values of the quantitative characterization module used for application I, II, III and IV. The same parameter values were used for all datasets.

Cell nuclei segmentation	Application number			
Parameter	I	II	III	IV

ImageZScalingFactor	3.5	3.5	3.5	3.5
ImageScalingFactor	0.5	0.75	0.5	0.75
NucleiSeedDetectionMinRadius	3	2	3	4
NucleiSeedDetectionMaxRadius	6	4	6	6
NucleiFilterRange	3	0	3	0
NucleiThresholdMethod	Otsu	Minimum error		
NucleiThresholdRange	12	8	10	10
NucleiSeedDilation	2	2	3	3
NucleiBackgroundFactor	0.25	0.25	0.25	0.5
Alpha	90	70	90	90
OutlierDistanceThreshold	20	23	20	30
EdgeDistanceThreshold	40	30	40	30
NucleiMinCount	250	100	150	200
NucleiMaxCount	42,500	50000	10,000	5000

3.5.9 Random cell position model

For Application I, we compared the generated proximity cell graphs to those of a mathematical random cell position (RCP) model. In this model, we performed Monte Carlo simulations with the following steps: (1) cells are distributed uniformly within a cell aggregate. (2) Cell nuclei are represented as spheres with positions drawn from a uniform random number distribution. The radius of the spheres was set to the median cell nucleus radius (~ 5.9 voxels, ~ 3.8 μm) from all datasets. (3) Cell nuclei are not allowed to overlap. (4) Cell nuclei positions are restricted to the volume of the cell aggregate (i.e. the alpha shape) and are not allowed outside this volume. (5) The number of randomly generated cell nuclei is the same as the number of cell nuclei determined for the real cell aggregate. For each cell dataset, we generated ten RCP models and obtained the cell graphs in the same way as for the real datasets.

3.6 Computer-aided cell tracking and lineage reconstruction

3.6.1 *TrackGen* combines cell tracking and lineage tracing

The program *TrackGen* (implemented in *Mathematica*) allowed for interactive visualization of Arabidopsis datasets as a function of time and was used to track the locations (referred to in Cartesian coordinates) and trace the lineages of cells. Using *TrackGen*, the user marked a cell at its first occurrence, i.e., right after the cell division in which the cell was born or the very first appearance of a cell, and at its last occurrence, i.e., right before the cell divided or the last time point analyzed. Daughter cells were assigned to their ancestral cell (precursor). After marking cell locations and assigning lineages, *TrackGen* stored for each cell a unique identifier, the spatial position, the time point of occurrence and a list holding the identification numbers of

precursors. The resulting dataset was exported to a file in tabular format for subsequent data processing and analysis. The spatial location of a cell in-between the two marked time points were later added by linear interpolation and the lineage information of the first occurrence was adopted.

3.6.2 Automated lineage reconstruction

A lineage tree is a directed graph $G = (V, E)$ consisting of a set of vertices $v \in V$, corresponding to cell occurrences and a set of directed edges $(u, v) \in E$ that represent the relationship between two vertices u and v . The root vertex r represents the first occurrence of the lineage founder cell. Each internal vertex has either one child representing the second occurrence of a cell or two children corresponding to the two daughter cells that were born in the cell division. Leaf vertices represent the cells that exist at the end of the lineage trace. An edge weight $w(u, v)$ holds the time span that passed between the cell occurrence u and v . All information about a cell occurrence is stored in the corresponding vertex of the lineage tree. The development of a cell lineage is represented in top to bottom direction. Thus, the constructed lineage tree encodes all information about cells and cell division events as a function of time. Lineage trees of all identified cells were reconstructed automatically using this principle.

3.6.3 Registration of the Arabidopsis datasets

Spatial registration

A comparison of the Arabidopsis datasets required spatial and temporal registration. For spatial registration, the cell nuclei locations were transformed into a common Cartesian coordinate system spanned by the principal component vectors of the cell nuclei locations in the last tracked time point of each dataset. The resulting coordinate system provides three orthogonal views on the lateral roots: front view (x-y, growth towards the observer), side view (z-y, along shoot-root axis), and radial view (x-z, transversal cut perpendicular to shoot/root axis). We empirically determined parameter values for a rigid transformation of the cell nuclei locations at the last time point and applied the same transformation to the cell nuclei locations at each time point. Additionally, the cell nuclei locations were centered at each time point. The standardized data was exported to a file in tabular format and used for all successive analysis.

Temporal registration

The time between gravistimulation of the plant and the initiation phase of lateral root formation differed substantially between the datasets. Thus, a temporal registration of the datasets based on the absolute time was inappropriate. Instead, we synchronized the datasets based on the total number of cells in the primordium. We therefore obtained the total number of cells at the first and last time point of the datasets. The intersection range of all datasets was [18, 143] cells. The resulting developmental synchronization allowed comparing the datasets at a given developmental stage given in number of cells.

3.6.4 Quantitative lateral root properties

Growth rate

Cellular growth rate and doubling time of the datasets were estimated by data fitting on the total number of cells assuming the exponential growth model

$$N(t) = N_0 e^{kt},$$

where $N(t)$ is the number of cells at time point t , N_0 the number of cells at the first time point t_0 and the growth rate k . Least squares fitting was used to fit the exponential growth model to the data (Supplementary Figure 11). The square roots of the total cell numbers at time point t were chosen as weights. The doubling time T was then computed as

$$T = \ln(2)/k.$$

Computation of lateral root volume, height, length and width

Based on the set of cell nuclei locations at a time point, Edelsbrunner algorithm for alpha shapes (Edelsbrunner et al., 1983) with parameter α was used to approximate the shape of the lateral root. The alpha shape is a generalization of the convex hull concept and is not necessarily convex or fully connected. The value of the parameter α was set to 50 voxels ($\sim 16 \mu\text{m}$) for all time points and all datasets. From the alpha shape, the volume of the lateral root at a time point could readily be obtained. The height of the lateral root was determined by computing the maximal distance between any two cell nuclei in the master cell file at each time point. Length and width of the lateral root were determined at 50% of the lateral root height at each time point.

3.6.5 Quantitative cell division properties

Spatial orientation of cell divisions relative to the primary root

Based on the relative orientation of a cell division to the primary root, cell divisions were classified as either anticlinal, periclinal or radial. Anticlinal cell divisions are parallel to the shoot-root axis. Periclinal cell divisions are normal to the surface of the primary root. Radial cell divisions are tangential to the surface of the primary root and orthogonal to the root-shoot axis. For the i^{th} cell division event, a vector \vec{d}_i is constructed from the spatial coordinates of the progeny (daughter cells) to represent the cell division axis. Thus, a list of vectors $\{\vec{d}_1, \dots, \vec{d}_k\}$ is obtained that represents the spatial orientation of the k cell divisions in the dataset. To measure the relative spatial orientation of the cell divisions, a reference system was constructed for each cell file. Each reference system is defined by two vectors \vec{n}_h and \vec{n}_l , where \vec{n}_h points along the principal direction of growth in the cell file (height axis) and \vec{n}_l points along the shoot-root axis of the primary root (length axis). The vectors \vec{n}_h and \vec{n}_l were determined using principal component analysis (PCA) of the cell nuclei locations in each cell file at the last time point. The

relative spatial orientation of each cell division event was determined by the angles φ and θ between the vector \vec{d}_i and the two vectors \vec{n}_h and \vec{n}_l of the respective cell file reference system. The values for φ and θ are given in the range $[0^\circ, 90^\circ]$. Based on the angles θ and φ , cell divisions were classified as either anticlinal, periclinal or radial. The following criteria were used for the classification:

- Cell divisions for which $\varphi \geq 45^\circ$ and $\theta < 45^\circ$ were classified as anticlinal. The spatial orientation of an ideal anticlinal cell division is $\varphi = 90^\circ$ and $\theta = 0^\circ$.
- Cell divisions for which $\varphi < 45^\circ$ and $\theta \geq 45^\circ$ were classified as periclinal. The spatial orientation of an ideal anticlinal cell division is $\varphi = 0^\circ$ and $\theta = 90^\circ$.
- Cell divisions for which $\varphi \geq 45^\circ$ and $\theta \geq 45^\circ$ were classified as radial. The spatial orientation of an ideal anticlinal cell division is $\varphi = 90^\circ$ and $\theta = 90^\circ$.

The classification of cell division types was validated manually and misclassified cell divisions were corrected.

Spatial orientation of cell divisions relative to the previous cell division

The spatial orientation of each cell division relative to the previous cell division was determined by computing the angle ω between the vector \vec{d}_i of the current cell division and \vec{d}_j of the previous cell division. The angle is given in the range $[0^\circ, 90^\circ]$.

4 Results

4.1 An automated data management module

Automated data management is of great importance in all applications where large amounts of image data are produced. The datasets have to be automatically organized and reduced to the relevant information. In collaboration with Frederic Strobl, I developed a data management module that automates multiple pre-processing steps for multidimensional image datasets: (1) renaming of the input files into a generalized file name format, (2) three-dimensional image cropping, (3) lossless image compression using the ZIP standard for TIF files (4) generation of a second dataset lossy compressed with the JPEG2000 standard, (5) automatic computation of maximum projections, (6) generation of time-lapse videos of maximum projections and (7) computation of basic image statistics of the complete dataset.

4.1.1 Workflow of the automated data management module

The module comprises a front end and a back end program (Figure 9). In the front end (*JobCreator*), a general user interface (GUI) allows the specification of processing settings. The user enters the parameter values for data location, three-dimensional cropping, computation of projections and image compression. The *JobCreator* writes a job description file in XLSX format that it sends to the back end (*JobProcessor*) that performs the bulk of image processing. The back end needs to run on a powerful workstation (i.e. a separate machine or server). The main workflow consists of the following steps:

1. **Job specification.** All settings such as the location of input and output directory, the name of the experiment, how to read the input data and further settings for processing are set up in the *JobCreator* (Figure 10a).
2. **Job submission.** The *JobCreator* writes all settings into a XLSX file and exports the file to the job input directory of the back end program *JobProcessor*.
3. **Image processing.** The *JobProcessor* will automatically check for new files in the job input directory. Once a new job description file is found, it is imported and data processing is performed as specified in the file. All output is written to a local drive in the first place.
4. **Termination.** Once all processing steps are finished, the *JobProcessor* moves the generated data to the target directory specified by the user and deletes the local files.

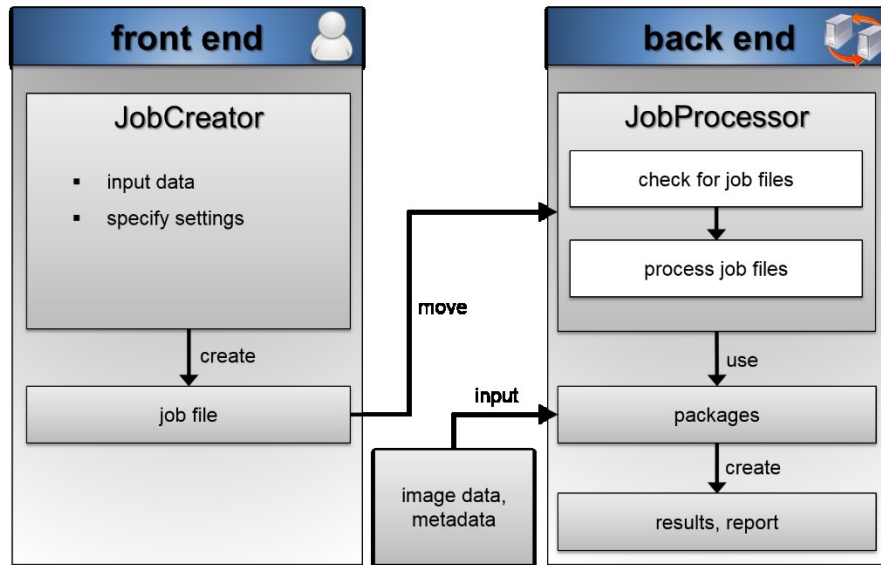


Figure 9 Framework of the automated data management module. The module comprises front end and back end. The front end comprises a user-friendly interface to specify settings of a job file. The *JobCreator* saves all job settings in a file in XLSX format and sends it to the input folder of the back end. The *JobProcessor* does the bulk of processing. All generated results are stored in the specified output folder.

4.1.2 Capabilities of the automated data management module

Image datasets usually explore many dimensions where the first (x) and second (y) spatial directions represent the number of pixels of a sectional plane image. The number of planes represents the third spatial dimension (z). For convenience, image planes are saved as a stack of images in multi-page tagged image file format (TIFF). All further dimensions (channels, views, and time points) are saved as individual files. The file name convention uses unique identifiers that have a fixed order in the file name. Each identifier consists of two characters followed by a fixed number of digits (Table 1). The file name indicates all subsequent processing and provides direct information about the contained data.

Cropping the raw data to the region of interest (ROI) along the dimensions x, y and z is an important pre-processing step for multidimensional image datasets. The developed module allows three-dimensional cropping with a user-friendly interface (Figure 10b). Settings for cropping can be specified in a user interface and are automatically applied to each stack of the dataset. After renaming and cropping the data, maximum projections along x, y and z are automatically computed and saved in a separate folder. Furthermore, the maximum projections are assembled into time-lapse videos. The generated files are in movie (MOV) or TIF format and compatible with other software.

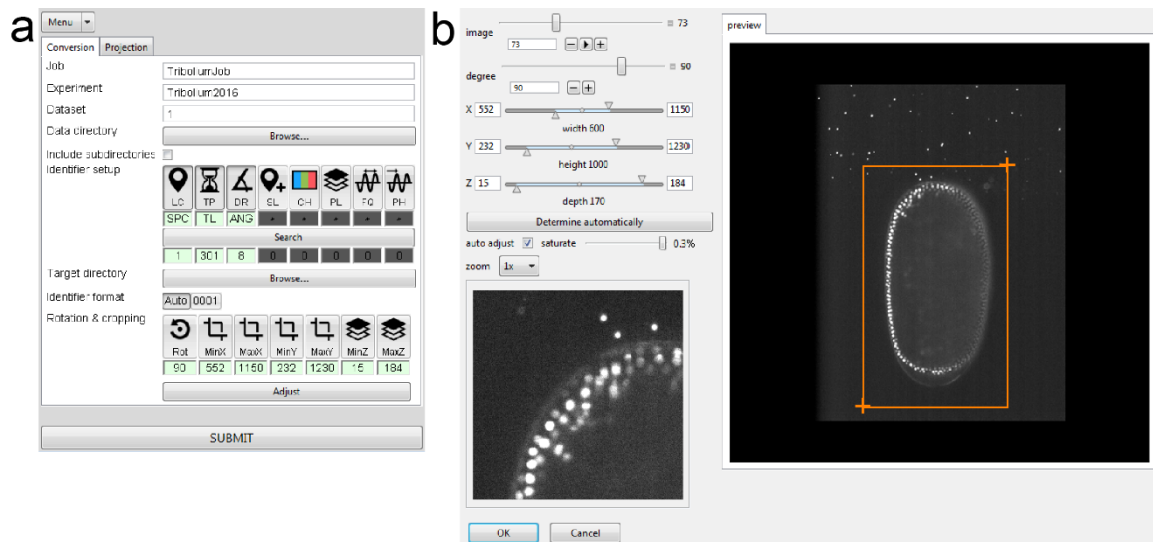


Figure 10 *JobCreator* user interface. (a) Main user interface of the *JobCreator*. (b) User interface for the setup of three-dimensional cropping parameter values. Both the sliders and boxes on the left side or an interactive box on the right side allow the specification of three-dimensional cropping.

4.1.3 Evaluation of ZIP and JPEG2000 image compression

Efficient image compression is indispensable to manage large amounts of data. Image compression should ideally yield high compression rates as well as high write and read speed, while maintaining the quality of the input image and ensuring compatibility with existing software. The data management module features two image compression standards to reduce the amount of data volume: (1) lossless ZIP compression of multipage TIF files and (2) lossy compression using the JPEG2000 standard. In collaboration with Frederic Strobl, I evaluated the performance of both compression schemes for three *Tribolium* datasets.

ZIP compression reduces the data volume without information loss

For 16-bit TIF files, the ZIP compression outperformed other alternatives for lossless compression (e.g. Lempel-Ziv-Welch (LZW) or PackedBits) in terms of compatibility with existing software, compression speed and efficiency. The ZIP compression yielded compression ratios of at least 2.4. For example, a *Tribolium* dataset consisting of 301 time points, eight views (2408 image stacks in total) with a raw data volume of 1486 Gigabyte was compressed to 629 Gigabyte with lossless ZIP compression (i.e. 42% of the raw data volume). The compression ratios were similar for the other test datasets (Table 6).

JPEG2000 compression drastically reduces the data amount

The data management module further generates a copy of the dataset compressed with the JPEG2000 standard. We evaluated the efficiency and quality of lossy JPEG2000 compression. The *Mathematica* built-in implementation of JPEG2000 is limited to two-dimensional images. Each plane of an image stack was therefore compressed separately and stored as a file in JP2 format. Thus, for a stack of 233 planes, 233 separate files in JP2 format were obtained.

LSFM datasets are typically stored in 16-bit multipage TIFF files. However, only a fraction of the dynamic range of the 16-bit images is actually used. For example, the *Tribolium* test datasets typically did not cover higher image intensity values than 2^{13} . We found that the *Mathematica* built-in algorithm for JPEG2000 compression was inappropriate for direct application to raw 16-bit images generated with LSFM. Applying the JPEG2000 compression directly resulted in the occurrence of compression artefacts. Thus, we evaluated a pre-processing step, which multiplies the image intensity values of the raw images with a constant scaling factor. This step effectively increased the dynamic range of the underlying images to span the 16-bit range, which resulted in an improved quality of the JPEG2000 compressed images (Figure 11).

Cell nuclei are accurately detected in JPEG2000 compressed images

To evaluate the quality of JPEG2000 compression with different scaling factors, we performed a segmentation of cell nuclei in a maximum projection image of a *Tribolium* embryo and compared it to the segmentation result for the maximum projection of the raw image (Figure 11a). We ensured that the threshold value used for the segmentation was identical for the different images. For low scaling factors, JPEG2000 compression resulted in the occurrence of ring and blur artefacts (Figure 11b, first row). Increasing the scaling factor to 32 resulted in an image quality that was indistinguishable from the raw image. This result is underpinned by the segmentation result obtained for cell nuclei (Figure 11b, second row). Low scaling factors led to a high amount of missed cell nuclei and a stronger deviation of the measured cell nucleus area from that obtained from the raw image (Figure 11c, d). For a scaling factor of 32, almost the same number of cell nuclei was found and the measured cell nucleus area matched the result obtained for the raw image.

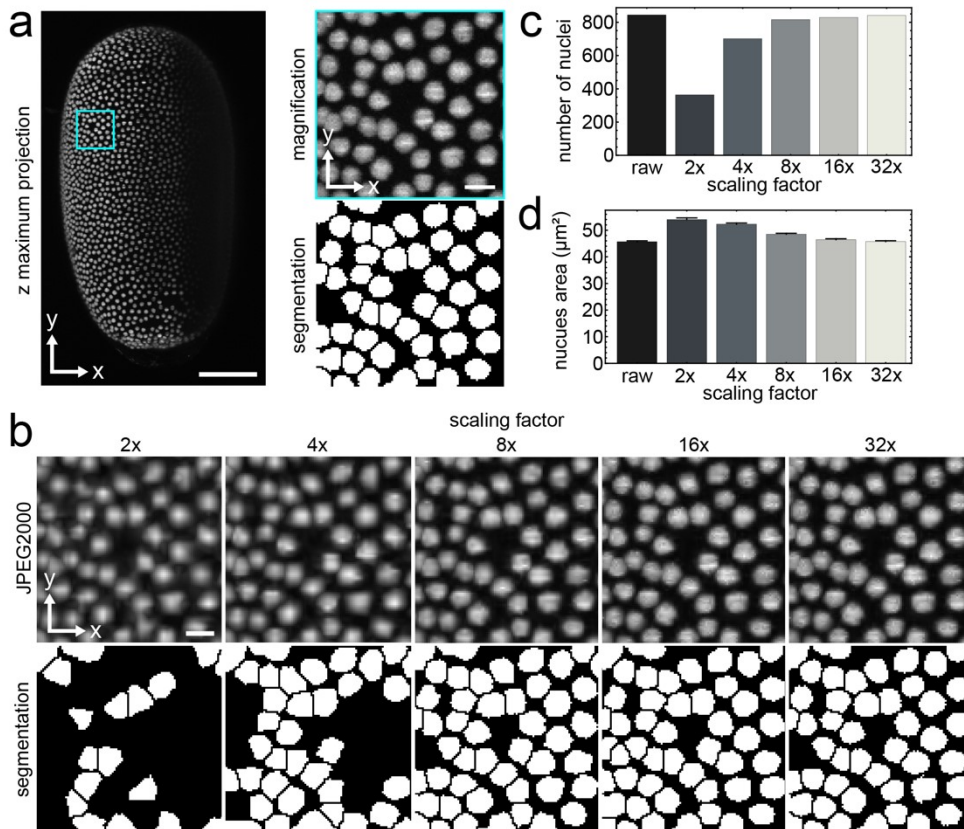


Figure 11 Evaluation of JPEG2000 compression for different scaling factors. (a) Maximum projection along z of the stack at the first time point and first view of a Tribolium dataset. Scale bar: $100\ \mu\text{m}$. region highlighted by the cyan box is magnified on the right with the corresponding segmentation result after applying a constant intensity threshold value. Scale bar: $10\ \mu\text{m}$. (b) The magnified region shown in panel a after JPEG2000 compression with different scaling factor values. The corresponding segmentation result is shown in the lower row. Note that the threshold value obtained for the raw image was multiplied with the appropriate scaling factor and used for segmentation (c) Plot of the identified number of cell nuclei in the complete image shown in panel a for different values of the scaling factor. (d) Plot of the mean cell nucleus area with standard error for different values of the scaling factor. Microscope: mDSLm, illumination objective: $2.5\times$ NA 0.06 Epiplan-Neofluar, detection objective: $10\times$ NA 0.3 N-Achroplan, camera: Andor Clara.

Based on these results, we evaluated the efficiency of JPEG2000 compression for the test datasets. We observed that JPEG2000 compressed datasets were drastically reduced with compression ratios of at least 500, i.e. to less than 1% of the raw data volume. For example, dataset 1 was compressed to 1.6 Gigabyte using lossy JPEG2000 compression (0.1% of the raw data volume) with a scaling factor of eight (Table 6).

Table 6 Performance of ZIP and JPEG2000 compression.

Number of	Dataset 1	Dataset 2	Dataset 3
pixels ($x \times y$)	1392×1040	1392×1040	1392×1040
planes per stack	229	233	210
time points	301	301	301
view directions	8	8	8
Total data volume (Gigabyte)			
raw	1486	1512	1364

ZIP	629	593	534
JPEG2000	1.6	1.7	2.3

JPEG2000 compression is suitable for long-term data storage

The image intensity levels of the test datasets varied as a function of time (Figure 12a). We observed that the JPEG2000 compression ratio adapted to the information content of the image. For example, file sizes of image planes that only contained background information were at least a factor of two smaller than file sizes for image planes that contained relevant information of the Tribolium embryo. For example, we observed that the compression ratio for dataset 1 varied between 600 and 1200 over time (Figure 12b), indicating that the variability in the exploited dynamic range is also reflected by the compression ratio of JPEG2000. A qualitative evaluation of the resulting image quality confirmed this result. We compared a time point at which the images had high intensity values and therefore more or less exploited the full 16-bit dynamic range with a time point that did not. We observed that the image quality was drastically worse for the image with lower intensity values (Figure 12c).

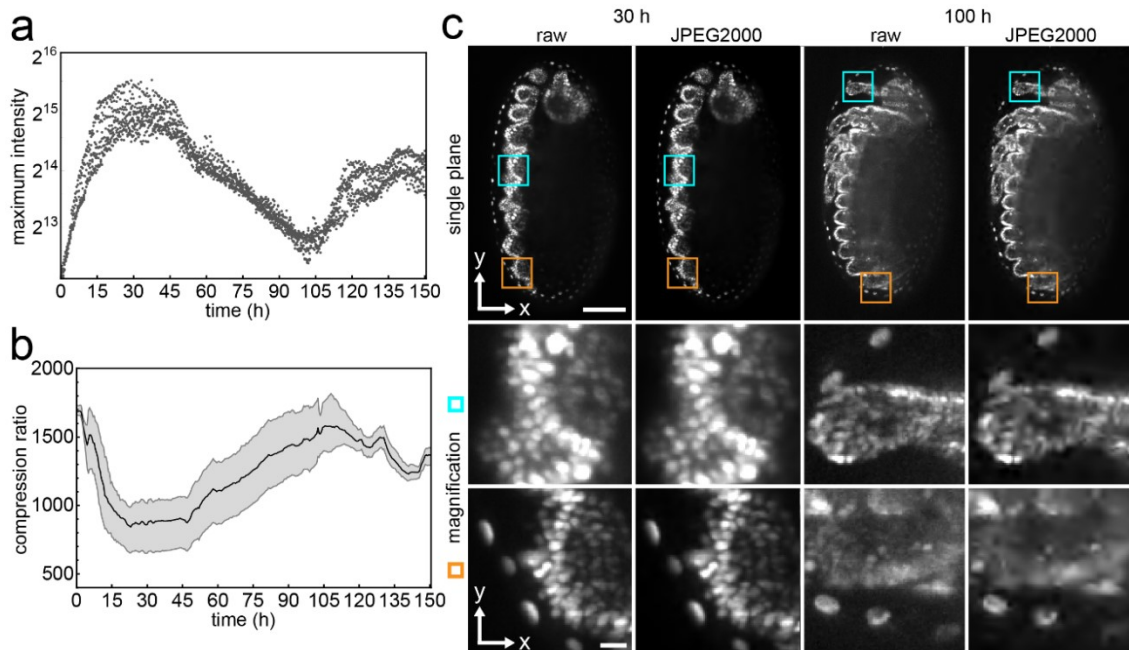


Figure 12 Qualitative evaluation of JPEG2000 compression for Tribolium dataset 1. (a) Plot of the maximum intensity measured as a function of time in the raw image stacks and multiplied with a scaling factor of eight. (b) Plot of the average (black line), minimum and maximum (gray lines) compression ratio as a function of time. The compression ratio is computed by the ratio between the file size of the raw image stack and the file size of the JPEG2000 compressed image stack. (c) Single planes of the raw and JPEG2000 compressed image stacks (first row) at time points 30 and 100 hours of the time-lapse. Scale bar: 100 μm . Magnifications of the two highlighted image regions (cyan and orange boxes) are shown in the second (cyan box) and third (orange box) row. Microscope: mDSLm, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 10x NA 0.3 N-Achroplan, camera: Andor Clara. Scale bar: 10 μm .

In summary, the developed data management module automates and facilitates several common steps of image pre-processing. The built-in ZIP compression of multipage TIFFs was suitable in

order to reduce the data volume. Because no intensity information is lost or altered, ZIP compression was considered suitable for data storage during the analysis phase of the data (e.g. direct measurements or segmentation). Lossy JPEG2000 compression of single planes led to a considerable reduction to less than 1% of the raw data volume. However, it required the adjustment of the dynamic range by applying a constant scaling factor to ensure high image quality after compression. A scaling factor that stretches the dynamic range of the image to the full 16-bit range yielded the highest image quality. Consequently, the relevant image features (e.g. cell nuclei) were maintained after JPEG2000 compression. For the investigated datasets, a scaling factor of eight was considered a good compromise between compression performance and resulting image quality.

4.2 Intensity normalization of time-lapse maximum projections

We developed a program to post-process the maximum projections generated by the data management module. The program removes background information and compensates fluorescence intensity changes that occur as a function of time. For example, the intensity levels of the *Tribolium* datasets varied as a function of time (Figure 13a, first row). To remove the background information a threshold is computed automatically and the resulting binary image is multiplied with the raw image to obtain the background corrected (BC) image (Figure 13a, third row). A reference image is determined and all images of the time series are scaled to match the mean intensity of the reference image yielding the background corrected and mean transformed (BC-MT) images (Figure 13a, fourth row). Using the developed program, intensity variations in a time-lapse of maximum projections were reduced by normalization (Figure 13b). The program was implemented in conjunction with a user-friendly interface (Figure 13c) and parts of it were published (Strobl et al., 2015).

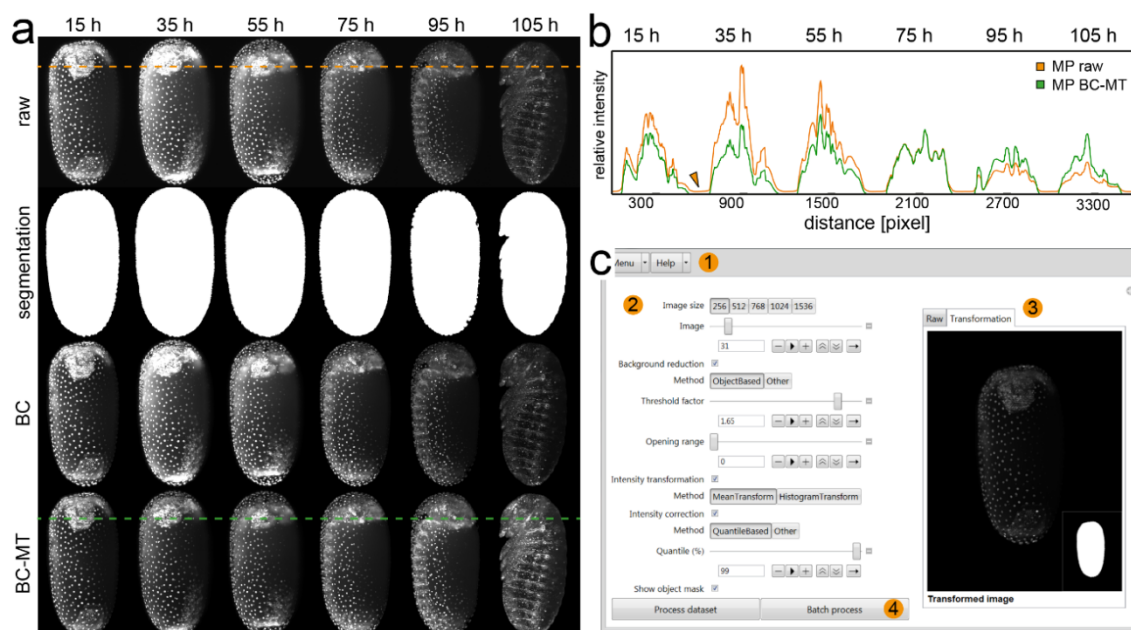


Figure 13 Intensity normalization of time-lapse maximum projections. (a) Background correction and mean intensity transformation shown for a time series of six maximum projections in 90° (lateral) views of a *Tribolium* dataset. The maximum projections contain background (raw). The background is distinguished from the embryo by automatic thresholding (segmentation). The background is set to zero and the mean background intensity is calculated and subtracted from the image (BC). For the time series, the images with minimum and maximum mean intensity are determined and their mean value is computed. The image for which the mean intensity is closest to that value is defined as the reference image. All images of the time series are adjusted to match the mean intensity of the reference image (BC-MT). Scale bar: 100 μm . (b) Intensity profile plot along the orange line in the raw images and along the green line in BC-MT images shown in panel a. The image intensity is either decreased (e.g. 15 h), remains almost unchanged (75 h), or is increased (105 h). The background information is removed (see orange arrow). MP, maximum projection. (c) Screenshot of the user interface: (1) data menu to import and export maximum projection time series, and help menu. (2) Settings menu to configure image data processing. Parameter functions are explained in the help menu. (3) Comparison of raw and transformed image. (4) Buttons to start processing and run batch mode for several files with the chosen settings. Microscope: mDSL, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 10x NA 0.3 N-Achroplan, camera: Andor Clara. Figure adapted from Strobl *et al.*, 2015.

4.3 *BugCube* enables web-based visualization and data sharing

The web-based program *BugCube* provides a platform for viewing and accessing the image data generated by the data management module. *BugCube* can be launched from the webpage <http://www.physikalischebiologie.de/bugcube>. After initialization, a dataset can be selected on the left upper panel (Figure 14a). The full dataset can also be downloaded via a button. *BugCube* automatically loads a description of the dataset and downsampled videos of z maximum projections as a function of time (Figure 14b). Using the slider above the videos, the user can scroll through the time-lapse and get a first glimpse at the dataset. The maximum projection videos can be downloaded in high quality via a save button. Downsampled videos of the z stack at the currently selected time point are loaded and displayed at the bottom (Figure 14c). The original z stacks can be downloaded via the save button below. Images can be magnified and displayed in one of the panels on the right (Figure 14c). The right panel displays single images or two images next to each other (Figure 14d). The contrast and brightness can be

adjusted for each image individually. A magnification view can be used to browse for details in the image. In sync-mode, the magnification views of the left and right image are synchronized.

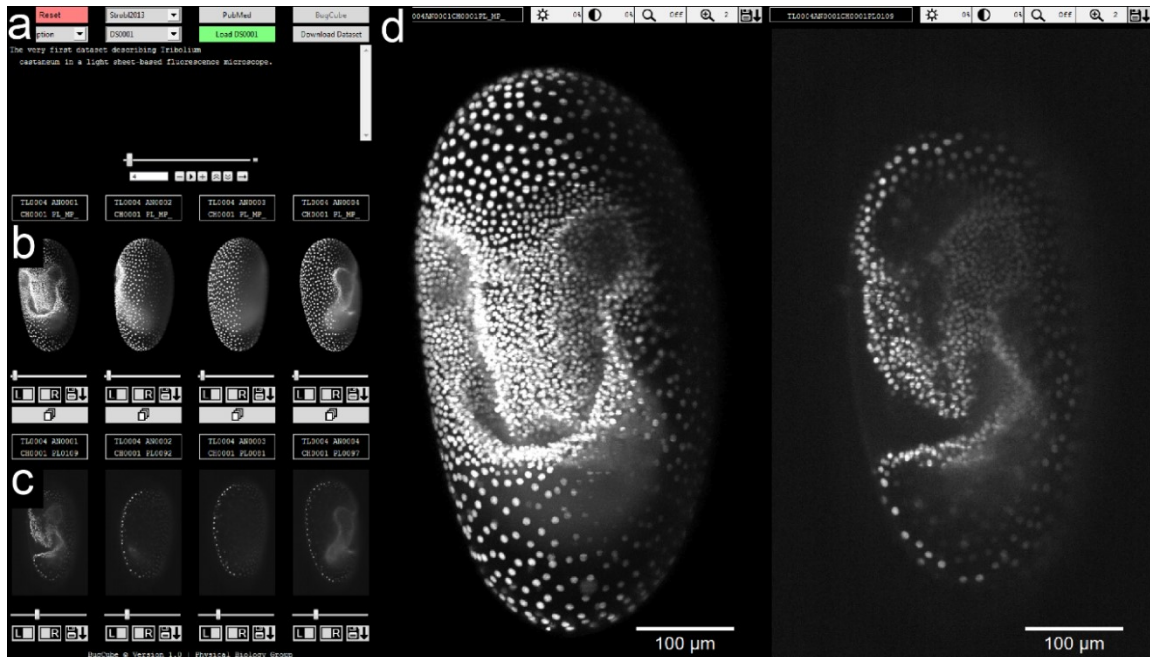


Figure 14 User interface of the program *BugCube*. The web-based program is available at <http://www.physikalischebiologie.de/bugcube>. After initialization, the program allows to browse, view and download image datasets. (a) Main menu of the program. (b) Display panel for t-stacks. (c) Display panel for z stacks. (d) Display panel for full resolution images. Shown here is a *Tribolium* dataset.

4.4 Reconstruction of multiview image datasets

In LSFM, datasets of large specimen can be acquired from multiple view directions. Image registration and fusion is required to combine the information content of all view directions into a single image. In the past, several approaches have been proposed to perform multiview reconstruction. In this thesis, the potential of an existing bead-based reconstruction approach as a module in the image analysis pipeline was evaluated. The multiview reconstruction was performed for time-lapse image datasets of developing *Tribolium* and *Gryllus* embryos. *Tribolium* embryos were fixed on an agarose half sphere containing fluorescent beads (Figure 5). Image stacks of *Tribolium* embryos were acquired along eight view directions for a total of 150 hours. Three *Tribolium* datasets were acquired following this procedure. Each of the raw datasets resulted in up to 1.5 Terabyte data volume. The *Gryllus* embryo was mounted in an agarose film containing fluorescent beads (Figure 6). Image stacks of *Gryllus* were acquired along four view directions for a total of 143 hours. The raw dataset resulted in three Terabyte data volume. Frederic Strobl and Alexander Ross performed sample preparation and imaging of *Tribolium* and *Gryllus* embryos and provided the resulting datasets.

4.4.1 Overview of the reconstruction procedure

Multiview reconstruction of all datasets was achieved using the plugin Multiview Reconstruction for the open-source software *FIJI* (Preibisch et al., 2010, 2014). In brief, the

plugin identifies the fluorescent beads in each view, computes a transformation based on the detected bead locations, registers the views and fuses them resulting in a single, three-dimensional image volume. For image fusion, either a weighted average fusion (WAF) or a fusion that uses multiview deconvolution (MVD) are available. We evaluated both methods based on the resulting image quality. For the Gryllus dataset, stitching of the tiled stacks with the *FIJI* plugin Image Stitching was a mandatory pre-processing step. The settings used for multiview reconstruction for Tribolium and Gryllus datasets are given in Table 3.

4.4.2 Multiview reconstruction of Tribolium datasets

For the detection of fluorescent beads, we used the difference of Gaussian (DoG) method as it outperformed the difference of mean (DoM) method in terms of accuracy. However, the DoG method needed more than three minutes per image stack on the central processing unit (CPU), which would have required several days of processing time for the bead detection step. Scaling down images prior to the bead detection resulted in fewer detected beads and lead to higher registration errors. Thus, we took advantage of the built-in Compute Unified Device Architecture (CUDA) implementation to speed up the multiview reconstruction. This reduced the required computation time by a factor of six to less than 30 seconds for an image stack using the DoG method. In total, bead detection took 17 hours per dataset.

Robust reconstruction of Tribolium datasets

We first qualitatively evaluated the registration performance by computing a maximum projection of the fused image along the y direction. We restricted the projected image volume to the agarose half sphere region containing the fluorescent beads. In the weighted average fusion, the fluorescent beads in the projection image mostly appeared star-shaped. This indicated that the point spread functions of the same bead observed in eight different views overlapped. Some of the beads were only observed in some of the view directions or even only in one view direction (Figure 15c).

The number of detected beads and the mean bead displacement provided a quantitative measure of the reconstruction quality. We determined the number of detected beads per view and pooled the numbers per view over all time points (Figure 15d). A different number of beads was detected in each view direction. However, this number was approximately constant over time. This indicated the robustness of the bead detection method in the plugin. The registration error varied by less than 0.3 μm . The best registration performance was achieved for dataset 3 with a registration error of about 1 μm (Figure 15e).

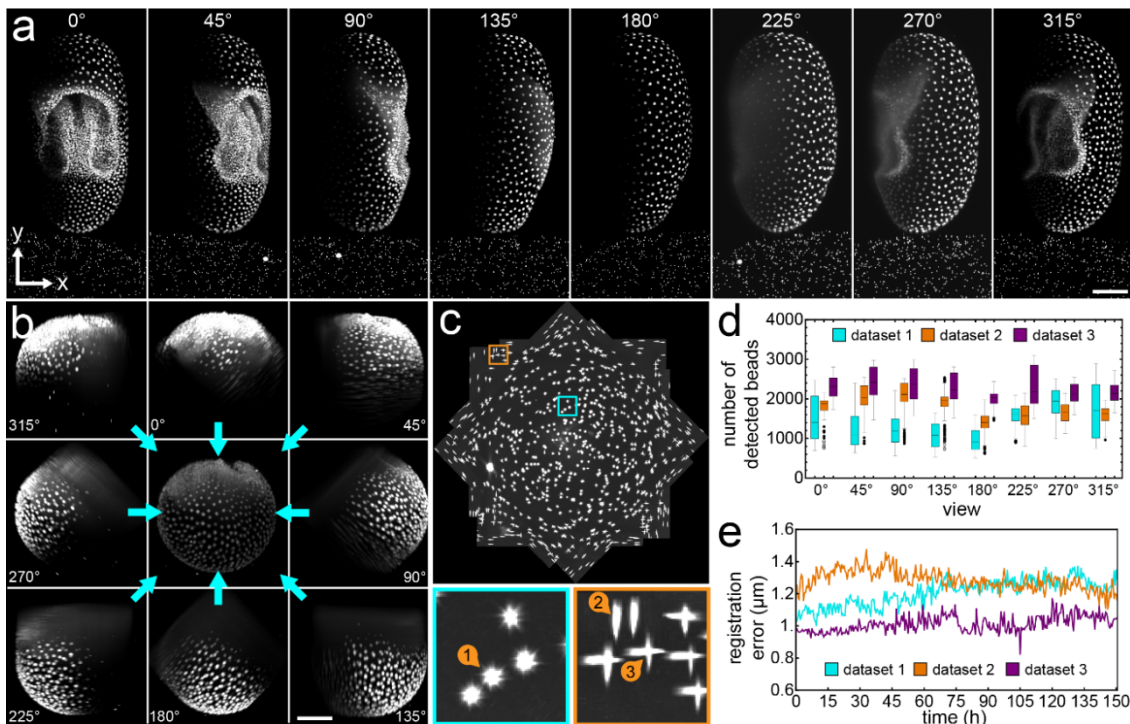


Figure 15 Multiview reconstruction of three Tribolium image datasets. (a) Maximum projections of image stacks along z shown for Tribolium dataset 1. The dataset was acquired along eight different view directions (0° - 315°). Depicted is a Tribolium embryo after 5h 30 min of recording. It was fixed on an agarose half sphere that was supplemented with fluorescent beads. Each view captures only a fraction of the embryo. Scale bar: $100\ \mu\text{m}$. (b) Maximum projections along y of the same time point and the same view directions as shown in panel a. Using the software *Multiview Reconstruction*, the information of all eight view directions is combined into a single three-dimensional image. Scale bar: $100\ \mu\text{m}$. (c) Maximum projection along y of the fluorescent beads in the agarose half sphere after weighted average fusion (WAF). Fluorescent beads that overlap in all eight view directions appear star shaped in the resulting fusion (cyan box, 1). Fluorescent beads that occur in a subset of view directions appear elongated or plus shaped (orange box, 2 and 3). (d) Boxplots of the number of beads in each view direction detected by the bead detection of the *Multiview Reconstruction* software at each time point of all datasets. (e) Plot of the mean registration error in μm as a function of time for all datasets. Microscope: mDSLm, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 10x NA 0.3 N-Achroplan, camera: Andor Clara. Boxplot parameters: the box contains 50% of the data points; the middle line of the box is the median. Whiskers and outliers represent the upper and lower 25% of the data. Outliers are outside the 1.5x interquartile range; far outliers are outside the 3x interquartile range. ●, outliers; ○, far outliers.

The weighted average fusion of a complete time-lapse consisting of 301 time points and eight views using took 15 hours using the CPU-based implementation. The weighted average fusion provided complete and high quality images of Tribolium embryogenesis (Figure 16).

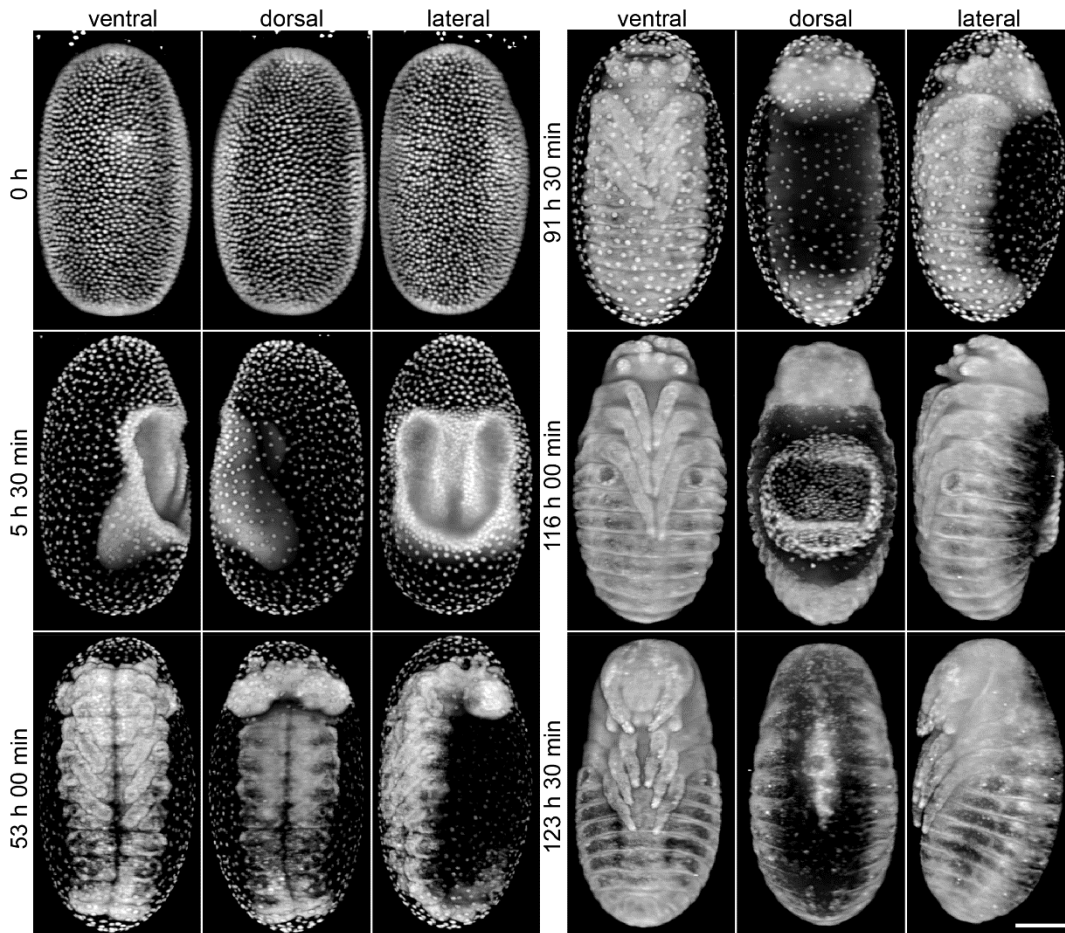


Figure 16 Renderings of a fused Tribolium image dataset. Depicted are three-dimensional renderings of selected time points in ventral, dorsal and lateral view on the embryo in Tribolium dataset 1. Images result from weighted average fusion (WAF) with parameter values listed in Table 3. Microscope: mDSLM, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 10x NA 0.3 N-Achroplan, camera: Andor Clara. Scale bar: 100 μm .

Multiview deconvolution improves the image quality of reconstructed Tribolium datasets

The WAF procedure treats all views identical and does not discriminate of high quality image regions against low quality regions. Consequently, the fusion incorporates background blur that corrupts the quality of the resulting fused image. The *Multiview Reconstruction* software supports a multiview deconvolution (MVD) procedure that addresses this problem using a Bayesian-based derivation of the Richardson Lucy deconvolution algorithm. We evaluated the performance of the MVD procedure for Tribolium image datasets.

The runtime of the MVD implementation on CPUs was high. For example, processing a single time point with eight iterations took more than six hours. Thus, we configured the software to use the available CUDA implementation, which reduced the computation time of the deconvolution. The parallelized CUDA implementation ran on graphics processing units (GPUs) and was about 15 times faster (on eight GPUs) than the standard CPU implementation (six hours on CPU compared to 24 minutes on GPU for one time point).

In the MVD procedure, a deconvolution process iteratively reduces the influence of the point spread function. We first identified the appropriate number of iterations (Figure 17). Eight iterations provided a good compromise between image quality and runtime. The application of the MVD enhanced the fused image quality (Figure 17a). Compared to the WAF procedure, the deconvolved images provided higher image contrast and the amount of background blur was lower than in the WAF image. Higher iteration numbers increased the occurrence of visible deconvolution artefacts (Figure 17b). Compared to the WAF procedure, the deconvolved images provided higher image contrast and the amount of background blur was reduced (Figure 17c).

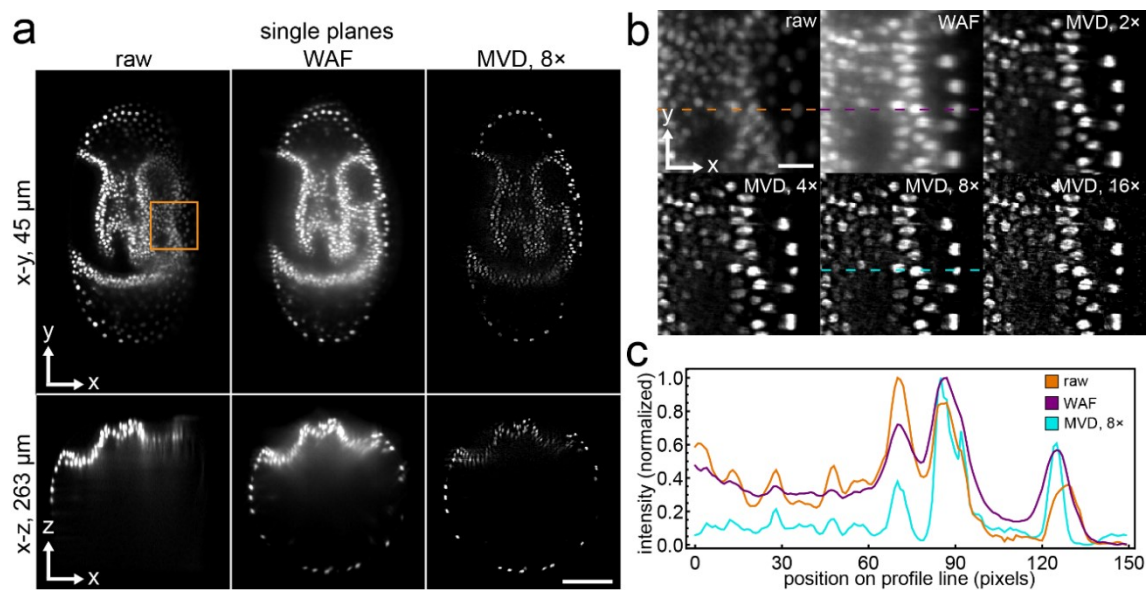


Figure 17 Image quality of images after WAF and MVD. (a) Single planes of a *Tribolium* embryo (dataset 1) after 5 h 30 min of imaging along x-y (45 μm depth) and x-z (263 μm depth from top). The single planes depict the quality of raw image, WAF and MVD after eight iterations. Scale bar: 100 μm . (b) Comparison of raw image quality with the result of WAF and MVD after 2, 4, 8 and 16 iterations for the region indicated by the orange box in panel a. Scale bar: 20 μm . (c) Intensity profile plot along the cyan, dashed line in panel b for the raw image, WAF and MVD (8 \times). The intensity values were normalized by scaling to the range [0.0, 1.0], where 0.0 corresponds to the minimum and 1.0 corresponds to the maximum intensity of the profile. Microscope: mDSLM, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 10x NA 0.3 N-Achroplan, camera: Andor Clara.

We qualitatively compared the image quality of the resulting images after fusion with the raw images (Figure 18). Compared to the raw images, both fusion procedures provided a complete, high quality three-dimensional image of the embryo by incorporating the information from all view directions. Due to the principle of the WAF procedure, higher amounts of background blur were observed. In contrast, the MVD achieved superior image quality.

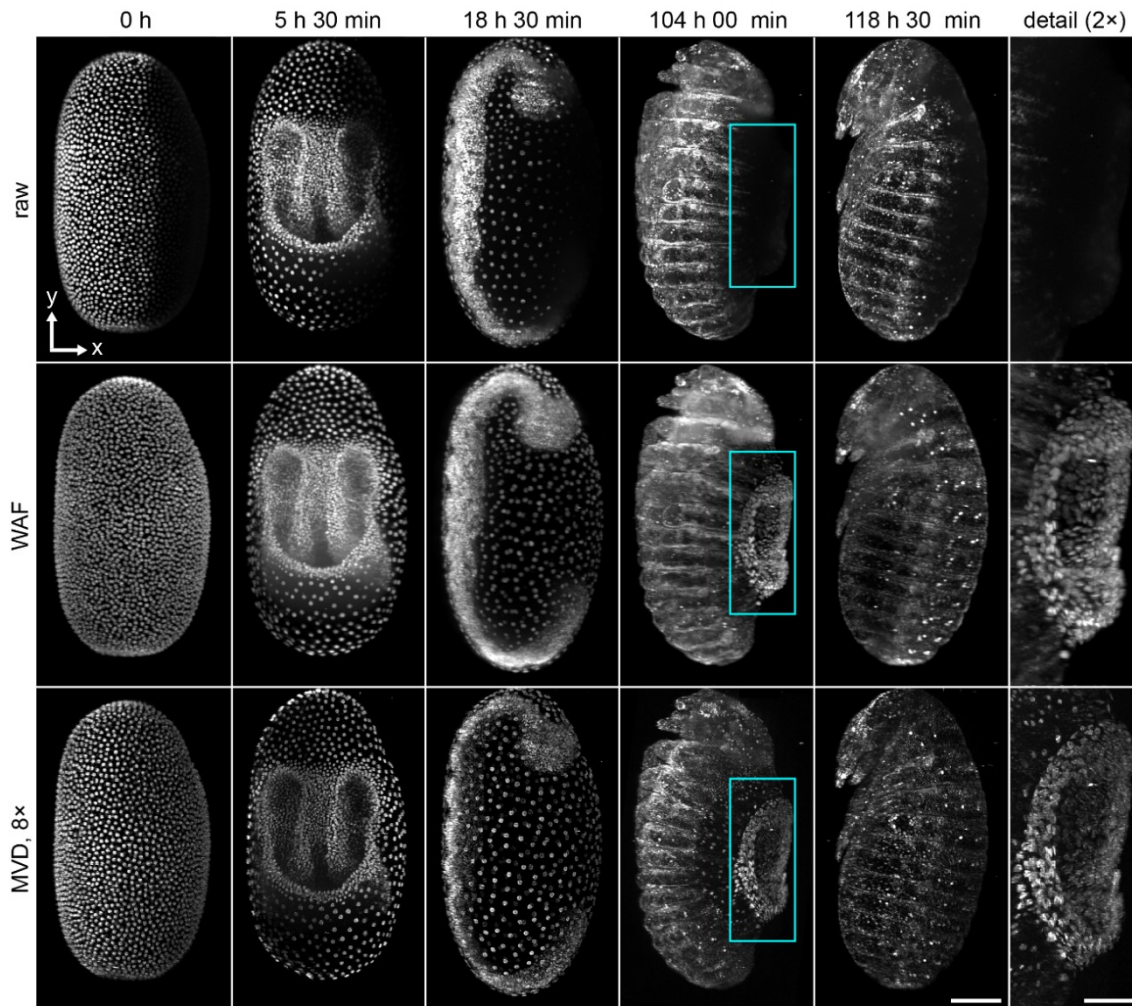


Figure 18 Qualitative comparison of raw images, WAF and MVD. Depicted are maximum projections along z of the 0° view of *Tribolium* dataset 1. Scale bar: $100\ \mu\text{m}$. The rightmost column depicts the regions outlined with cyan boxes, magnified by a factor of two. Microscope: mDSLML, illumination objective: $2.5\times$ NA 0.06 Epiplan-Neofluar, detection objective: $10\times$ NA 0.3 N-Achroplan, camera: Andor Clara. Scale bar: $50\ \mu\text{m}$.

4.4.3 Multiview reconstruction of *Gryllus* datasets

The *Gryllus* dataset was acquired along four view directions with tiled stacks. The tiled stacks were reconstructed into a single image using the *Image Stitching* software. The stitched image stacks were then registered and fused into a single three-dimensional image using the WAF procedure of the *Multiview Reconstruction* software (Figure 19a). The MVD procedure was not applicable due to the dimensions of the image stacks. Bead detection with CUDA speedup took 16 hours 16 minutes, registration took 7 minutes and WAF of the complete dataset took 99 hours 46 minutes. Similar to the *Tribolium* datasets, the number of beads detected per view was different and varied between 2700 (90° view) and 10200 (0° view) detected beads (Figure 19b). The mean registration error for the complete time-lapse was $1.7\ \mu\text{m}$ with a minimum error of $1.1\ \mu\text{m}$ for time point 52 hours (Figure 19c). For several time points, the registration quality was rather low for several time points and we observed registration errors of up to $8.5\ \mu\text{m}$ (e.g. time point 92 hours). Still, the quality of the resulting fused images provides a first three-dimensional view on embryogenesis of *Gryllus* (Figure 20).

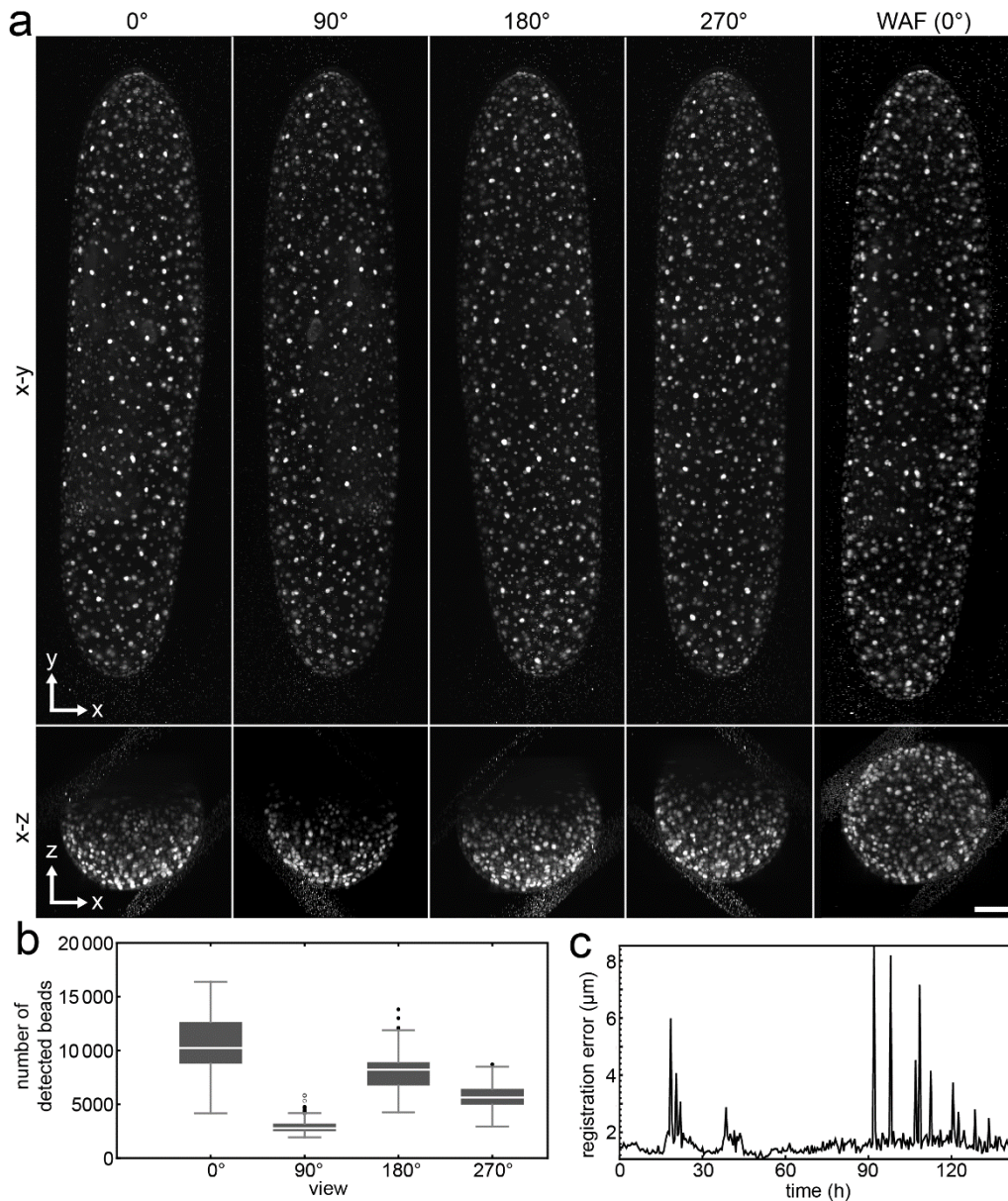


Figure 19 Multiview reconstruction of a *Gryllus* image dataset. (a) Maximum projections along z of image stacks (x - y , upper row) and along y (x - z , lower row) at the first time point. The dataset was acquired along four different view directions (0° , 90° , 180° and 270°). The last column shows the result after WAF in the same orientation as the first view (0°). Scale bar: $200\ \mu\text{m}$. (b) Boxplots of the number of beads in each view direction detected by the bead detection of the *Multiview Reconstruction* software at each time point. (c) Plot of the mean registration error in μm . Microscope: LightSheet Z.1, illumination objective: $10\times$ NA0.2, detection objective: $20\times$ NA 1.0, camera: Sony ICX 285. Boxplot parameters: the box contains 50% of the data points; the middle line of the box is the median. Whiskers and outliers represent the upper and lower 25% of the data. Outliers are outside the $1.5\times$ interquartile range; far outliers are outside the $3\times$ interquartile range. ●, outliers; ○, far outliers.

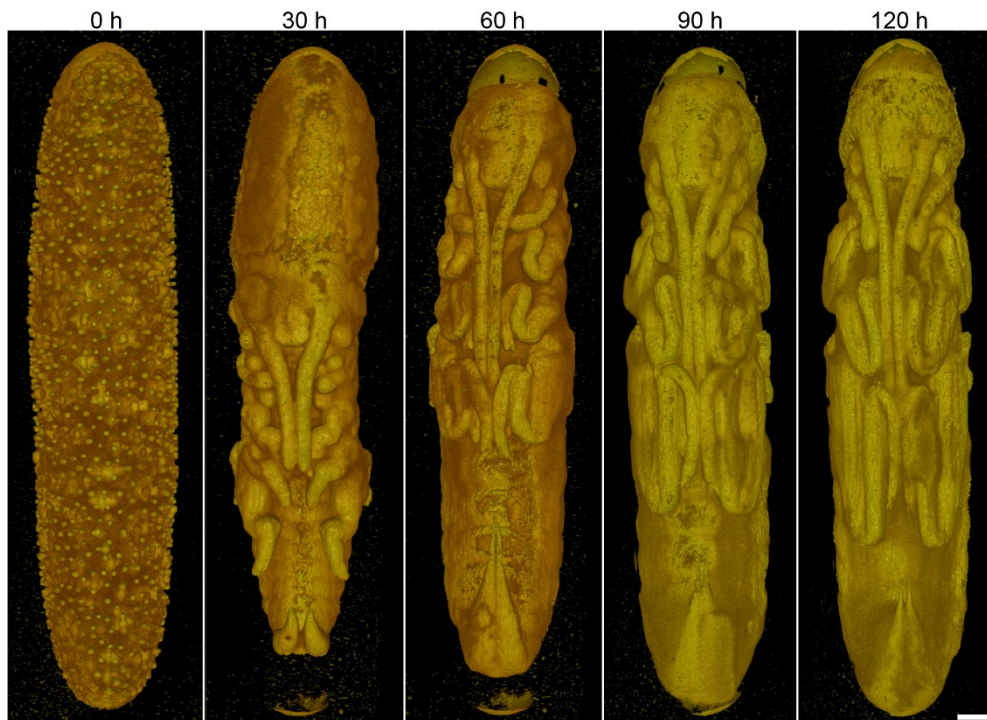


Figure 20 Three-dimensional renderings of a reconstructed *Gryllus* embryo. Depicted are three-dimensional renderings of five time points of a reconstructed *Gryllus* embryo in ventral view. Renderings were produced with the *Arivis* software. Microscope: LightSheet Z.1, illumination objective: 10x NA0.2, detection objective: 20x NA 1.0, camera: Sony ICX 285. Scale bar: 200 μm .

In summary, three multiview image datasets of *Tribolium* embryogenesis (EFA-nGFP transgenic line) and one dataset of *Gryllus* embryogenesis (H2B-eGFP transgenic line) were successfully reconstructed using the software *Image Stitching* and *Multiview Reconstruction* available as plugins for *FIJI*. For the *Tribolium* datasets, weighted average fusion and multiview deconvolution were evaluated. The resulting image quality after multiview deconvolution was superior to the raw images and provided a full three-dimensional view on *Tribolium* embryogenesis. For the *Gryllus* dataset a similar procedure resulted in the first multiview reconstruction of *Gryllus* embryogenesis. In conclusion, the bead-based approach in conjunction with the existing *FIJI* plugin has proven to be an efficient and accurate solution for multiview reconstruction and is therefore considered suitable as a module of the image analysis pipeline.

4.5 Multiscale quantitative characterization

In modern cell biology approaches, solutions for extracting quantitative features of individual objects and the relationship between objects are needed. In this thesis, a quantitative characterization module was developed based on spheroids datasets. A manuscript that describes the developed module and its application was accepted for publication (Schmitz et al., 2017). Please note that written parts (sections 1.1.2, 1.2.2, 3.5, 4.5, 4.6.1, 5.5, 6.1.3), tables and figures (indicated appropriately in the captions) of this thesis were adapted and modified from the publication.

4.5.1 Overview of the multiscale quantitative characterization module

The developed module accurately identifies cell nuclei in image stacks and extracts a multitude of features for the characterization of these spheroids (Figure 21). A pre-processing step ensures the conversion of the input image stack to a three-dimensional volume with isotropic voxels (Figure 21b). To speed up the processing, an optional downsampling of the image is available in this step. For the cell nuclei segmentation, local thresholding provides an initial segmentation of candidate regions (Figure 21c). To improve the segmentation result, a marker controlled three-dimensional watershed algorithm was included. For the marker point detection, a multiscale Laplacian of Gaussian (LoG) filter algorithm was developed (Figure 21d). The final segmentation of cell nuclei is then used for a subsequent series of post-processing steps (Figure 21e). The cell nuclei centroids are used to compute the shape of the spheroid using the alpha shapes approach (Figure 21f). The resulting alpha shape provides a geometrical model of the spheroid. The alpha shape surface facilitates the determination of the relative location of cell nuclei within a spheroid. Further, cell graphs were extended to describe the three-dimensional spatial cell network (Figure 21g). Features of individual cell nuclei and the spheroid are automatically extracted and stored by the quantitative characterization module (Figure 21h).

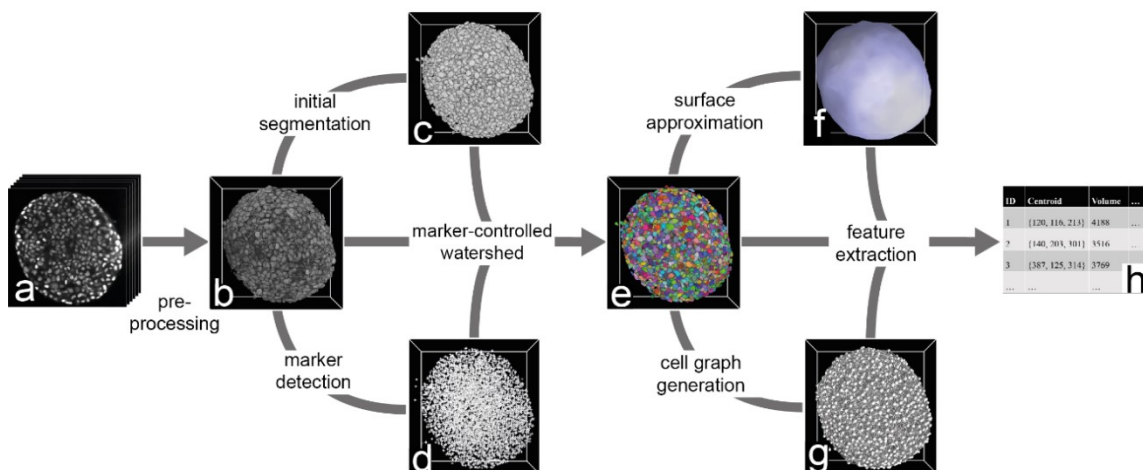


Figure 21 Illustration of the main steps of the quantitative characterization module. (a) The raw data comprises a TIF image stack of two-dimensional optical section images of a spheroid with fluorescently labelled cell nuclei. (b) The image stack is interpolated to obtain isotropic voxels. (c) Multiscale Laplacian of Gaussian (LoG) filtering of the raw image identifies marker points (indicated as white crosses) are identified by. (d) Local thresholding by of the raw image separates regions that contain cell nuclei from the image background. (e) Using the identified seeds, the pre-processed raw image and the initial segmentation a three-dimensional marker-controlled watershed algorithm separates clusters of apparently touching cell nuclei. Different colors indicate individual cell nuclei. (f) An alpha shape is constructed from the cell nuclei locations and the boundary region is extracted as the surface (light blue). (g) Cell graphs are generated where vertices correspond to the cell nuclei (white spheres) and edges indicate the neighborhood relation between two cell nuclei (white lines). (h) Features of each individual cell nucleus and the cell aggregate are extracted and stored in tabular format. Figure adapted from Schmitz *et al.*, 2017.

Several parameters control the procedure of the quantitative characterization (please refer to the Methods section for a complete description of all available parameters). For ease of use, I developed a user interface for setting up the parameter values of the quantitative

characterization. The program comprises two major steps: (1) cell nuclei segmentation (Figure 22) and (2) post-processing which includes alpha shape and cell graph generation (Figure 23). The current version of the program supports image stacks in TIF format and allows processing single of image stacks as well as batch processing of multiple image stacks.

4.5.2 Interactive control of cell nuclei segmentation

In the first part of the program, the user selects an image stack in TIF format from a file browser. The image stack can be interpolated to obtain isotropic voxels and downsampled for fast processing. The control elements on the left side allow adjusting the values of the cell nuclei segmentation parameters. The segmentation parameters are intuitive and mainly associated with the size and intensity of the cell nuclei captured in the image stack. During adjustment of the parameter values, a corresponding segmentation is computed immediately for the selected plane of the image stack and displayed on the right side of the program (Figure 22a). I implemented different visualization modes for displaying the segmentation result including the initial segmentation, the detected marker points and the final segmentation (Figure 22b). Once all parameter values are set up, the program exports the values to a settings file in XLSX format. The user starts the cell nuclei segmentation by selecting a single TIF file or a complete directory of TIF files and the exported settings file. After cell nuclei segmentation, the program exports several files that provide an overview of the segmentation result. These files include visualizations of the segmentation result and files that store the extracted cell nuclei features.

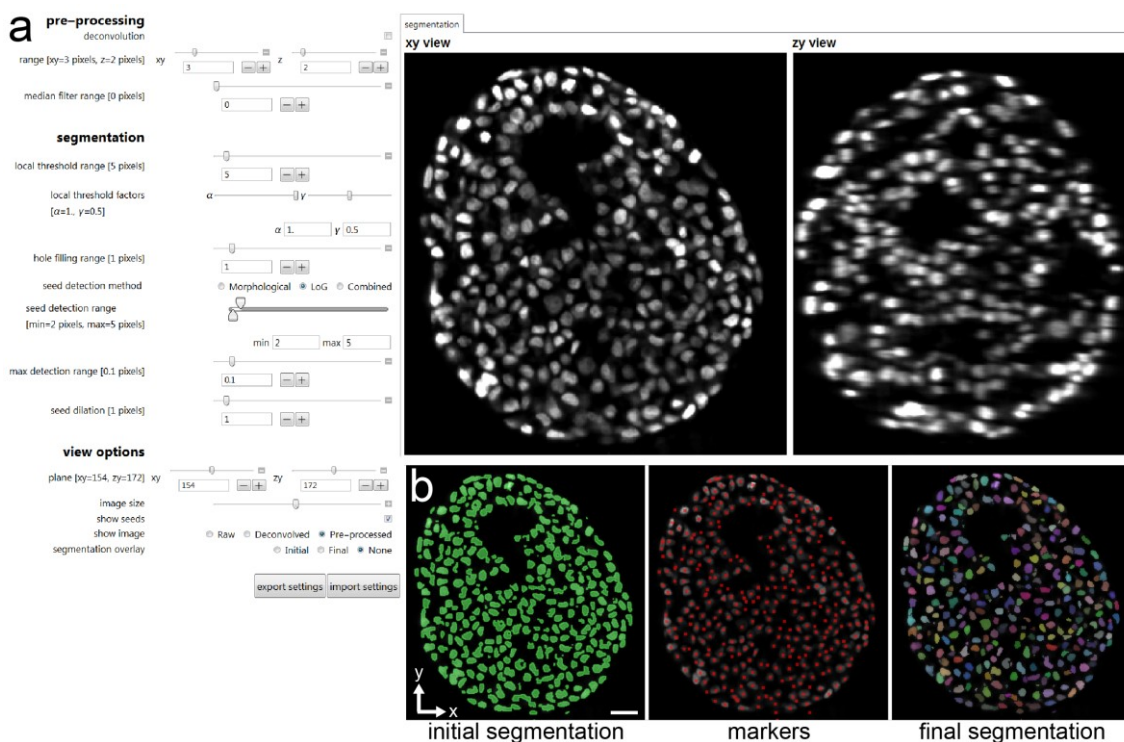


Figure 22 User interface for cell nuclei segmentation. (a) Main user interface with control elements to adjust values of the segmentation parameters and two panels of the current image plane in x-y and z-Y. (b) Different modes of visualization overlaid on the raw image. The initial segmentation, the detected

marker points and the final segmentation result after watershed. Note that, the program will display the segmentation result for the current image plane and continuously update the displayed image. The chosen settings are then used for three-dimensional segmentation. The dataset shown is S9 from application I. Scale bar: 25 μm . Figure adapted from Schmitz *et al.*, 2017.

4.5.3 Post-processing the cell nuclei segmentation results

In the second step, a sequence of post-processing steps yield the final cell nuclei segmentation result (Figure 23). A file containing the extracted cell nuclei features has to be chosen in a file browser. After importing the file, the control elements on the left side allow adjusting the values of the post-processing parameters. The post-processing includes a selection of cell nuclei based on the volume, the surface approximation and the cell graph generation. During adjustment of the parameter values, the corresponding result of post-processing is immediately displayed on the right side of the program (Figure 23a). Different visualization modes can be selected that include a model of the cell nuclei, the surface and a model of the generated cell graphs that is displayed for cell nuclei on the surface of the spheroid (Figure 23b).

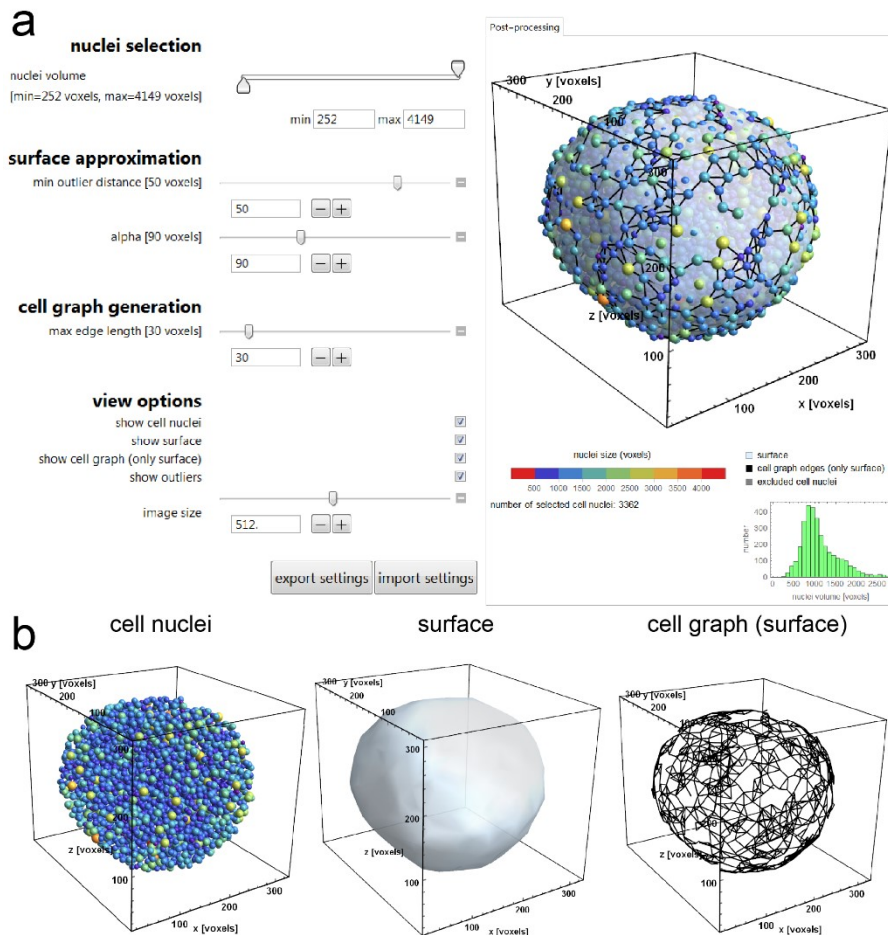


Figure 23 User interface for cell nuclei post-processing. (a) Main user interface with control elements to adjust the values of the post-processing parameters. The program continuously updates the resulting selection of cell nuclei as well as the surface of the alpha shape and the cell graph (shown on the right). (b) Different modes of visualization. The images depict cell nuclei rendered as spheres and colored according to the volume, the alpha shape surface and the cell graph for cell nuclei on the surface. The dataset shown is S9 from application I. Figure adapted from Schmitz *et al.*, 2017.

4.6 Quantitative characterization of spheroids under stress conditions

The developed quantitative characterization module was employed in a number of research projects to demonstrate its broad applicability. The following sections present the results obtained in each application.

4.6.1 Application I: nutrient and oxygen deficiency

High quality image datasets of differently sized T47D spheroids

In this application, we used the quantitative characterization module to profile the morphometric organization of cells in differently sized spheroids generated from T47D breast cancer cells. The project was a collaboration with Christian Mattheyer (PhD student), responsible for spheroid preparation and imaging and Dr. Sabine Fischer (postdoc) who supervised the development of the module and the subsequent data analysis. The datasets served as the test data to develop and evaluate the developed module. Thus, the following paragraphs contain results for the quantitative characterization module and its application to the datasets.

The dataset comprised sixteen T47D spheroids that were seeded from 500 to 10,000 cells, grown for two weeks, optically cleared and, finally, imaged *in toto* with LSMF (Figure 24). Spheroid diameters ranged from 150 μm to more than 500 μm . Small spheroids show a uniform internal morphology (Figure 24, upper row), whereas large spheroids exhibit a visible concentric layered structure (Figure 24, lower row).

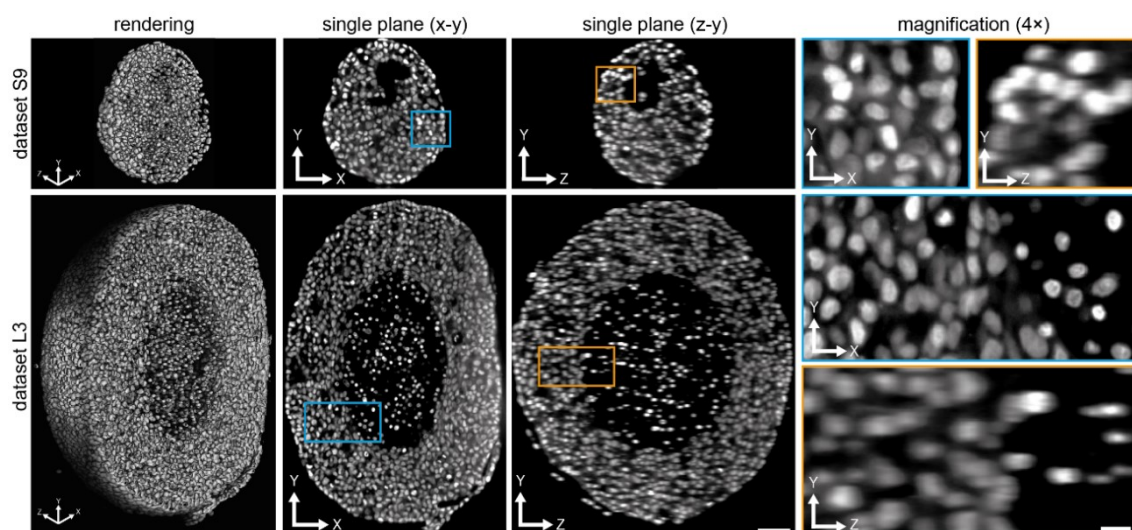


Figure 24 Image quality of three-dimensional datasets. Three-dimensional volume rendering (first column), single planes along x-y (second column), single planes along z-y (third column) and magnification (fourth column) of dataset S9 (upper row) and dataset L3 (lower row). Renderings in the first column clipped at about the center of the spheroids and single planes taken at the same position. Boxes indicate the parts of the images magnified in the fourth column. Scale bars: 50 μm in the second and the third column, 25 μm in the fourth column. Microscope: mDSLm, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 20x NA 0.5 N-Achroplan, camera: Andor Neo, Draq5: 638 nm laser and 680/60 bandpass filter. Figure adapted from Schmitz *et al.*, 2017.

Local thresholding accurately identifies cell nuclei candidate regions

Limitations in imaging large spheroids lead to apparently touching cell nuclei in the images that are difficult to separate using common segmentation approaches (Figure 24, magnification; Figure 25, first row). First, global thresholding algorithms were evaluated for their segmentation performance. We found that most of these algorithms were not capable of accurately detecting cell nuclei in the underlying images. Global thresholding approaches generally suffered from the fact that a single, fixed intensity threshold was computed for the segmentation. Thus, the approaches did not consider local image intensity variations or intensity differences between different cell nuclei. For example, using Otsu's method we observed an insufficient accuracy in the detection of cell nuclei and reconstruction of cell nuclei borders (Figure 25, second row). Besides, the usage of a fixed threshold value resulted in a high proportion of undetected cell nuclei that exhibited low signal intensities (results not shown). After connected component labelling, cell nuclei appeared in large connected clusters (Figure 25, second row, three-dimensional rendering). In contrast to global thresholding, local thresholding algorithms determine the threshold value at a certain location by considering pixel intensities in a window of a fixed range. A threshold algorithm that selects the mean of the local intensity distribution accurately detected cell nuclei in the images independent of the fluorescence intensity distribution of the cell nucleus (Figure 25, third row). For the range parameter, a value that corresponds to the mean radius of a cell nucleus yielded the best segmentation results. Lower values led to over-segmentation of cell nuclei whereas higher values increased the number of clustered objects. An estimation procedure yielded the mean background intensity. The computed value was introduced as an additional factor that increased the locally computed threshold eliminating false detections in the background. The computation of a local threshold in three dimensions is computationally cost expensive. Thus, the runtime efficiency of the algorithm was improved by performing the local thresholding operation per sectional plane along the dimensions x , y and z and multiplying the resulting three-dimensional binary images.

The enhanced thresholding algorithm compensated local intensity variations in the image, was capable of detecting cell nuclei at various intensity levels and reconstructed cell nuclei borders accurately. Connected component labelling of the obtained binary image revealed a better separation of cell nuclei clusters compared to global thresholding (Figure 25, third row, three-dimensional rendering).

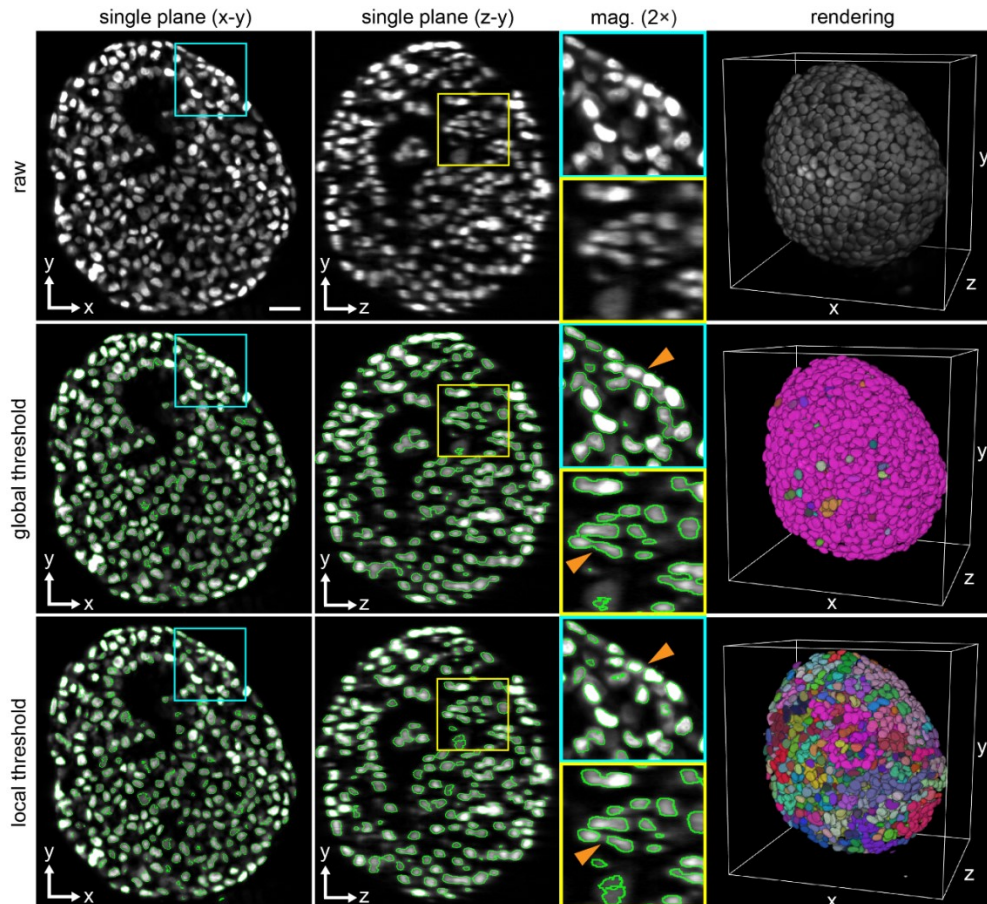


Figure 25 Efficiency of local thresholding in the initial cell nuclei segmentation step. Single image planes along x-y (first column), z-y (second column), magnification and three-dimensional renderings. The green outlines represent the obtained segmentation. Orange arrowheads indicate cell nuclei that are connected in the segmentation resulting from global thresholding (Otsu's method), whereas these clusters are separated in the segmentation resulting from local thresholding. Three-dimensional renderings show the raw image, and colorized components after connected component labelling of the segmentation results for global and local thresholding. Scale bar: 25 μm . Microscope: see caption of Figure 24. Figure adapted from Schmitz *et al.*, 2017.

Reliable detection of marker points by multiscale Laplacian of Gaussian filtering

Many cell nuclei remained connected and appeared in large clusters of multiple cell nuclei. For a complete separation of these clusters, we developed a marker-controlled watershed segmentation. The marker point detection uses multiscale Laplacian of Gaussian (LoG) filtering, followed by the detection of local maxima and subsequent selection of the detected maxima (Figure 26). A general property of the LoG filter is that the magnitude of the LoG response is maximal at the center of a cell nucleus given that the scale of the LoG matches the size of the cell nucleus (Figure 26b). By incorporating multiple scales, the seed detection algorithm is more robust towards variations in cell nuclei volume.

The minimal and maximal radii (r_{min} and r_{max}) of cell nuclei were measured in the images and used to determine the minimal and maximal scales σ_{min} and σ_{max} . In the underlying approach, LoG filtering with a range of scales $\{\sigma_{min}, \dots, \sigma_i, \dots, \sigma_{max}\}$ was then applied and a LoG scale-space maximum intensity projection was computed (Figure 26c). A subsequent detection of local maxima resulted in a number of candidate marker points (Figure 26d). To remove

detections that occurred in the background region of the image, the image containing the detected maxima was multiplied with the initial binary image (Figure 26e). The multiscale LoG filter algorithm needs careful adjustment of minimal and maximal scales σ_{min} and σ_{max} . Filter ranges that did not correspond to the size of the cell nuclei lead to under- or over-segmentation.

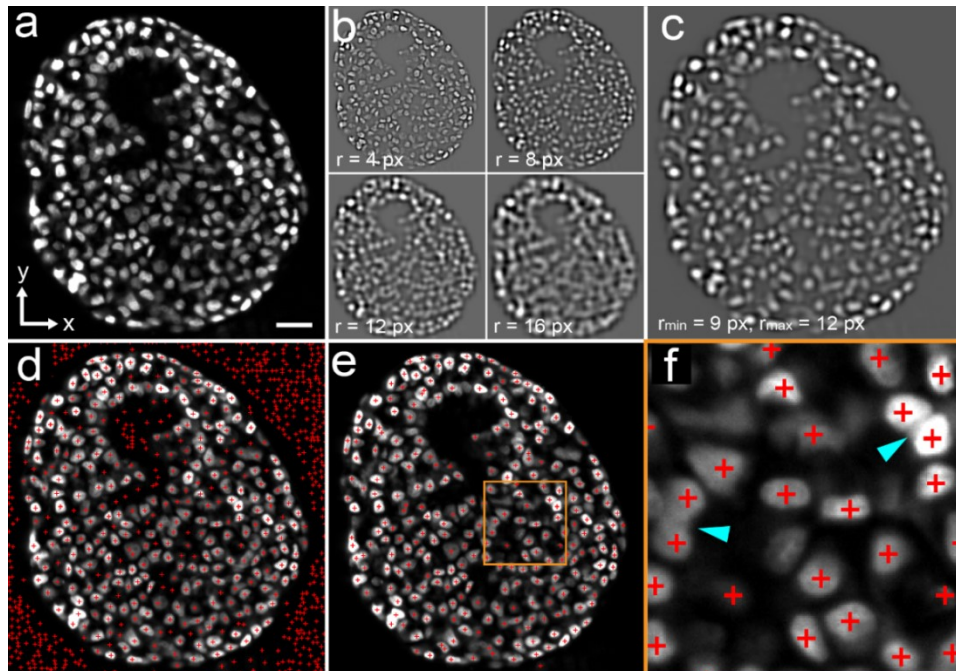


Figure 26 Marker point detection with the Laplacian of Gaussian filter. (a) Sub-region of a single plane of the raw volume along X-Y. (b) Resulting images of the Laplacian of Gaussian (LoG) filter algorithm applied to the image in (a) for different values of the LoG filter range r in pixels. (c) The response images for different values of r are combined to provide a maximum response image. (d) Marker points detected by an extended local maxima search displayed as red crosses in the raw image. (e) Detections in the background region are discarded by multiplying the detected maxima image with the initial segmentation result. (f) Magnification of the indicated region (orange box) in panel e. The marker points accurately reflect the locations of cell nuclei. For apparently touching cell nuclei (cyan arrowheads), unique marker points are identified that define starting points for the subsequent watershed algorithm. Please note that for the purpose of this illustration, the marker points were computed in two dimensions. Marker points are computed in three dimensions. Scale bar: 25 μm . Microscope: see caption of Figure 24. Figure adapted from Schmitz *et al.*, 2017.

We found that the multiscale variant of the LoG filter approach is robust towards nuclei volume and intensity variations (Figure 27a, second row). The obtained marker points were incorporated into a marker controlled watershed immersion algorithm to improve the separation of cell nuclei. The marker points guide the watershed segmentation, thereby eliminating the problems of over-segmentation commonly associated with watershed algorithms. Due to the operation principle of the algorithm, the number of markers determines the number of objects that are extracted by the watershed immersion algorithm. The intrinsic formation of watershed surfaces during the immersion process results in a subdivision of clusters into individual components (Figure 27a, third row). Compared to the cell nuclei clusters obtained after the initial segmentation step (Figure 27b), the final segmentation result provides a reliable identification of the cell nuclei that can be used to extract features from individual cell nuclei (Figure 27c).

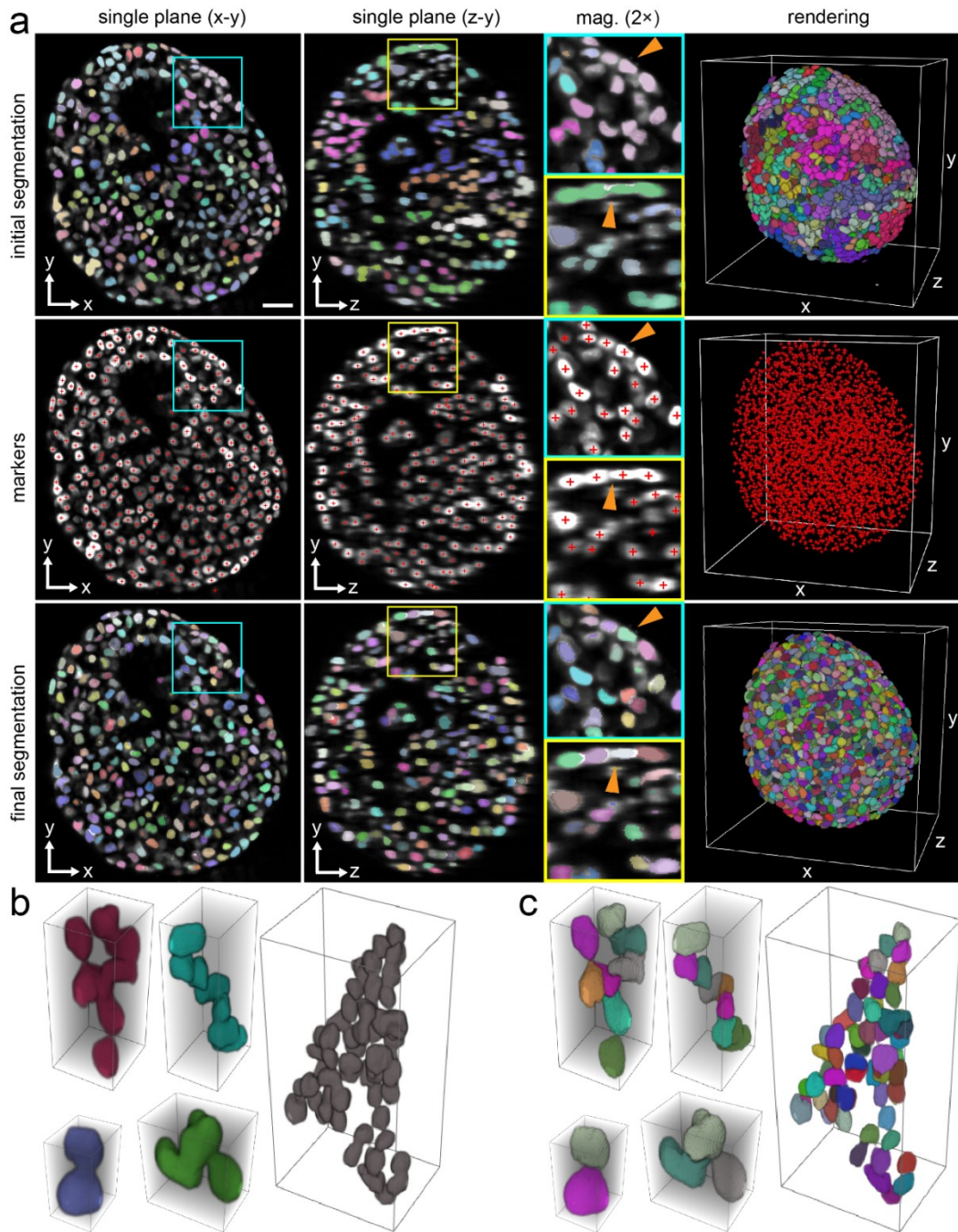


Figure 27 Marker-controlled watershed segmentation improves the separation of cell nuclei. (a) Single planes along x-y (first column), z-y (second column), two-fold magnifications (mag. 2×) of cyan and yellow boxes (third column) and three-dimensional renderings (fourth column). The individual steps of the cell nuclei segmentation are shown for dataset S9. Shown are single planes along x-y (first column), z-y (second column) at about the center of the spheroid, two-fold magnifications (mag. 2×) of regions indicated by cyan and orange boxes (third column) and three-dimensional renderings (fourth column). The obtained components after the initial segmentation are colored and overlaid on the raw image (first row). The range parameter r for the initial segmentation was set to 12 pixels. The initial segmentation identifies the foreground region in the image, but fails to separate apparently touching cell nuclei (orange arrowheads). Second row: the LoG filter identifies unique marker points (indicated by red crosses). Parameter values for r_{min} and r_{max} were set to three and six, respectively. In the final segmentation (third row), apparently touching cell nuclei are separated (orange arrowheads). Three-dimensional renderings of clusters of connected cell nuclei after connected component labelling of the initial segmentation (b) and after the final segmentation (c). Different colors represent individual components. Please note that to generate the images in the third row, the marker points were computed in two dimensions. In the quantitative characterization module, the marker detection is three-dimensional. Scale bar: 25 μm . Microscope: see caption of Figure 24. Figure adapted from Schmitz *et al.*, 2017.

Suitable parameter values for the quantitative characterization module were empirically determined and the same set of parameter values was applied for processing all 16 datasets (Table 5). All parameters of the cell nuclei segmentation are intuitive and mainly associated with the size of the cell nuclei. We found that the range parameters for the initial segmentation and marker point detection were the most important parameters that needed careful adjustment. The average diameter of the cell nuclei provided a good starting value for these parameters. However, we observed high robustness of the segmentation results with respect to changes in the parameter values (Figure 28). The processing time for cell nuclei segmentation was less than one hour per dataset (Supplementary Table 1).

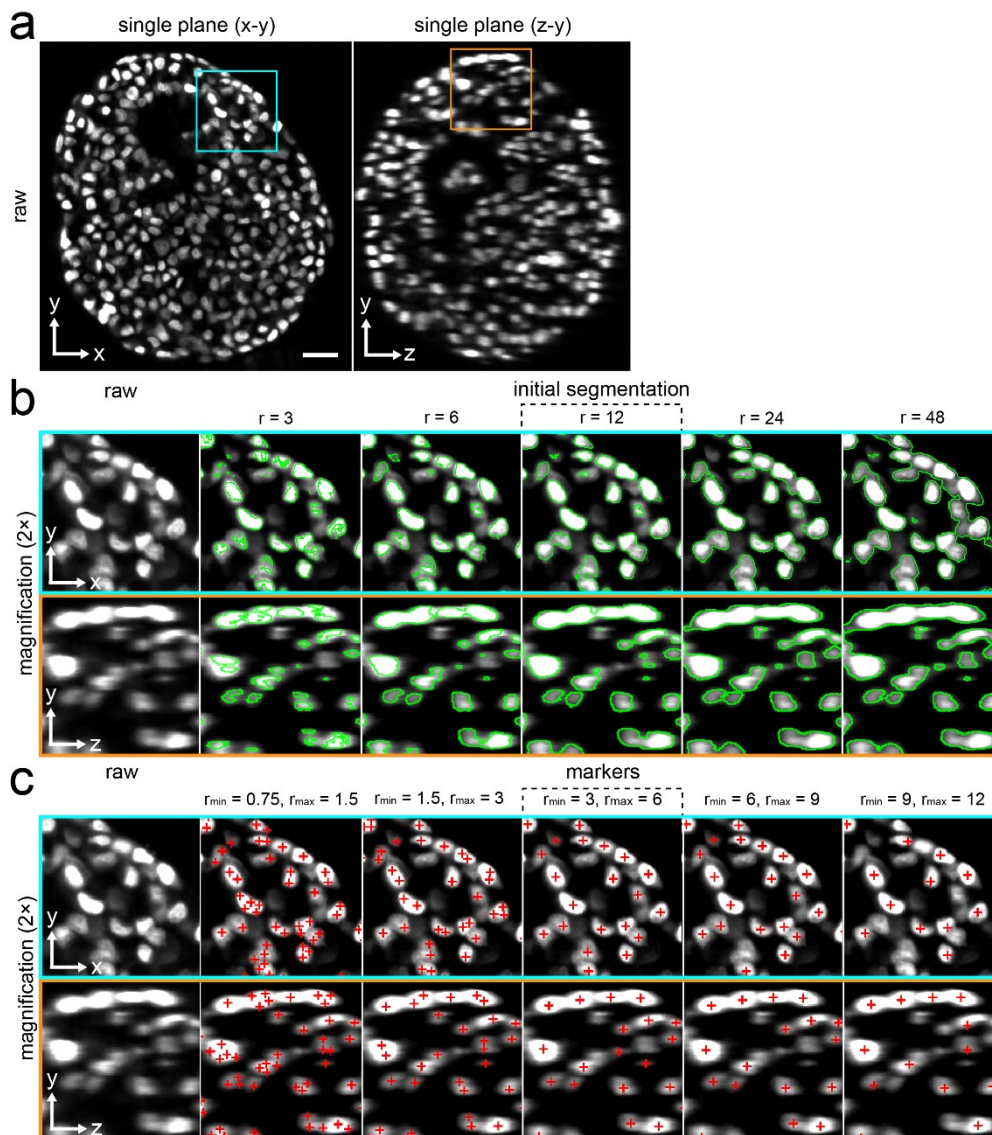


Figure 28 Evaluation of parameter values for initial segmentation and marker point detection. (a) Representative single planes along x-y and z-y of dataset S9 at the center of the spheroid. Scale bar: 25 μm . The results of initial segmentation (outlined in green) for different values of the range parameter r (b) and the results of marker point detection (depicted by red crosses) for different values of the parameters r_{min} and r_{max} (c) are shown in two-fold magnification for the regions indicated by cyan and orange boxes in panel a. The parameter values that were used for in our study are indicated by a dashed box. Please note that the initial segmentation was applied as described in the Methods section, whereas for illustration purposes the marker point detection was applied in two dimensions. In the quantitative characterization module, the marker point detection is performed in three dimensions. A starting value for

the range parameter of the initial segmentation is the average diameter of cell nuclei (about 12 pixels in the underlying datasets). However, the initial segmentation proved to be robust towards small deviations. Over-segmentation and under-segmentation occur for large deviations. For the marker point detection, reliable values for the parameters for r_{min} and r_{max} are the minimal and maximal radius of cell nuclei in the dataset (three pixels for r_{min} and six pixels for r_{max}). The marker point detection is more sensitive to the choice of the parameter values. For example, lower values (e.g. $r_{min}=1.5$, $r_{max} = 3$) lead to the detection of parts of cell nuclei, whereas for higher values (e.g. $r_{min}=9$, $r_{max} = 12$) clusters of cell nuclei are detected instead of individual cell nuclei. Microscope: see caption of Figure 24. Figure adapted from Schmitz *et al.*, 2017.

In summary, local thresholding provided an improvement in the segmentation of cell nuclei compared to global thresholding. Less clustered cell nuclei were obtained and the borders of cell nuclei were accurately reconstructed. The adaptable multiscale Laplacian of Gaussian (LoG) filter algorithm detects marker points reliably and reproducibly. The marker-controlled watershed is capable of separating the remaining clusters of cell nuclei into individual components.

High segmentation performance in large, densely packed spheroids

The segmentation performance was evaluated for three different regions *outer*, *interface* and *core* (Figure 29a). These regions represent the variability in cell nuclei morphology and density of cells. Independent of cell density and morphology, the cell nuclei segmentation accurately identified cell nuclei in all regions.

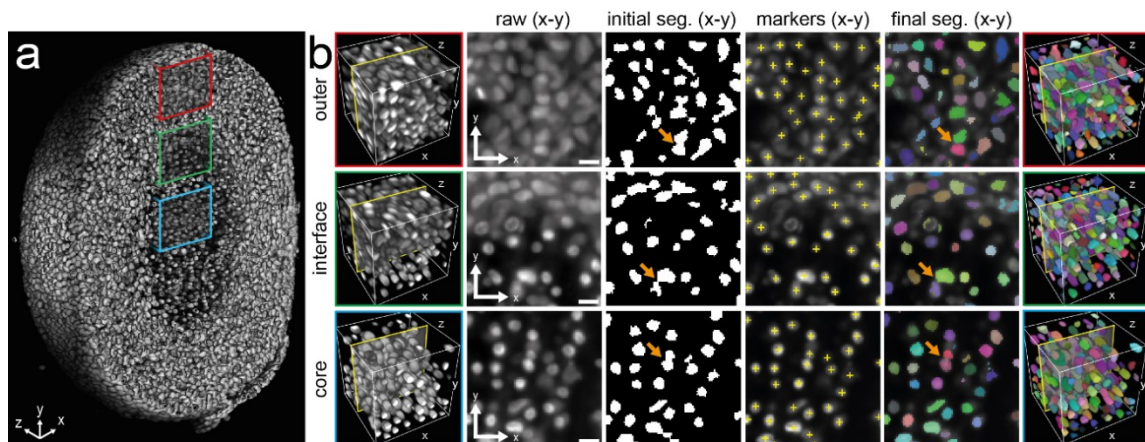


Figure 29 Quality of the cell nuclei segmentation in different regions of a large T47D spheroid. (a) Three-dimensional rendering of cell nuclei clipped at about the center of a large spheroid (dataset L3) that comprises 38,783 cells. Colored boxes indicate three regions that exhibit visually distinguishable properties. Red box: Cell nuclei in the outer region (*outer*) appear tightly packed. Green box: cell nuclei at the transition zone between outer and inner region (*interface*) exhibit diverse morphologies and the intercellular distances differ. Blue box: cell nuclei in the core region of the spheroid (*core*) are small, spherical and appear well separated. (b) Exemplary cell nuclei segmentation results for the three regions *outer*, *interface* and *core*. First column: sub-regions of $100 \times 100 \times 100$ voxels were copied out of the pre-processed raw image. Second column: plane 58 of each sub-region in x-y view. Third column: the result of the initial segmentation for plane 58. The initial segmentation accurately identifies the foreground in the image. Locations marked with orange arrows show cell nuclei clusters that are not separated. Fourth column: seeds detected by the multiscale LoG filter are overlaid as yellow crosses. Note that for illustration purposes, the seed detection was performed in two dimensions. Fifth column: overlay of the final cell nuclei segmentation of plane 58 after three-dimensional marker-controlled watershed. This step effectively separates apparently touching cell nuclei (orange arrows). Sixth column: the final segmentation result after three-dimensional marker-controlled watershed of the sub-regions shown in the

first column. Different colors represent individual cell nuclei. Scale bar: 10 μm . Microscope: see caption of Figure 24. Figure adapted from Schmitz *et al.*, 2017.

Next, we conducted a quantitative evaluation of the segmentation performance (Table 7). We used sub-regions (shown in Figure 29b) and generated ground truth datasets (GT) by visually extracting the locations of all cell nuclei with a custom program (Mathew *et al.*, 2015). We matched the centroids in the GT with those of the segmentation and determined the number of correctly detected (true positives), falsely detected (false positives) and undetected cell nuclei (false negatives). The metrics recall, precision and F score summarize the segmentation performance. The segmentation identified 216 out of 252 cell nuclei and 9 false detections occurred in the core region (recall: 0.86, precision: 0.96, F score: 0.91). In the interface region, 230 out of 269 cell nuclei were correctly identified and the number of false detections was 32 (recall: 0.86, precision: 0.88, F score: 0.87). In the outer region, 198 out of 233 cell nuclei were correctly identified and 24 false detections occurred (recall: 0.85, precision: 0.89, F score: 0.87). Overall, the segmentation achieved an average recall of 0.86 an average precision of 0.91 and an average F score of 0.88. The segmentation performance exceeds that of established approaches that were evaluated for T47D spheroids in previous studies (Mathew *et al.*, 2015).

Table 7 Evaluation of segmentation performance for different regions within the spheroid. The performance was measured against a manually determined ground truth for the regions outer, interface and core. The performance measures recall, precision and F score are determined from the number of true positives (TP), false negatives (FN) and false positives (FP). Values range from 0 (worst performance) to 1 (best performance). GT, number of cell nuclei in the ground truth; SC, number of cell nuclei determined by the segmentation; TP, true positives; FN, false negatives; FP, false positives. Table adapted from Schmitz *et al.*, 2017.

region	GT	SC	TP	FN	FP	recall	precision	F score
outer	233	222	198	35	24	0.85	0.89	0.87
interface	269	262	230	39	32	0.86	0.88	0.87
core	252	225	216	36	9	0.86	0.96	0.91

Alpha shapes as a geometrical model for spheroids

We distinguished low-level features of the cell nuclei directly obtained from the segmentation and higher order features of the cell neighborhood and the whole spheroid. The output of our quantitative characterization is a comprehensive set of features (Table 4). We used the alpha shape approach based on the cell nuclei centroids to compute a geometrical shape model of the spheroid. The alpha shape approach is based on a Delaunay triangulation and works well for point sets P of high density and uniform distribution. For the corresponding parameter $\alpha \rightarrow 0$, the alpha shape converges to P , whereas for $\alpha \rightarrow \infty$ the alpha shape converges to the convex hull of P . For the underlying datasets, the alpha shape smoothly approximated the spheroid surface using an alpha value of 90 voxels (Figure 30, Supplementary Figure 1).

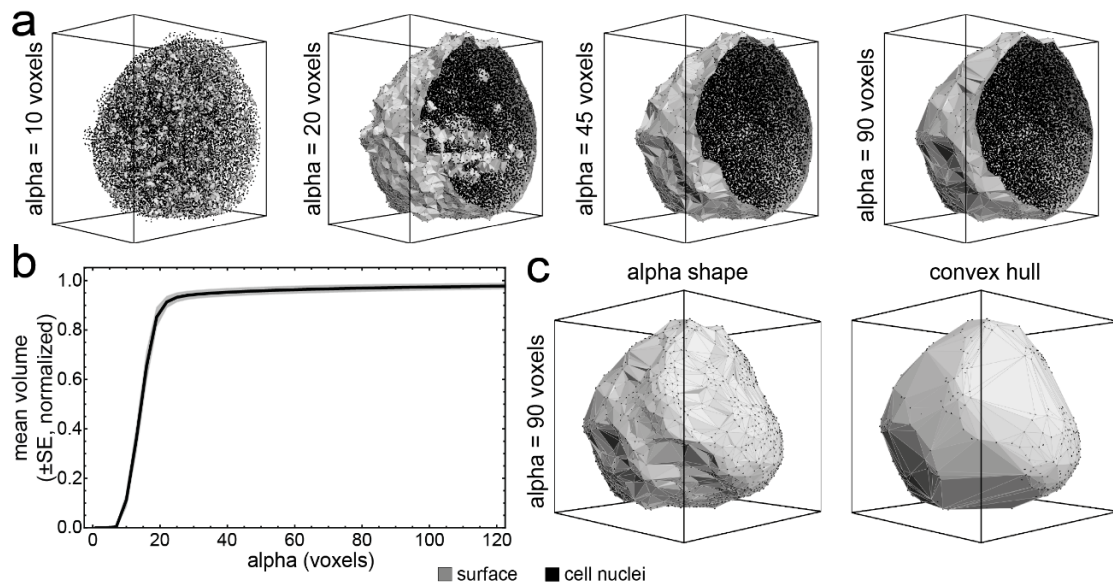


Figure 30 Surface approximation using alpha shapes. (a) Surface approximation of dataset M3 for different values of alpha. Black points show the cell nuclei locations. (b) Mean volume of all datasets normalized to the volume of the convex hull as a function of the alpha parameter value. The shaded region indicates the standard error of the mean (SE). (c) Alpha shape approximation with alpha set to 90 voxels (left) and convex hull (right) for dataset M3. Figure adapted from Schmitz *et al.*, 2017.

From the alpha shape and its surface, we determined the volume, surface area, centroid and the minimal distance of the centroid to the surface. The quantitative characterization module identified a spheroid volume that was proportional to the number of cells with a slope of $1,127 \mu\text{m}^3$ per cell (Figure 31b). An automated cluster analysis (partitioning around medoids with squared Euclidean distance) of spheroid volume and cell number enabled a separation of the datasets into groups of small ($n=9$), medium-sized ($n=3$) and large ($n=4$) spheroids. Each group member was given a unique identifier, S1-S9 for small, M1-M3 for medium and L1-L4 for large spheroids.

For each group we determined the mean and standard deviation of the cell numbers (small: $4,977 \pm 2,766$, medium: $25,582 \pm 454$, large: $34,742 \pm 2,941$) and the mean and standard deviation of the spheroid volume (small: $6.0 \times 10^6 \pm 3.1 \times 10^6 \mu\text{m}^3$, medium: $2.7 \times 10^7 \pm 0.04 \times 10^7 \mu\text{m}^3$, large: $4.0 \times 10^7 \pm 0.2 \times 10^7 \mu\text{m}^3$, Supplementary Table 2). The cell nuclei volumes of each group did not follow a normal distribution (Figure 31c). The distributions were with positive skewness (small: 1.77, medium: 1.71, large: 1.71) and high kurtosis values (small: 8.48, medium: 7.63, large: 7.70). The median cell nucleus volume in medium and large spheroids was slightly lower compared to small spheroids (small: $267 \pm 68 \mu\text{m}^3$, medium: $234 \pm 51 \mu\text{m}^3$, large: $225 \pm 61 \mu\text{m}^3$). All distributions showed a peak around $250 \mu\text{m}^3$ and a prominent shoulder region at around $450 \mu\text{m}^3$. We examined cell nuclei with volumes between 300 and $600 \mu\text{m}^3$, but could not identify a pattern in their spatial localization (Supplementary Figure 2).

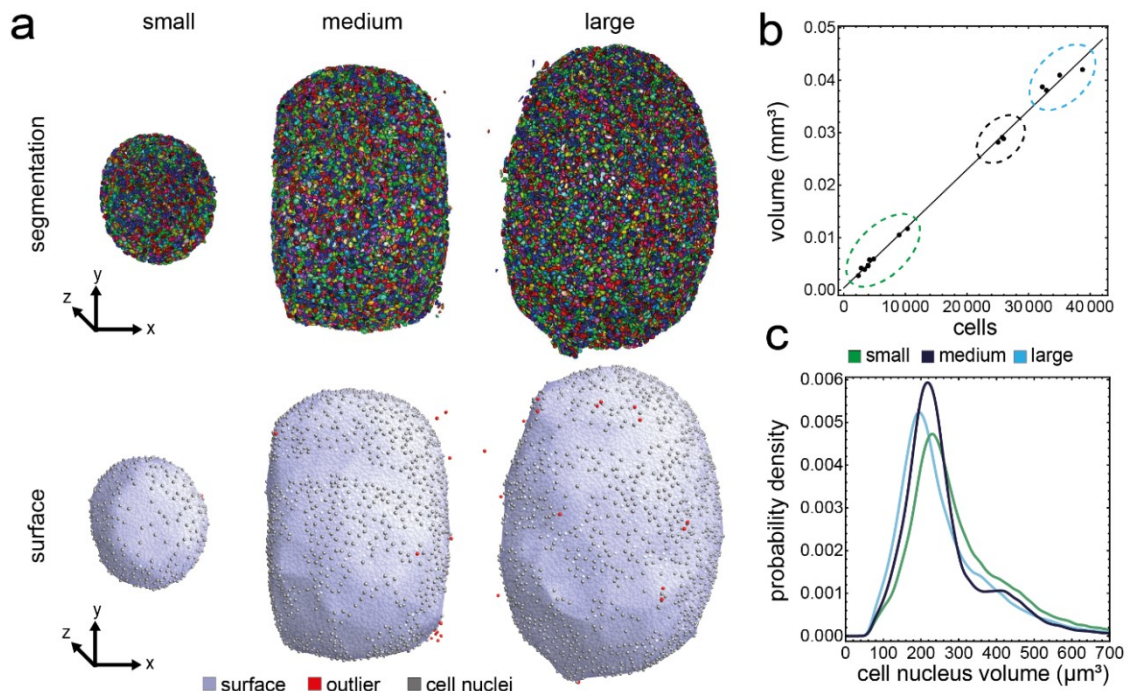


Figure 31 Surface approximation and characterization of T47D spheroids. (a) Results of cell nuclei segmentation (first row) for a small (first column, dataset S7), medium (second column, dataset M2) and large (third column, dataset L3) spheroid. Different colors indicate individual cell nuclei found by the segmentation. Second row: the obtained surface approximation using the alpha shape approach. Outliers that are removed prior to surface approximation are depicted in red, whereas cell nuclei are shown in gray. (b) Plot of spheroid volume versus number of cells detected in the spheroid for all datasets. The measurements are fitted well by a linear model with slope $1,127 \mu\text{m}^3$. According to cell number and spheroid volume, clustering analysis of the datasets resulted in three groups: small (nine datasets, green ellipse), medium (three datasets, dark blue ellipse) and large spheroids (four datasets, light blue ellipse). (c) Smoothed histogram of the cell nuclei volume distribution in small, medium and large spheroids. Figure modified from Schmitz *et al.*, 2017.

A comprehensive set of features for the quantitative characterization of spheroids

For each cell nucleus, we computed the shortest distance between its centroid and the surface of the spheroid as a measure for its relative position. The distance was normalized to the maximum distance of a cell nucleus yielding the normalized distance to the surface (NDS). Consequently, cell nuclei with an NDS of 0 are at the surface and those with an NDS of 1 are in the center of a spheroid.

We computed two cell graphs in order to analyze the spatial arrangement of cells in spheroids. In both graphs, the vertices correspond to cell nuclei and edges, i.e. pairs of vertices, represent a neighborhood relationship. In the proximity cell graph (PCG), edges are solely created according to the Euclidean distance. In the Delaunay cell graph (DCG), edges are further restricted to vertices that are connected by a line in the corresponding Delaunay triangulation. Both cell graphs share a distance threshold of 40 voxels ($\sim 25 \mu\text{m}$) for edges. The number of neighbors (i.e. the vertex degree) and the distance to the neighbors (i.e. the weights of all incident edges) of a vertex are measures of cell density. In the following, the number of neighbors for a vertex in the PCG is referred to as the local cell density in cells/unit volume (cells/u.v.), where a unit volume corresponds to $65,450 \mu\text{m}^3$.

Cell graph features reveal structural heterogeneity at multiple scales

Cell nuclei features directly obtained from the segmentation did not show any variation along the radial direction of the spheroids. Features derived from the cell graphs by contrast varied along the radial direction in medium and large spheroids (Figure 32a, Supplementary Figure 3).

To assess the randomness of such a pattern, we placed cells randomly into the alpha shape of each dataset, computed the cell graphs and compared the derived features with those found for the real datasets. The curves of the cell density for medium and large spheroids strongly deviated from the random cell position (RCP) model, whereas the curves for small spheroids was similar to that of the RCP model (Figure 32b). Thus, the spatial distribution of cells in medium and large spheroids deviates from randomness, whereas randomly positioning the cells could generate the spatial distribution in small spheroids. We therefore analyzed local cell density features in medium and large spheroids. Cells at the surface have only few neighboring cells (mean cell density ≈ 45 cells/u.v.). At 0.5 NDS, the cell density decreases from 65 to 40 cells/u.v. in large spheroids, whereas a similar decrease from 60 to 45 cells/u.v. is observed at 0.75 NDS in medium spheroids (Figure 32a).

For a comparison of our results to the concentric cell layering found in large spheroids (Kunz-Schughart et al., 2004), we subdivided the spheroids into three distinct regions *surface*, *outer* and *core* based on the computed NDS. For the separation of surface and outer region, we took the transition point between the initial rise and the approximately constant region, resulting in a threshold of 0.1 NDS for medium and large spheroids. For the separation of outer and core region we took the transition point between the approximately constant region and the decreasing part resulting in 0.75 NDS for medium and 0.5 NDS for large spheroids. We could identify a difference in the mean cell density for the outer and core region in medium and large spheroids (Figure 32c). The mean cell density between outer and core region differs by 9.4% (outer: 61 cells/u.v., core: 56 cells/u.v.) in medium spheroids, whereas in large spheroids, the difference is 17.7% (outer: 65 cells/u.v., core: 53 cells/u.v.). A coloring of the segmented cell nuclei according to the cell density underpins this finding (Supplementary Figure 4). In the central and surface regions of the spheroids, the cell density is low, whereas the cell density is higher for cells between these two regions (Figure 32d). The cell density within the regions of all spheroids was inhomogeneous with patches of higher and lower cell density (Supplementary Figure 4). The mean distance to neighbors determined from the DCG was constant (~ 13 μm) up to 0.75 NDS in medium spheroids and increased to 15.5 μm in the core. A similar increase was observed in large spheroids at 0.5 NDS (Figure 32e). Based on the quantitative measures, we determined the average radius of the spheroids and the average thickness of the outer region in micrometers (Figure 32f). Medium spheroids have an average radius of 191.3 μm and large spheroids have an average radius of 213.7 μm . The thickness of the outer region is 143.5 μm in

medium spheroids and 106.8 μm in large spheroids. We highlighted the detected border between outer and core region in the raw data and found that the quantitatively identified location of the border fits well with the visually observable boundary (Figure 32g).

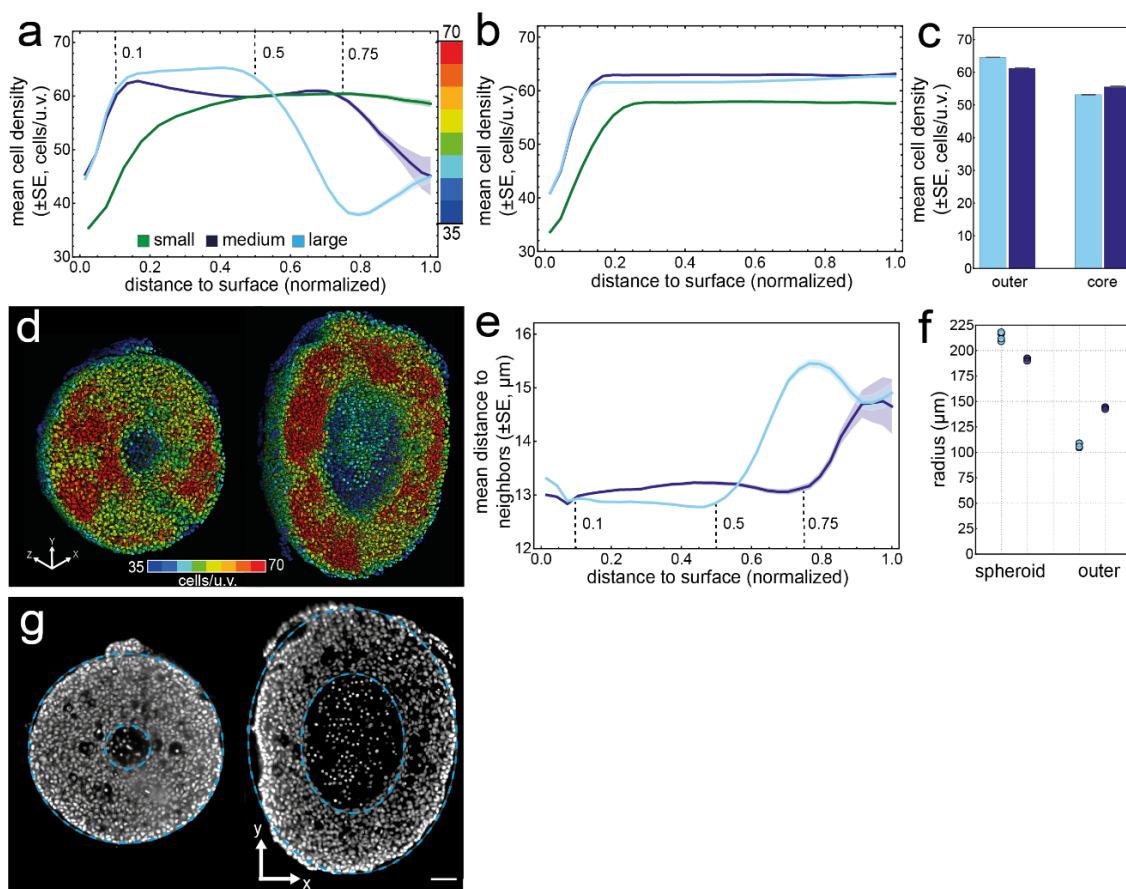


Figure 32 Medium and large spheroids show differences in internal structure. Plot of mean cell density versus the normalized distance to the surface for small, medium and large spheroids (a) and the corresponding random cell position (RCP) models (b). The RCP models represent an exact analogue of each spheroid with the only difference being that cell nuclei are randomly positioned. Details about RCP model generation can be found in the Materials and Methods section. The shaded regions indicate the standard error of the mean (SE). Based on the normalized distance to the surface, cell nuclei in the range [0.1, 0.75] for medium spheroids, and [0.1, 0.5] for large spheroids are assigned to the outer region, whereas cell nuclei in the range [0.75, 0.1] for medium and [0.5, 0.1] for large spheroids are assigned to the core region. (c) Mean cell density in the outer and core region for medium and large spheroids. (d) Three-dimensional rendering of segmented cell nuclei colored according to their corresponding cell density value for a medium (dataset M3) and a large spheroid (dataset L2), ranging from blue (35 cells) to red (70 cells). Renderings were clipped at about the center of the spheroids. (e) Plot of the mean distance to neighbors versus the normalized distance to the surface. The shaded regions indicate the standard error of the mean (SE). (f) Radius of medium and large spheroids and thickness of the outer region. (g) Raw single planes at the center of the spheroids shown in D with cyan dashed lines indicating the thickness of the outer region. Scale bar: 50 μm . Figure modified from Schmitz *et al.*, 2017.

In summary, features capturing local cell packing showed variations along the radial direction and can be used to identify structurally different zones within medium and large spheroids.

4.6.2 Application II: apoptosis induction

In this application, the quantitative characterization module enabled the evaluation the long-term phenotypic effects of the apoptosis-inducing drug Staurosporine on spheroids generated from T47D breast cancer cells. Differently sized spheroids were generated from cell

suspensions containing a seed number of 500, 1,000, 2,000 or 5,000 T47D breast carcinoma cells. After twelve days of formation, the spheroids were treated with Staurosporine. Spheroids were fixed, stained and optically cleared after 0, 24, 48 and 72 hours post treatment (p.t.). Spheroids cultivated for 72 hours in 0.1% DMSO containing medium served as negative control (NC) for time point 72 hours. Negative controls for other time points were not available. Image stacks were acquired using a Carl Zeiss LSM 780 laser scanning confocal microscope. Nariman Ansari (postdoc) performed spheroid preparation, treatment and imaging. The acquired image stacks provided an *in toto* view on the morphology of spheroids and individual cell nuclei (Figure 33). For some conditions, an insufficient number of datasets was available.

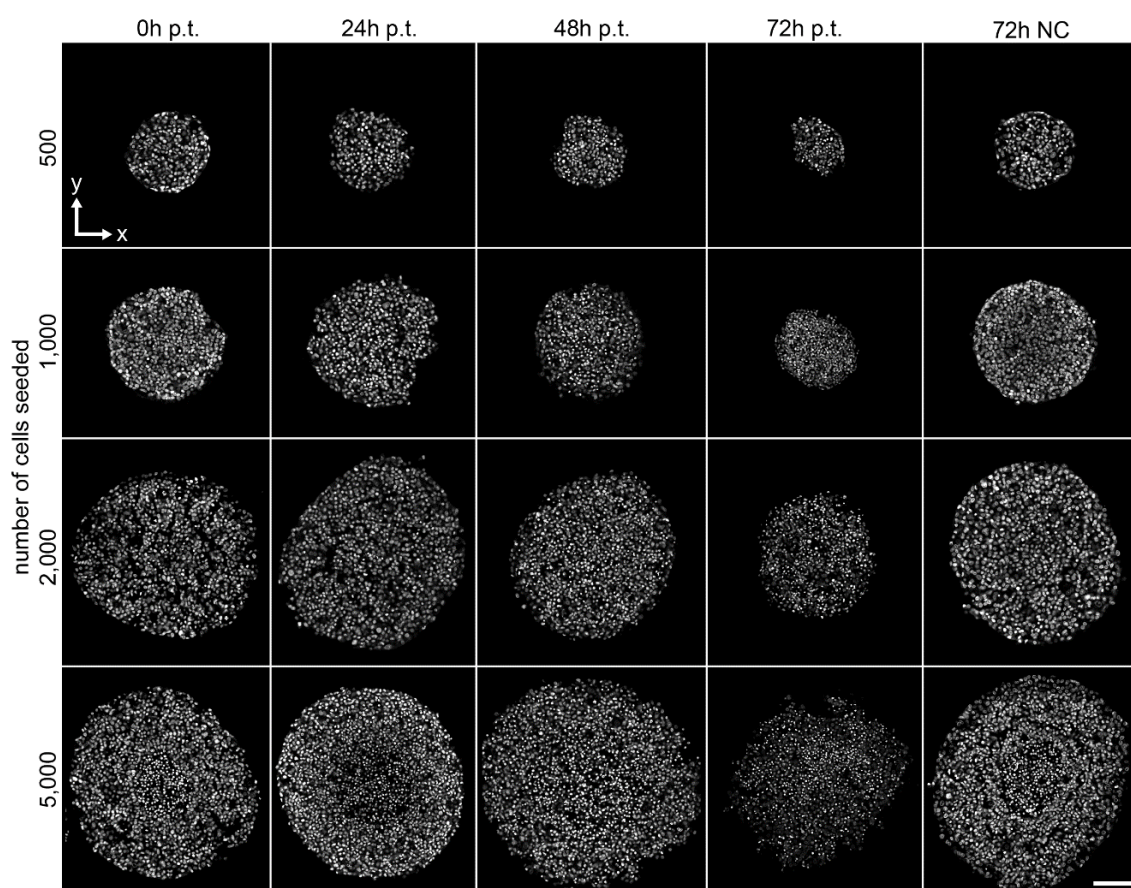


Figure 33 Examples of a Staurosporine-treated and untreated T47D spheroids. Depicted are single planes in x-y of selected spheroids for different numbers of cells seeded and time points after Staurosporine treatment. Single planes were taken at about the center of the spheroids. h, hours; p.t., post treatment; NC, negative control. Scale bar: 100 μ m. Microscope: Carl Zeiss LSM 780, objective lens: 25x0.8 NA LD LCI Plan-Apochromat.

We quantitatively analyzed 80 spheroids using the quantitative characterization module (Supplementary Table 3). Suitable parameter values were empirically determined and the same set of parameter values was applied for processing all datasets (Table 5). The segmentation results were empirically evaluated (Figure 34).

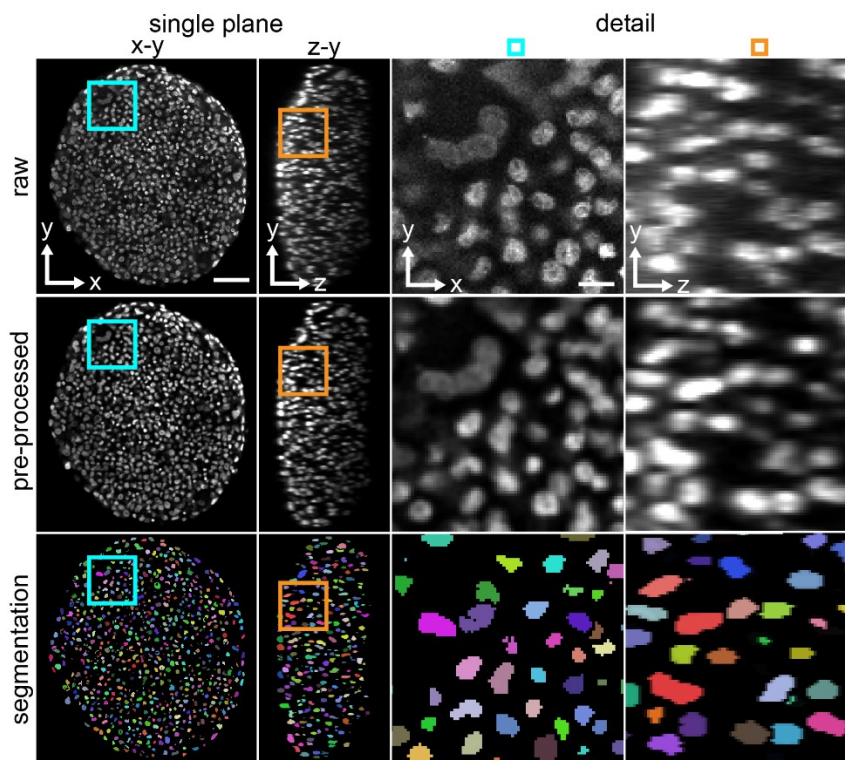


Figure 34 Cell nuclei segmentation performance in T47D spheroids. Single planes along x-y and z-y of raw images, pre-processed images and the corresponding cell nuclei segmentation. Different colors indicate individual cell nuclei found by the segmentation. Scale bar 50 μm . Detail images in the right column show the regions in panel a, indicated by cyan and orange boxes. Scale bar: 10 μm . Microscope: see caption of Figure 33.

Spheroid volume and cell number decrease upon Staurosporine treatment

We first evaluated the effects of Staurosporine on the level of the complete spheroid (Figure 35). The alpha shape generated by the quantitative characterization module provided a geometrical model of the spheroids that could be readily used to analyze shape and volume. We observed an overall size reduction of spheroids that were treated with Staurosporine compared to untreated spheroids (Supplementary Figure 5). We determined the average spheroid volume per condition (Figure 35a, Supplementary Table 4). After 72 hours p.t., the spheroid volume for 2,000 and 5,000 seeded cells was lower compared to spheroids of the negative control (Figure 35b). Treated spheroids seeded from 2,000 cells had 32% and spheroids seeded from 5,000 cells had 57% of the volume of the negative control. For spheroids seeded from 500 cells, not enough datasets were available after 72 hours.

The cell number in spheroids treated with Staurosporine also decreased as a function of time post treatment (Figure 35b, Supplementary Table 5). For 1,000 and 2,000 seeded cells, the number of cells in each spheroid after 72 hours of Staurosporine treatment was lower than for spheroids in the negative control. For example, the number of cells in treated spheroids seeded from 1,000 cells was only about one third of the final cell number in untreated spheroids (2,800 vs. 9,078 cells). No differences were detected for 500 and 5,000 seeded cells.

Spheroid shape is not affected by apoptosis induction

From a visual inspection of the alpha shapes generated by the quantitative characterization module, we hypothesized that Staurosporine treatment could also affect the spheroid shape. Thus, we performed a principal component analysis (PCA) of the cell nuclei locations and determined the principal directions ($p1$, $p2$ and $p3$) and their absolute extension (Figure 35c, Supplementary Table 6). For a comparison of differently sized spheroids, we computed the relative extension by normalizing the absolute extension of $p2$ and $p3$ to the absolute extension of the first component (Figure 35d). All spheroids had an oblate shape, i.e. the relative extension along $p2$ was almost identical to $p1$, whereas the relative extension along $p3$ was reduced to about 0.5. However, no shape differences were detected between spheroids treated for 72 hours with Staurosporine and the negative control. We concluded that Staurosporine treatment induced a reduction of spheroid volume and cell number, but it did not affect the spheroid shape.

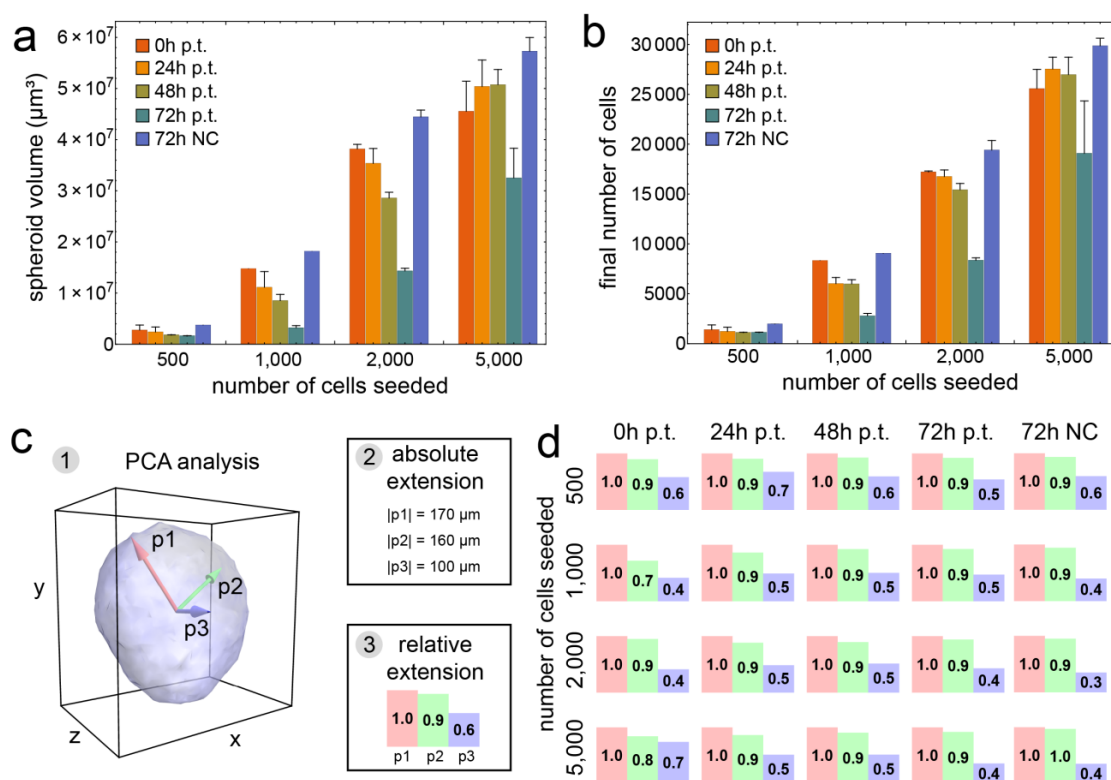


Figure 35 Comparison of T47D spheroid volume, cell numbers and shape for all conditions. (a) Plot of the average spheroid volume as a function of cell seed number and time point compared to the negative control. (b) Plot of the average final cell number as a function of cell seed number and time point compared to the negative control. Error bars in plots indicate the standard error of the mean. Hypothesis testing was performed for condition 72h p.t. against 72h NC. h, hours, p.t., post treatment. NC, negative control. (c) Shape analysis methodology. Principal component analysis (PCA) gives the principal directions $p1$, $p2$ and $p3$ of the spheroid (1). The absolute extension measures the extension of the spheroid in μm (2). Normalization to the absolute extension of the first principal direction gives the relative extension (3). (d) Relative extension of principal directions for all conditions averaged over spheroids of the same condition. Please note that for some conditions only a few spheroids were available for the analysis (Supplementary Table 3).

The volume of cell nuclei decreases upon apoptosis induction in the entire spheroid

We evaluated the effects of Staurosporine treatment on cell nuclei morphology. From the quantitative characterization, we obtained the cell nucleus volume. The distributions of cell nuclei volumes were asymmetric and deviated from a normal distribution (result not shown). Thus, we obtained the median and the median absolute deviation values for each condition (Supplementary Table 7). The median cell nucleus volume in spheroids treated with Staurosporine decreased as a function of time. Compared to the negative control, the median cell nucleus volume was decreased for spheroids treated with Staurosporine after 72 hours (Figure 36a). Compared to the negative control, the volume of cell nuclei was reduced by at most 43% when treated with Staurosporine (27%, 43%, 41% and 36% for 500, 1,000, 2,000 and 5,000 cells seeded).

We evaluated whether the effect of Staurosporine on cell nuclei morphology depends on the location of a cell within the spheroid. Thus, we used the distance of the cell nucleus to the spheroid surface and normalized it to the range [0.0, 1.0]. Cell nuclei located on the surface of the spheroid have a normalized distance (NDS) of zero whereas cell nuclei in the center have a normalized distance of one (Figure 36b). We plotted the cell nucleus volume as a function of the normalized distance to the surface for all conditions. For spheroids generated from a seed number of 500 cells, the curves showed high variability along the radial direction. However, this could be due to the low number of datasets available for 500 cells seeded. For spheroids generated from cell seed numbers of 1,000 and 2,000, the cell nucleus volume was constant along the radial direction in all spheroids. For 5,000 seeded cells, the cell nucleus volume slightly decreased towards the center of spheroids for all conditions. We crosschecked these findings by plotting the locations of cell nuclei as spheres and colored them according to their respective cell nucleus volume (Supplementary Figure 6). Consistent with the quantification, cell nuclei volume after 72 hours of Staurosporine treatment was lower than for the negative control. However, we could not observe any dependency of the cell nuclei volume on the relative location within the spheroids (Figure 36c-f).

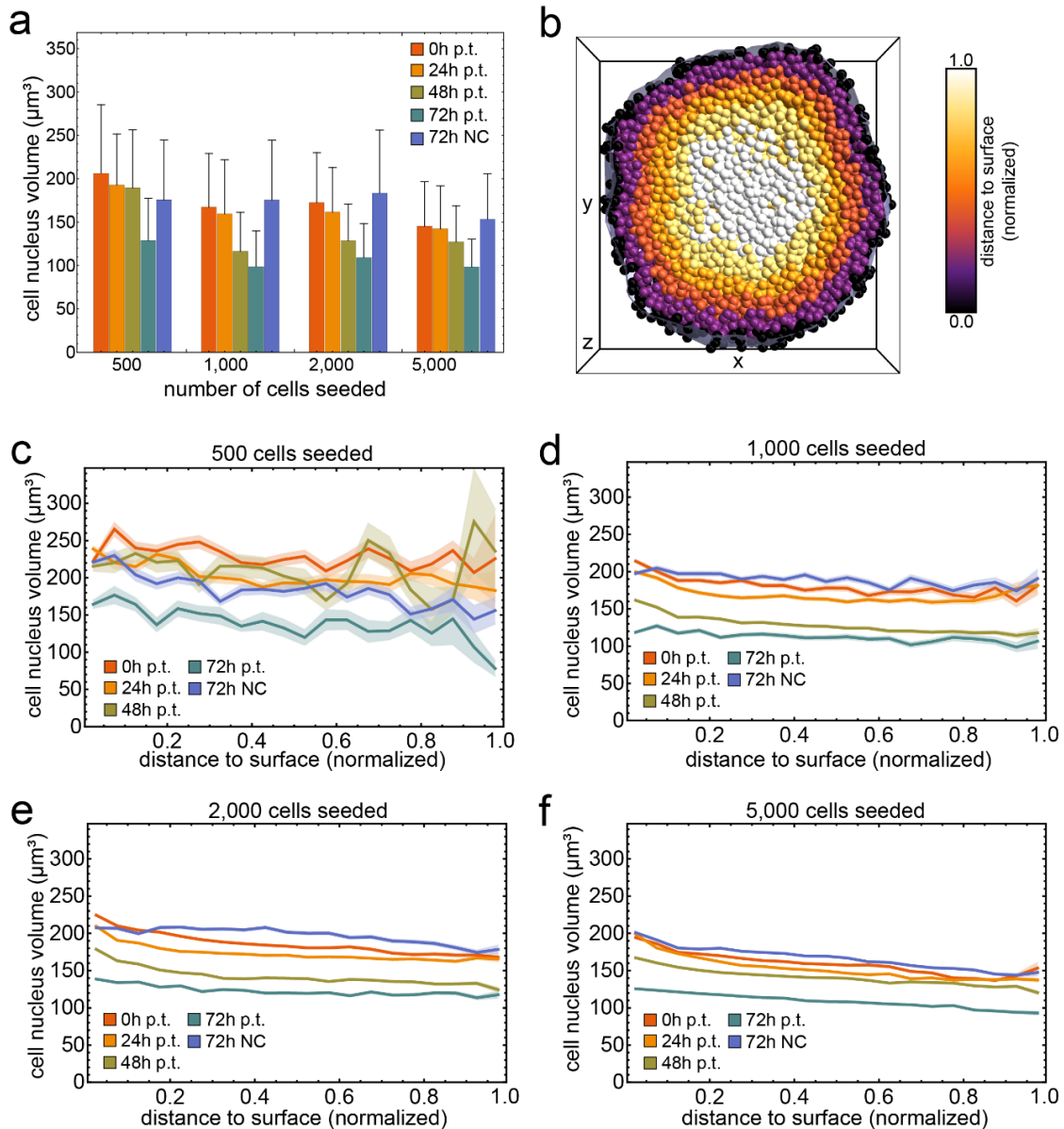


Figure 36 Cell nucleus volume in T47D spheroids for all conditions. (a) Plot of the median cell nucleus volume as a function of cell seed number and time point. Error bars indicate the median absolute deviation. Hypothesis testing was performed for condition 72 h p.t. against 72 h NC. (b) Illustration of normalization of distance to spheroid surface. The cell nuclei were binned into shells of 0.05 NDS and the mean cell nucleus volume was measured in each shell. Plot of mean cell nucleus volume versus the normalized distance to the surface for spheroids generated from 500 (c), 1,000 (d), 2,000 (e) and 5,000 (f) T47D cells. h, hours, p.t., post treatment. NC, negative control.

In summary, we investigated the long-term effects of Staurosporine on spheroid and cell nuclei morphology. Spheroids treated with Staurosporine showed a decrease in volume and cell number over time. Independent of the Staurosporine treatment, the overall shape of all spheroids was oblate. Cell nuclei volume was reduced after 72 hours compared to the negative control. We did not detect a dependency of the cell nucleus volume on the relative location of a cell within the spheroid.

4.6.3 Application III: autophagy induction

The quantitative characterization module was applied to evaluate the effect of long-term morphometric effects of autophagy modulation on spheroids generated from T47D breast cancer cells. We compared the effect of long-term autophagy induction and long-term autophagy inhibition. Spheroids were generated from cell suspensions containing a seed number of 2,000 T47D breast carcinoma cells. After three days of formation, autophagy was induced for seven days by either starvation with EBSS medium (positive control) or treatment with Rapamycin in two concentrations. Autophagy was inhibited by treatment with Bafilomycin in two concentrations. As a negative control, spheroids were cultivated for the same time in DMSO. After seven days of stimulation, spheroids were fixed, stained for nuclei, β -catenin or GM130 (protein of the cis-Golgi complex), optically cleared and imaged with a light sheet-based fluorescence microscope. Katharina Hötte (PhD student) and Michael Koch (master student) performed spheroid preparation, treatment and imaging. The acquired datasets allow a detailed investigation of morphometric effects of autophagy modulation on cell nuclei and cells in T47D spheroids (Figure 37).

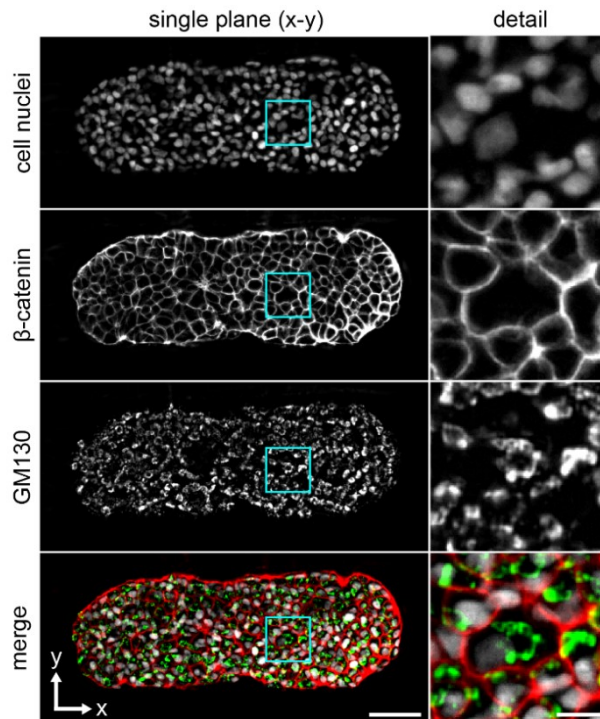


Figure 37 Single planes along x-y of an optically cleared and triple-stained T47D spheroid. Scale bar 50 μm . Detail images in the right column show an acinus during the initiation phase. Scale bar: 10 μm . Spheroid were formed from 2,000 seeded human breast cancer cells (T47D) grown for ten days. Cell nuclei are stained with DAPI, cell membranes (red) and Golgi apparatus (green) are stained against β -catenin and GM130, respectively. Close-up of indicated region (cyan box). Microscope: mDSLM, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 20x NA 0.5 N-Achroplan, camera: Andor Clara. GM130: 561 nm, bandpass filter 607/70. β -catenin: 488 nm, bandpass filter 525/50. Cell nuclei: 405 nm, bandpass filter 447/55. Figure adapted from Strobl *et al.*, 2017.

We used the module to analyze 95 spheroids with at least ten spheroids available per condition (Supplementary Table 8). Parameter values for cell nuclei segmentation, alpha shape and cell

graph generation were empirically determined and the same set of parameter values was applied for processing all datasets (Table 5). The high quality images allowed the accurate segmentation of cell nuclei (Figure 38a, b). The generated alpha shape yielded a geometrical model of the spheroids (Figure 38c). We observed that the alpha shape provided an accurate geometrical model for spheroids of arbitrary shapes (Supplementary Figure 7).

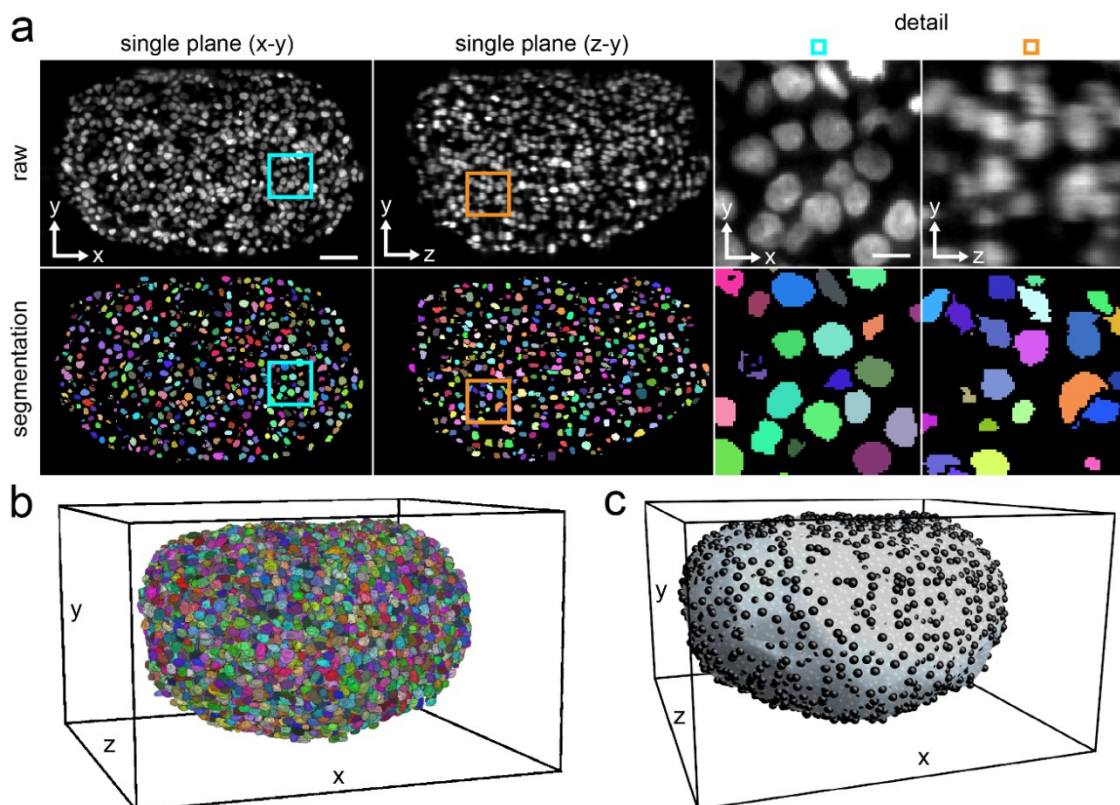


Figure 38 Cell nuclei segmentation performance in T47D spheroids. (a) Single planes along x-y and z-y of raw images and the corresponding segmentation. Scale bar 50 μm . Detail images in the right column show the regions in panel a, indicated by cyan and orange boxes. Scale bar: 10 μm . (b) Rendering of the corresponding cell nuclei segmentation. Different colors indicate individual cell nuclei found by the segmentation. (c) Rendering of cell nuclei locations (black spheres) and the alpha shape surface (light blue). Microscope: see caption of Figure 37.

Autophagy induction leads to a decrease in spheroid volume and cell number

Spheroid volumes and cell numbers differed between the individual experiments. We therefore normalized the results for spheroid volume and the number of cells to the average of the DMSO control in each experiment (Supplementary Figure 8). The normalization allowed pooling the datasets from different experiments and thus a comparison of the different conditions (Supplementary Table 9).

We observed that starvation with EBSS medium or autophagy induction by Rapamycin treatment led to a significant reduction of the spheroid volume (Figure 39a). Spheroids treated under these conditions had about 52% of the volume of untreated spheroids (DMSO). Similar results were observed for the number of cells (Figure 39b). The number of cells in spheroids treated with EBSS was reduced to 57% and in spheroids treated with Rapamycin (0.16 μM) it

was reduced to 66% compared to spheroids treated with DMSO. Long-term autophagy inhibition by treatment with Bafilomycin (0.016 μM) did neither affect the spheroid volume nor the number of cells in a spheroid. Spheroids treated with Bafilomycin (0.016 μM) had about 75% of the spheroid volume and 75% of the number of cells compared to untreated spheroids (Figure 39b).

Autophagy induction leads to an increase of the cell density in spheroids

The average volume of cell nuclei in EBSS treated spheroids was reduced by 2% compared to the DMSO control (Figure 39c). A volume reduction of 18% was observed for cell nuclei in spheroids treated with Rapamycin (0.16 μM). Cell nuclei in Bafilomycin treated spheroids were 2% smaller compared to the DMSO control. The cell density was increased by 13% in EBSS and 15% in Rapamycin treated spheroids compared to untreated spheroids (Figure 39d).

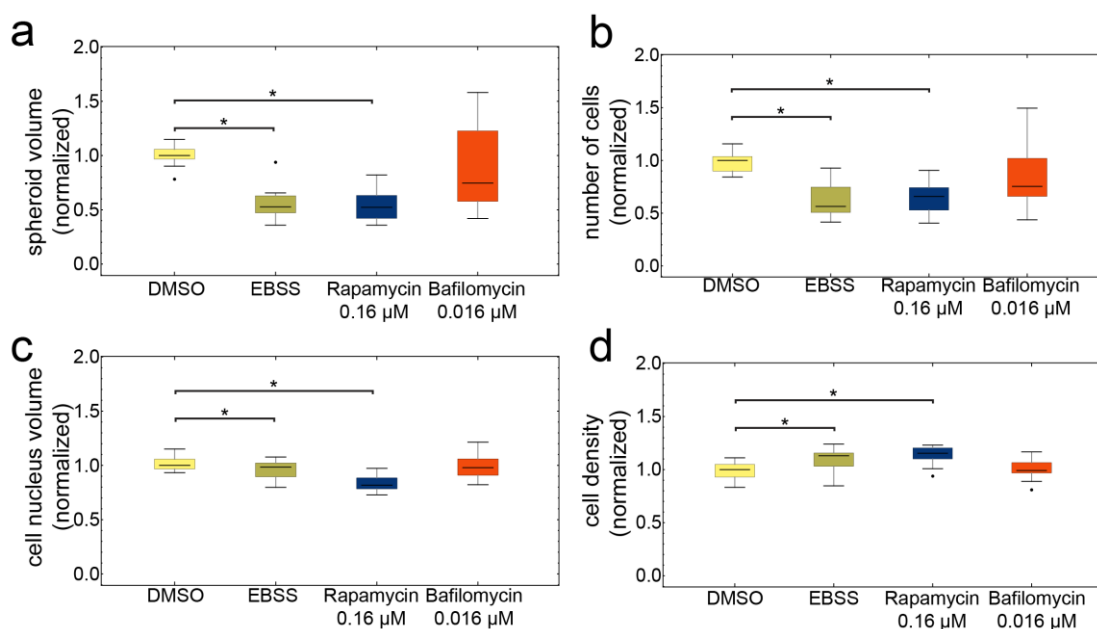


Figure 39 Cellular density and nuclear volume are reduced following prolonged autophagy modulation. Boxplots of the normalized spheroid volume (A), the number of cell nuclei (B), the average nucleus volume (C) and the average local cell density (D) for T47D spheroids treated with 0.08 μM and 0.16 μM Rapamycin, DMSO or starved in EBSS medium. Boxplot parameters: the box contains 50% of the data points; the middle line of the box is the median. Whiskers and outliers represent the upper and lower 25% of the data. Outliers are outside the 1.5x interquartile range; far outliers are outside the 3x interquartile range, ●, outliers; ○, far outliers. Asterisks indicate significant differences (p-value < 0.05, Wilcoxon-Mann-Whitney-Test followed by Bonferroni correction, n=5).

In summary, long-term induction of autophagy led to a significant change in spheroid morphology. The spheroids treated with Rapamycin were smaller and contained fewer cells. Similarly, the cell nucleus volume was reduced and at the same time, the cellular density was increased upon autophagy induction. These effects were not observed for spheroids in which autophagy was inhibited (e.g. by treating the spheroids with Bafilomycin).

4.6.4 Application IV: effects of mechanical perturbation

We applied the quantitative characterization module to evaluate the effect of induced physical constriction of spheroids that were generated from 5,000 HC11 (human breast) cells and then grown for ten days. The diameters of the final spheroids was about 250 μm . Physical constriction was achieved by placing spheroids into agarose capillaries with a diameter of 200 μm for 1 hour, 24 hours or 96 hours. As a negative control, spheroids were placed into agarose capillaries with a diameter of 300 μm for the same time. Spheroids were then fixed, removed from the agarose capillaries and stained for f-actin, cleaved-caspase-3 and cell nuclei. After staining, spheroids were optically cleared and image stacks were acquired with LSMF. All experimental work was performed by Isabell Smyrek (PhD student). The acquired images provided a detailed view on the internal spheroid morphology at the cellular level (Figure 40).

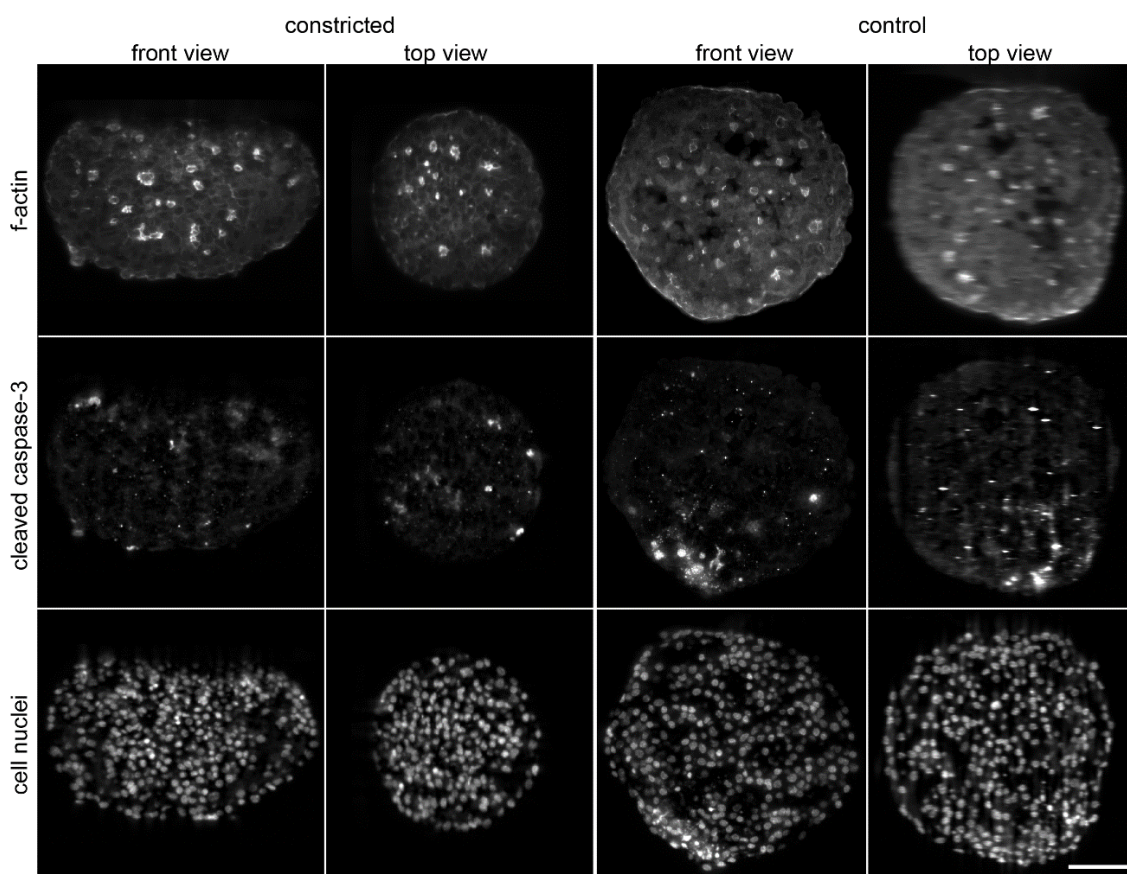


Figure 40 Example of a constricted and a non-constricted (control) HC11 spheroid. Depicted are single plane images at about the center of a constricted and a control spheroid. Please note that the images were transformed, such that views match with the horizontal orientation of the capillary shown in Figure 7d. The front view thus corresponds to the side view on the capillary, whereas the top view corresponds to the view from the top into the capillary. Microscope: mDSLML, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 20x NA 0.5 N-Achroplan, camera: Andor Clara. GM130: 561 nm, bandpass filter 607/70, β -catenin: 488 nm, bandpass filter 525/50, cell nuclei: 405 nm, bandpass filter 447/55. Scale bar: 50 μm .

Robust and accurate cell nuclei segmentation in HC11 spheroids

We employed the developed module to obtain a quantitative characterization of the spheroids. The alpha shape was used to measure the spheroid volume and the deformation of the spheroids

induced by the treatment. Moreover, we used the staining against cleaved-caspase-3 to obtain an objective measure for the proportion of apoptotic cells within the spheroids. Due to the high quality of the acquired images, cell nuclei could be accurately identified in a comprehensive dataset of 115 spheroids (Figure 41).

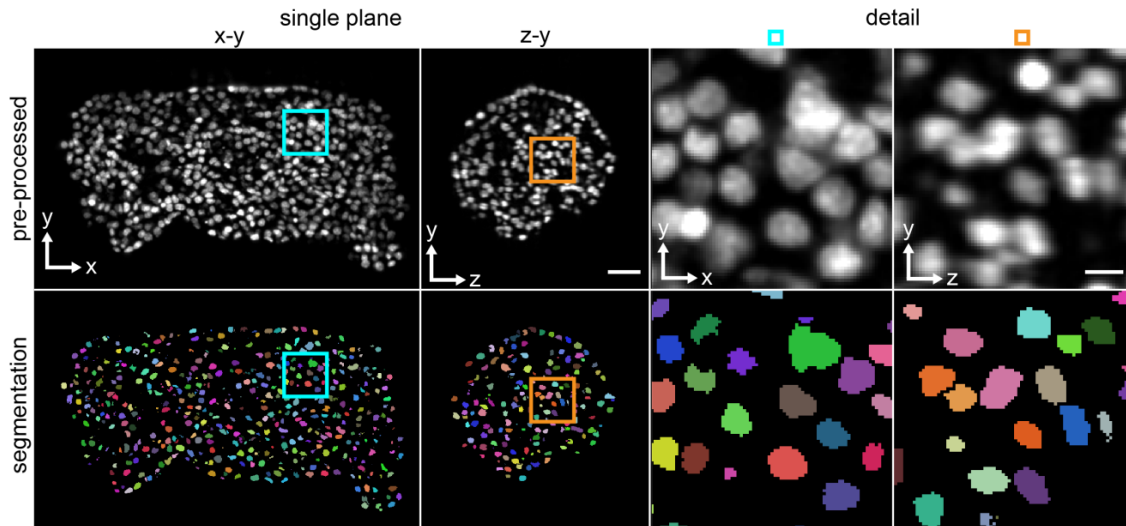


Figure 41 Cell nuclei segmentation performance in HC11 spheroids. Single plane images along x-y and z-y and magnified images of the regions indicated by the cyan and orange boxes. Different colors indicate individual cell nuclei found by the segmentation. Scale bar for single plane images: 25 μm . Scale bar for detail images: 10 μm . Microscope: see caption of Figure 40.

Due to experimental variations in the number of seeded cells, we observed that some spheroids in the group of constricted spheroids were too small. Thus, the transfer into the agarose capillary with an inner diameter of 0.2 mm did not induce mechanical constriction of the spheroid. We used the spheroid shape as a selection criterion in order to consider only spheroids that were actually constricted. A PCA analysis on the cell nuclei positions yielded the principal directions p_1 , p_2 and p_3 and their absolute extension. We then computed the relative extension by generating pairwise ratios between p_1 , p_2 and p_3 (Figure 42a). Threshold values for the ratios were empirically determined and used to select spheroids for the control and constricted groups. For the control group, all spheroids that showed an elongation of at most 15% (i.e. the ratio between $|p_1|$ and $|p_2|$ had to be less or equal to 1.15) were selected. For the constricted group, we selected all spheroids that showed an elongation of at least 15% (i.e. the ratio between $|p_1|$ and $|p_2|$ had to be greater or equal to 1.15). The spheroids that passed the selection (at least 6 per condition) were used for a subsequent analysis (Supplementary Table 10).

Mechanical constriction affects the volume of cell nuclei and the entire spheroid

We first analyzed the shape of the selected spheroids using the elongation along the principal directions (Figure 42b). Constricted spheroids had a prolate shape. We observed that the elongation along p_1 increased for constricted spheroids at time points 24 hours and 96 hours compared to the control. The spheroids in the control group had a more spherical shape.

However, some spheroids of the control group at 24 hours had an oblate shape, indicated by the high ratio of $p1$ to $p3$ and $p2$ to $p3$.

Due to variations of the spheroid sizes in each batch, we normalized spheroid volumes and cell numbers to the median of the control group at each time point (Supplementary Figure 9). The normalized volume of constricted spheroids after one hour was only 1% less than the volume of the control group, 3.4% less after 24 hours and 12.6% less after 96 hours (Figure 42c). The normalized number of cells varied by 7% (after one hour and 24 hours) between constricted and non-constricted spheroids and was almost identical after 96 hours (Figure 42d). The cell density was similar between constricted and non-constricted spheroids after 1 hour and 96 hours. After 24 hours, the cell density was decreased (Figure 42e).

Mechanical construction does not affect the induction of apoptosis

The cell nuclei volume after one hour of constriction was similar to the control group. After 24 and 96 hours, we observed that cell nuclei were slightly smaller in the constricted spheroids compared to the control (Figure 42f). We pooled cell nuclei from each condition and plotted the distribution of the cell nucleus volume (Figure 42g). The mean cell nucleus volume after 1 hour was almost identical ($76 \pm 35 \mu\text{m}^3$ in control, $76 \pm 36 \mu\text{m}^3$ in constricted). After 24 hours, the mean cell nucleus volume differed significantly by 21% (control: $94 \pm 36 \mu\text{m}^3$, constricted: $77 \pm 35 \mu\text{m}^3$). A significant difference in cell nucleus volume by 18% was observed after 96 hours (control: $80 \pm 36 \mu\text{m}^3$, constricted: $68 \pm 34 \mu\text{m}^3$). The amount of apoptotic cells was similar in constricted and non-constricted spheroids (Figure 42h). We measured a proportion of 10% apoptotic cells after 1 hour and 96 hours and 5% apoptotic cells after 24 hours.

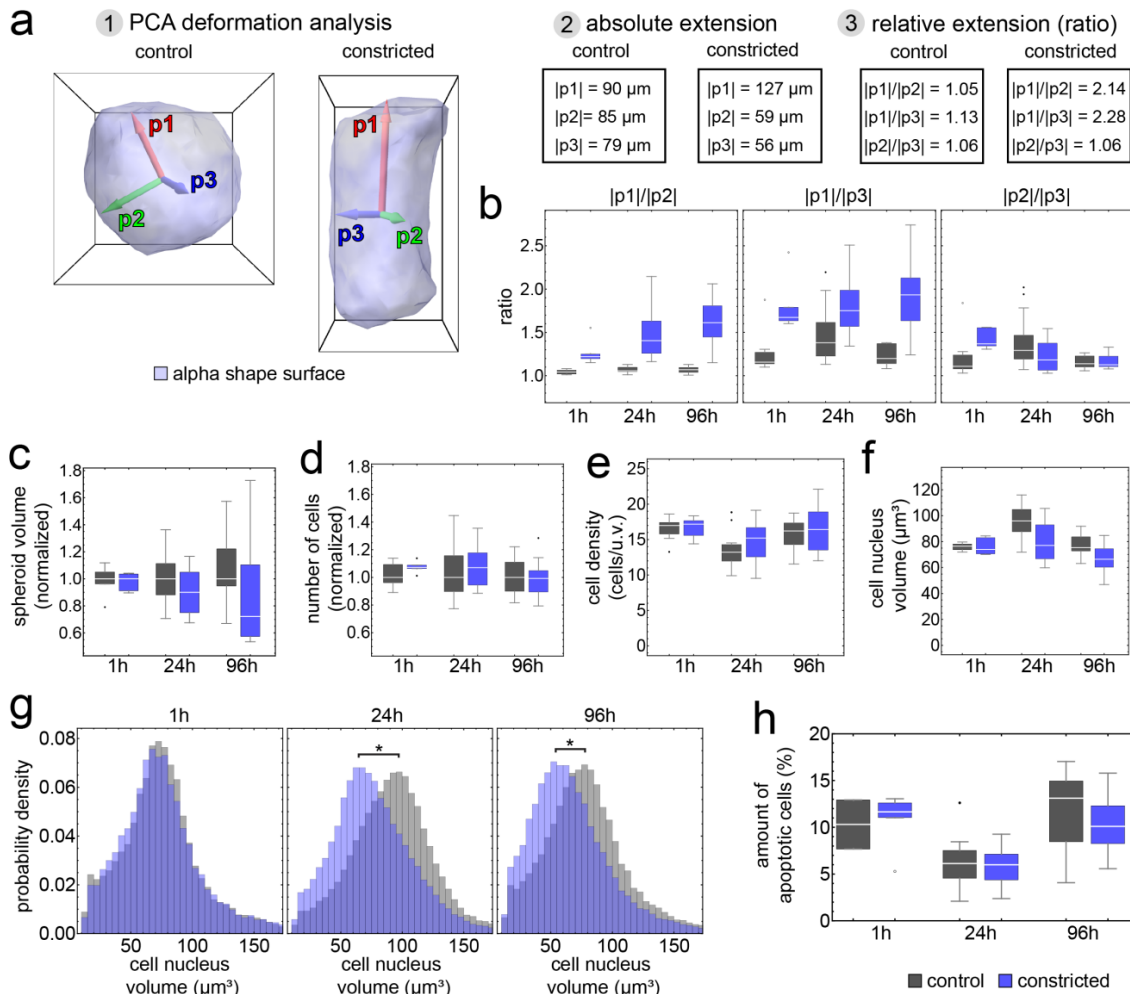


Figure 42 Quantitative results obtained for constricted and non-constricted HC11 spheroids. (a) Shape analysis methodology. Principal component analysis (PCA) gives the principal directions p1, p2 and p3 of the spheroid (1). The absolute extension measures the extension of the spheroid in μm (2). Normalization to the absolute extension of the first principal direction gives the relative extension (3). (b) Boxplots of the pairwise ratios between p1, p2 and p3 for all conditions. Boxplots of the normalized spheroid volume (c) the normalized number of cells (d) the cell density in cells per unit volume (u.v.) (e) and the cell nucleus volume (f) for all conditions. (g) Histograms showing the distributions of the cell nucleus volume of nuclei pooled for each condition. Asterisks indicate significant differences (p -value < 0.05 , Mann-Whitney test). (h). Boxplots of the proportion of apoptotic cells in spheroids for all conditions. Boxplot parameters: the box contains 50% of the data points; the middle line of the box is the median. Whiskers and outliers represent the upper and lower 25% of the data. Outliers are outside the 1.5x interquartile range; far outliers are outside the 3x interquartile range. ●, outliers; ○, far outliers. h, hours.

After analyzing these general quantitative properties, we also investigated the cell nucleus volume and cell density relative to the distance to the surface (Figure 43). We first plotted both features according to the normalized distance to the surface (NDS). Since we were interested in the overall trend, we normalized each curve to the first value and compared the curve progression for each time point. Thus, we focused on analyzing the trend of the curve.

The impact of mechanical constriction is similar in the entire spheroid

After one hour, the cell nucleus volume linearly decreased along the radial direction of non-constricted spheroids. Cell nuclei in the center are about 20% smaller than in the outer regions. For constricted spheroids, the cell nucleus volume decreased slower than for non-constricted

spheroids. Thus, cell nuclei in the center of constricted spheroids were about 10% smaller than in the outer regions. After 24 hours, the curves were almost constant, indicating a consistent cell nucleus volume throughout the entire spheroid. For 96 hours, the cell nucleus volume in both conditions decreased towards the center (Figure 43a). For the cell density, we observed a lower density close to the surface of the spheroids as expected (Figure 43b).

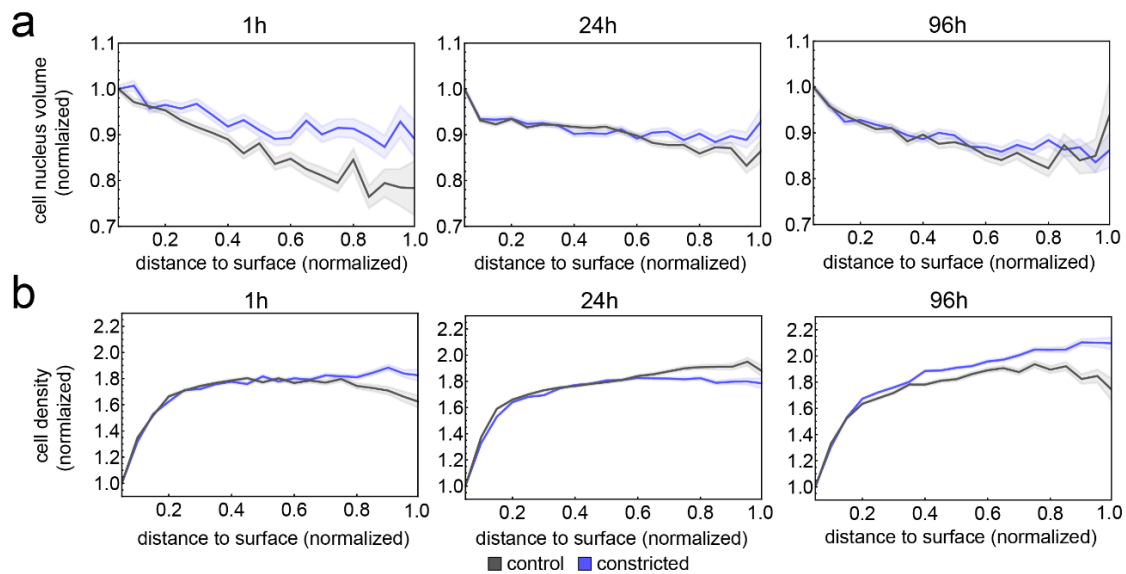


Figure 43 Cell nucleus volume and cell density as a function of depth in HC11 spheroids. Plots of normalized mean cell nucleus volume (a) and the normalized mean cell density (b) versus the normalized distance to the surface for constricted and control spheroids after 1 hour, 24 hours and 96 hours. h, hours.

In summary, we observed an elongation and hence a prolate shape of HC11 spheroids after transferring them into agarose capillaries of 200 μm diameter. In contrast, control spheroids retained a spherical shape. The volume of spheroids decreased after 96 hours of constant constriction. The constriction did not affect the number of cells and cell density. A strong deviation between constricted spheroids and the control was measured for the cell nucleus volume after 24 hour and 96 hours. We did not detect any differences in the proportion of apoptotic cells.

4.7 Computer-aided cell tracking and lineage reconstruction

Time-resolved datasets allow the analysis of cell dynamics and lineages. Computer-aided solutions simplify the task of tracking cells and perform lineage reconstruction. This section presents the module *TrackGen* that was developed for this task and applied to datasets capturing lateral root development in *Arabidopsis thaliana*. Please note that parts of this section have been published (von Wangenheim et al., 2016).

4.7.1 Five datasets capturing lateral root development

The plasticity of post-embryonic organogenesis in plants requires the sampling of the variance in cell behavior. Therefore, five datasets of lateral root development in five different plants were

acquired under identical conditions by Daniel von Wangenheim. Each recording consists of image stacks acquired at an interval of five minutes for up to three days. The resulting datasets were named according to the date of acquisition as #120830, #121204, #121211, #130508 and #130607. The recordings capture the complete lateral root formation from the first cell division up to the emergence of the lateral root primordium (LRP) (Figure 44). Three canonical views were defined to observe lateral root formation (Figure 44a). In the front view (x-y), the primordium grows towards the observer. In the side view (z-y), the primordium grows to the right, away from the axis of the primary root. The third view perspective is the radial view (x-z) that represents a transversal cut through the primary root, perpendicular to the shoot-root axis. In this work, the acquired datasets were used for tracking cells, perform lineage reconstruction and subsequent data analysis.

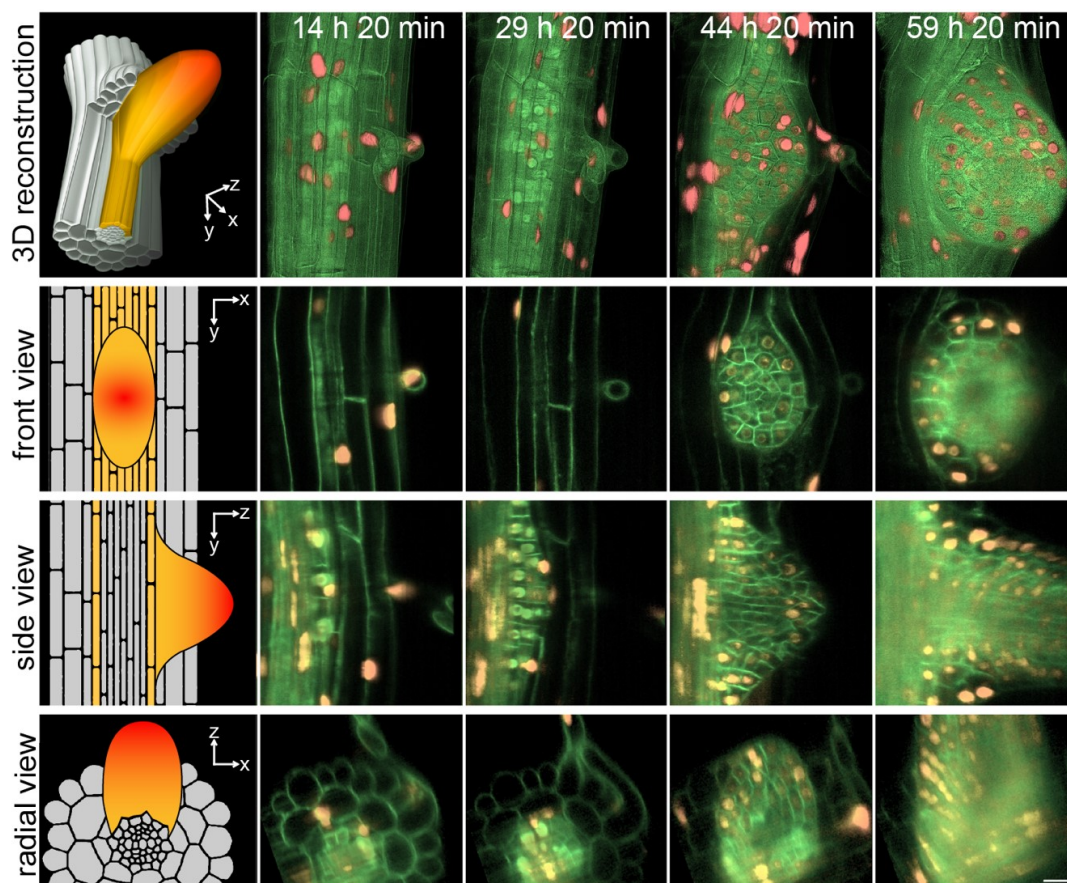


Figure 44 Live recording of lateral root development from initiation to emergence. Schematics in the first column describe the different perspective views on the lateral root. The primary axis of the root is oriented vertically such that the plant shoot is located at the top and the root tip is located at the bottom. Three view directions are obtained: front, side and radial view. First row: three-dimensional reconstruction of the lateral root growing out of the primary root is shown. Second row: single slices along x-y (front view), 10 μm inside the epidermis cell layer. Third row: single slices along z-y (side view), 80 μm inside the primary root. Fourth row: single slices along x-z (radial view) through the center of the primordium. Time points indicated at the top are relative to gravistimulation. The plant stably expressed a pan-plasma membrane marker (pUBQ10::YFP-PIP1;4, green), a pan-nuclear marker (pUBQ10::H2B-RFP, red) and a nuclear reporter (pGATA23::GUS-GFP, green). Time points are given relative to the time point of gravistimulation. Microscope: mDSLM; illumination objective: 5 \times /0.16 detection objective: 40 \times /0.75; camera: Andor Clara (binning 2 \times 2). Scale bar: 20 μm . Images were taken from dataset #121204. Figure adapted from von Wangenheim et al., 2016.

4.7.2 *TrackGen* enables cell tracking and lineage tracing

Growth and division of cells predominantly define the shape of the lateral root primordium. The high spatio-temporal resolution of the recordings enabled tracking the spatial location and tracing the lineage of lateral root founder cells and their progeny through up to six rounds of cell division. We focused our analysis on the localization of cell nuclei as representation of the location of cells. We avoided automatic segmentation of cell nuclei due to three main reasons: first, the signal to noise ratio of the fluorescence signal of cells located deep inside the primary root was rather low. Second, the fluorescence intensity of the proliferating cells in the lateral root was weak compared to surrounding cells. Third, the fluorescence intensity varied with the expression level and strongly depended on various factors such as the stage or growth conditions of the plant. Additionally, the number of cells to track was moderate and therefore manual tracking of individual cells feasible. To facilitate the tracking process, I developed a user-friendly program for simplified manual identification of cell nuclei locations and tracing the cell lineage as a function of time. Since we were mostly interested in the cell division behavior during lateral root development, we used *TrackGen* and only marked the location of each cell at the first and last occurrence. In this context, the first occurrence was either right after the cell division (birth of the cell) or the very first appearance of a cell (i.e. a founder cell). The last occurrence corresponded to the time point before a cell divided or the last time point analyzed (Table 8). Thus, the manual effort to extract the relevant lineage information of each cell and the amount of data for each dataset was kept minimal. For each marked cell nucleus, the program stored a unique identifier, the spatial position, the time point and a list holding the identification numbers of its precursors.

Table 8 Excerpt of the file for dataset #120830 generated by the program *TrackGen*. A unique identification number (id), the location of the cell nucleus in x, y and z, the time point of occurrence and a list representing the lineage trace (lineage) given by the cell ids are stored for each cell.

cell id	x position	y position	z position	time point	lineage
1	320	72	112	1	{}
1	299	70	111	19	{}
⋮	⋮	⋮	⋮	⋮	⋮
146	307	69	111	20	{1}
147	294	71	111	20	{1}
⋮	⋮	⋮	⋮	⋮	⋮
178	291	46	95	293	{1, 146, 155, 158, 176}
180	295	76	93	283	{1, 146, 155, 158, 175}

4.7.3 Complete reconstruction of cell lineages

Cell nuclei locations and cell division events could be tracked for up to 30 hours until emergence of the lateral roots (Figure 45a). Thereby, we captured lateral root formation from the first divisions of founder cells in the pericycle cell layer up to the emergence of the lateral root through the overlaying cortex and epidermis cell layers of the primary root. We obtained the locations of all cells and cell divisions as a function of time for five datasets (Figure 45b). We then programmatically reconstructed the cell lineages of founder cells using a lineage tree data structure (Figure 45c). The lineage tree encodes all information about the cells and cell division events as a function of time.

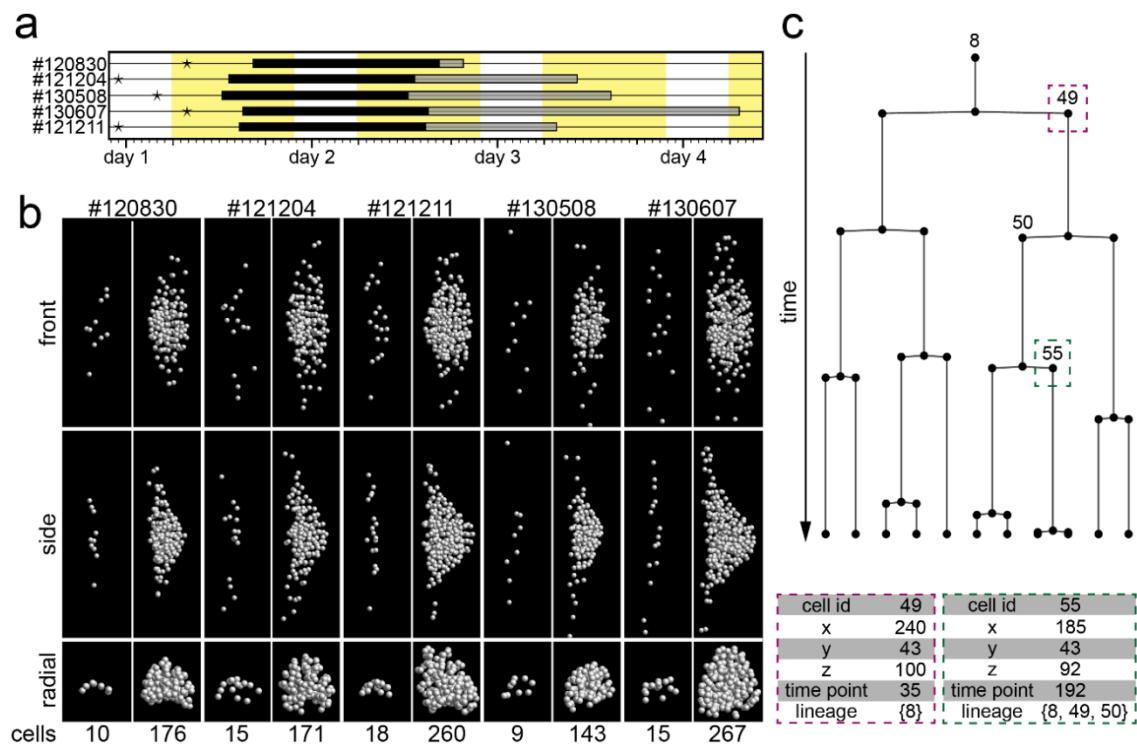


Figure 45 Tracking and lineage tracing of cells in lateral root development. (a) Time scales of the five analyzed lateral root datasets. The total recording time of each dataset is indicated in gray, the time span of cell nuclei tracking in black. The time point of gravistimulation is indicated with an asterisk. (b) Cell nuclei locations of the five analyzed datasets at the first (left) and last (right) time point in front, side and radial view. (c) Exemplary lineage tree of founder cell #8 in dataset #120830. Note that each vertex in the lineage tree stores all information about the cell (see magenta and green boxes for examples). Panel a adapted from von Wangenheim *et al.*, 2016.

Table 9 Summary of all analyzed lateral root datasets.

	number of founder cells	final number of cells	tracking time (h)	number of cell files	number of cell divisions
#120830	10	176	25	6	166
#121204	15	171	25	8	156
#121211	18*	260	25	6	242
#130508	9	143	30	5	134

#130607	15	267	25	6	252
---------	----	-----	----	---	-----

*The first cell divisions in dataset #121211 had already occurred when the recording started. The true number of founder cells could not be reconstructed and the number of cells at the first time point of the recording was used as default value.

4.7.4 Assignment of cell files

Lateral root founder cells are arranged in a field of parallel pericycle cell files that face the xylem pole of the primary root. We identified founder cells and assigned identification numbers (cell id). Visual inspection of the datasets revealed that in datasets #121204, #130607 and #121211 the first cell divisions of founder cells had already occurred before the image acquisition started. Thus, virtual cells were introduced to reconstruct the cell lineage. The reconstruction of the first cell divisions was based on the time point at which the next cell division occurred and considering the shape of the cell wall between neighboring cells. Virtual cells were introduced for datasets #121204 and #130607, whereas the reconstruction of the first cell divisions was not possible for #121211.

Cell lineages were assigned to cell files and alternating colors were given to the lineages of each cell file (Figure 46). A recent publication reported the increase in volume of cells before the first division of a founder cell (Vermeer et al., 2014). This event represents the onset of lateral root formation and we marked the corresponding cell file with index #0. We termed this cell file the master cell file, as it adopts a leading role during the development of the lateral root. Cell files to the left of the master cell file were given negative indices (#-1, #-2, #-3, #-4), whereas cell files to the right were given positive indices (#+1, #+2, #+3, #+4). Thus, the indices express the position of a cell file relative to the master cell file. The master cell file (#0) and its two flanking cell files (#-1 and #+1) were defined as the core region and the remaining cell files were defined as the periphery of the lateral root.

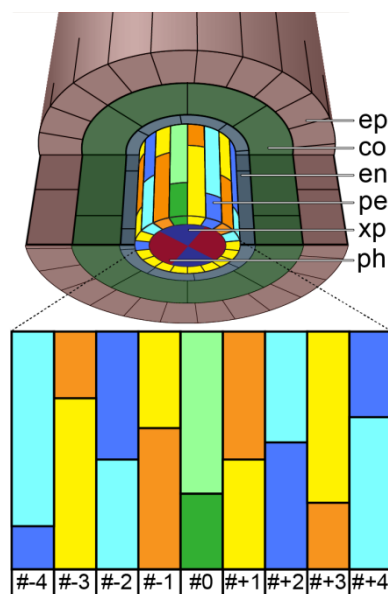


Figure 46 Schematic illustration of the cellular disposition in lateral roots and assignment of cell files. Lateral root founder cells originate from the pericycle cell layer located several cell layers within

the primary root. These founder cells face the xylem pole of the vasculature in the primary root. Cell lineages were assigned to cell files ranging from #-4 to #+4, where cell file #0 corresponds to the master cell file. The cell files with indices #-1, #0 and #+1 represent the core region and the remaining cell files the periphery of the lateral root. Cell lineages are represented with alternating colors. co: cortex, en: endodermis, ep: epidermis, pe: pericycle, xp: xylem pole, ph: phloem. Figure adapted and modified from von Wangenheim et al., 2016.

4.7.5 Classification of cell division types

To study patterns of cell divisions in lateral root development, we analyzed the relative orientation of cell divisions to the primary root (Figure 47). Anticlinal cell divisions are parallel to the shoot-root axis and increase the number of cells in a cell file. Periclinal cell divisions are normal to the surface of the primary root. A periclinal division results in the generation of a new cell layer. By visual inspection of the cell division orientations, we observed that next to anticlinal and periclinal cell divisions a third type of cell division was involved in the formation of the lateral root. The number of cell files and thus the thickness of the LRPs was increased by cell divisions that were tangential to the surface of the primary root and orthogonal to the root-shoot axis. Consequently, we termed these cell divisions radial (Figure 47a).

Based on the relative orientation of cell divisions, we automatically classified a cell division either as an anticlinal, periclinal or a radial cell division. In brief, a cell file based reference system was defined by the principal directions of growth in each cell file. We introduced two vectors that point along the height axis of the cell file and the length axis of the primary root and measured the spatial orientation of each cell division to these vectors (please refer to the Material and Methods section for more details). Misclassified cell divisions were corrected manually. We used the classified cell divisions to analyze spatio-temporal patterns in the occurrence of the three cell division types.

To evaluate the accuracy of the classification method, we compared the classified cell division type with the true cell division type after manual correction. The accuracy was determined as the proportion of correctly classified cell divisions.

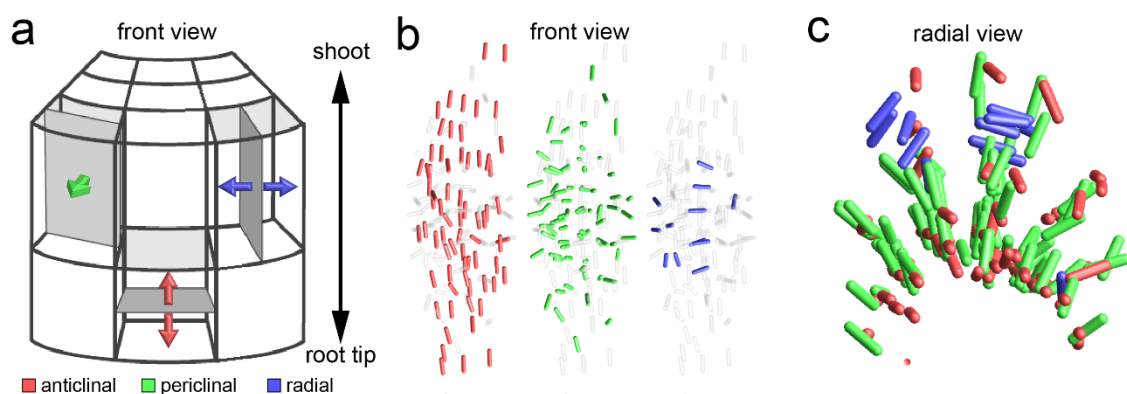


Figure 47 Classification of cell division types in lateral root development. (a) Schematic illustration of cell division types. The orientation of the cell division axis is shown as arrows and a gray plane represents the created cell wall. Anticlinal cell divisions (red) are parallel to the shoot-root axis of the root. Periclinal

cell divisions (green) are normal to the surface of the primary root. Radial cell divisions are tangential to the surface and normal to the shoot-root axis of the primary root. (b) Front view of the spatial arrangement of the classified cell divisions types anticlinal (red), periclinal (green) and radial (blue) for dataset #121204 up to the last time point tracked. Sticks colored according to the classified cell division type visualize the orientations of cell divisions. Cell divisions were classified based on their relative orientation in a cell file based reference system (see Methods section). (c) Radial view of all cell divisions that occurred in dataset #121204 up to the last time point tracked.

Overall, the classification method identified 84.8% of the cell divisions correctly (806 out of 950 cell divisions). The accuracy of the cell division type was different for the cell division types. The classification method correctly identified anticlinal cell divisions in 84.8% (328 out of 387 cell divisions) and periclinal cell divisions in 87.8% (353 out of 402 cell divisions) of the cases. The accuracy for radial cell divisions was considerably lower with 77.6% (125 out of 161 cell divisions).

Furthermore, we found that the accuracy of the classification was high for early cell division rounds and decreases for later cell division rounds. The classification accuracy for cell divisions of the first cell division round was 98.5% (66 out of 67), whereas the accuracy dropped to 76.9% (60 out of 78) for the sixth round (Table 10). This was also reflected by the relative deviation in percent from the ideal orientation of anticlinal, periclinal and radial cell divisions, which increased by a factor of 3.5 between the first and the sixth cell division round.

Table 10 Accuracy of the cell division classification for all cell division rounds. The relative deviation was computed as the percentage deviation of the determined orientation angles of a cell division from the orientation angles of an ideal cell division of the same type.

cell division round	number of cell divisions	accuracy (%)	relative deviation (%)
1	67	98.5	7.9
2	109	88.1	22.5
3	171	88.8	19.9
4	258	85.7	20.6
5	267	79.0	23.6
6	78	76.9	28.4

4.7.6 Complete atlas of cell lineages for five lateral root datasets

The module *TrackGen* enabled computer-aided cell tracking and lineage tracing in time-lapse image datasets. Using the module, we were able to reconstruct cell lineages over several days in lateral root development. By integrating all retrieved information into the reconstructed lineage trees, a comprehensive atlas of all cells that contribute to lateral root formation was obtained (Supplementary Figure 10). The atlas captures location, time and type of all cell divisions throughout lateral root formation in five datasets.

4.7.7 Growth dynamics of lateral root development

Lateral root growth profiles and developmental synchronization

We first analyzed the lateral root datasets for similarities in their general growth behavior. The first founder cell division occurred in a range of 7 to 14 hours after gravistimulation of the plant ($10:41 \pm 3:10$; hh:mm; mean \pm SD) when cell nuclei in the pericycle cell layer had already migrated towards each other. Although there was a high variability in the onset of lateral root formation, the growth curves of the different lateral root datasets all followed an exponential growth profile (Figure 48a). We fitted an exponential growth model to the curves and determined growth parameters of the lateral root datasets (Figure 48, Supplementary Figure 11). The doubling time is the period required for a lateral root to double its total number of cells. The average doubling time for the lateral root datasets was 7 hours 23 minutes ($07:23 \pm 00:27$; hh:mm; mean \pm SD).

Although the general growth behavior of lateral root development was similar, we observed a high variation in the onset of proliferation. For example, 24 hours after gravistimulation, dataset #121204 comprised 33 cells whereas #130607 already consisted of 134 cells. Due to this large deviation in the onset of cell proliferation after stimulation of the plants, a comparison of the datasets based on the absolute time is error-prone. Attempts of temporal registration of the datasets were regarded as inappropriate and we instead used a developmental synchronization based on the total number of cells. The developmentally synchronized datasets were used to compare the datasets at a fixed developmental stage given in total number of cells in the LRP.

Shape analysis of the lateral roots using alpha shapes

We obtained a shape approximation based on the locations of cell nuclei at each developmental stage, using the alpha shapes approach with alpha set to 50 voxels ($\sim 16 \mu\text{m}$). This approach gave an accurate representation of the shape of the developing lateral root (Figure 48b). From the alpha shape approximation, the volume of each dataset could be readily obtained (Figure 48c). The height of the primordium at each stage was determined as the maximal distance between two cell nuclei in the master cell file. The length and width of the primordium was computed at 50% of the primordium height at each time point (Figure 48d).

We observed an overall linear increase of the lateral root volume and height as a function of the total number of cells (Figure 48c, e). We measured an average increase in lateral root volume of $563 \pm 94 \mu\text{m}^3$ per cell, whereas the lateral root height increased by $0.21 \pm 0.02 \mu\text{m}$ per cell (Figure 48, Supplementary Figure 12). We plotted the ratio of the measured lateral root length to the width as a function of the number of cells (Figure 48f). We found that the ratio decreases roughly by a factor of three (from a maximum of 4.5 to about 1.5) during lateral root development.

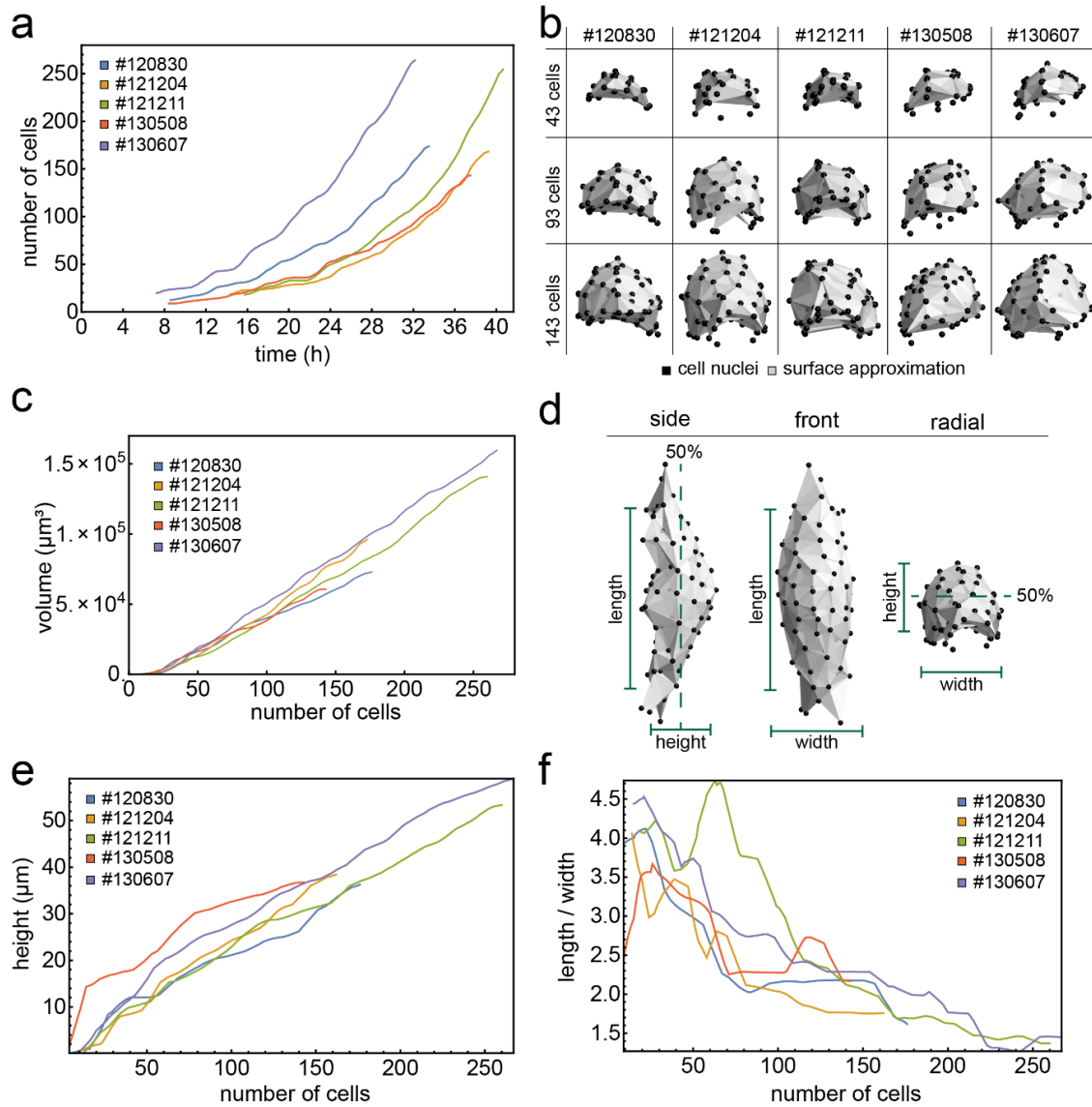


Figure 48 Growth and shape dynamics in lateral root development. (a) Plot of the number of cells as a function of time. (b) Alpha shape approximation of the lateral root based the cell nuclei locations for all datasets at a developmental stage of 43, 93 and 143 cells. (c) Plot of the volume of the lateral root datasets as a function of the number of cells. (d) Illustration of the height, length and width measurement principle. The length and width are measured at 50% of the height at each developmental stage. (e) Plots of the height (e) and the ratio between length to width (f) of the lateral root datasets as a function of the number of cells.

Table 11 Growth rates, doubling times, height growth and volume growth for all lateral root datasets.

	growth rate (%)	doubling time (hh:mm)	height growth ($\mu\text{m}/\text{cell}$)	volume growth ($\mu\text{m}^3/\text{cell}$)
#120830	0.78	07:23	0.19	449
#121204	0.80	07:14	0.25	635
#121211	0.84	06:52	0.20	607
#130508	0.71	08:06	0.21	476
#130607	0.79	07:19	0.22	650
mean	0.78	07:23	0.21	563

SD	0.05	00:27	0.02	94
-----------	------	-------	------	----

In summary, the analyzed lateral root datasets followed a similar growth profile. Since we observed a large variance in the onset of proliferation, a developmental synchronization based on the number of cells in the LRP allowed for a direct comparison of the datasets at a developmental stage given in the number of cells in the primordium. The similar profiles for volume and height growth indicate that the overall shape of the LRPs was conserved. The decrease in the ratio of length to width suggests the occurrence of a dome-shaped primordium.

4.7.8 Cellular composition of the lateral root

The core region of cell files dominates the cellular contribution to the lateral root

Next, we analyzed the cellular composition of the lateral roots by investigating the contribution of cell files and founder cells. As a point of reference, we compared the contribution of cells at the latest common developmental stage of all datasets, which was 143 cells. Lateral root founder cells were located in 6 ± 1 (mean \pm SD) pericycle cell files with a minimum number of five (dataset #130508) and a maximum number of eight (dataset #121204) pericycle cell files (Figure 49a, Supplementary Table 12). The number of cell files to the left (negative indices) and to the right (positive indices) of the master cell file and their cellular contribution varied between the different datasets.

We measured the duration of the interphase in each cell file (Figure 49b). Cells in the core region (cell files -1, 0, +1) divided faster than cells in the periphery (core region: 5.68 ± 1.85 hr; mean \pm SD, periphery: 6.77 ± 2.73 hr; $p = 2.014 \times 10^{-12}$; Mann-Whitney test). This in turn led to a dominant contribution of the master cell file to the total cell mass of the primordium ($31 \pm 3\%$; 44 ± 6 cells for a primordium at the 143 cells stage). In total, the core region of the primordium contributed 75% (107 ± 9 cells) and the periphery contributed 25% (36 ± 9 cells) of the total cell mass at the 143 cells stage (Figure 49c, Supplementary Table 12).

Cellular contribution of lateral root founder cells is not conserved

By tracing back the lineage of each cell, we quantified the contribution of founder cells to the lateral root at the stage of 143 cells (Figure 49d). The number of founder cells that initiated lateral root formation was between 9 to 15 (12 ± 3 , mean \pm SD, Table 9). The number of founder cells and their contribution to the cell mass was highly variable. In dataset #130508, 60% of the lateral root was comprised of the progeny of only two founder cells at the stage of 143 cells, whereas in dataset #130607 it required five founder cells to reach 60% of the cell mass (Figure 49e).

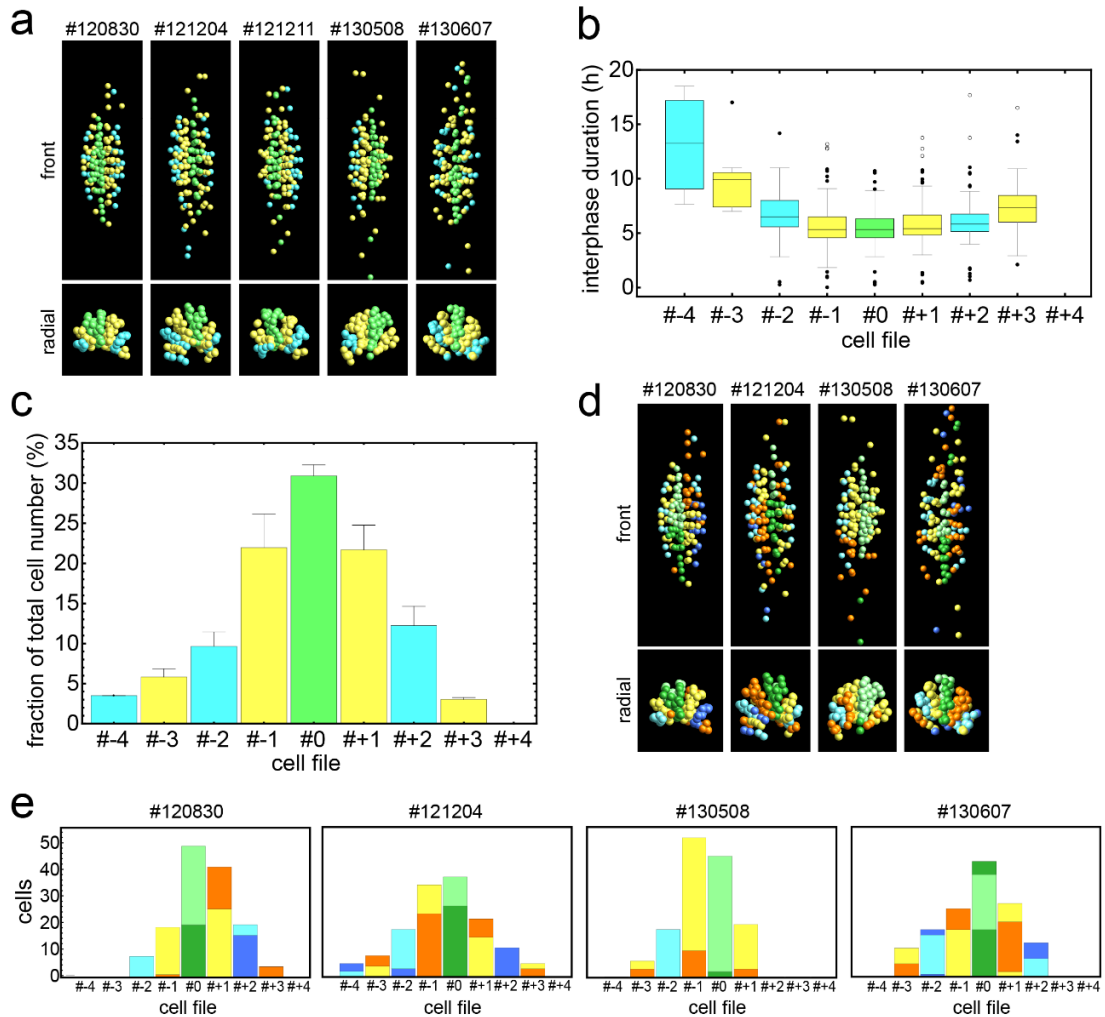


Figure 49 Contribution of cell files and cell lineages to the lateral root. (a) Cell nuclei locations of the five analyzed datasets at the 143 cells stage in front view (first row) and radial view (second row). Cell nuclei are colored according to the corresponding cell file (alternating between cyan and yellow, green for the master cell file). (b) Boxplots of the interphase duration (time span between two consecutive divisions) in each cell file. (c) Contribution of cell files to the lateral root in percent at the 143 cells stage. (d) Cell nuclei locations of the five analyzed datasets at the 143 cells stage in front (first row) and radial (second row) view. Clonally related cell nuclei (i.e. cells of the same founder cell lineage) are depicted in the same color. The first cell divisions of dataset #121211 could not be reconstructed and the dataset was thus omitted from the analysis. (e) Plots of the number of cells that derived from the same founder cell lineage for each dataset. Note that bar colors correspond to the founder cells of the same color as in panel d. Boxplot parameters: the box contains 50% of the data points; the middle line of the box is the median. Whiskers and outliers represent the upper and lower 25% of the data. Outliers are outside the 1.5x interquartile range; far outliers are outside the 3x interquartile range. ●, outliers; ○, far outliers.

In summary, the number of cell files that contribute to the LRP is not conserved. The master cell file and its two flanking cell files contribute the bulk of the cell mass of the LRP up to the developmental stage of 143 cells. We could not detect any patterns in the contribution of individual founder cells to the LRP.

4.7.9 Spatio-temporal patterns of cell divisions

The development of lateral roots is based on a series of cell divisions that leads to a defined structure. Thus, the spatio-temporal distribution of cell divisions is an important feature to study. We employed the classified cell division types to search for underlying regularities in the

cell division behavior. Particularly, we were interested in the generation of new cell layers (periclinal cell divisions) and the increase in thickness by new cell files (radial cell divisions). Thus, we analyzed the spatio-temporal distribution of anticlinal, periclinal and radial cell divisions (Figure 50).

Analysis of the occurrence of cell division types

We first compared the number of occurrences of each cell division type up to a developmental stage of 143 cells between the datasets (Supplementary Table 13). Up to this stage, an average number of 55 anticlinal (55 ± 5 , mean \pm SD), 59 periclinal (59 ± 3 , mean \pm SD) and 16 radial (16 ± 7 , mean \pm SD) cell divisions occurred. Thus, the occurrence of anticlinal and periclinal cell divisions was balanced up to this stage. The number of radial cell divisions varied between the analyzed datasets by a factor of up to three, i.e., nine radial divisions occurred in dataset #130607, whereas already 27 radial cell divisions occurred in dataset #130508.

Consistent with previous studies, we confirmed that the first formative cell division is always anticlinal (Figure 50a). All founder cells underwent an anticlinal, asymmetric cell division that generated two smaller cells flanked by two larger cells. Additionally, we could observe that this first cell division always occurred in the core region (cell files indexed #-1, #0, #+1) of the LRP (Figure 50b, Supplementary Table 14). Similarly, the first periclinal division in all datasets occurred in the core region of the LRP (Figure 50b, Supplementary Table 14). Radial cell divisions increase the thickness of the LRP by creating new cell files. The first radial cell division was observed in the third cell division round in dataset #120830, whereas in the other datasets, it occurred in the fourth cell division round. The first radial cell division was always observed in the master cell file.

The first occurrence of anticlinal and periclinal cell divisions correlated with the developmental stage, i.e. the number of cells in the primordium. The first anticlinal cell division occurred at 13 ± 4 cells, the first periclinal cell division at 24 ± 6 cells. In contrast, the developmental stage at which the first radial cell division occurred varied substantially between the datasets (60 ± 27 cells).

We further analyzed the spatial distribution of the different cell division types (Figure 50c). To summarize the spatial distribution of cell division types, we computed heatmaps for each cell division type in the front view (Figure 50d). We observed a predominant occurrence of periclinal and radial cell divisions close to the center of the LRPs, whereas anticlinal cell divisions are more dispersed.

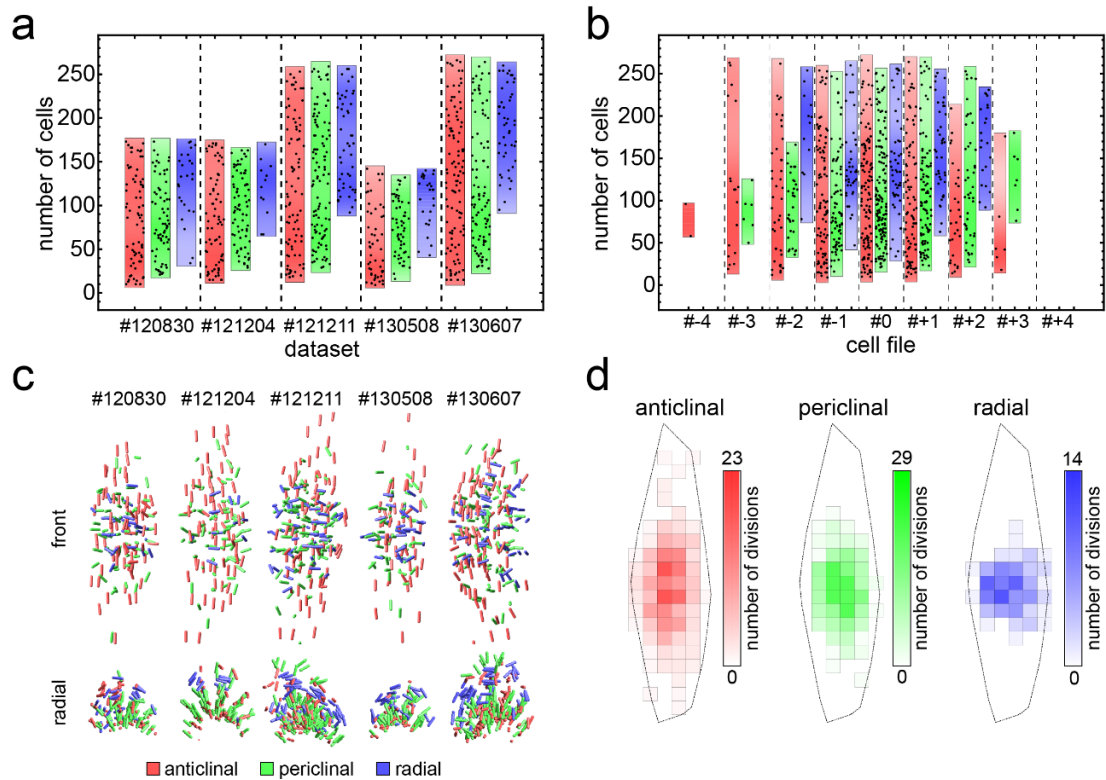


Figure 50 Spatio-temporal occurrence of the three cell division types anticlinal, periclinal and radial. (a) Plot of anticlinal (red bar), periclinal (green bar) and radial (blue bar) cell divisions in each dataset as a function of the developmental stage (i.e. the number of cells). Black points represent individual cell divisions. (b) Plot of all anticlinal (red bar), periclinal (green bar) and radial (blue bar) cell divisions in each cell file as a function of the developmental stage. (c) Stick representations of anticlinal (red), periclinal (green) and radial (blue) cell divisions in each dataset in front and radial view. (d) Heatmaps summarizing the spatial occurrence of all anticlinal, periclinal and radial cell divisions in front view. Cell divisions of the same type were grouped into bins. The color saturation indicates the number of cell divisions in each bin. The black line represents the average outline of the LRP in front view.

Random sequences of cell divisions

Our analysis of the occurrence of anticlinal, periclinal and radial cell divisions suggests that lateral root development is conserved on multiple levels. The cell division round at which a certain cell division type occurs is almost invariant. The first cell division of each cell division type always occurred in the core region of the LRP. We observed a tendency of periclinal and radial cell divisions to occur close to the center of the LRP in the front view.

We hypothesized that the occurrence of the cell division types might be influenced by the type of the previous cell division. As mentioned above, all cells divided anticlinal in the first cell division round. Two possible scenarios were observed for the second cell division round: either cells underwent another anticlinal cell division (AA sequence) or the number of cell layers was increased by undergoing a periclinal cell division (AP sequence). However, we could not identify a pattern that explained whether cells follow the AA sequence or the AP sequence. For example, in the master cell file, all possible cell division scenarios for the daughter cells could be observed (see Supplementary Figure 10): in dataset #120830, all daughters of cell #2 and #3 divided periclinal. In dataset #121204, one daughter cell of cell #4 divided periclinal, the other one anticlinal, and for cell #5566 both daughter cells divided anticlinal. In dataset #130508,

both daughter cells of cell #6 divided anticlinal, whereas cell #5 did not divide again. In dataset #130607, we could also observe a mixed pattern of periclinal and anticlinal cell divisions of the daughter cells. Thus, the second cell division round seems to arbitrarily follow one of the two sequences AA or AP.

We checked whether in subsequent cell division rounds the cell division type was also randomly selected. Interestingly, we found that an alternating cell division behavior was predominant in all datasets. For example, the pattern APAP was the dominant cell division sequence before the first radial cell division (Supplementary Figure 10). Out of 161 analyzed sequences, 43 (25%) followed the APAP division sequence before the first radial cell division. Other observed sequences were AP (four occurrences, e.g. in cell file #0, dataset #120830), AAP (eight occurrences, e.g. in cell file #0, dataset #121204) and APP (19 occurrences, e.g. cell file #0, dataset #130607).

Alternating pattern of cell division orientation

We analyzed whether the alternation of cell division orientation is a general feature of lateral root development. Therefore, we neglected the classified type of the cell divisions and focused computed the relative orientation of two consecutive cell division axes. We therefore measured the spatial orientation of two consecutive cell divisions $n - 1$ and n by computing the planar angle ω between the underlying cell division axes (Figure 51a). A cell division is termed alternated, when the angle ω to the previous division is greater than 45 degrees. Using this method, we could identify that most cells (77.8%) rotate their cell division plane by 90 degrees between two consecutive divisions (Figure 51b). We analyzed whether this pattern could be observed in all cell division rounds. Interestingly, between the first and second cell division round, only 46% of the cell divisions were found to rotate their division plane by 90° (e.g. anticlinal to periclinal), whereas 54% divided along the previous orientation (Figure 51c). In all subsequent cell division rounds, the alternating pattern was dominant with more than 75% alternated cell divisions. This observation could also explain the dominant alternating cell division sequence (e.g., the APAP sequence). Consequently, for a cell that divided anticlinal, their daughter cells had a high probability to divide either periclinal or radial in the next division round.

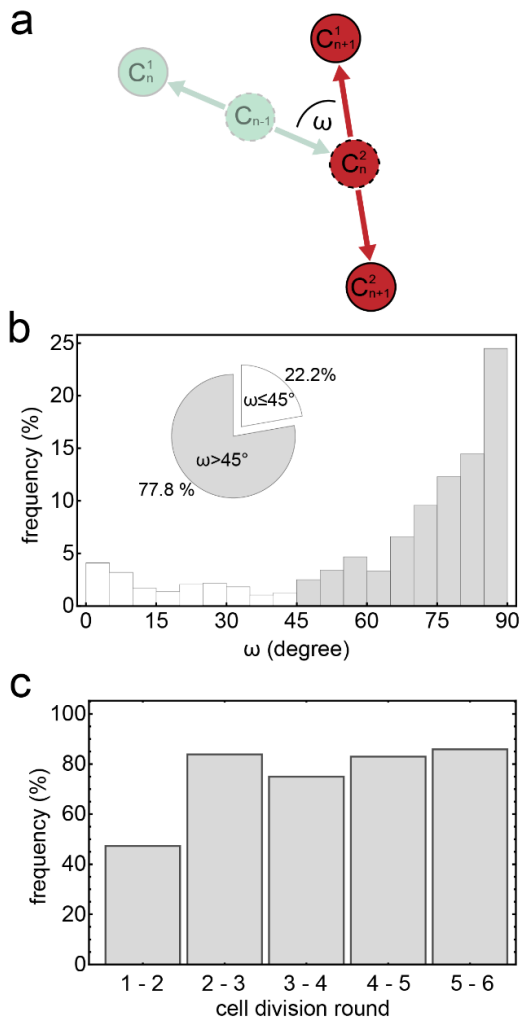


Figure 51 Alternation of cell division orientation in lateral root development. (a) Schematic illustrating the principle of measuring the angle ω between two consecutive divisions. In cell division round $n - 1$ the division of cell C_{n-1} generated the daughter cells C_n^1 and C_n^2 . In the next cell division round, the division of cell C_n^2 generated the daughter cells C_{n+1}^1 and C_{n+1}^2 . The angle ω is computed between the cell division axes $\{C_{n-1}, C_n^2\}$ and $\{C_n^1, C_{n+1}^2\}$. (b) Histogram showing the distribution of the angle ω for all cell divisions in all datasets. The piechart shows the proportion of alternated cell divisions ($\omega > 45^\circ$) (c) Chart of the proportion of alternated cell divisions in each cell division round.

The generation of a layered lateral root primordium is conserved

The primary root in *Arabidopsis* is organized in concentric tissue layers. This layered organization also has to be established in the LRP. The formation of cell layers is the direct consequence of periclinal cell divisions. Based on the classified cell divisions, we assigned the cells to cell layers (Figure 52). The assignment started by initializing all cells in the first time point to be members of pericycle cell layer “P”. After a periclinal division occurred, the progeny was assigned to new cell layers. Upon division of a founder cell located in cell layer “P”, the daughter cell that was closer to the lateral root base was assigned to the inner layer “I”, whereas the daughter cell closer to the lateral root tip was assigned to the outer layer “O”. Following this principle, the daughter cells of a cell division in layer “I” were assigned to the cell layers “II” (closer to the lateral root base) and “IO” (closer to the lateral root tip) (Figure 52a). We applied the assignment algorithm to all datasets and reconstructed up to eight cell layers (Figure 52b, c).

We could identify a spatio-temporal pattern in the sequence of periclinal cell divisions in all analyzed datasets (Figure 52d). The timing of development and occurrence of periclinal cell divisions in the lateral root was consistent across all datasets and linked to the developmental stage given in cell numbers (Supplementary Table 15). The first periclinal cell division creates

an inner (“I”) and an outer layer (“O”) at a stage of 26 ± 6 cells. Except for dataset #130607, the outer layer always divided prior to the inner layer at 57 ± 6 cells. For dataset #130607, the developmental stage at which the divisions occurred, differed only by one cell. On average, the inner layer divided at a stage of 68 ± 14 cells.

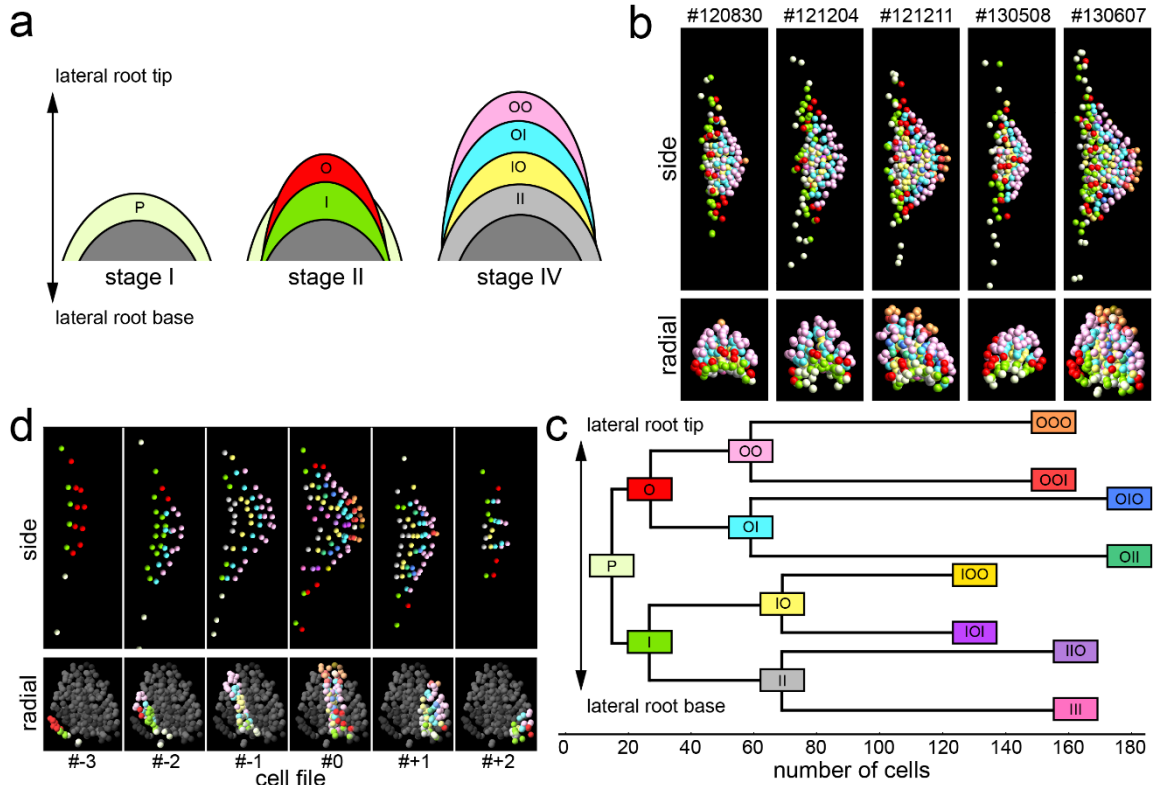


Figure 52 Occurrence of cell layers in lateral root development. (a) Scheme illustrating the layer assignment algorithm. The stage of the lateral root is linked to the number of cell layers (compare Supplementary Figure 10). (b) The sequence of periclinal cell divisions in all datasets in side and radial view. Cell nuclei that emerge through the same number of periclinal cell divisions are shown in the same color. (c) The sequence of periclinal cell divisions in dataset #130607 shown for each cell file. (d) Average occurrence of the different cell layers in number of cells.

In summary, the developed module for computer-aided cell tracking and lineage reconstruction allowed the generation of the first four-dimensional atlas of lateral root development. We observed a correlation of the occurrence of formative cell divisions with the developmental stage of the primordium and the cell division round of the founder cells. An alternating behavior in the orientation of cell division orientation was a prominent feature in lateral root development. After the second cell division round most cells switched the orientation of their division plane by 90 degrees to the previous cell division. The emergence of cell layers is a consequence of a stereotypic sequence of periclinal cell divisions.

5 Discussion

Image-based systems biology approaches rely on the application of advanced light microscopy. Huge amounts of complex multidimensional image datasets are produced. In order to extract the relevant information from the acquired data, efficient modules for an image analysis pipeline are required. The aim of this work was to develop such modules and to evaluate existing tools for an image analysis pipeline. The development was guided by the underlying interdisciplinary research projects. The benefit of this application-driven approach was two-fold: first, the developed solutions could be directly applied and validated on real datasets. Second, the modules were successfully used in these applications to address scientific questions.

5.1 Multidimensional image datasets require automated data management

Automated data management has gained an important role in science. The developed module for automated data management simplifies (1) structuring and naming of multidimensional image datasets, (2) three-dimensional cropping, (3) image compression and (4) the computation of maximum projections. The generalized nomenclature facilitates the usage of a common data importer interface that is applicable to various processing modules. The developed file name organization generalizes the file nomenclature, which is independent of the experimental or microscopy setup. The file names further describe the underlying data structure (e.g. if multiple views were acquired) and indicate already applied processing steps. This allows for a clear identification and communication of the data by different users but also supports the interoperability of modules in the image analysis pipeline.

Efficient processing of image datasets is only possible when data has been reduced to an amount that is manageable by current computers and workstations. A reduction to the relevant information content of the images further improves runtime and accuracy of subsequent image analysis steps. For example, a segmentation algorithm will detect less outliers in the background region if the dataset is cropped to the region of interest beforehand. The developed module simplifies the cropping of the data to the region of interest with a user-friendly interface and automatically applies the settings to image stacks of an entire time series.

The performance of lossless ZIP and lossy JPEG2000 image compression was evaluated for large multidimensional image datasets produced with LSFM. The built-in ZIP compression of

the TIF format yielded compression ratios of at least 2.4. No information was lost and intensity values remain unchanged with ZIP compression. It is therefore considered a suitable default compression method for large image datasets and should be used through the entire lifetime of the datasets. JPEG2000 compression yielded even higher compression ratios and reduced the amount of data to less than 1% of the raw data volume. However, because the JPEG2000 algorithms are optimized to preserve visual quality, the integrity of the underlying data might get destroyed (Bernas et al., 2006). To preserve image quality, we introduced a constant scaling of the image intensities, which made artifacts of JPEG2000 compression become almost imperceptible and undetectable. However, varying intensity levels (e.g. as a function of time) are problematic for constant scaling factors. For example, high scaling factors could lead to information loss whereas too low scaling factors could result in the introduction of compression artefacts. A solution could be the application of a dynamic scaling factor that adapts to the intensity level of the underlying images. However, this would also imply that the intensity information is not directly available (e.g. in a subsequent analysis step). Furthermore, the applied JPEG2000 compression per single plane resulted in a large amount of single files. Although a description for three-dimensional compression exists for the JPEG2000 standard, only few (mostly commercial) implementations actually exist (e.g. from the companies Comprinato and Pictools). Moreover, state-of-the-art image processing and rendering software (e.g. *FIJI*, *Arivis* or *Mathematica*) mostly do not support the three-dimensional JPEG2000 codec. In its present form, JPEG2000 is therefore considered useful for long-term storage of datasets, when intensities can be changed and the data can be stored as single images. Keeping track of the scaling factors (i.e. storing them in separate files) would allow the complete reconstruction of the original data whenever required. Besides, one should keep alternative solutions in mind. For example, Amat *et al.* presented an alternative approach by combining a lossless compression similar to JPEG2000 with the HDF5 container format. They reported compression ratios and timings superior to existing three-dimensional compression algorithms (Amat et al., 2015). Other approaches used anisotropic diffusion algorithms for compression and report high performance in terms of quality and speed (Peter and Weickert, 2014; Schmaltz et al., 2009).

5.2 BugCube enables web-based visualization and sharing

Efficient solutions for visualization and sharing large image datasets are required. In this context, platforms where multi-dimensional image datasets can be uploaded and shared have become an important topic (Burel et al., 2015; Hill, 2008; Kvilekval et al., 2010). The developed web-based application *BugCube* provides a platform for viewing and accessing images. The image dataset can be quickly browsed for features of interest. This is facilitated by displaying images that were downsampled to 1% of the original data volume. High quality

images (raw data) can be viewed upon request. Such a repository supports the exchange of knowledge between collaborating scientists and further increases the potential of acquired datasets.

5.3 Reconstruction of multiview image datasets

5.3.1 The first multiview reconstruction of *Tribolium* and *Gryllus* embryogenesis

Any quantitative analysis (e.g. segmentation or cell tracking) benefits from the availability of full three-dimensional information. We configured and applied the software *Image Stitching* and *Multi Reconstruction* for the reconstruction of *Tribolium* and *Gryllus* datasets that were acquired together with fluorescent beads. In contrast to the single views, the reconstructed images provided the complete three-dimensional information of an embryo in a single image. Moreover, the reconstruction in conjunction with image deconvolution (MVD) improved the image quality and the resolution of cell nuclei in *Tribolium* datasets (Preibisch et al., 2010). Due to light scattering, the reconstructed datasets did not provide high image quality in the center of the embryos. Structured illumination allows to discriminate against background and blur using striped illumination patterns (Keller et al., 2010). With structured illumination, the image contrast is improved, which is beneficial for the raw image quality, the bead detection step (i.e. beads are detected more accurately) and thus the resulting quality of the fused image could be improved.

For the *Gryllus* dataset, we observed high registration errors for later time points. By visual inspection of the images, we found that muscular movement began early in the embryogenesis of *Gryllus*. This led to a noticeable motion blur in the acquired images. The *Gryllus* embryo was placed on a thin film of agarose. Motion of the embryo could have caused a displacement of the fluorescent beads in the sample. The displacement could have impaired the bead detection accuracy of the software. Whether similar effects occur in other datasets of *Gryllus* needs to be clarified in future applications of the module. An alternative sample preparation similar to that of *Tribolium* could be considered prevent this issue.

5.3.2 A general workflow for reconstruction of multiview image datasets

The bead-based approach in conjunction with the existing *FIJI* plugin provides an efficient and accurate way to obtain a reconstruction of multiview image datasets. Under the premise that a sufficient number of detectable fluorescent beads is available, the approach is completely independent of the underlying specimen (Preibisch et al., 2014). The reconstructed datasets provide an improved resolution (Wu et al., 2016) and allow a drastic reduction of the data amount. The approach is therefore considered suitable as a module for multiview reconstruction in the image analysis pipeline.

The three-dimensional acquisition of image datasets with LSM followed by multiview reconstruction provides new opportunities for developmental biology. Instead of generating a tailored dataset for a specific scientific question, the systematic acquisition and reconstruction of image datasets allow the generation of a comprehensive collection of datasets. The collection can be used by the scientific community and provides a resource to answer various scientific questions (Strobl et al., 2017). For example, the high quality of the datasets allows a systematic quantitative characterization of *Tribolium* development similar to *Drosophila* (Amat et al., 2014). The combination of the module with subsequent segmentation and cell tracking will provide the complete reconstruction of cell lineages and cellular dynamics (Chhetri et al., 2015). This opens up a broad range of possibilities for high-content analysis of cellular and organismal dynamics in entire developing organisms. This enables the analysis of wild type or mutated embryogenesis in its entirety (Strobl and Stelzer, 2016). The complete workflow is not limited to *Tribolium* or *Gryllus* embryos but readily applicable to other specimen (e.g. three-dimensional cell cultures) as well.

5.4 Multiscale quantitative characterization

The developed module for quantitative characterization complements cell nuclei segmentation by powerful methods from graph theory and computational topology. The module outputs cell nuclei features obtained from the segmentation extended by features obtained from graph and topological analysis. Thus, complex image datasets are reduced to simple lists that can be used in subsequent data analysis. The number of extracted features goes well beyond most existing programs (Bilgin et al., 2013; Friebel et al., 2015; Lin et al., 2007; Morales-Navarrete et al., 2015; Stegmaier et al., 2014). The quantitative characterization enables the assessment of cellular heterogeneity and synchrony at the single cell level. Such a level of detail is important to understand mechanisms like cellular differentiation or local control mechanisms exerted by neighboring cells.

5.4.1 Densely packed cell nuclei in diverse spheroid datasets are accurately identified

The cell nuclei segmentation is objective, robust, insensitive towards intensity variations and capable of separating apparently touching cell nuclei. The initial local segmentation is less sensitive towards heterogeneous intensity distributions than global thresholding (Uchida, 2013). The multiscale LoG filter algorithm is capable of detecting marker points for irregular nuclear shapes, varying intensity distributions and in regions of high cell density. In accordance with findings from other studies (Chittajallu et al., 2015; Morales-Navarrete et al., 2015), the identified combination of methods is a promising candidate for a universally applicable cell nuclei segmentation approach. This is also elucidated by the cell nuclei segmentation results that were obtained for diverse spheroid datasets in the four applications. Independent of cell

nuclei size, nuclei labelling, microscope setup and cell line accurate cell nuclei segmentation was achieved.

5.4.2 Alpha shapes provide a geometrical model for arbitrarily shaped cell aggregates

Based on the positions of cells and a suitable parameter value for alpha, the computed alpha shapes provide a geometrical model of arbitrarily shaped cell aggregates. Compared to spherical harmonics (Khairy et al., 2008a, 2008b), the alpha shape excels with low computational costs, applicability to arbitrarily shaped cell clusters and provide an accurate geometrical model of solid three-dimensional objects.

5.4.3 Cell graphs investigate the spatial topology of cell aggregates

Graphs give a formal description of complex, multidimensional networks. Cell graphs inherently contain the information to compute global and local features that define the cell aggregate topology and relationships between individual cells (Schäfer et al., 2015). To investigate local cell structures in a spheroid, two implementations extend cell graphs to three-dimensional spatial networks: a purely distance-based variant (proximity cell graph, PCG) and a variant that is based on Delaunay triangulation (Delaunay cell graph, DCG). In conjunction with the alpha shape surface, these graphs capture structural patterns as a function of depth. PCG and DCG model different modes of cell neighborhood. The PCG provides a measure of local cell density and connects all cells that interact across long ranges. Neighborhood in the DCG is an approximation of which cells are in direct contact with each other. For the PCG, the number of edges and hence the computation time increases with the distance threshold. The DCG is less sensitive to the chosen distance threshold, since the number of edges is restricted by the Delaunay triangulation. Since the Delaunay triangulation is the dual form of the Voronoi tessellation, an approximation of cellular shape can be readily obtained from the DCG (Friebel et al., 2015; Schaller and Meyer-Hermann, 2005). Features extracted from the segmentation can be incorporated into the cell graphs to analyze the spatial correlation of cell morphology, cell type or expression profiles.

5.5 Quantitative characterization of spheroids under stress conditions

Spheroids are important model systems to investigate questions in cell and translational biology. Experiments rely on chemical (Friedrich et al., 2009) or mechanical (Delarue et al., 2014) perturbations of spheroids. Thus, stress conditions such as the influence of mechanical constriction on spheroid morphology or the impact of chemical compounds on cell proliferation, differentiation and viability can be studied in a three-dimensional context. Apart from technical advances of three-dimensional cell culturing and microscopy, automated image analysis is still

limited. In this work, a combination of three-dimensional cell culture, optical clearing, LSMF imaging and three-dimensional image analysis was employed. The combination of optical clearing and LSMF resulted in full penetration depth, homogeneous intensity distribution and good axial resolution. Thus, we obtained the first comprehensive dataset of high quality images for subsequent quantitative characterization of spheroids at the single cell level. Such data is required to progress from population averages. The developed approach was employed to investigate the impact of different stress conditions on spheroids in four applications.

5.5.1 Application I: nutrient and oxygen deficiency

Structural heterogeneity in differently sized T47D spheroids

The complete approach was applied to differently sized spheroids generated from T47D breast carcinoma cells. The extension of global spheroid properties such as volume, cell number or the distribution of cell nuclei volumes by objective features derived cell graphs and alpha shapes allowed the quantitative investigation of structural features in spheroids. The derived cell locations serve as quantitative data for the development of advanced statistical analysis (Dini et al., 2016). The boundary between outer and core regions could be identified and we could show that an extensive core region in breast carcinoma spheroids arises when the spheroid features at least 30,000 cells. The extent of the core region is not proportional to the diameter of the spheroid. Larger spheroids exhibit a thinner outer region (50% of the spheroid radius) compared to medium-sized spheroids (75% of the spheroid radius). Earlier studies have shown, that the thickness of the outer region strongly depends on nutrient and oxygen supply (Sutherland, 1988) that vary with spheroid size (Mueller-Klieser, 1984). Our results suggest a conservation of the outer region cell mass.

Furthermore, we found structural inhomogeneity in the different concentric cell layers. All spheroids exhibited cell densities between 35 and 70 cells/u.v. with patches of higher and lower cell density in the outer region. Recent work with two-dimensional cell cultures revealed the importance of the cell density on cellular behavior and function (Mo et al., 2014) including cell differentiation (Gorman et al., 2014), cell proliferation and the response to compounds (Greene et al., 2016). Similar findings were obtained for real tissues (Mateus et al., 2015). We suppose that the observed differences in cell density affect individual cells and thus, such heterogeneity needs to be incorporated into existing models. Furthermore, cell nuclei volume, cell number and the spheroid diameter allow validation and refinement of mathematical models for spheroid growth dynamics (Drasdo and Höhme, 2005).

5.5.2 Application II: apoptosis induction

Drug screening applications require objective, three-dimensional characterization

Many publications provide new methods and protocols for cell culturing, spheroid preparation and imaging in the context of drug screening. However, there is currently only a limited amount

of work focusing on reliable quantitative characterization. Many analysis concepts for two-dimensional image datasets have not yet been implemented on a broad scale and not adapted for three-dimensional applications. Consequently, previous studies mostly employed low resolution imaging of spheroids and single two-dimensional images to analyze the overall spheroid size or intensities.

Apoptosis induction is associated with morphological changes in spheroids

The quantitative characterization enabled the association of apoptosis induction with changes in spheroid and nuclear morphology in three-dimensional images. Spheroid and cell nuclei morphology could be objectively quantified in differently sized spheroids, generated from T47D cells. Previous analysis of two-dimensional images showed that apoptosis induction led to a decrease in cell nuclei size in immortalized retinal pigment epithelial cells (Eidet et al., 2014), primary neuronal cells (Daniel and DeCoster, 2004) or Hep2 squamous carcinoma cells (Helmy and Azim, 2012). In accordance with these results, we observed that in three-dimensional images the cell nuclei volume of T47D cells was reduced to about 57% after 72 hours of apoptosis induction. In future experiments, it needs to be clarified to what extent the reduction of spheroid and cell nuclei volume is due to the dehydration and optical clearing procedure.

We could investigate for the first time whether cell nuclei morphology differs as a function of depth into differently sized T47D spheroids. We did not detect a dependency of the cell nucleus volume on the relative location of the cell within the spheroid. This suggests that there is no regional difference in the response of cells to apoptosis induction. However, the evaluation of a temporal difference in the response of cells to apoptosis induction requires suitable controls at each time point, which were not available for this work.

The quantitative characterization was efficient in the detection of bright blob-like cell nuclei. However, cell nuclei exhibit different phases of structural change during apoptosis (Toné et al., 2007). The chromatin first condenses into a ring-like structure at the interior surface of the nuclear envelope. Then, the ring structure develops discontinuities and the nucleus begins to shrink. Finally, the nucleus forms apoptotic bodies and collapses. Mooney and coworkers also identified diverse cell nuclei morphologies of T47D cells treated with 1 μ M Staurosporine and stained with DAPI in two dimensional images (Mooney et al., 2002). In accordance, we observed that T47D cell nuclei morphologies upon apoptosis induction strongly deviated from bright, blob-like objects in three-dimensional images. Thus, the segmentation procedure requires optimization in order to cover these different cell nuclei morphologies.

Towards quantitative drug screening with three-dimensional model systems

Confocal microscopy delivers high quality images of spheroids. The combination of optical clearing, confocal microscopy and the quantitative characterization module enabled the

extraction of valuable quantitative data from spheroids that exceeds existing approaches. However, the image quality (e.g. the resolution along the z dimension) was rarely good enough to obtain high quality cell nuclei segmentation results directly. Several pre-processing (i.e. filtering, background subtraction) and post-processing steps (selection of cell nuclei) were required to obtain reliable segmentation results. Higher image quality with an improved resolution along the z dimension could be obtained with LSFM. Recently, the high throughput LSFM (HT-LSFM) has been described that would allow the combination of LSFM with multiwell plates and thus high throughput and high image quality could be obtained (Pampaloni et al., 2015). Such experiments also provide the temporal resolution in order to study the cellular dynamics upon apoptosis induction.

In this project, we have focused on the combination of three-dimensional imaging with confocal microscopy and quantitative analysis at cellular detail in entire spheroids. The quantitative characterization delivered a number of informative phenotypic features of spheroid and cell nuclei that enable assessing the effectiveness of apoptosis inducing drugs based on cytometric measurements. For solid results, the number of spheroids per condition needs to be increased in future experiments and control experiments need to be performed for each time point. Due to the generality of the approach, it is then readily applicable to other compounds and cell lines.

5.5.3 Application III: autophagy induction

Targeting alternative cell death pathways is an attractive strategy for improving anti-tumor therapy (Schleicher et al., 2010). Autophagy is essential for cellular homeostasis and plays important roles in developmental and pathogenic processes. Deregulation of autophagy signaling is associated with cancer progression. The mTOR pathway is key regulatory factor of autophagy signaling and therefore affects cell cycle progression, proliferation and survival. The mTOR pathway is upregulated in many breast cancers (Kerekatte et al., 1995) leading to inhibition of autophagy. Thus, mTOR inhibition is tested as a novel target for breast cancer therapy. Rapamycin and its derivatives induce autophagy by inhibition of mTOR activity and are commonly used as immunosuppressive drugs (e.g. after organ transplantation). However, long-term effects of autophagy induction on breast cancer tissues have not been studied.

In this project, we investigated the effects of long-term treatment with Rapamycin on T47D spheroids. For quantitative characterization of the spheroids, we took advantage of the developed module for quantitative characterization and compared the effects of autophagy induction, autophagy inhibition with suitable controls on multiple scales including properties of the complete spheroid, the local cell neighborhood and individual cells.

Autophagy induction reduces spheroid volume and cell number

Consistent with previous studies we found that Rapamycin inhibits cell growth and proliferation through induction of autophagy in breast cancer spheroids (Shapira et al., 2006). Consequently, we observed a significantly reduced spheroid volume and decreased number of cells in breast cancer spheroids treated with Rapamycin. The effects were identical for the positive control EBSS. On the other side, Bafilomycin treatment (autophagy inhibition) did not affect cell number or spheroid volume.

We observed that the cell nuclei volume significantly decreased, whereas the cell density in spheroids increased upon autophagy induction. Under nutrient-deprived conditions, cells compensate the lack of nutrients by autophagy induction. We hypothesize that the reduced cell nuclei volume also indicates an overall reduced volume of the cells, which induces a higher cell density. In a recent study, similar observations were made for growth factor-deprived lymphocytes (Hecht et al., 2016). Hecht and coworkers conclude that from a biophysical point of view, the cell maximizes its metabolic efficiency by decreasing the diffusion time of molecules and on the other hand increasing the probability of interaction between certain molecules.

5.5.4 Application IV: acute mechanical constriction

Cell nuclei act as a mechanosensor for applied mechanical forces

Mechanical forces play a fundamental role in normal tissue and malignant tissue. Many studies have focused on the analysis of compressive stress in the context of tumor biology. In this project, we investigated the impact of acute mechanical constriction on mammary epithelial spheroids.

At short term, we did not measure an impact of acute mechanical constriction of spheroids on the internal morphology. Instead, we observed a delayed reduction of cell nuclei volume after 24 hours. Cell nuclei increased again after 96 hours. In a tissue context, physical forces arise from forces generated by the extracellular matrix, neighboring cells or intracellular processes. These forces are transmitted through the cytoskeleton to which the cell nucleus is attached to and therewith directly exposed to mechanical forces from the outside. Thus, it acts as a direct mechanosensor and extracellular forces can be measured by nuclear deformation and volume reduction (Guilak et al., 2000). Nuclear shape changes can directly influence transcription independent of other signals from the cytoplasm (Dahl et al., 2008). Our observations imply that the cells actively adapt to the mechanical stress situation with a certain delay. The observed increase of the cell nucleus volume after 96 hours might indicate that the applied force was compensated by reorganization of the cells. Similar minimization of the acting mechanical force on the nucleus has been observed for example in vascular endothelial cells (Hazel and Pedley, 2000).

Recent studies showed that continuous compressive pressure only weakly affected the apoptosis rate in spheroids (Delarue et al., 2014). Continuous compressive pressure increased the proliferation rate of the cells that depended on the duration of the treatment but impaired the mitotic progression of the cells (Desmaison et al., 2013). Our results indirectly suggest that acute mechanical constriction does not affect the proliferation rate (i.e. the cell numbers between constricted and control spheroids are similar). The apoptosis rate in constricted and control spheroids was similar and no dependency on the duration of the constriction was detectable. An analysis of the proliferation rate in HC11 spheroids could be performed to confirm these results.

Cell density is homogeneous in spheroids under acute mechanical constriction

The cell nucleus volume was observed to be homogeneous throughout the entire spheroid at the investigated time points. Cells on the periphery of the spheroid did not show a different response to the mechanical force than those cells in the core of the spheroid. We also did not detect differences in the cell density as a function of depth. In contrast to T47D spheroids, no necrotic core was detectable. Compressive stress has been shown to have different impact on cells in the core and periphery of spheroids (Montel et al., 2012). We cannot exclude the possibility of having missed differences in the distribution of cell nucleus volume or cell density. However, this would require analyzing the cellular dynamics during or fixing the spheroids directly after the application of the mechanical constriction. A fixation in the minute range after constriction is experimentally possible.

Acute mechanical constriction is relevant in other tissue contexts

We showed an example of measuring the deformation in the context of breast tissue (mammary gland) in a three-dimensional cell culture model system. This phenomenon is found *in vivo* during the weaning process when milk accumulates in the mammary gland due to decreased suckling of pups. This becomes more prominent when suckling is ceased and involution of the mammary gland is forced. Acute mechanical stress is also relevant in other tissue contexts. For example, traumatic brain injuries induce an elevated intracranial pressure that can be fatal to the patient. Our system can be applied in these contexts, if a suitable three-dimensional spheroid model is available.

5.6 Computer-aided cell tracking and lineage reconstruction

5.6.1 Simplified cell tracking and lineage reconstruction

Tracking cells and analyzing their lineage is required in many applications in developmental biology. A variety of approaches exist to fulfill this task in an automated manner (Amat et al., 2014; Fernandez et al., 2010). However, the quality of image datasets often restricts the applicability of automated solutions (Meijering et al., 2009). Interactive programs are required

to correct automated tracking results or to perform the tracking manually. The developed module *TrackGen* provides an interactive user interface that simplifies the task of cell tracking and lineage reconstruction in time-lapse image datasets.

5.6.2 The first four-dimensional atlas of lateral root development

We applied the module for cell tracking in image datasets of lateral root development in *Arabidopsis thaliana*. This enabled the reconstruction of the first curated and complete four-dimensional atlas of cell lineages in lateral root development. Our atlas captures spatial locations and divisions of all contributing cells in five lateral root primordia (LRP) for up to 30 hours of development. The data analysis includes the necessary pre-processing steps of spatial and temporal synchronization necessary to compare the datasets. We analyzed the synchronized datasets for patterns that govern lateral root development on the level of the complete organ (e.g., the number of cells and lateral root volume), local structures (e.g. the contribution of cell files and the occurrence of cell layers) and individual cells (e.g. the cell divisions).

5.6.3 Lateral root development is conserved on the organ level

Previous studies reported that initiated primordia did not develop at uniform rates, the development of some LRPs was arrested (Dubrovsky et al., 2006) or delayed (Moreno-Risueno et al., 2010). We observed that the onset of proliferation in lateral root development differs, but once proliferation started, the profiles for cell number, volume and height growth of the LRPs follow a similar trend. We did not detect any correlation between the proliferation rates and the day/night cycle as observed for emerged lateral roots (Voß et al., 2015). The decrease in the length-to-width-ratio in all LRPs indicates the occurrence of a dome-shaped primordium and coincides with previous observations (Lucas et al., 2013).

5.6.4 Cellular contribution of cell lineages and cell files is variable

Previous studies lacked the spatio-temporal information to investigate the contribution of cell lineages and cell files to the LRP. For the first time, we showed that LRPs are formed from 5 to 8 cell files. The contribution of cell files was dominant in the core region and declined to the left and right of the master cell file. Lateral roots were formed from a variable number of 9 to 15 founder cells of which some emerged stochastically and showed a dominant contribution to the LRP, whereas the contribution of others was negligible.

In conclusion, the contribution behavior of founder cells depends on several factors. First, founder cells are located arbitrarily in the cell files along the primary root and thus, the extent to which a founder cell contributes to the LRP depends on its distance to the lateral root center (Kurup et al., 2005). Second, the formation of lateral roots was induced by gravistimulation (i.e. rotation by 90°). This stimulation results in a bending of the primary root inducing. Both factors result in a different microenvironment for cells in the pericycle cell layer, which promotes or

impedes cellular behavior such as proliferation or differentiation. For example, plant growth is fundamentally regulated by phytohormones such as auxin (Friml, 2003). Local gradients of auxin are altered (Ditengou et al., 2008) and lead to a different behavior of the cells that respond to these gradients (De Smet et al., 2007; Vieten et al., 2007).

5.6.5 Organ-level coupling in the occurrence of formative cell divisions

Cell divisions at early stages of lateral root development were clearly distinguishable as anticlinal, periclinal or radial. We observed a correlation of the occurrence of formative cell divisions with the developmental stage of the primordium and the cell division round of the founder cells. In later stages, the orientation of cell divisions was more variable. During development, the lateral root transforms from a bilateral symmetric system with a reference system defined by the primary root into a structure with radial symmetry. This radialization process of the lateral root (Lucas et al., 2013) explains the higher variability in the orientation of cell divisions. Except for the first anticlinal cell division, occurrence and orientation of individual cell divisions was not conserved. However, a dominant feature observed in all LRPs was the tendency of cell divisions orientation to be perpendicular to the orientation of the previous cell division.

The Arabidopsis root consists of a defined number of tissue cell layers (Dolan et al., 1993). Our results demonstrate that the emergence of cell layers is a consequence of a stereotypic sequence of periclinal cell divisions. These results are consistent the previous model for the generation of tissue layers (Malamy and Benfey, 1997).

5.6.6 Developmental instability and self-organization in lateral root development

The formation of lateral roots is an interplay of various factors. Environmental factors such as the availability of nutrients, water and light or the properties of surrounding tissues. The identity and behavior of a cell is strongly affected by its surrounding neighboring cells (Scheres et al., 1994). Several studies reported the important role of cell-to-cell communication in the coordination of lateral root development (Yue and Beeckman, 2014). Finally, the mechanical properties of the surrounding tissues guide the morphogenesis of lateral roots (Vermeer et al., 2014). Moreover, ‘developmental instability’ (Forde, 2009) was introduced as an additional parameter suggesting that intrinsic stochastic processes perturb the development and interplay with genetic and environmental factors resulting in phenotypic variation (Polak, 2003). In accordance, studies based on microscopic images and mathematical modelling propose that lateral root growth is manifested by local symmetry breaking events including local variations in growth rates, oblique cell divisions, auxin-related mutations or atypically shaped cells (Szymanowska-Pulka, 2013; Szymanowska-Pulka et al., 2012, 2014). Such perturbing effects are compensated and do not impair the final, stereotyped morphology of the lateral root. Our results suggest that an organ-level coupling of LRP growth and cell divisions could account for

the controlled emergence of the dome-shaped primordium. Such self-organization and non-determinism are widespread phenomena and typical for developing systems.

Atlases of reconstructed cell lineages and their analysis represent a tool for developmental biology and have been generated to study flower development in *Arabidopsis thaliana* (Fernandez et al., 2010) or even whole animal embryogenesis (Amat et al., 2014; Keller et al., 2008; Long et al., 2009). We generated and published the first, comprehensive atlas of cell lineages in lateral root formation (von Wangenheim et al., 2016). The atlas will be a valuable resource in order to understand the role and interplay of the different factors involved in LRP formation and to study how self-organization and non-deterministic behavior compares between different, organs, or even organisms.

6 Outlook and Conclusion

6.1 Future directions of the image analysis pipeline

In this thesis, several modules of an image analysis pipeline for three-dimensional fluorescence microscopy were developed or existing solutions were evaluated and optimized. The results of this thesis can be extrapolated towards the following topics:

- Multiview reconstruction as a basis for cell tracking and lineage tracing
- Generalized usage of the multiview reconstruction procedure
- Extension and application of the multiscale quantitative characterization
- Extension of the pipeline towards automated high-throughput processing

6.1.1 Multiview reconstruction as a basis for cell tracking and lineage tracing

In this thesis, the reconstruction of multiview image datasets with cell nuclei labelling was successfully achieved. In conjunction with cell tracking, this allows in-depth analysis of cellular dynamics (e.g. movement patterns), cell lineages and thus the generation of complete virtual embryos. In this context, the reconstructed datasets of *Tribolium* and *Gryllus* provide a basic resource.

First preliminary results towards automated tracking could already be obtained for cell nuclei in maximum projections along z of a H2B-GFP labeled *Tribolium* embryo (Figure 53). For the identification of the cells, the quantitative characterization module was suitable. Tracking was then performed by a greedy tracking algorithm that solely identified the continuation of the tracks by the distance to the closest objects in the next frame. The tracking worked well for the first hours of the recording until cells of the uniform blastoderm began to migrate during the gastrulation phase. Complete cell lineaging would require higher temporal resolution such that the daughter cells of a cell division can be accurately identified. Moreover, the embryo would have to be acquired from multiple views with subsequent multiview reconstruction to obtain the information of all cells in the embryo. Matching both requirements (high temporal resolution and multiview imaging) at the same time is difficult to fulfill. In specialized systems such as *SimView* (Tomer et al., 2012) or *MuVi-SPIM* (Krzic et al., 2012), the embryo is simultaneously acquired from multiple views. These systems were successfully applied in the reconstruction of cell lineages in *Drosophila melanogaster*. The applicability of such systems for *Tribolium* and the impact on the viability of the specimen need to be clarified.

With additional labels (i.e. for cell membranes), the combined approach of multiview reconstruction, tracking and lineaging provides comprehensive data about cell shape or expression patterns in the embryo as a function of time. Investigations at the cellular level and at the level of the entire embryo become feasible and system-related scientific questions can be addressed. Moreover, the morphogenetic variability in development could be compared systematically between different species.

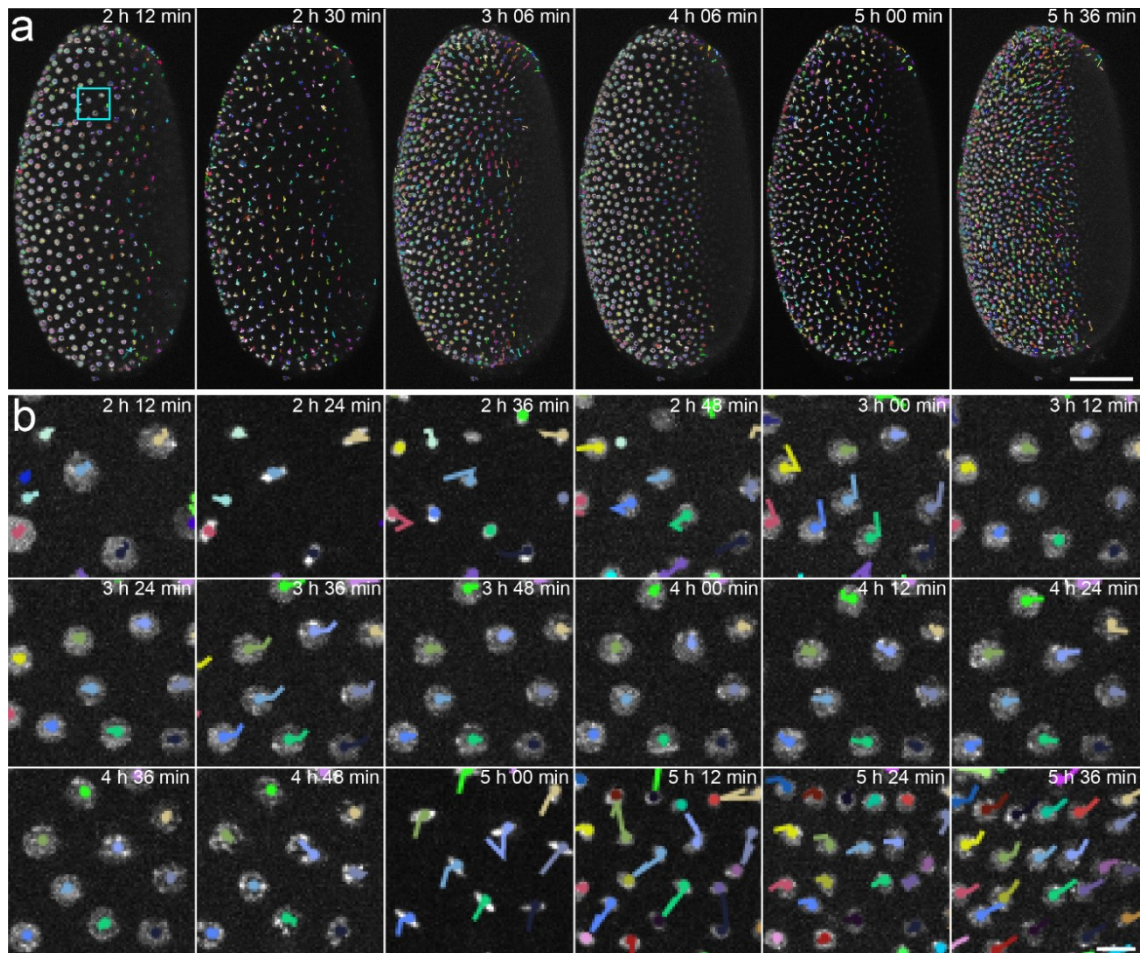


Figure 53 Cell nuclei tracking in *Tribolium castaneum*. (a) Maximum projections along z with cell nuclei tracks overlaid of a *Tribolium* dataset at six time points. Colored dots indicate the current location of a cell, whereas colored lines indicate the path from the previous three time points. The *Tribolium* transgenic line expresses H2B-GFP. The total duration of the time-lapse was 153 h at a time interval of 6 minutes. Scale bar: 100 μm . (b) Close-ups of the region indicated by a cyan box in panel a at different time points in steps of 12 minutes arranged in a montage. Scale bar: 10 μm . Microscope: mDSLm, illumination objective: 2.5x NA 0.06 Epiplan-Neofluar, detection objective: 10x NA 0.3 N-Achroplan, camera: Andor Clara, axial pitch: 2.58 μm , number of planes: 229, exposure time: 50 ms, laser power: 135 μW . Frederic Strobl provided the dataset.

6.1.2 Generalized usage of the multiview reconstruction procedure

The procedures for preparation of *Tribolium* and *Gryllus* embryos together with fluorescent beads and the subsequent multiview reconstruction can be extended to other samples such as three-dimensional cell cultures. In the Bachelor project of Elena Gosau, the sample preparation principle used for *Gryllus* embryos was optimized for T47D spheroids. Elena acquired spheroids from four views and the images were reconstructed with the *Multiview*

Reconstruction software. Interestingly, only few parameters (e.g. the concentration of fluorescent beads and some settings of the *Multiview Reconstruction* software) needed to be adjusted in order to obtain a reconstruction of the datasets. Due to the built-in image deconvolution procedure, the resulting images provided an image quality that was superior to the raw image quality. In conclusion, multiview reconstruction could be a valuable pre-processing step in order to obtain high quality images of entire spheroids.

6.1.3 Extension and application of the multiscale quantitative characterization

The multiscale quantitative characterization was developed based on cell nuclei in three-dimensional cell cultures. Cell nuclei represent a simple, yet powerful sensor for the state of the cell. For example, nuclear shape or texture characteristics are altered in cancer cells (Zink et al., 2004). Classical pathology relies on visual examination of these characteristics by experienced persons. The developed module bridges the gap between classical examination and an automated characterization using image analysis. It allows the extraction of quantitative data at multiple scales including individual cell nuclei, the cell neighborhood and the complete cell aggregate. The module proved its potential in analyzing the impact of four different stress conditions on spheroids. In the future, this approach could be extended towards the investigation of cell differentiation, the comparison between different cell lines, the arrangement of cells in heterotypic spheroids, spheroids under confined growth as well as mechanical or chemical perturbation (Figure 54).

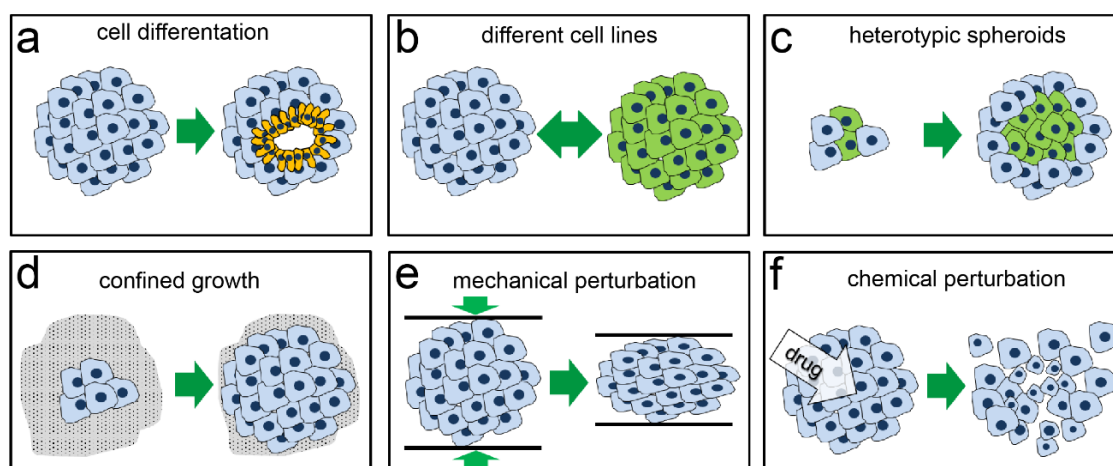


Figure 54 Future application scenarios of the quantitative characterization module. The developed module could provide valuable data in the analysis of (a) cell differentiation (b) different cell lines (c) heterotypic spheroids (d) confined spheroid growth (e) mechanical perturbation of spheroids (f) chemical perturbation of spheroids.

These applications would benefit from information about the cell shape. To obtain morphological information about cells, suitable cell membrane markers and an extension of the module towards cell segmentation are required (Meijering, 2012). An adaption of the module using the same principle of a marker-controlled watershed algorithm for cell segmentation already yielded promising results (result not shown).

In summary, future studies in three-dimensional cell and developmental biology require an image-based systems biology approach with measurements at the level of the single cell, the cell neighborhood and the whole system (Figure 55). The developed quantitative characterization module is not restricted to spheroids. For any set of high-quality images with nuclear staining, the module provides a multiscale characterization ranging from the single cell level to the cell microenvironment and the whole system.

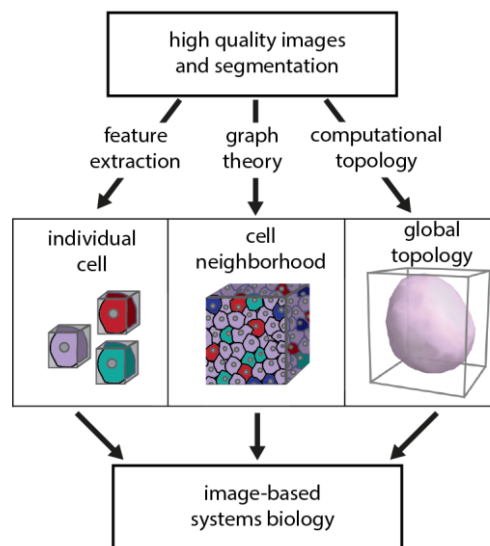


Figure 55 A multiscale approach for image-based systems biology. The established methods for advanced three-dimensional microscopy and image segmentation need to be extended by concepts from other fields including graph theory and computational topology. This will provide a system-level understanding of the architecture and function of tissues, organs and organisms. Figure adapted from Schmitz *et al.*, 2017.

6.1.4 Extension of the pipeline towards automated high-throughput processing

Future, high-throughput image analysis assays will require an all-in-one pipeline that is fast, efficient, adaptable yet easy to use and outputs the results in suitable formats, such as reports. Such a pipeline includes a setup platform in which specialized processing modules can be chosen from a predefined palette and parameters adjusted according to individual needs. After the setup, the user ideally launches the pipeline and picks up the results after processing. From a computer scientist's point of view, this idea of an all-in-one pipeline poses several tough requirements that need to be addressed: (1) The input data have to follow a standard, which encompasses the data quality, the file, name nomenclature and data structure. (2) The processing steps of the pipeline need to be modularized with a high level of parameterization. (3) These modules need to be executable independent from each other and interfaces between the modules need to be clearly defined. (4) The setup platform has to be user-friendly allowing people with varying levels of expertise to use the pipeline successfully.

In this thesis, I addressed several of these requirements of an all-in-one solution and developed or optimized modules for data management and visualization, multiview reconstruction, multiscale quantitative characterization, computer-aided cell tracking and lineage

reconstruction. The transition between modules still requires transformation steps such that the output from one module fits the input of the other module. This might require refinement and optimization of the present modules.

In the future, a module for automated data analysis and reporting should extend the pipeline. Such a module should be at the end of the pipeline and allow the generation of standard reports. For example, a report of a spheroid datasets includes metadata about the experimental setup, the acquisition, the applied pipeline modules (including parameter settings), global quantitative measurements (e.g. cell number, volume or cell density), detailed plots (e.g. cell density as a function of depth) and visualizations (e.g. maximum projections or cross sections).

With the increasing amount of data generated per experiment, the execution of the pipeline should furthermore be directly connected to the image acquisition phase. For example, during long-term experiments, image compression, deconvolution or other processing steps could be performed during the idle time of the microscope workstation in-between two subsequent acquisition processes. In this context, the efficiency of the pipeline modules would largely benefit from implementations on GPUs (Narayanaswamy et al., 2010) or the usage of high performance computing clusters (Schmied et al., 2016).

6.2 Concluding remarks

The quality of the underlying image data determines success or failure of any image analysis process. Besides, the datasets produced are typically much larger and complex than actually needed to address a specific research question. These points highlight the importance of interdisciplinary collaborations with active dialogues and synergy between computer scientists and experimental biologists throughout all steps of the pipeline. Thereby, the generation of datasets with suitable data volume and image quality for subsequent image analysis is ensured.

In the last years, advanced light microscopy has rapidly developed. Especially, with time-resolved imaging of large specimen, the amount and complexity of the generated data has increased dramatically. In conjunction, the fields of developmental and cell biology undergo a transition from purely qualitative to demanding quantitative analysis (Pantazis and Supatto, 2014). This development raises the need for efficient management, processing and analysis of the data. In practice, it turns out to be very tough to come up with one image analysis pipeline that works across multiple applications. The higher the diversity of datasets in the applications the more parameters will have to be included and fine-tuned in order to obtain reasonable output from the pipeline.

As outlined in a recent article in *Nature Biotechnology*, the main goal of automated image analysis pipelines is to increase our knowledge of biological systems and to develop new

hypotheses from the resulting data (Meijering et al., 2016). Depending on the application, not all modules of a pipeline are required for each application and a feedback from individual modules will guide the application of other modules. Thus, an image analysis pipeline cannot be seen as an unidirectional flow of modules, in which one follows the other. Instead, it is important to have a strong modular architecture with clearly defined interfaces at the input and output sites.

This work provided and discussed efficient, user-friendly solutions to these requirements. The modules were successfully applied in interdisciplinary research projects in cell and developmental biology. In future work, the individual modules should be extended to incorporate them into a holistic platform. Such a platform could comprise an open source framework that joins the knowledge and effort of the scientific community. Under the premise of a controlled process (Cardona and Tomancak, 2012; Ince et al., 2012), scientists with expertise in bioimage informatics contribute to this platform, by developing, using and optimizing modules for image processing and analysis. This approach is not restricted to image-based systems biology, but can also be applied in medical diagnostics or industrial high-throughput applications.

7 References

- Alessandri, K., Sarangi, B.R., Gurchenkov, V.V., Sinha, B., Kießling, T.R., Fetler, L., Rico, F., Scheuring, S., Lamaze, C., Simon, A., et al. (2013). Cellular capsules as a tool for multicellular spheroid production and for investigating the mechanics of tumor progression in vitro. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 14843–14848.
- Amat, F., Lemon, W., Mossing, D.P., McDole, K., Wan, Y., Branson, K., Myers, E.W., and Keller, P.J. (2014). Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data. *Nat. Methods* *11*, 951–958.
- Amat, F., Höckendorf, B., Wan, Y., Lemon, W.C., McDole, K., and Keller, P.J. (2015). Efficient processing and analysis of large-scale light-sheet microscopy data. *Nat. Protoc.* *10*, 1679–1696.
- Andre, F., Berrada, N., and Desmedt, C. (2010). Implication of tumor microenvironment in the resistance to chemotherapy in breast cancer patients. *Curr. Opin. Oncol.* *22*, 547–551.
- Ansari, N., Hardung, S., Hötte, K., Rakel, S., Antonietti, P., Kögel, D., Stelzer, E.H.K., and Pampaloni, F. (2014). Quantifying the autophagy-triggering effects of drugs in cell spheroids with live fluorescence microscopy. *Methods Mol. Biol.* *1165*, 19–29.
- Antony, P.M.A., Trefois, C., Stojanovic, A., Baumuratov, A.S., and Kozak, K. (2013). Light microscopy applications in systems biology: opportunities and challenges. *Cell Commun. Signal.* *11*, 24.
- Aredia, F., Guamán Ortiz, L.M., Giansanti, V., and Scovassi, A.I. (2012). Autophagy and Cancer. *Cells* *1*, 520–534.
- Bacus, S.S., Goldschmidt, R., Chin, D., Moran, G., Weinberg, D., and Bacus, J.W. (1989). Biological grading of breast cancer using antibodies to proliferating cells and other markers. *Am. J. Pathol.* *135*, 783–792.
- Bartholomä, P., Gorjup, E., Monz, D., Reininger-Mack, A., Thielecke, H., and Robitzki, A. (2005). Three-dimensional in vitro reaggregates of embryonic cardiomyocytes: a potential model system for monitoring effects of bioactive agents. *J. Biomol. Screen.* *10*, 814–822.
- Bergeest, J.-P., and Rohr, K. (2012). Efficient globally optimal segmentation of cells in fluorescence microscopy images using level sets and convex energy functionals. *Med. Image Anal.* *16*, 1436–1444.
- Bernas, T., Asem, E.K., Robinson, J.P., and Rajwa, B. (2006). Compression of fluorescence microscopy images based on the signal-to-noise estimation. *Microsc. Res. Tech.* *69*, 1–9.
- Bilgin, C., Demir, C., Nagi, C., and Yener, B. (2007). Cell-Graph Mining for Breast Tissue Modeling and Classification. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (IEEE)*, pp. 5311–5314.
- Bilgin, C.C., Kim, S., Leung, E., Chang, H., and Parvin, B. (2013). Integrated profiling of three dimensional cell culture models and 3D microscopy. *Bioinformatics* *29*, 3087–3093.
- Böttger, K., Hatzikirou, H., Voss-Böhme, A., Cavalcanti-Adam, E.A., Herrero, M.A., and Deutsch, A. (2015). An Emerging Allee Effect Is Critical for Tumor Initiation and Persistence. *PLoS Comput. Biol.* *11*, e1004366.
- Bradshaw, A.D. (2006). Unravelling phenotypic plasticity -- why should we bother? *New Phytol.* *170*, 644–648.
- Bredel-Geissler, A., Karbach, U., Walenta, S., Vollrath, L., and Mueller-Klieser, W. (1992). Proliferation-associated oxygen consumption and morphology of tumor cells in monolayer and spheroid culture. *J. Cell. Physiol.* *153*, 44–52.
- van den Brink, S.C., Baillie-Johnson, P., Balayo, T., Hadjantonakis, A.-K., Nowotschin, S., Turner, D.A., and Martinez Arias, A. (2014). Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* *141*, 4231–4242.

- Buck, T.E., Li, J., Rohde, G.K., and Murphy, R.F. (2012). Toward the virtual cell: Automated approaches to building models of subcellular organization “learned” from microscopy images. *BioEssays* 34, 791–799.
- Burel, J.M., Besson, S., Blackburn, C., Carroll, M., Ferguson, R.K., Flynn, H., Gillen, K., Leigh, R., Li, S., Lindner, D., et al. (2015). Publishing and sharing multi-dimensional image data with OMERO. *Mamm. Genome* 26, 441–447.
- Cardona, A., and Tomancak, P. (2012). Current challenges in open-source bioimage informatics. *Nat. Methods* 9, 661–665.
- Cheng, J., and Rajapakse, J.C. (2009). Segmentation of clustered nuclei with shape markers and marking function. *IEEE Trans. Biomed. Eng.* 56, 741–748.
- Cheng, G., Tse, J., Jain, R.K., and Munn, L.L. (2009). Micro-environmental mechanical stress controls tumor spheroid size and morphology by suppressing proliferation and inducing apoptosis in cancer cells. *PLoS One* 4, e4632.
- Chessell, A. (2017). An Overview of data science uses in bioimage informatics. *Methods* (in press).
- Chhetri, R.K., Amat, F., Wan, Y., Höckendorf, B., Lemon, W.C., and Keller, P.J. (2015). Whole-animal functional and developmental imaging with isotropic spatial resolution. *Nat. Methods* 12, 1171–1178.
- Chittajallu, D.R., Florian, S., Kohler, R.H., Iwamoto, Y., Orth, J.D., Weissleder, R., Danuser, G., and Mitchison, T.J. (2015). In vivo cell-cycle profiling in xenograft tumors by quantitative intravital microscopy. *Nat. Methods* 12, 577–585.
- Cho, S., Haralick, R., and Yi, S. (1989). Improvement of Kittler and Illingworth’s minimum error thresholding. *Pattern Recognit.* 22, 609–617.
- Chuang, H.-Y., Hofree, M., and Ideker, T. (2010). A decade of systems biology. *Annu. Rev. Cell Dev. Biol.* 26, 721–744.
- Ciarletta, P., Ambrosi, D., Maugin, G.A., and Preziosi, L. (2013). Mechano-transduction in tumour growth modelling. *Eur. Phys. J. E* 36, 23.
- Dahl, K.N., Ribeiro, A.J.S., and Lammerding, J. (2008). Nuclear shape, mechanics, and mechanotransduction. *Circ. Res.* 102, 1307–1318.
- Daniel, B., and DeCoster, M.A. (2004). Quantification of sPLA2-induced early and late apoptosis changes in neuronal cell cultures using combined TUNEL and DAPI staining. *Brain Res. Protoc.* 13, 144–150.
- Degenhardt, K., Mathew, R., Beaudoin, B., Bray, K., Anderson, D., Chen, G., Mukherjee, C., Shi, Y., Gélinas, C., Fan, Y., et al. (2006). Autophagy promotes tumor cell survival and restricts necrosis, inflammation, and tumorigenesis. *Cancer Cell* 10, 51–64.
- Delarue, M., Montel, F., Vignjevic, D., Prost, J., Joanny, J.-F., and Cappello, G. (2014). Compressive stress inhibits proliferation in tumor spheroids through a volume limitation. *Biophys. J.* 107, 1821–1828.
- Demir, C., Gultekin, S.H., and Yener, B. (2005). Augmented cell-graphs for automated cancer diagnosis. *Bioinformatics* 21, ii7-12.
- Depeursinge, A., Foncubierta-Rodriguez, A., Van De Ville, D., and Müller, H. (2014). Three-dimensional solid texture analysis in biomedical imaging: Review and opportunities. *Med. Image Anal.* 18, 176–196.
- Desmaison, A., Frongia, C., Grenier, K., Ducommun, B., and Lobjois, V. (2013). Mechanical stress impairs mitosis progression in multi-cellular tumor spheroids. *PLoS One* 8, e80447.
- Dini, S., Binder, B.J., Fischer, S.C., Mattheyer, C., Schmitz, A., Stelzer, E.H.K., Bean, N.G., and Green, J.E.F. (2016). Identifying the necrotic zone boundary in tumour spheroids with pair-correlation functions. *J. R. Soc. Interface* 13, 123.
- Ditengou, F.A., Teale, W.D., Kochersperger, P., Flittner, K.A., Kneuper, I., van der Graaff, E., Nziengui, H., Pinosa, F., Li, X., Nitschke, R., et al. (2008). Mechanical induction of lateral root initiation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 105, 18818–18823.
- Dotd, H.-U., Leischner, U., Schierloh, A., Jährling, N., Mauch, C.P., Deininger, K., Deussing, J.M., Eder, M., Zieglgänsberger, W., and Becker, K. (2007). Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nat. Methods* 4, 331–336.

- Dolan, L., Janmaat, K., Willemsen, V., Linstead, P., Poethig, S., Roberts, K., and Scheres, B. (1993). Cellular organisation of the *Arabidopsis thaliana* root. *Development* *119*, 71–84.
- Drasdo, D., and Höhme, S. (2005). A single-cell-based model of tumor growth in vitro: monolayers and spheroids. *Phys. Biol.* *2*, 133–147.
- Dubrovsky, J.G., Gambetta, G.A., Hernández-Barrera, A., Shishkova, S., and González, I. (2006). Lateral root initiation in *Arabidopsis*: developmental window, spatial patterning, density and predictability. *Ann. Bot.* *97*, 903–915.
- Edelsbrunner, H., and Harer, J. (2010). *Computational topology: an introduction* (American Mathematical Society).
- Edelsbrunner, H., and Mücke, E.P. (1994). Three-Dimensional Alpha Shapes. *ACM Trans. Graph.* *13*, 43–72.
- Edelsbrunner, H., Kirkpatrick, D., and Seidel, R. (1983). On the shape of a set of points in the plane. *IEEE Trans. Inf. Theory* *29*, 551–559.
- Efeyan, A., and Sabatini, D.M. (2010). MTOR and cancer: Many loops in one pathway. *Curr. Opin. Cell Biol.* *22*, 169–176.
- Eidet, J.R., Pasovic, L., Maria, R., Jackson, C.J., and Utheim, T.P. (2014). Objective assessment of changes in nuclear morphology and cell distribution following induction of apoptosis. *Diagn. Pathol.* *9*, 92.
- Eliceiri, K.W., Berthold, M.R., Goldberg, I.G., Ibáñez, L., Manjunath, B.S., Martone, M.E., Murphy, R.F., Peng, H., Plant, A.L., Roysam, B., et al. (2012). Biological imaging software tools. *Nat. Methods* *9*, 697–710.
- Fernandez, R., Das, P., Mirabet, V., Moscardi, E., Traas, J., Verdeil, J.-L., Malandain, G., and Godin, C. (2010). Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution. *Nat. Methods* *7*, 547–553.
- Figge, M.T., and Murphy, R.F. (2015). Image-based systems biology. *Cytom. Part A* *87*, 459–461.
- Forde, B.G. (2009). Is it good noise? The role of developmental instability in the shaping of a root system. *J. Exp. Bot.* *60*, 3989–4002.
- Friebel, A., Neitsch, J., Johann, T., Hammad, S., Hengstler, J.G., Drasdo, D., and Hoehme, S. (2015). TiQuant: Software for tissue analysis, quantification and surface reconstruction. *Bioinformatics* *31*, 3234–3236.
- Frieboes, H.B., Edgerton, M.E., Fruehauf, J.P., Rose, F.R. a J., Worrall, L.K., Gatenby, R. a., Ferrari, M., and Cristini, V. (2009). Prediction of drug response in breast cancer using integrative experimental/computational modeling. *Cancer Res.* *69*, 4484–4492.
- Friedrich, J., Seidel, C., Ebner, R., and Kunz-Schughart, L.A. (2009). Spheroid-based drug screen: considerations and practical approach. *Nat. Protoc.* *4*, 309–324.
- Friml, J. (2003). Auxin transport — shaping the plant. *Curr. Opin. Plant Biol.* *6*, 7–12.
- González, R., and Woods, R. (2010). *Digital image processing* (Prentice Hall).
- Gorman, B.R., Lu, J., Baccei, A., Lowry, N.C., Purvis, J.E., Mangoubi, R.S., and Lerou, P.H. (2014). Multi-scale imaging and informatics pipeline for in situ pluripotent stem cell analysis. *PLoS One* *9*, e116037.
- Greene, J.M., Levy, D., Herrada, S.P., Gottesman, M.M., and Lavi, O. (2016). Mathematical modeling reveals that changes to local cell density dynamically modulate baseline variations in cell growth and drug response. *Cancer Res.* *76*, 2882–2890.
- Greger, K., Swoger, J., and Stelzer, E.H.K. (2007). Basic building units and properties of a fluorescence single plane illumination microscope. *Rev. Sci. Instrum.* *78*, 23705.
- Guilak, F., Tedrow, J.R., and Burgkart, R. (2000). Viscoelastic properties of the cell nucleus. *Biochem. Biophys. Res. Commun.* *269*, 781–786.
- Guo, J.Y., Xia, B., and White, E. (2013). Autophagy-Mediated tumor Promotion. *Cell* *155*, 1216–1219.
- Hazel, A.L., and Pedley, T.J. (2000). Vascular endothelial cells minimize the total force on their nuclei. *Biophys. J.* *78*, 47–54.

- Hecht, V.C., Sullivan, L.B., Kimmerling, R.J., Kim, D.H., Hosios, A.M., Stockslager, M.A., Stevens, M.M., Kang, J.H., Wirtz, D., Vander Heiden, M.G., et al. (2016). Biophysical changes reduce energetic demand in growth factor-deprived lymphocytes. *J. Cell Biol.* *212*, 439–447.
- Helmy, I.M., and Azim, A.M.A. (2012). Efficacy of ImageJ in the assessment of apoptosis. *Diagn. Pathol.* *7*, 15.
- Hill, E. (2008). Announcing the JCB dataviewer, a browser-based application for viewing original image files. *J. Cell Biol.* *183*, 969–970.
- Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., and Stelzer, E.H.K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science* *305*, 1007–1009.
- Ichikawa, T., Nakazato, K., Keller, P.J., Kajiura-Kobayashi, H., Stelzer, E.H.K., Mochizuki, A., and Nonaka, S. (2013). Live imaging of whole mouse embryos during gastrulation: migration analyses of epiblast and mesodermal cells. *PLoS One* *8*, e64506.
- Imamura, Y., Mukohara, T., Shimono, Y., Funakoshi, Y., Chayahara, N., Toyoda, M., Kiyota, N., Takao, S., Kono, S., Nakatsura, T., et al. (2015). Comparison of 2D- and 3D-culture models as drug-testing platforms in breast cancer. *Oncol. Rep.* *33*, 1837–1843.
- Ince, D.C., Hatton, L., and Graham-Cumming, J. (2012). The case for open computer programs. *Nature* *482*, 485–488.
- Jagiella, N., Müller, B., Müller, M., Vignon-Clementel, I.E., and Drasdo, D. (2016). Inferring Growth Control Mechanisms in Growing Multi-cellular Spheroids of NSCLC Cells from Spatial-Temporal Image Data. *PLoS Comput. Biol.* *12*, e1004412.
- Kalinka, A.T., and Tomancak, P. (2012). The evolution of early animal embryos: Conservation or divergence? *Trends Ecol. Evol.* *27*, 385–393.
- Keller, P.J., and Dodt, H.U. (2012). Light sheet microscopy of living or cleared specimens. *Curr. Opin. Neurobiol.* *22*, 138–143.
- Keller, P.J., and Stelzer, E.H. (2008). Quantitative in vivo imaging of entire embryos with Digital Scanned Laser Light Sheet Fluorescence Microscopy. *Curr. Opin. Neurobiol.* *18*, 624–632.
- Keller, P.J., Schmidt, A.D., Wittbrodt, J., and Stelzer, E.H.K. (2008). Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* *322*, 1065–1069.
- Keller, P.J., Schmidt, A.D., Santella, A., Khairy, K., Bao, Z., Wittbrodt, J., and Stelzer, E.H.K. (2010). Fast, high-contrast imaging of animal development with scanned light sheet-based structured-illumination microscopy. *Nat. Methods* *7*, 637–642.
- Keller, P.J., Schmidt, A.D., Wittbrodt, J., and Stelzer, E.H.K. (2011). Digital scanned laser light-sheet fluorescence microscopy (DSLIM) of zebrafish and drosophila embryonic development. *Cold Spring Harb. Protoc.* *6*, 1235–1243.
- Kempf, H., Hatzikirou, H., Bleicher, M., and Meyer-Hermann, M. (2013). In Silico Analysis of Cell Cycle Synchronisation Effects in Radiotherapy of Tumour Spheroids. *PLoS Comput. Biol.* *9*, e1003295.
- Kerekatte, V., Smiley, K., Hu, B., Smith, A., Gelder, F., and De Benedetti, A. (1995). The proto-oncogene/translation factor eIF4E: A survey of its expression in breast carcinomas. *Int. J. Cancer* *64*, 27–31.
- Khairy, K., and Keller, P.J. (2011). Reconstructing embryonic development. *Genesis* *49*, 488–513.
- Khairy, K., Reynaud, E., and Stelzer, E. (2008a). Detection of deformable objects in 3D images using Markov-Chain Monte Carlo and spherical harmonics. *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* *5242 LNCS*, 1075–1082.
- Khairy, K., Foo, J., and Howard, J. (2008b). Shapes of Red Blood Cells: Comparison of 3D Confocal Images with the Bilayer-Couple Model. *Cell. Mol. Bioeng.* *1*, 173–181.
- Kherlopian, A.R., Song, T., Duan, Q., Neimark, M.A., Po, M.J., Gohagan, J.K., and Laine, A.F. (2008). A review of imaging techniques for systems biology. *BMC Syst. Biol.* *2*, 74.
- Kim, S.-H., Kuh, H.-J., and Dass, C.R. (2011). The reciprocal interaction: chemotherapy and tumor microenvironment. *Curr. Drug Discov. Technol.* *8*, 102–106.
- Kittler, J., and Illingworth, J. (1986). Minimum error thresholding. *Pattern Recognit.* *19*, 41–47.

- Krzic, U., Gunther, S., Saunders, T.E., Streichan, S.J., and Hufnagel, L. (2012). Multiview light-sheet microscope for rapid in toto imaging. *Nat. Methods* 9, 730–733.
- Kunz-Schughart, L.A., Freyer, J.P., Hofstaedter, F., and Ebner, R. (2004). The use of 3-D cultures for high-throughput screening: the multicellular spheroid model. *J. Biomol. Screen.* 9, 273–285.
- Kurup, S., Runions, J., Köhler, U., Laplaze, L., Hodge, S., and Haseloff, J. (2005). Marking cell lineages in living tissues. *Plant J.* 42, 444–453.
- Kvilekval, K., Fedorov, D., Obara, B., Singh, A., and Manjunath, B.S. (2010). Bisque: a platform for bioimage analysis and management. *Bioinformatics* 26, 544–552.
- Kyle, A.H., Baker, J.H.E., and Minchinton, A.I. (2012). Targeting quiescent tumor cells via oxygen and IGF-I supplementation. *Cancer Res.* 72, 801–809.
- Laplante, M., and Sabatini, D.M. (2012). MTOR signaling in growth control and disease. *Cell* 149, 274–293.
- Levsky, J.M., and Singer, R.H. (2003). Gene expression and the myth of the average cell. *Trends Cell Biol.* 13, 4–6.
- Li, G., Liu, T., Tarokh, A., Nie, J., Guo, L., Mara, A., Holley, S., and Wong, S.T.C. (2007). 3D cell nuclei segmentation based on gradient flow tracking. *BMC Cell Biol.* 8, 40.
- Lichtman, J.W., and Conchello, J.-A. (2005). Fluorescence microscopy. *Nat. Methods* 2, 910–919.
- Liedtke, C., Wang, J., Tordai, A., Symmans, W.F., Hortobagyi, G.N., Kiesel, L., Hess, K., Baggerly, K.A., Coombes, K.R., and Pusztai, L. (2010). Clinical evaluation of chemotherapy response predictors developed from breast cancer cell lines. *Breast Cancer Res. Treat.* 121, 301–309.
- Lin, R.Z., and Chang, H.Y. (2008). Recent advances in three-dimensional multicellular spheroid culture for biomedical research. *Biotechnol. J.* 3, 1172–1184.
- Lin, G., Chawla, M.K., Olson, K., Barnes, C.A., Guzowski, J.F., Bjornsson, C., Shain, W., and Roysam, B. (2007). A multi-model approach to simultaneous segmentation and classification of heterogeneous populations of cell nuclei in 3D confocal microscope images. *Cytom. Part A* 71, 724–736.
- Lindeberg, T. (1998). Feature detection with automatic scale selection. *Int. J. Comput. Vis.* 30, 79–116.
- Long, F., Peng, H., Liu, X., Kim, S.S.K., and Myers, E. (2009). A 3D digital atlas of *C. elegans* and its application to single-cell analyses. *Nat. Methods* 6, 667–672.
- Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* 60, 91–110.
- Lucas, M., Kenobi, K., von Wangenheim, D., Voß, U., Swarup, K., De Smet, I., Van Damme, D., Lawrence, T., Péret, B., Moscardi, E., et al. (2013). Lateral root morphogenesis is dependent on the mechanical properties of the overlaying tissues. *Proc. Natl. Acad. Sci. U. S. A.* 110, 5229–5234.
- Lum, J.J., Bauer, D.E., Kong, M., Harris, M.H., Li, C., Lindsten, T., and Thompson, C.B. (2005). Growth factor regulation of autophagy and cell survival in the absence of apoptosis. *Cell* 120, 237–248.
- Malamy, J.E., and Benfey, P.N. (1997). Organization and cell differentiation in lateral roots of *Arabidopsis thaliana*. *Development* 124, 33–44.
- Mateus, R., Lourenco, R., Fang, Y., Brito, G., Farinho, A., Valerio, F., and Jacinto, A. (2015). Control of tissue growth by Yap relies on cell density and F-actin in zebrafish fin regeneration. *Development* 142, 2752–2763.
- Mathew, B., Schmitz, A., Muñoz-Descalzo, S., Ansari, N., Pampaloni, F., Stelzer, E.H.K., and Fischer, S.C. (2015). Robust and automated three-dimensional segmentation of densely packed cell nuclei in different biological specimens with Lines-of-Sight decomposition. *BMC Bioinformatics* 16, 187.
- McIlwain, D.R., Berger, T., and Mak, T.W. (2013). Caspase functions in cell death and disease. *Cold Spring Harb. Perspect. Biol.* 5, a008656.
- Meijering, E. (2012). Cell segmentation: 50 Years down the road. *IEEE Signal Process. Mag.* 29, 140–145.
- Meijering, E., Dzyubachyk, O., Smal, I., and van Cappellen, W. a. (2009). Tracking in cell and developmental biology. *Semin. Cell Dev. Biol.* 20, 894–902.
- Meijering, E., Carpenter, A.E., Peng, H., Hamprecht, F.A., and Olivo-Marin, J.-C. (2016). Imagining the

- future of bioimage analysis. *Nat. Biotechnol.* *34*, 1250–1255.
- Mikula, K., Peyri ras, N., Remeřikov, M., and Stařov, O. (2011). Segmentation of 3D cell membrane images by PDE methods and its applications. *Comput. Biol. Med.* *41*, 326–339.
- Mo, J.-S., Park, H.W., and Guan, K.-L. (2014). The Hippo signaling pathway in stem cell biology and cancer. *EMBO Rep.* *15*, 642–656.
- Montel, F., Delarue, M., Elgeti, J., Malaquin, L., Basan, M., Risler, T., Cabane, B., Vignjevic, D., Prost, J., Cappello, G., et al. (2011). Stress Clamp Experiments on Multicellular Tumor Spheroids. *Phys. Rev. Lett.* *107*, 188102.
- Montel, F., Delarue, M., Elgeti, J., Vignjevic, D., Cappello, G., and Prost, J. (2012). Isotropic stress reduces cell proliferation in tumor spheroids. *New J. Phys.* *14*, 55008.
- Mooney, L.M., Al-Sakkaf, K.A., Brown, B.L., and Dobson, P.R.M. (2002). Apoptotic mechanisms in T47D and MCF-7 human breast cancer cells. *Br. J. Cancer* *87*, 909–917.
- Morales-Navarrete, H., Segovia-Miranda, F., Klukowski, P., Meyer, K., Nonaka, H., Marsico, G., Chernykh, M., Kalaidzidis, A., Zerial, M., and Kalaidzidis, Y. (2015). A versatile pipeline for the multi-scale digital reconstruction and quantitative analysis of 3D tissue architecture. *Elife* *4*, e11214.
- Moreno-Risueno, M.A., Van Norman, J.M., Moreno, A., Zhang, J., Ahnert, S.E., and Benfey, P.N. (2010). Oscillating gene expression determines competence for periodic Arabidopsis root branching. *Science* *329*, 1306–1311.
- Mueller-Klieser, W. (1984). Method for the determination of oxygen consumption rates and diffusion coefficients in multicellular spheroids. *Biophys. J.* *46*, 343–348.
- Myers, G. (2012). Why bioimage informatics matters. *Nat. Methods* *9*, 659–660.
- Nakamura, T., Yoshizaki, M., Ogawa, S., Okamoto, H., Shinmyo, Y., Bando, T., Ohuchi, H., Noji, S., and Mito, T. (2010). Imaging of transgenic cricket embryos reveals cell movements consistent with a syncytial patterning mechanism. *Curr. Biol.* *20*, 1641–1647.
- Narayanaswamy, A., Dwarakapuram, S., Bjornsson, C.S., Cutler, B.M., Shain, W., and Roysam, B. (2010). Robust adaptive 3-D segmentation of vessel laminae from fluorescence confocal microscope images and parallel GPU implementation. *IEEE Trans. Med. Imaging* *29*, 583–597.
- Van Norman, J.M., Xuan, W., Beeckman, T., and Benfey, P.N. (2013). To branch or not to branch: the role of pre-patterning in lateral root formation. *Development* *140*, 4301–4310.
- Ntziachristos, V. (2010). Going deeper than microscopy: the optical imaging frontier in biology. *Nat. Methods* *7*, 603–614.
- Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man. Cybern.* *9*, 62–66.
- Pampaloni, F., Reynaud, E.G., and Stelzer, E.H.K. (2007). The third dimension bridges the gap between cell culture and live tissue. *Nat. Rev. Mol. Cell Biol.* *8*, 839–845.
- Pampaloni, F., Chang, B.J., and Stelzer, E.H.K. (2015). Light sheet-based fluorescence microscopy (LSFM) for the quantitative imaging of cells and tissues. *Cell Tissue Res.* *360*, 129–141.
- Pantazis, P., and Supatto, W. (2014). Advances in whole-embryo imaging: a quantitative transition is underway. *Nat. Rev. Mol. Cell Biol.* *15*, 327–339.
- Peter, P., and Weickert, J. (2014). Colour image compression with anisotropic diffusion. In 2014 IEEE International Conference on Image Processing (ICIP), (IEEE), pp. 4822–4826.
- Pincus, Z., and Theriot, J.A. (2007). Comparison of quantitative methods for cell-shape analysis. *J. Microsc.* *227*, 140–156.
- Polak, M. (2003). *Developmental Instability: Causes and Consequences* (Oxford University Press).
- Powathil, G.G., Swat, M., and Chaplain, M.A.J. (2014). Systems oncology: Towards patient-specific treatment regimes informed by multiscale mathematical modelling. *Semin. Cancer Biol.* *30*, 13–20.
- Preibisch, S., Ejsmont, R., Rohlfing, T., and Tomancak, P. (2008). Towards digital representation of Drosophila embryogenesis. In 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, (IEEE), pp. 324–327.
- Preibisch, S., Saalfeld, S., and Tomancak, P. (2009). Globally optimal stitching of tiled 3D microscopic

- image acquisitions. *Bioinformatics* 25, 1463–1465.
- Preibisch, S., Saalfeld, S., Schindelin, J., and Tomancak, P. (2010). Software for bead-based registration of selective plane illumination microscopy data. *Nat. Methods* 7, 418–419.
- Preibisch, S., Amat, F., Stamatakis, E., Sarov, M., Singer, R.H., Myers, E., and Tomancak, P. (2014). Efficient Bayesian-based multiview deconvolution. *Nat. Methods* 11, 645–648.
- Qi, J. (2014). Dense nuclei segmentation based on graph cut and convexity-concavity analysis. *J. Microsc.* 253, 42–53.
- Quaglino, A., Salierno, M., Pellegrotti, J., Rubinstein, N., and Kordon, E.C. (2009). Mechanical strain induces involution-associated events in mammary epithelial cells. *BMC Cell Biol.* 10, 55.
- Rabinowitz, J.D., and White, E. (2010). Autophagy and metabolism. *Science* 330, 1344–1348.
- R Development Core Team (2016). R: A language and environment for statistical computing.
- Ravi, M., Paramesh, V., Kaviya, S.R., Anuradha, E., and Solomon, F.D.P. (2015). 3D cell culture systems: advantages and applications. *J. Cell. Physiol.* 230, 16–26.
- Ravikumar, B., Vacher, C., Berger, Z., Davies, J.E., Luo, S., Oroz, L.G., Scaravilli, F., Easton, D.F., Duden, R., O’Kane, C.J., et al. (2004). Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. *Nat. Genet.* 36, 585–595.
- Reid, B.G., Jerjian, T., Patel, P., Zhou, Q., Yoo, B.H., Kabos, P., Sartorius, C. a., and LaBarbera, D. V. (2014). Live Multicellular Tumor Spheroid Models For High-Content Imaging and Screening In Cancer Drug Discovery. *Curr. Chem. Genomics Transl. Med.* 8, 27–35.
- Reynaud, E.G., Krzic, U., Greger, K., and Stelzer, E.H.K. (2008). Light sheet-based fluorescence microscopy: more dimensions, more photons, and less photodamage. *HFSP J.* 2, 266–275.
- Rozbicki, E., Chuai, M., Karjalainen, A.I., Song, F., Sang, H.M., Martin, R., Knölker, H.-J., MacDonald, M.P., and Weijer, C.J. (2015). Myosin-II-mediated cell shape changes and cell intercalation contribute to primitive streak formation. *Nat. Cell Biol.* 17, 397–408.
- Rubio-Guivernau, J.L., Gurchenkov, V., Luengo-Oroz, M.A., Duloquin, L., Bourguin, P., Santos, A., Peyrieras, N., and Ledesma-Carbayo, M.J. (2012). Wavelet-based image fusion in multi-view three-dimensional microscopy. *Bioinformatics* 28, 238–245.
- De Rybel, B., Vassileva, V., Parizot, B., Demeulenaere, M., Grunewald, W., Audenaert, D., Van Campenhout, J., Overvoorde, P., Jansen, L., Vanneste, S., et al. (2010). A novel Aux/IAA28 signaling cascade activates GATA23-dependent specification of lateral root founder cell identity. *Curr. Biol.* 20, 1697–1706.
- Saito, A., Numata, Y., Hamada, T., Horisawa, T., Cosatto, E., Graf, H.-P., Kuroda, M., and Yamamoto, Y. (2016). A novel method for morphological pleomorphism and heterogeneity quantitative measurement: Named cell feature level co-occurrence matrix. *J. Pathol. Inform.* 7, 36.
- Sarrazin, A.F., Peel, A.D., and Averof, M. (2012). A segmentation clock with two-segment periodicity in insects. *Science* 336, 338–341.
- Schäfer, H., Schäfer, T., Ackermann, J., Dichter, N., Döring, C., Hartmann, S., Hansmann, M.L., and Koch, I. (2015). CD30 cell graphs of Hodgkin lymphoma are not scale-free - An image analysis approach. *Bioinformatics* 32, 122–129.
- Schaller, G., and Meyer-Hermann, M. (2005). Multicellular tumor spheroid in an off-lattice Voronoi-Delaunay cell model. *Phys. Rev. E* 71, 51910.
- Scheres, B., Wolkenfelt, H., Willemsen, V., Terlouw, M., Lawson, E., Dean, C., and Weisbeek, P. (1994). Embryonic origin of the Arabidopsis primary root and root meristem initials. *Development* 2487, 2475–2487.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682.
- Schleicher, S.M., Moretti, L., Varki, V., and Lu, B. (2010). Progress in the unraveling of the endoplasmic reticulum stress/autophagy pathway and cancer: Implications for future therapeutic approaches. *Drug Resist. Updat.* 13, 79–86.
- Schmaltz, C., Weickert, J., and Bruhn, A. (2009). Beating the quality of jpeg 2000 with anisotropic

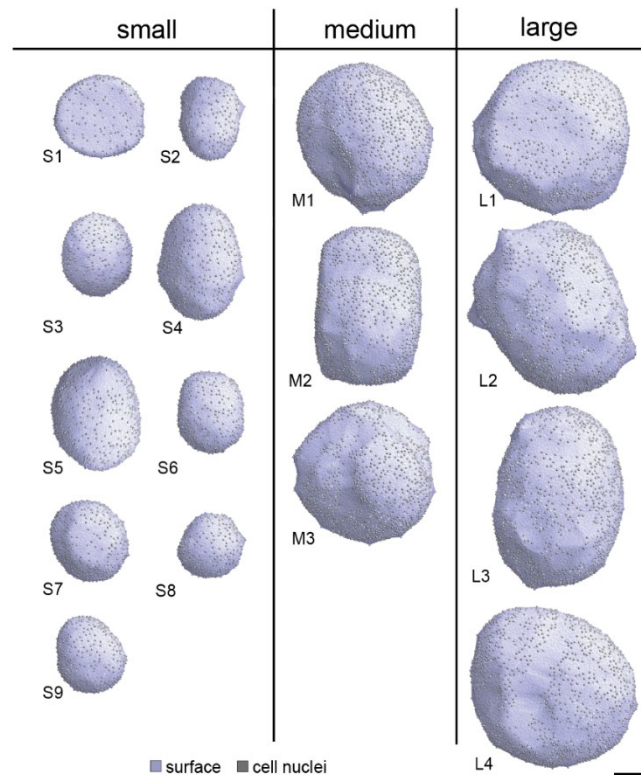
- diffusion. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, (Springer Berlin Heidelberg), pp. 452–461.
- Schmied, C., Steinbach, P., Pietzsch, T., Preibisch, S., and Tomancak, P. (2016). An automated workflow for parallel processing of large multiview SPIM recordings. *Bioinformatics* *32*, 1112–1114.
- Schmitz, A., Fischer, S.C., Mattheyer, C., Pampaloni, F., and Stelzer, E.H.K. (2017). Multiscale image analysis reveals structural heterogeneity of the cell microenvironment in homotypic spheroids. *Sci. Rep.* (accepted).
- Schroeder, W.J., Martin, K.M., and Lorensen, W.E. (1996). The design and implementation of an object-oriented toolkit for 3D graphics and visualization. *Proc. Seventh Annu. IEEE Vis. '96* *1*, 93–100.
- Seto, B. (2012). Rapamycin and mTOR: a serendipitous discovery and implications for breast cancer. *Clin. Transl. Med.* *1*, 29.
- Shamir, L., Delaney, J.D., Orlov, N., Eckley, D.M., and Goldberg, I.G. (2010). Pattern recognition software and techniques for biological image analysis. *PLoS Comput. Biol.* *6*, e1000974.
- Shapira, M., Kakiashvili, E., Rosenberg, T., and Hershko, D.D. (2006). The mTOR inhibitor rapamycin down-regulates the expression of the ubiquitin ligase subunit Skp2 in breast cancer cells. *Breast Cancer Res.* *8*, R46.
- Shaw, P.J. (2006). Comparison of Widefield Deconvolution and Confocal Microscopy for Three-Dimensional Imaging. *Handb. Biol. Confocal Microsc.* 453–467.
- Shekhar, M.P. V (2011). Drug resistance: challenges to effective therapy. *Curr. Cancer Drug Targets* *11*, 613–623.
- De Smet, I. (2012). Lateral root initiation: one step at a time. *New Phytol.* *193*, 867–873.
- De Smet, I., Tetsumura, T., De Rybel, B., Frey, N.F. dit, Laplaze, L., Casimiro, I., Swarup, R., Naudts, M., Vanneste, S., Audenaert, D., et al. (2007). Auxin-dependent regulation of lateral root positioning in the basal meristem of *Arabidopsis*. *Development* *134*, 681–690.
- De Smet, I., Lau, S., Mayer, U., and Jürgens, G. (2010). Embryogenesis - the humble beginnings of plant life. *Plant J.* *61*, 959–970.
- Smolle, J., Hofmann-Wellenhof, R., Soyer, H.P., Stettner, H., and Kerl, H. (1989). Nuclear size and shape parameters correlate with proliferative activity in cutaneous melanocytic tumors. *J. Invest. Dermatol.* *93*, 178–182.
- Stegmaier, J., Otte, J.C., Kobitski, A., Bartschat, A., Garcia, A., Nienhaus, G.U., Strähle, U., and Mikut, R. (2014). Fast segmentation of stained nuclei in terabyte-scale, time resolved 3D microscopy image stacks. *PLoS One* *9*, e90036.
- Stelzer, E.H.K. (2015). Light-sheet fluorescence microscopy for quantitative biology. *Nat. Methods* *12*, 23–26.
- Stelzer, E.H.K., and Lindek, S. (1994). Fundamental reduction of the observation volume in far-field light microscopy by detection orthogonal to the illumination axis: confocal theta microscopy. *Opt. Commun.* *111*, 536–547.
- Strobl, F., and Stelzer, E.H. (2016). Long-term fluorescence live imaging of *Tribolium castaneum* embryos: principles, resources, scientific challenges and the comparative approach. *Curr. Opin. Insect Sci.* *18*, 17–26.
- Strobl, F., and Stelzer, E.H.K. (2014). Non-invasive long-term fluorescence live imaging of *Tribolium castaneum* embryos. *Development* *141*, 2331–2338.
- Strobl, F., Schmitz, A., and Stelzer, E.H.K. (2015). Live imaging of *Tribolium castaneum* embryonic development using light-sheet-based fluorescence microscopy. *Nat. Protoc.* *10*, 1486–1507.
- Strobl, F., Schmitz, A., and Stelzer, E.H.K. (2017). Improve your four-dimensional image – traveling ten innovative and insightful years in the multiverse of light sheet-based fluorescence microscopy. *Nat. Protoc.* (submitted).
- Sutherland, R.M. (1988). Cell and Environment Interactions in Tumor Microregions: The Multicell Spheroid Model. *Science* *240*, 177–184.
- Swoger, J., Verveer, P., Greger, K., Huiskens, J., and Stelzer, E.H.K. (2007). Multi-view image fusion improves resolution in three-dimensional microscopy. *Opt. Express* *15*, 8029–8042.

- Szymanowska-Pulka, J. (2013). Form matters: morphological aspects of lateral root development. *Ann. Bot.* *112*, 1643–1654.
- Szymanowska-Pułka, J., Potocka, I., Karczewski, J., Jiang, K., Nakielski, J., and Feldman, L.J. (2012). Principal growth directions in development of the lateral root in *Arabidopsis thaliana*. *Ann. Bot.* *110*, 491–501.
- Szymanowska-Pułka, J., Lipowczan, M., Szymanowska-Pulka, J., and Lipowczan, M. (2014). Growth rate distribution in the forming lateral root of *Arabidopsis*. *Ann. Bot.* *114*, 913–921.
- Tanramluk, D., Schreyer, A., Pitt, W.R., and Blundell, T.L. (2009). On the origins of enzyme inhibitor selectivity and promiscuity: A case study of protein kinase binding to staurosporine. *Chem. Biol. Drug Des.* *74*, 16–24.
- Thoma, C.R., Zimmermann, M., Agarkova, I., Kelm, J.M., and Krek, W. (2014). 3D cell culture systems modeling tumor growth determinants in cancer target discovery. *Adv. Drug Deliv. Rev.* *69–70*, 29–41.
- Tomer, R., Khairy, K., Amat, F., and Keller, P.J. (2012). Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nat. Methods* *9*, 755–763.
- Toné, S., Sugimoto, K., Tanda, K., Suda, T., Uehira, K., Kanouchi, H., Samejima, K., Minatogawa, Y., and Earnshaw, W.C. (2007). Three distinct stages of apoptotic nuclear condensation revealed by time-lapse imaging, biochemical and electron microscopy analysis of cell-free apoptosis. *Exp. Cell Res.* *313*, 3635–3644.
- Tse, J.M., Cheng, G., Tyrrell, J.A., Wilcox-Adelman, S.A., Boucher, Y., Jain, R.K., and Munn, L.L. (2012). Mechanical compression drives cancer cells toward invasive phenotype. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 911–916.
- Uchida, S. (2013). Image processing and recognition for biological images. *Dev. Growth Differ.* *55*, 523–549.
- Vermeer, J.E.M., von Wangenheim, D., Barberon, M., Lee, Y., Stelzer, E.H.K., Maizel, A., and Geldner, N. (2014). A spatial accommodation by neighboring cells is required for organ initiation in *Arabidopsis*. *Science* *343*, 178–183.
- Verveer, P.J., Swoger, J., Pampaloni, F., Greger, K., Marcello, M., and Stelzer, E.H.K. (2007). High-resolution three-dimensional imaging of large specimens with light sheet-based microscopy. *Nat. Methods* *4*, 311–313.
- Vieten, A., Sauer, M., Brewer, P.B., and Friml, J. (2007). Molecular and cellular aspects of auxin-transport-mediated development. *Trends Plant Sci.* *12*, 160–168.
- Vincent, L., and Soille, P. (1991). Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* *13*, 583–598.
- Vinci, M., Gowan, S., Boxall, F., Patterson, L., Zimmermann, M., Court, W., Lomas, C., Mendiola, M., Hardisson, D., and Eccles, S. a (2012). Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation. *BMC Biol.* *10*, 29.
- von Wangenheim, D., Fangerau, J., Schmitz, A., Smith, R.S., Leitte, H., Stelzer, E.H.K., and Maizel, A. (2016). Rules and Self-Organizing Properties of Post-embryonic Plant Organ Cell Division Patterns. *Curr. Biol.* *26*, 439–449.
- Voß, U., Wilson, M.H., Kenobi, K., Gould, P.D., Robertson, F.C., Peer, W.A., Lucas, M., Swarup, K., Casimiro, I., Holman, T.J., et al. (2015). The circadian clock rephases during lateral root organ initiation in *Arabidopsis thaliana*. *Nat. Commun.* *6*, 7641.
- Wagner, M., Weber, P., Bruns, T., Strauss, W.S.L., Wittig, R., and Schneckenburger, H. (2010). Light dose is a limiting factor to maintain cell viability in fluorescence microscopy and single molecule detection. *Int. J. Mol. Sci.* *11*, 956–966.
- Wait, E., Winter, M., Björnsson, C., Kokovay, E., Wang, Y., Goderie, S., Temple, S., and Cohen, A.R. (2014). Visualization and correction of automated segmentation, tracking and lineaging from 5-D stem cell image sequences. *BMC Bioinformatics* *15*, 328.
- Wallace, W., Schaefer, L.H., and Swedlow, J.R. (2001). A workingperson's guide to deconvolution in light microscopy. *Biotechniques* *31*, 1076–8, 1080, 1082 passim.
- Walter, T., Shattuck, D.W., Baldock, R., Bastin, M.E., Carpenter, A.E., Duce, S., Ellenberg, J., Fraser, A., Hamilton, N., Pieper, S., et al. (2010). Visualization of image data from cells to organisms. *Nat.*

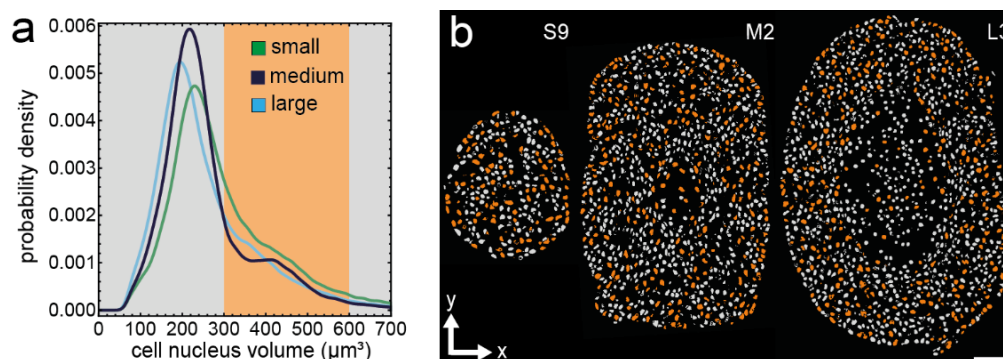
- Methods 7, S26–S41.
- Wang, Y.C., Chao, T.K., Chang, C.C., Yo, Y. Te, Yu, M.H., and Lai, H.C. (2013). Drug Screening Identifies Niclosamide as an Inhibitor of Breast Cancer Stem-Like Cells. *PLoS One* 8, e74538.
- Wenzel, C., Riefke, B., Gründemann, S., Krebs, A., Christian, S., Prinz, F., Osterland, M., Golfier, S., Räse, S., Ansari, N., et al. (2014). 3D high-content screening for the identification of compounds that target cells in dormant tumor spheroid regions. *Exp. Cell Res.* 323, 131–143.
- Wu, Y., Chandris, P., Winter, P.W., Kim, E.Y., Jaumouillé, V., Kumar, A., Guo, M., Leung, J.M., Smith, C., Rey-Suarez, I., et al. (2016). Simultaneous multiview capture and fusion improves spatial resolution in wide-field and light-sheet microscopy. *Optica* 3, 897.
- Yoo, T.S., Ackerman, M.J., Lorensen, W.E., Schroeder, W., Chalana, V., Aylward, S., Metaxas, D., and Whitaker, R. (2002). Engineering and algorithm design for an image processing API: A technical report on ITK - The Insight Toolkit. In *Studies in Health Technology and Informatics*, pp. 586–592.
- Yoshida, S., Barbier de Reuille, P., Lane, B., Bassel, G.W., Prusinkiewicz, P., Smith, R.S., Weijers, D., BarbierdeReuille, P., Lane, B., Bassel, G.W., et al. (2014). Genetic control of plant development by overriding a geometric division rule. *Dev. Cell* 29, 75–87.
- Yue, K., and Beeckman, T. (2014). Cell-to-cell communication during lateral root development. *Mol. Plant* 7, 758–760.
- Zanella, F., Lorens, J.B., and Link, W. (2010). High content screening: Seeing is believing. *Trends Biotechnol.* 28, 237–245.
- Zink, D., Fischer, A.H., and Nickerson, J. a (2004). Nuclear structure in cancer cells. *Nat. Rev. Cancer* 4, 677–687.

8 Supplementary Material

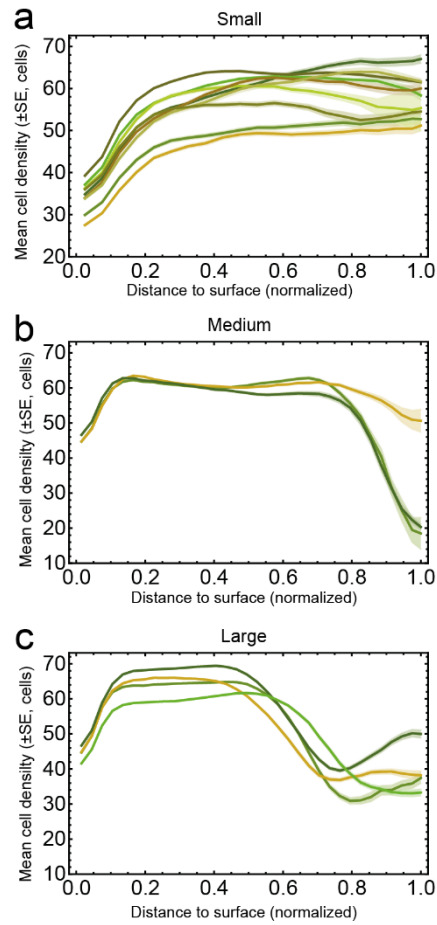
Multiscale quantitative characterization



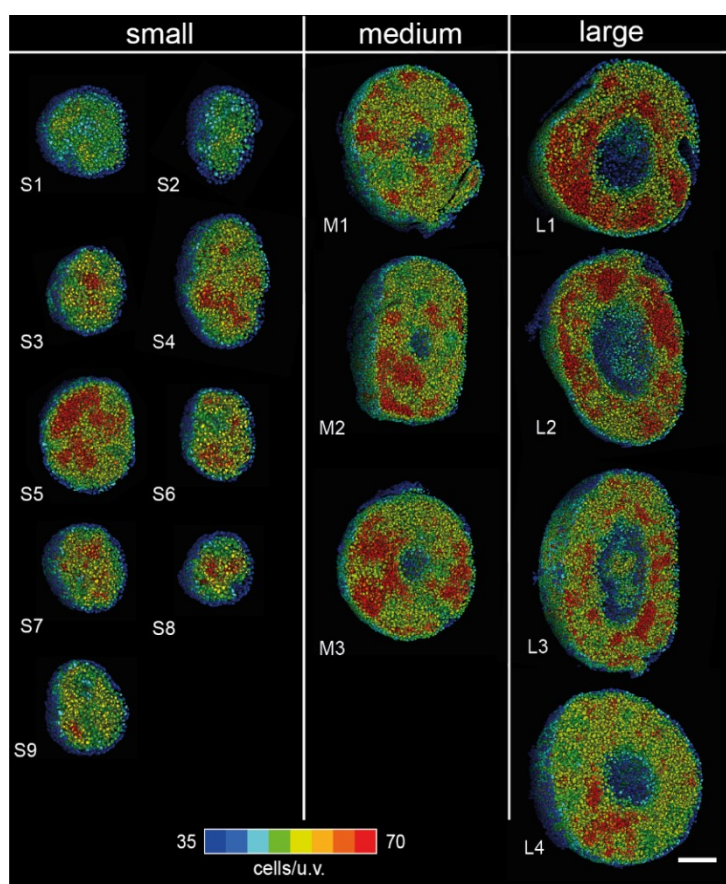
Supplementary Figure 1 Renderings of the surface approximation. Depicted are the approximated surfaces and cell nuclei locations of all T47D spheroid datasets in application I. Scale bar: approximately 50 μm . Figure adapted from Schmitz *et al.*, 2017.



Supplementary Figure 2 Spatial localization of cell nuclei with a volume between 300 and 600 μm^3 . (a) Smoothed histogram of the cell nuclei volume distribution of small, medium and large spheroids. Cell nuclei with volumes in the region highlighted in orange (300 to 600 μm^3) were colored in the segmentation images to analyze their spatial localization. (b) Single planes of datasets S9, M2 and L3 at the center of the spheroid. All cell nuclei with a volume of less than 300 μm^3 or greater than 600 μm^3 are depicted in gray. The cell nuclei with a volume between 300 μm^3 and 600 μm^3 are colored in orange. Figure adapted from Schmitz *et al.*, 2017.



Supplementary Figure 3 Plots of the mean cell density versus the normalized distance to the surface for small (a), medium (b) and large (c) spheroids. Figure adapted from Schmitz *et al.*, 2017.



Supplementary Figure 4 Local cell density in all datasets. Three-dimensional rendering of segmented cell nuclei colored according to their corresponding cell density value for all small, medium and large spheroids, ranging from blue (35 cells) to red (70 cells). Renderings were clipped at about the center of the spheroids. Scale bar: 50 μm . Figure adapted from Schmitz *et al.*, 2017.

Supplementary Table 1 Runtime of the quantitative characterization module. Runtimes are shown for cell nuclei segmentation and post-processing of T47D spheroid datasets in application I.

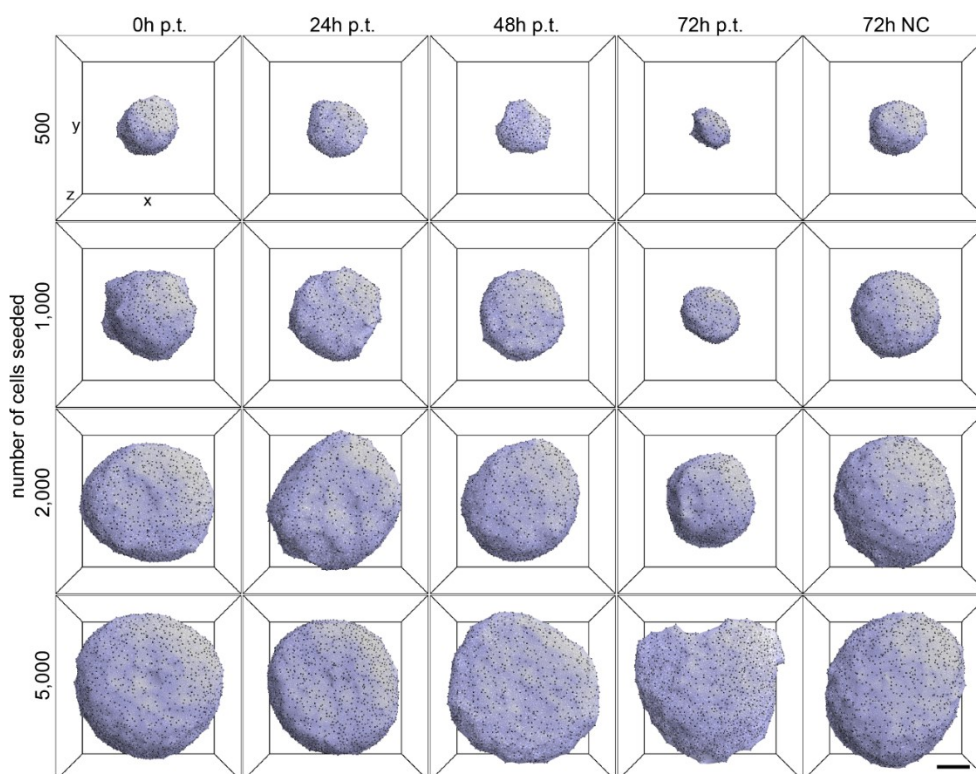
	File size (MB)	initial segmentation (s)	marker point detection (s)	marker- controlled watershed (s)	alpha shape/cell graphs (s)	total (min)
S1	258	10	258	80	25	6
S2	237	9	237	72	17	6
S3	189	7	189	58	21	5
S4	565	22	565	179	58	14
S5	512	21	512	182	72	13
S6	229	9	229	75	21	6
S7	271	10	270	89	27	7
S8	148	5	148	45	11	3
S9	168	6	168	52	18	4
M1	1376	57	1376	454	355	37
M2	1299	50	1299	436	357	36
M3	1444	55	1444	468	324	38
L1	1451	57	1451	539	582	44

L2	2009	78	2009	740	630	58
L3	1829	71	1830	667	750	55
L4	1552	61	1552	519	563	45

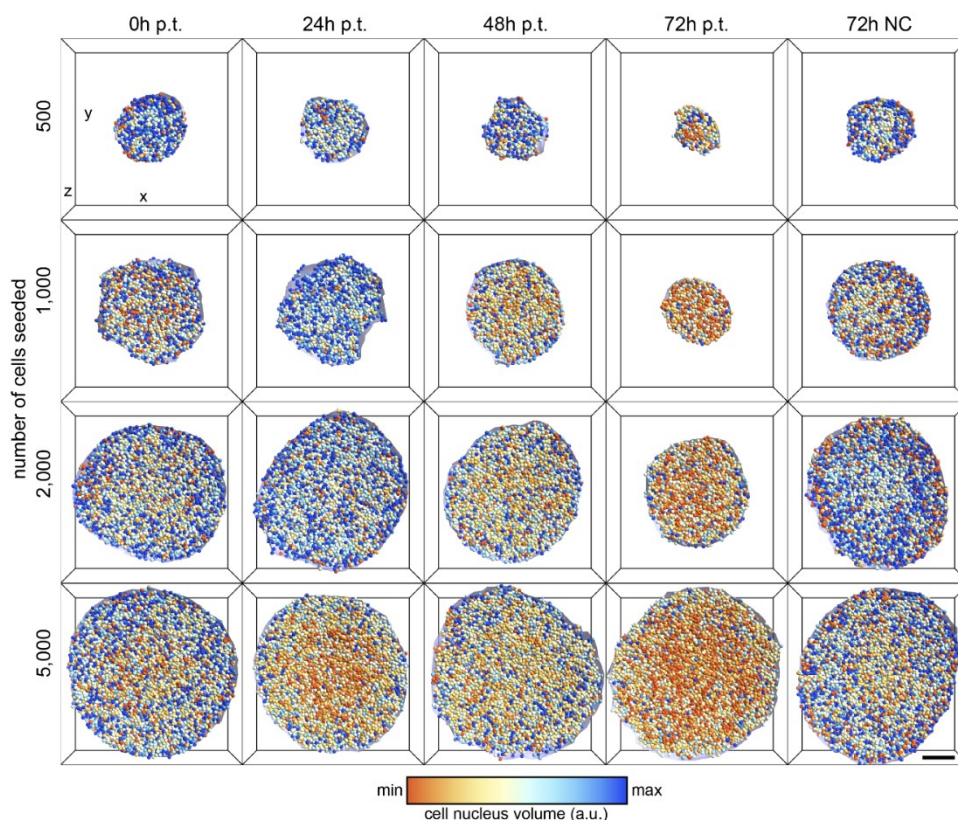
* Note that the raw image stacks were processed according to the settings given in Table 5. Computations were conducted on a workstation that comprises two quad-core CPUs (Xeon E5630 at 2.53 GHz, Intel Corporation), 96 Gigabyte DDR3 memory, running 64 bit Windows 7 Professional.

Supplementary Table 2 Quantitative measures for all datasets. Shown are the quantitative measures for T47D spheroid datasets in application I. Table adapted from Schmitz *et al.*, 2017.

Dataset	Group	Volume (μm^3)	# of cells	Median cell nucleus volume (μm^3)	Median absolute deviation of cell nucleus volume (μm^3)
S1	Small	5.81×10^6	4,155	333	77
S2	Small	4.15×10^6	2,883	341	78
S3	Small	4.68×10^6	3,981	263	57
S4	Small	1.06×10^7	8,975	245	57
S5	Small	1.18×10^7	10,311	249	55
S6	Small	4.58×10^6	3,938	268	71
S7	Small	5.96×10^6	4,866	260	66
S8	Small	2.71×10^6	2,326	261	77
S9	Small	3.98×10^6	3,359	280	66
M1	Medium	2.91×10^7	25,775	234	51
M2	Medium	2.88×10^7	25,908	235	48
M3	Medium	2.82×10^7	25,064	232	53
L1	Large	3.81×10^7	32,898	211	50
L2	Large	4.10×10^7	35,097	211	53
L3	Large	4.21×10^7	38,752	209	54
L4	Large	3.87×10^7	32,222	279	62



Supplementary Figure 5 Rendering of cell nuclei locations and approximated surface of a selection of spheroids, one per cell seed number and time point after treatment. Scale bar: 100 μm .



Supplementary Figure 6 Comparison of cell nucleus volume in T47D spheroids for all conditions. The images depict three-dimensional renderings of the segmented cell nuclei represented as spheres and color coded according to their corresponding cell nucleus volume. Please note that the cell nucleus volume was normalized to the minimum (min) and maximum (max) cell nucleus volume determined in all datasets. To discriminate against outliers, the 1/10-quantile and the 9/10-quantile were chosen for min and max. Renderings were clipped at about the center of the spheroids. Scale bar: 100 μm . h, hours, p.t., post treatment. NC, negative control.

Supplementary Table 3 Number of T47D spheroids per condition and time point. Spheroids were generated from 500, 1,000, 2,000 and 5,000 cells seeded. Spheroids grown in 0.1% DMSO were used as negative control. h, hours; NC, negative control; p.t., post treatment.

	0h p.t.	24h p.t.	48h p.t.	72h p.t.	72h NC
500	3	3	1	1	2
1,000	2	4	5	4	2
2,000	4	10	6	5	3
5,000	3	9	6	3	4

Supplementary Table 4 Volume of T47D spheroids per condition and time point given as mean±standard error × 10⁷ μm³. Spheroids were generated from 500, 1,000, 2,000 and 5,000 cells seeded. For 500 cells seeded, only one spheroid was available for time points 48h and 72h. h, hours; NC, negative control; p.t., post treatment.

	0h p.t.	24h p.t.	48h p.t.	72h p.t.	72h NC
500	0.28±0.10	0.24±0.10	0.18	0.17	0.38±0.04
1,000	1.48±0.07	1.11±0.31	0.85±0.12	0.32±0.04	1.82±0.46
2,000	3.82±0.09	3.54±0.29	2.86 ±0.11	1.43±0.05	4.45±0.13
5,000	4.55±0.59	5.04±0.51	5.08±0.29	3.25±0.58	5.73±0.26

Supplementary Table 5 Number of cells in T47D spheroids (mean±standard error) per condition and time point. For 500 cells seeded, only one spheroid was available for time points 48h and 72h. h, hours; NC, negative control; p.t., post treatment.

	0h p.t.	24h p.t.	48h p.t.	72h p.t.	72h NC
500	1,414±461	1237±404	1,138	1,149	1,991±55
1,000	8,350±232	6020±607	5,997±414	2,800±237	9,078±1,120
2,000	17,221±96	16,764±654	15,441±621	8,365±240	19,425±949
5,000	25,584±1,918	27,547±1176	26,979±1742	19,091±5251	29,873±750

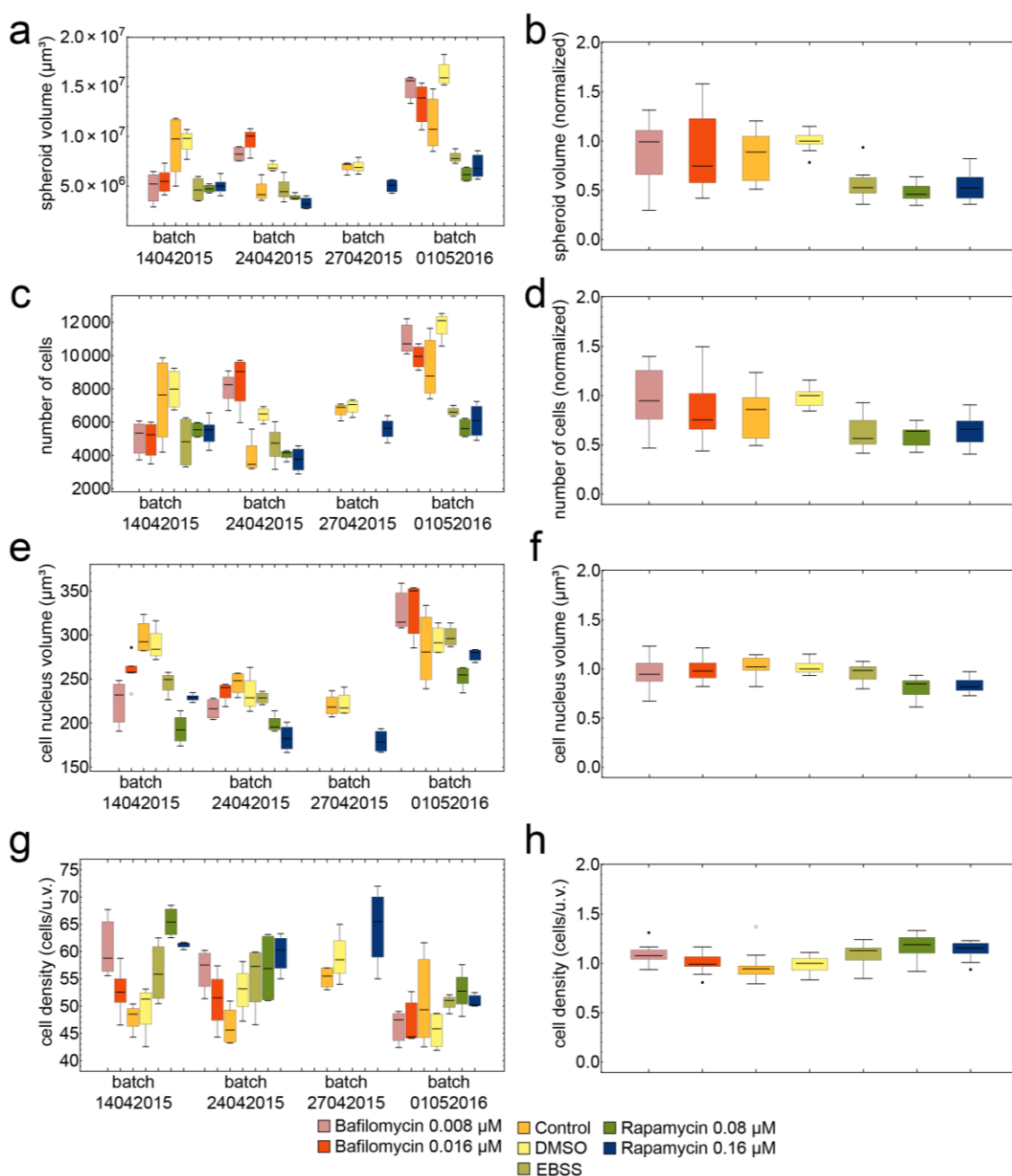
Supplementary Table 6 Average diameter of T47D spheroids (in μm) along the principal component directions p1, p2 and p3 determined by PCA analysis.

	0h p.t.			24h p.t. NC			48h p.t.			72h p.t.			72h NC		
	p1	p2	p3	p1	p2	p3	p1	p2	p3	p1	p2	p3	p1	p2	p3
500	170	153	99	200	144	83	169	161	69	136	115	101	189	175	108
1,000	282	255	190	281	242	138	259	228	125	189	164	98	304	278	199
2,000	392	363	232	407	376	204	373	334	192	296	268	153	426	378	248
5,000	425	393	229	463	428	218	474	445	203	430	395	155	460	421	252

Supplementary Table 7 Average cell nucleus volume in T47D spheroids (median±median absolute deviation) per condition and time point. h, hours; p.t., post treatment.

number of cells seeded	0h p.t.	24h p.t.	48h p.t.	72h p.t.	72h NC
500	206±79	193±59	189±67	129±48	176±69
1,000	167±62	160±62	116±45	99±41	175±69
2,000	173±58	160±51	129±42	109±39	184±72
5,000	145±51	142±49	127±41	98±32	153±53

**Supplementary Figure 7 Alpha shape surface for selected T47D spheroids.** Rap., Rapamycin, Baf., Bafilomycin. Scale bar: approximately 50 μm .



Supplementary Figure 8 Quantitative results for all four experiments. Boxplots of the spheroid volume for each experiment (a) and pooled after normalization (b), the number of cell nuclei for each experiment (c) and pooled after normalization (d), the average nucleus volume for each experiment (e) and pooled (f), the average cell density for each experiment (g) and pooled (h). T47D spheroids were treated with 0.08 μM and 0.16 μM Bafilomycin for autophagy inhibition, 0.08 μM and 0.16 μM Rapamycin for autophagy induction, cell culture medium (control), DMSO, or starved in EBSS medium. Boxplot parameters: the box contains 50% of the data points; the middle line (black) of the box is the median. Whiskers and outliers represent the upper and lower 25% of the data. Outliers are outside the 1.5x interquartile range; far outliers are outside the 3x interquartile range. ●, outliers; ○, far outliers.

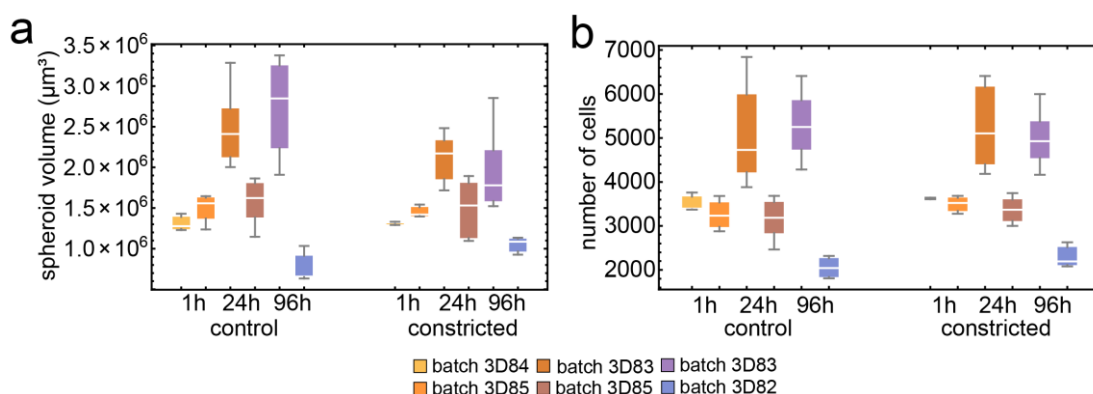
Supplementary Table 8 Number of T47D spheroids per condition.

	batch 14042015	batch 24042015	batch 27042015	batch 01052016	total
Bafilomycin 0.008 μM	3	4	0	3	10
Bafilomycin 0.016 μM	6	4	0	3	13

control	4	4	4	3	15
DMSO	4	4	4	4	16
EBSS	4	4	0	4	12
Rapamycin 0.08 μM	4	4	0	5	13
Rapamycin 0.16 μM	5	4	4	3	16

Supplementary Table 9 Quantitative results for pooled datasets. Results are given as the median and the median absolute deviation.

	spheroid volume (normalized)	number of cells (normalized)	cell nucleus volume (μm^3)	cell density (cells/u.v.)
Bafilomycin 0.008 μM	0.97 \pm 0.22	0.97 \pm 0.25	230 \pm 24	59 \pm 6
Bafilomycin 0.016 μM	0.77 \pm 0.24	0.77 \pm 0.13	257 \pm 19	55 \pm 3
control	0.88 \pm 0.22	0.86 \pm 0.16	255 \pm 27	52 \pm 3
DMSO	0.99 \pm 0.05	1.02 \pm 0.05	267 \pm 35	54 \pm 4
EBSS	0.52 \pm 0.09	0.58 \pm 0.14	249 \pm 24	56 \pm 4
Rapamycin 0.08 μM	0.47 \pm 0.06	0.64 \pm 0.10	214 \pm 23	61 \pm 7
Rapamycin 0.16 μM	0.53 \pm 0.11	0.67 \pm 0.13	212 \pm 24	64 \pm 3



Supplementary Figure 9 Quantitative results for all experiments. Boxplots of the spheroid volume (a) and the number of cells (b) for each experiment. The data was obtained from four different experiments (batch 3D82, 3D83, 3D84 and 3D85). Boxplot parameters: the box contains 50% of the data points; the middle line (black) of the box is the median. Whiskers and outliers represent the upper and lower 25% of the data.

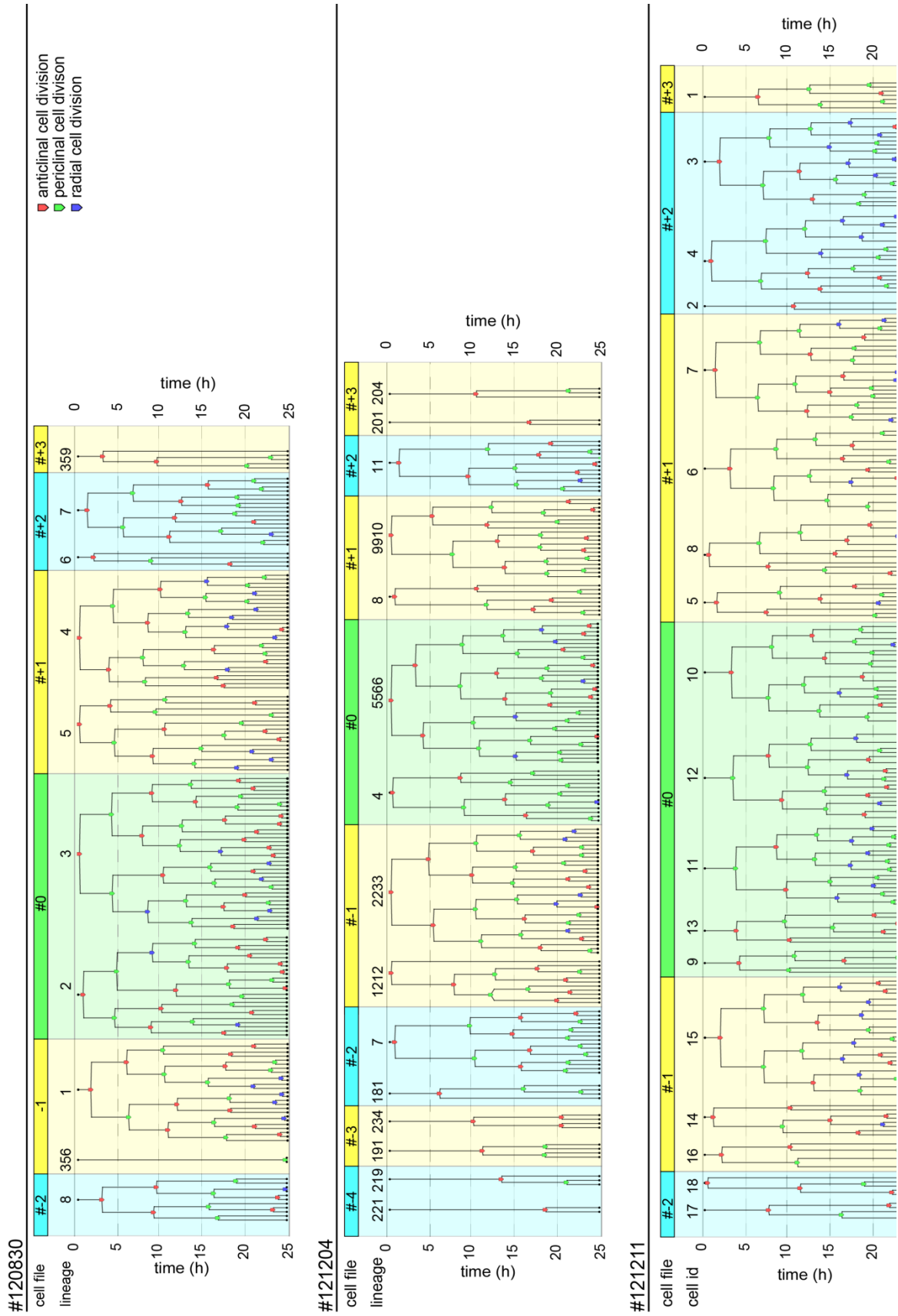
Supplementary Table 10 Number of HC11 spheroids per condition. Spheroids were generated from 5,000 cells seeded and transferred into agarose capillaries with a diameter of 0.3 mm (control) or 0.2 mm (constricted).

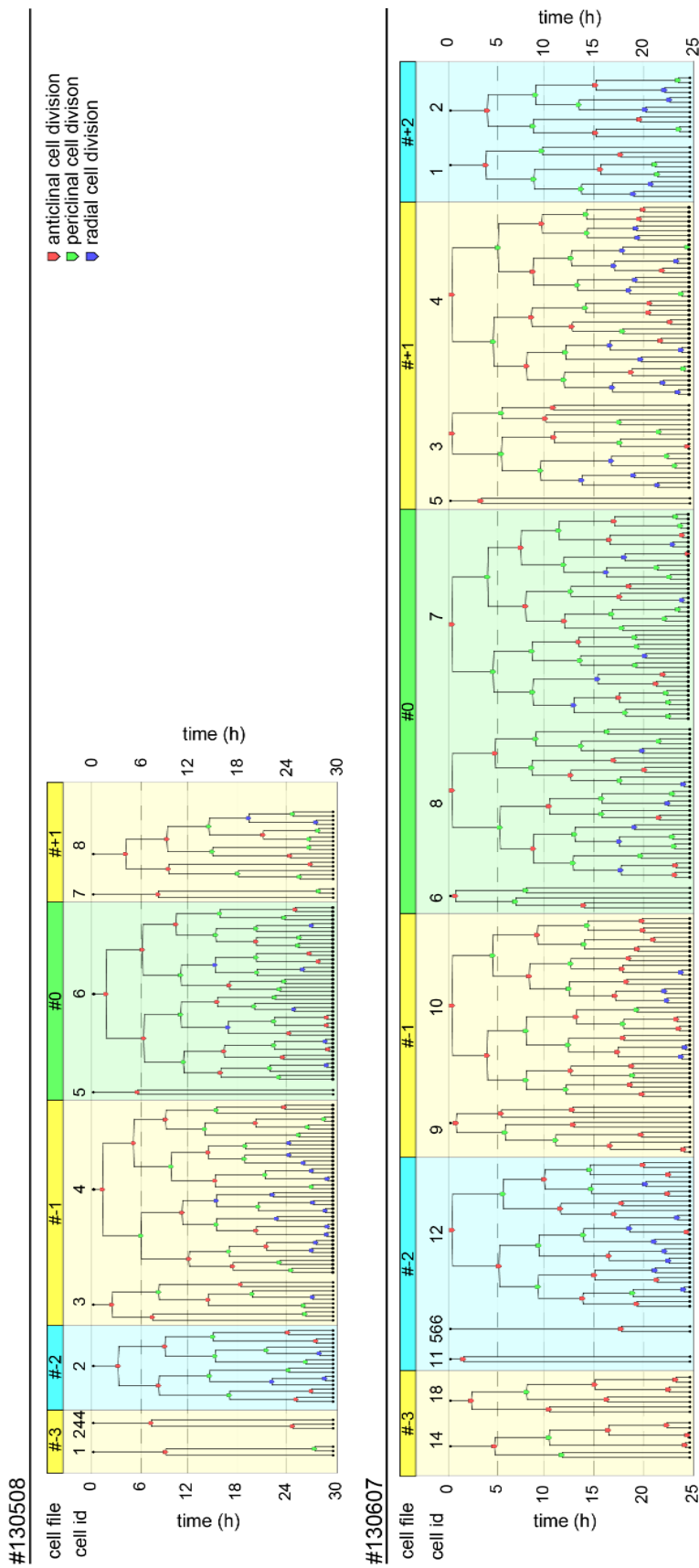
	1h p.t.	24h p.t.	72h p.t.
control	7	19	9
constricted	6	17	13

Supplementary Table 11 Quantitative results for pooled datasets. Results are given as the median and the median absolute deviation.

	spheroid volume (normalized)	number of cells (normalized)	mean cell nucleus volume (μm^3)	cell density (cells/u.v.)
control, 1h p.t.	1.00±0.04	1.00±0.05	73±18	17±3
constricted, 1h p.t.	1.00±0.04	1.07±0.01	71±19	17±4
control, 24h p.t.	1.00±0.12	1.00±0.11	93±21	13±3
constricted, 24h p.t.	0.90±0.14	1.07±0.11	73±21	15±4
control, 96h p.t.	1.00±0.10	1.00±0.10	76±20	15±4
constricted, 96h p.t.	0.72±0.17	0.99±0.08	62±20	17±5

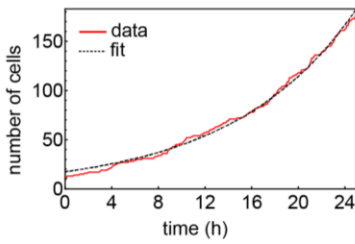
Computer-aided cell tracking and lineage reconstruction





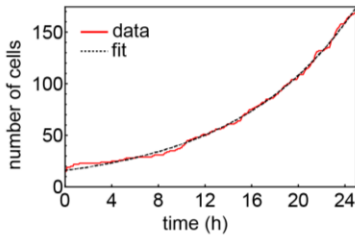
Supplementary Figure 10 Lineage trees of lateral root founder cells in all datasets. The cell file and cell id are indicated above each lineage tree. All lineage trees are given in top to bottom orientation. Each pair of vertices represents the first and last occurrence of a cell. Thus, the root vertex of each tree corresponds to the first occurrence of a founder cell, whereas the leaves represent the progenitor cells present at the last time point tracked. The cell division types anticlinal, periclinal, radial are indicated for each cell division.

#120830



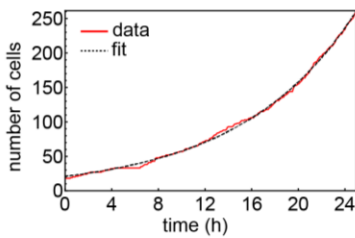
Model	$N_0 e^{kt}$		
Best fit	$17.4918 e^{0.00782631t}$		
Confidence intervals	Estimate	Standard Error	Confidence Interval
	k	0.00782631	0.0000400236 {0.00774755, 0.00790508}
	N_0	17.4918	0.172263 {17.1528, 17.8308}

#121204



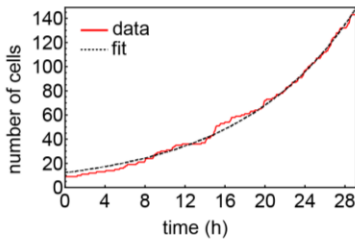
Model	$N_0 e^{kt}$		
Best fit	$15.9077 e^{0.00798417t}$		
Confidence intervals	Estimate	Standard Error	Confidence Interval
	k	0.00798417	0.0000292333 {0.00792664, 0.0080417}
	N_0	15.9077	0.114806 {15.6817, 16.1336}

#121211



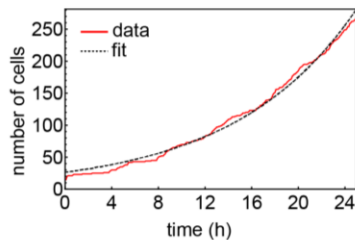
Model	$N_0 e^{kt}$		
Best fit	$20.9916 e^{0.00840563t}$		
Confidence intervals	Estimate	Standard Error	Confidence Interval
	k	0.00840563	0.0000242466 {0.00835792, 0.00845335}
	N_0	20.9916	0.126714 {20.7423, 21.241}

#130508



Model	$N_0 e^{kt}$		
Best fit	$12.309 e^{0.00712489t}$		
Confidence intervals	Estimate	Standard Error	Confidence Interval
	k	0.00712489	0.0000363864 {0.00705332, 0.00719645}
	N_0	12.309	0.129818 {12.0537, 12.5644}

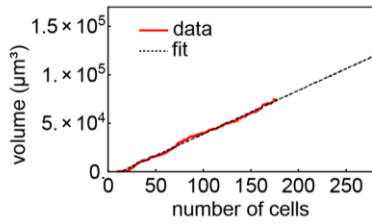
#130607



Model	$N_0 e^{kt}$		
Best fit	$26.2874 e^{0.00790168t}$		
Confidence intervals	Estimate	Standard Error	Confidence Interval
	k	0.00790168	0.0000514495 {0.00780043, 0.00800293}
	N_0	26.2874	0.33332 {25.6315, 26.9434}

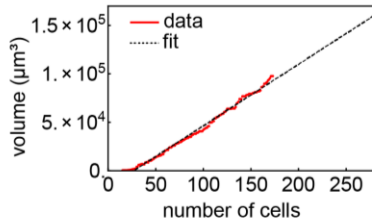
Supplementary Figure 11 Plot of the exponential growth profiles and fitting parameters. Left column: plots of the number of cells as a function of time since gravistimulation of the plant (red curve). The black dashed line represents the fitted exponential growth curve. Right column: parameters table of the exponential growth model fit, where N_0 the number of cells at the first time point t_0 and k is the growth rate.

#120830



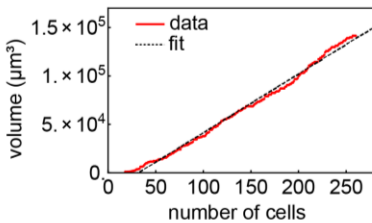
Model	$ax + b$			
Best fit	448.865 x - 5851.4			
Confidence intervals		Estimate	Standard Error	Confidence Interval
	a	448.865	2.19275	{444.536, 453.194}
	b	-5851.4	229.695	{-6304.92, -5397.88}

#121204



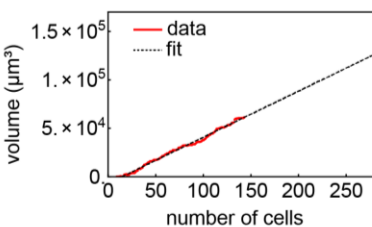
Model	$ax + b$			
Best fit	634.693 x - 17028.5			
Confidence intervals		Estimate	Standard Error	Confidence Interval
	a	634.693	4.41763	{625.967, 643.418}
	b	-17028.5	462.116	{-17941.3, -16115.8}

#121211



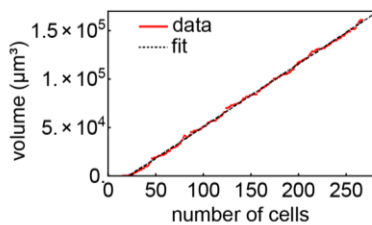
Model	$ax + b$			
Best fit	607.109 x - 19291.6			
Confidence intervals		Estimate	Standard Error	Confidence Interval
	a	607.109	3.18936	{600.827, 613.392}
	b	-19291.6	496.575	{-20269.8, -18313.4}

#130508



Model	$ax + b$			
Best fit	475.992 x - 6821.65			
Confidence intervals		Estimate	Standard Error	Confidence Interval
	a	475.992	3.14595	{469.769, 482.215}
	b	-6821.65	268.692	{-7353.11, -6290.18}

#130607



Model	$ax + b$			
Best fit	649.465 x - 14203.2			
Confidence intervals		Estimate	Standard Error	Confidence Interval
	a	649.465	1.31216	{646.881, 652.049}
	b	-14203.2	208.36	{-14613.5, -13792.8}

Supplementary Figure 12 Plot of the volume as a function of cell number and fitting parameters.

Left column: plots of the number of cells as a function of time since gravistimulation of the plant (red curve). The black dashed line represents the fitted exponential growth curve. Right column: parameters table of the exponential growth model fit, where N_0 the number of cells at the first time point t_0 and k is the growth rate.

Supplementary Table 12 Contribution of cell files to the lateral root. The contribution is given in number of cells in each cell file at the 143 cells stage for each dataset.

	cell file								
	#-4	#-3	#-2	#-1	#0	#+1	#+2	#+3	#+4
#120830	-	-	8	19	50	42	20	4	-
#121204	5	8	18	35	38	22	11	5	-
#121211	-	-	7	24	43	39	26	4	-
#130508	-	6	18	53	46	20	-	-	-
#130607	-	-	13	32	44	26	18	11	-

Supplementary Table 13 Number of anticlinal, periclinal and radial cell divisions. The numbers comprise all cell divisions up to the developmental stage of 143 cells.

	number of anticlinal divisions	number of periclinal divisions	number of radial divisions
#120830	59	60	15
#121204	55	64	10
#121211	50	57	19
#130508	51	56	27
#130607	60	60	9
mean	55	59	16
SD	5	3	7

Supplementary Table 14 First occurrence of anticlinal, periclinal and radial cell divisions. Note that the first cell division round for dataset #121211 could not be properly reconstructed. The dataset was therefore omitted from the analysis.

	anticlinal			periclinal			radial		
	cell division round	number of cells	cell file	cell division round	number of cells	cell file	cell division round	number of cells	cell file
#120830	1	10	0	2	21	0	3	34	0
#121204	1	16	-1	2	29	+1	4	65	0
#130508	1	9	-1	2	16	-1	4	43	0
#130607	1	15	-1	2	28	0	4	95	0
mean		13			24			60	
SD		4			6			27	

Supplementary Table 15 Occurrence of cell layers. The occurrence of each layer is given in number of cells.

dataset	cell layer						
	P	O	I	OO	OI	IO	II
#120830	22	58	66	146	-	119	152
#121204	30	60	73	-	-	147	154
#121211	30	64	89	157	172	145	165
#130508	17	48	56	-	-	-	-
#130607	29	56	55	158	174	102	172
mean	26	57	68	154	173	128	161
SD	6	6	14	7	1	22	9



Curriculum Vitae - Alexander Schmitz

Personal information

Place of birth Daun (Eifel)
Date of birth 27/06/1985

Education

Aug 2012 – Aug 2016 **Ph. D. student in Bioinformatics**
Physical Biology Group, Prof. Ernst H. K. Stelzer,
Buchmann Institute for Molecular Life Sciences (BMLS),
Goethe Universität Frankfurt am Main

Aug 2011 – May 2012 **Diploma thesis in Bioinformatics**
Thesis: „*Bioinformatical Analyses of Medical Image Data for Cell Recognition*”
Molecular Bioinformatics Group, Prof. Ina Koch,
Institute for Computer Science and Mathematics, Goethe
Universität Frankfurt am Main
Graduation mark: 1.0

Oct 2006 – Aug 2011 **Studies in Bioinformatics (Diploma)**
Goethe Universität Frankfurt am Main
Major subjects (graduation mark):

- Practical Bioinformatics (1.0)
- Biochemistry (1.0)
- Theoretical Bioinformatics (1.0)
- Structure of Biomolecules (1.3)

Jun 2005 **Graduation from high school**
Salvatorkolleg, Bad Wurzach
Graduation mark: 2.1

Professional experience

Since Nov 2016 **IT Consultant**
Solveno GmbH

Kirchstraße 21b, 65239 Hochheim am Main

- Aug 2012 – Aug 2016 **Scientific employee**
Computational image analysis
Physical Biology Group of Prof. Ernst H. K. Stelzer,
Buchmann Institute for Molecular Life Sciences (BMLS),
Max-von-Laue-Straße 15, Goethe Universität Frankfurt am
Main
- May 2006 – Aug 2006 **Biological laboratory assistant (internship)**
Production of antibodies, *in vitro* diagnostics test kits and
reagents
PROGEN Biotechnik GmbH, Maaßstraße 30, 69123 Heidelberg

Publications

Peer-reviewed

- (1) Mathew, B.*, Schmitz, A., Muñoz-Descalzo, S., Ansari, N., Pampaloni, F. and Stelzer, E.H.K. (2015) Robust and automated three-dimensional segmentation of densely packed cell nuclei in different biological specimens with Lines-of-Sight decomposition. *BMC Bioinformatics* **16**:187. doi: 10.1186/s12859-015-0617-x.
- (2) Strobl, F.*, Schmitz, A. and Stelzer, E.H.K. (2015) Live imaging of *Tribolium castaneum* embryonic development using light-sheet-based fluorescence microscopy. *Nature Protocols* **10**:1486–507. doi:10.1038/nprot.2015.093.
- (3) von Wangenheim, D.*, Fangerau, J., Schmitz, A., Smith, R.S., Leitte, H., Stelzer, E.H.K. and Maizel, A. (2016) Rules and Self-Organizing Properties of Post-embryonic Plant Organ Cell Division Patterns. *Current Biology* **26**:439–49. doi:10.1016/j.cub.2015.12.047.
- (4) Dini, S.*, Binder, B. J., Fischer, S.C., Mattheyer, C., Schmitz, A. and Stelzer, E.H.K., Bean, N.G. and Green, J.E.F (2016). Identifying the necrotic zone boundary in tumour spheroids with pair-correlation functions. *J. R. Soc. Interface* **13**:123. doi: 10.1098/rsif.2016.0649.
- (5) Schmitz, A.*, Fischer, S.C.*, Mattheyer, C., Pampaloni, F. and Stelzer, E.H.K. (2017) Multiscale image analysis reveals structural heterogeneity of the cell microenvironment in homotypic spheroids (accepted for publication, *Scientific Reports*).
- (6) Strobl, F.*, Schmitz, A. and Stelzer, E.H.K. (2017) Improve your four-dimensional image – traveling ten innovative and insightful years in the multiverse of light sheet-based fluorescence microscopy (invited commentary, submitted).

Non-peer-reviewed

- (1) Smyrek, I.*, Hötte, K.*, Strobl, F.*, Schmitz, A.* and Stelzer, E.H.K. (2015) Die Macht der dunklen Seite. *Forschung Frankfurt*, **2**: 16-20.
- (2) Smyrek, I.*, Hötte, K.*, Strobl, F.*, Schmitz, A.* and Stelzer, E.H.K. (2015) Scheibchenweise statt am Stück. *Laborjournal*, **9**: 72-74

*first author

Conferences

- (1) Quantitative BioImaging Conference (QBI), 13-15 Jan 2016, Delft. Presentation and poster: “Scalable quantitative characterization of three-dimensional multicellular aggregates at single cell resolution”.
- (2) Focus on Microscopy Conference (FOM), 29 Mar – 1 Apr 2015, Göttingen. Presentation: “Robust quantitative analysis of structural morphology in three-dimensional tumor spheroids”.
- (3) European Wolfram Technology Conference (EWTC), 2-3 Jun 2015, Frankfurt am Main. Attended.
- (4) Light Sheet Fluorescence Microscopy International Conference, 25-26 Sep 2014, Barcelona. Presentation: “*Robust quantitative analysis of cellular dynamics in 3D tumor spheroids*”.
- (5) European Wolfram Technology Conference (EWTC), 27-28 May 2014, Frankfurt am Main. Presentation: “*The virtual plant: Mathematica in 4D cell tracking of lateral root development in Arabidopsis thaliana*”.
- (6) Molecular Life Sciences Conference (MLSC), 3-6 Oct 2013, Frankfurt am Main. Attended.
- (7) European Wolfram Technology Conference (EWTC), 11-12 Jun 2013, Frankfurt am Main. Attended.

Erklärung

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung im Mathematisch-Naturwissenschaftlichen Bereich unterzogen habe.

Frankfurt am Main, den _____

Alexander Schmitz

Eidesstattliche Versicherung

Ich erkläre hiermit an Eides Statt, dass ich die vorgelegte Dissertation über „*Application-driven development of modules for an image analysis pipeline in three-dimensional fluorescence microscopy*“ selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind. Ich versichere, die Grundsätze der guten wissenschaftlichen Praxis beachtet, und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Frankfurt am Main, den _____

Alexander Schmitz