

HelmholtzZentrum münchen
German Research Center for Environmental Health



Quantification and analysis of single-cell protein
dynamics in stem cells using time-lapse microscopy

Michael Schwarzfischer

December 2013

TECHNISCHE UNIVERSITÄT MÜNCHEN

Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung
und Umwelt

Lehrstuhl für Mathematische Modelle biologischer Systeme (M12)

Quantification and analysis of single-cell protein dynamics in stem cells using time-lapse microscopy

Michael Schwarzfischer

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende: Univ.-Prof. Dr. I. Antes

Prüfer der Dissertation: 1. Univ.-Prof. Dr. Dr. F. J. Theis
2. Univ.-Prof. Dr. H.-W. Mewes

Die Dissertation wurde am 19.12.2013 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 24.06.2014 angenommen.

Danksagung

Als erstes möchte ich mich bei meinem Betreuer Carsten bedanken. Für deinen guten Rat, die Motivation, deine Geduld und den ganzen Spaß an dieser Arbeit!

Vielen Dank Fabian, dass du mir dieses Projekt anvertraut und ermöglicht hast. Es war mir eine Ehre unter dir zu promovieren.

Ein weiteres Dankeschön geht an meine Kollegen Felix, Michi, Sabine und Ivan mit denen ich die großen und kleinen Probleme der Wissenschaft lösen konnte, naja zumindest haben wir sie diskutiert. Außerdem danke an das ganze QSCD, CMB und ICB Team für die tolle Arbeitsatmosphäre.

Danke auch an meine Kollaborationspartner Philipp, Adam und Timm für die erfolgreiche Zusammenarbeit. Ich hoffe wir trinken bald unser längst überfälliges Bier!

Bleiben noch meine Roomies Jan und Steffen. Wie langweilig wäre der Alltag ohne euch...

Ein wichtiger Dank geht an meine Eltern und Familie für eure Unterstützung über all die Jahre. Ebenso möchte ich mich bei meinen Freunden und meiner Band bedanken, die immer zu mir gehalten haben und es bis zu Letzt geschafft haben, dass ich auch mal abschalten konnte.

Vielen Dank an Julia für das geniale QTFy Logo.

Zu guter Letzt: Ein großes Dankeschön an Theresa, dass du mich durch alle Hochs und Tiefs dieser Arbeit begleitet und unterstützt hast. Du hast mehr zu dieser Arbeit beigetragen als du denkst :*

Abstract

Single-cell analysis based on time-lapse microscopy is an emerging field in molecular biology. Microscopy has been revolutionized in recent years and allows capturing digital images of biological samples in a high throughput manner. In collaboration with the Research Unit Stem Cell Dynamics, Helmholtz Zentrum München, we develop methods and tools to efficiently process single-cell data based on long-term time-lapse microscopy movies.

To resolve issues of fluorescence imaging we develop methods to normalize fluorescence images using machine-learning algorithms. In combination with image processing methods we accurately measure cellular fluorescence of individual cells. We incorporate all developed methods in a software tool which is optimized to provide an efficient and intuitive workflow to quantify and analyze cellular fluorescence. We validate our quantifications based on well-established flow cytometry data.

The two transcription factors PU.1 and Gata1 play a major role in myeloid lineage decision. We investigate a new mouse model incorporating these two factors expressed as fusion proteins together with fluorescence proteins and analyze their explicit function on a single-cell level by fluorescence quantification. Interestingly, we find no detectable Gata1 expression during granulocyte/macrophage differentiation. On the other hand every Gata1 expression leads to megakaryocyte/erythrocyte differentiation. This finding is in clear contrast with differentiation models of current literature.

We investigate Nanog expression, which is a key factor of pluripotency in embryonic stem cells, and reproduce the reported heterogeneity using single-cell protein dynamics. By quantifying immunofluorescence of further pluripotency factors we identify two emerging subpopulation with differences in their actual Nanog dynamics and distinct partial correlation networks indicating a different inherent transcription factor wiring. Furthermore we find indications that an interaction between the transcription factors Rex1 and Klf4 is missing in the current network of pluripotency.

In a further application of our tool, we unravel potential cell-to-cell variability of protein half-life between individual cells. After evaluating the estimation of protein half-life by fluorescence imaging on a population level we quantify this molecular parameter on a single-cell level. The protein half-life of Nanog is heterogeneously distributed among cells, which is not correlated to the heterogeneous protein abundance. With this technology we are also able to measure protein half-life in rare cell types such as primary hematopoietic stem cells, a feature impossible up to now.

By publishing our tools and methods we contribute to a growing community and provide the basic concepts supporting experimentalist to answer biological questions requiring single-cell quantification.

Zusammenfassung

Einzelzellanalyse mittels zeitaufgelöster Mikroskopie ist eine vielversprechende Technologie in der heutigen molekularbiologischen Forschung. Mikroskopie hat sich in den letzten Jahren so stark entwickelt, dass nun digitale Aufnahmen mit dem Mikroskop einfach und schnell zu erstellen sind. Im Rahmen dieser Arbeit entwickeln wir Methoden und Programme in einer Kollaboration mit der Forschungseinheit Stem Cell Dynamics am Helmholtz Zentrum München, die das effiziente Prozessieren der Einzelzelldaten aus Zeitraffer-Filmen ermöglicht.

Um die Probleme bei der Quantifizierung von Fluoreszenzbildern zu bewältigen, entwickeln wir Normalisierungsmethoden mittels maschinellen Lernens. In Kombination mit Bildverarbeitungsmethoden vermessen wir so Fluoreszenzsignale auf Einzelzellebene. Wir fassen alle Methoden in einem Programm zusammen, das darauf optimiert ist, dem Anwender einen möglichst effizienten und intuitiven Arbeitsablauf zu ermöglichen. Wir validieren unsere zellulären Fluoreszenzmessungen mit Daten der etablierten Durchflussszytometrie.

Die Interaktion von zwei Transkriptionsfaktoren PU.1 und Gata1 spielt eine große Rolle bei der myeloischen Zelldifferenzierung. Wir wenden unser Programm auf Zeitrafferfilme mit hämatopoetischen Stammzellen an, in denen die zwei Faktoren als fluoreszierende Fusionsproteine exprimiert werden. Einerseits ist für differenzierende Zellen Richtung Granulozyt/Makrophagen keine Gata1 Expression festzustellen, andererseits führt jede erkannte Gata1 Expression zu einer Linienentscheidung Richtung Megakaryozyt/Erythrozyt. Dieses Ergebnis stellt einen klaren Konflikt zu bisherigen Differenzierungsmodellen dar.

Da der Transkriptionsfaktor Nanog eine zentrale Rolle für die Pluripotenz von embryonalen Stammzellen spielt, untersuchen wir die Expression von Nanog, ebenfalls in Form eines fluoreszierenden Fusionsproteins, und vollziehen die bekannte Heterogenität des Faktors nach. Wir analysieren die Zusammenhänge von zusätzlichen Transkriptionsfaktoren mittels Immunfluoreszenzmessung und partieller Korrelation. Es stellt sich heraus, dass komplette Stammbäume, die über 4 Tage keinerlei Nanog Expression zeigen, ein anderes Korrelationsnetzwerk aufzeigen verglichen mit Stammbäumen mit heterogener Nanog Expression. Außerdem finden wir Indizien für eine Interaktion zwischen den Transkriptionsfaktoren Klf4 und Rex1, die im regulatorischen Netzwerk der Pluripotenz bisher unbekannt ist.

In einer weiteren Anwendung unseres Programms überprüfen wir die Variabilität der Protein Halbwertszeit zwischen individuellen Zellen. Nachdem wir anhand von Western Blot Ergebnissen validieren, dass die Halbwertszeit in unseren Zeitrafferfilmen auf Populationsebene korrekt messbar ist, finden wir für Nanog eine heterogen verteilte Halbwertszeit. Interessanterweise stellen wir keinen Zusammenhang zwischen Nanog Expression und Halb-

wertszeit fest. Im Gegensatz zu etablierten Methoden sind wir mit dieser Technik in der Lage die Protein Halbwertszeiten von primären Stammzellen zu messen.

Durch die Veröffentlichung unserer Programme bieten wir eine Grundlage für die weitere Forschung auf Einzelzellebene an und ermöglichen weitere Erkenntnisse für biologische Fragestellungen, die eine Quantifizierung auf Einzelzellebene voraussetzen.

Publications

The results presented in this thesis are partly based on the following papers, which are already published or within the publication process:

- **M. Schwarzfischer***, C. Marr*, J. Krumsiek, P. S. Hoppe, T. Schroeder, and F. J. Theis. Efficient fluorescence image normalization for time lapse movies. In *Microscopic Image Analysis with Applications in Biology 2011*, 2011.
- I. Burtscher, W. Barkey, **M. Schwarzfischer**, F. J. Theis, and H. Lickert. The Sox17-mCherry fusion mouse line allows visualization of endoderm and vascular endothelial development. *Genesis*, 2012.
- C. Marr*, M. Strasser*, **M. Schwarzfischer**, T. Schroeder, and F. J. Theis. Multi-scale modeling of GMP differentiation based on single-cell genealogies. *The FEBS Journal*, 2012.
- F. Buggenthin, C. Marr, **M. Schwarzfischer**, P. S. Hoppe, O. Hilsenbeck, T. Schroeder, and F. J. Theis. An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC Bioinformatics*, 2013.
- **M. Schwarzfischer***, O. Hilsenbeck*, B. Schauburger, S. Hug, A. Filipczyk, P. S. Hoppe, M. Strasser, F. Buggenthin, J. S. Feigelman, J. Krumsiek, D. Loeffler, K. D. Kokkaliaris, A. J. J. van den Berg, M. Endeke, S. Hastreiter, C. Marr, F. J. Theis, and T. Schroeder. Single-cell quantification of cellular and molecular behavior in long-term time-lapse microscopy. Under review at *Nature methods*.
- P. S. Hoppe, **M. Schwarzfischer**, D. Loeffler, K. D. Kokkaliaris, O. Hilsenbeck, N. Moritz, M. Endeke, A. Filipczyk, M. A. Rieger, C. Marr, M. Strasser, B. Schauburger, I. Burtscher, O. Ermakova, A. Bürger, H. Lickert, C. Nerlov, F. J. Theis and T. Schroeder. Random PU.1 / Gata1 protein ratios do not induce early myeloid lineage choice. Under review at *Nature*.
- A. Filipczyk*, C. Marr*, **M. Schwarzfischer**, J. S. Feigelman, S. Hastreiter, P. S. Hoppe, D. Loeffler, K. D. Kokkaliaris, M. Endeke, B. Schauburger, O. Hilsenbeck, J. Hasenauer, K. Anastassiadis, F. J. Theis, and T. Schroeder. Network Plasticity of Pluripotency Transcription Factors in Embryonic Stem Cells. Under review at *Cell*.

* These authors contributed equally.

Supervised theses

During my work as a doctoral candidate I supervised the following theses:

- Felix Buggenthin, Computational prediction of hematopoietic cell fates using single cell time lapse imaging. Diploma thesis 2011, supervised together with Carsten Marr (Buggenthin, 2011).
- Oliver Hilsenbeck, Automated construction of cell lineage trees from time-lapse microscopy data. Bachelor thesis 2011 (Hilsenbeck, 2011)
- Sven Punga, Supervised lipid raft cluster identification in hematopoietic stem cells. Bachelor thesis 2013 (Punga, 2013)
- André Seitz, Development and Application of a Single Cell Quantification Tool for Microscopy Images. Bachelor thesis 2013 (Seitz, 2013)
- Sandra Fischer, Quantification of marker expression for endoderm and mesoderm segregation. Bachelor thesis 2013, supervised together with Carsten Marr and Ingo Burtscher (Fischer, 2013)
- Oliver Hilsenbeck, Maximally interesting extremal regions (MINER) and a probabilistic model for cell tracking. Master thesis 2014.
- Zijing Liu, Fluorescence microscopy background estimation using regularization methods. Master thesis 2014, supervised together with Laurent Demaret.

Contents

1	Introduction	1
1.1	Biological context	2
1.1.1	Embryonic stem cells	2
1.1.2	Hematopoiesis	3
1.2	Technical background	6
1.2.1	Western Blot analysis	6
1.2.2	Fluorescent proteins	6
1.2.3	Flow cytometry	7
1.2.4	Acquisition of microscopy images	7
1.2.5	Time-lapse microscopy challenges	9
1.3	Research questions	10
1.4	Overview of this thesis	10
2	Methods	13
2.1	Machine-learning	13
2.1.1	DBSCAN	13
2.1.2	Random forest classification	14
2.1.3	Active batch learning	14
2.2	Image processing	15
2.2.1	Definitions	16
2.2.2	Segmentation by thresholding	17
2.2.3	Watershed segmentation	18
2.2.4	Tiling method for background estimation	21
2.2.5	Fitting the time-independent gain	26
2.2.6	Normalization of fluorescence images	26
2.2.7	Quantification	29
2.3	Correlation analysis	29
2.3.1	Pearson correlation coefficient	29
2.3.2	Partial correlation coefficient	29
2.3.3	Coefficient of determination	30
2.4	Maximum likelihood	31
2.5	Model comparison	32
2.5.1	Bayesian information criterion	32

3	Software tools	33
3.1	Existing tracking and segmentation software	33
3.1.1	Single-cell tracking	34
3.1.2	Timm’s Tracking Tool	34
3.1.3	Segmentation and quantification	35
3.1.4	Conclusion	37
3.2	Background estimation tools	38
3.2.1	Unsupervised background estimation	38
3.2.2	Supervised background estimation	38
3.3	Quantification of tracked fluorescence (QTFy)	41
3.3.1	QTFy workflow	41
3.3.2	Quantifying tracked cells	41
3.3.3	Visualization of time courses	44
3.3.4	Tree browsing	47
3.3.5	Heat tree	47
3.3.6	Cell inspector	48
3.3.7	Summary	49
3.4	Quantification of fluorescence snapshot images (sQTFy)	50
3.5	Annotated single-cell tracking	52
4	Benchmarking and evaluation of QTFy	55
4.1	Fluorescence normalization	55
4.1.1	Existing methods	57
4.1.2	Our method	57
4.1.3	Evaluation based on bead fold-changes	58
4.1.4	Discussion	58
4.2	Determination and evaluation of a negative gate	61
4.2.1	Background level of <i>in silico</i> cells	61
4.2.2	Background level of control cells	62
4.2.3	Autofluorescence of cells vanishes after 12 hours	62
5	Single-cell kinetics refute the stochastic PU.1 / Gata1 switch as the basis of myeloid lineage choice	65
5.1	Biological background: hematopoiesis	66
5.2	The PU.1eYFP/Gata1mCHERRY double knock-in mouse	69
5.3	Phenotypical comparison of wild type and double knock-in mice	69
5.3.1	Blood counts	71
5.3.2	Composition of bone marrow	71
5.3.3	Protein half-life	71
5.3.4	Colony assay contribution	72
5.4	Simultaneous detection of PU.1eYFP and Gata1mCHERRY	72
5.5	Comparison to flow cytometry	74
5.6	Inference of protein abundance	75
5.6.1	Gata1mCHERRY	75

5.6.2	PU.1eYFP	80
5.6.3	Mapping protein numbers to imaging intensity	81
5.6.4	Detection limit in protein numbers	83
5.7	Tracking and quantification of differentiating HSCs until lineage marker expression	84
5.8	Analysis of GM and MegE differentiating branches	88
5.9	Population analysis of GM and MegE differentiation	91
5.10	PU.1 dynamics before lineage marker onsets	96
5.11	GemM trees show no interdependence of PU.1 and Gata1	100
5.12	Discussion	102
6	Network heterogeneity of pluripotency transcription factors in embryonic stem cells	105
6.1	Biological background: embryonic stem cells	106
6.2	Novel pluripotency transcription factor protein reporter ESC lines	108
6.3	Live continuous long-term single-cell quantification of Nanog-VENUS protein expression in ESCs	109
6.4	Comparison to flow cytometry	114
6.5	ESCs maintain Nanog concentration on average	117
6.6	Nanog negative ESCs are a heterogeneous cell population	118
6.7	Pluripotency factor correlation networks reveal different Oct4 and Klf4 wiring in Nanog negative versus re-expressing ESC clones	119
6.8	Rex1 influences the interplay between Klf4 and Nanog	122
6.9	Nanog expression is not coordinated with the pluripotency network	125
6.10	NanogVENUS intensity is independent of spatial position within ESC colonies	127
6.11	Discussion	131
7	Single-cell quantification of protein stability	135
7.1	Biological background	136
7.2	Population average analysis	136
7.3	Single-cell snapshot analysis	139
7.4	Single-cell trajectory analysis	140
7.5	Cell-cycle dependence	141
7.6	Protein half-life in siblings	143
7.7	Protein half-lives in rare cell types	145
7.8	Discussion	147
8	Identification of two distinguishable cell populations during gastrulation based on cell movement	149
8.1	Biological background: embryogenesis	149

8.2	Application and results	150
9	Summary and outlook	153
9.1	Summary	153
9.2	Outlook	155
9.3	Conclusion	158

Chapter 1

Introduction

Microscopy began with the first microscope prototype in the late 16th century invented by Hans and Zacharias Jansen and with the first publication in the mid 17th century by Robert Hooke (1667). Its relevance and its importance has quickly been recognized and further promoted by pioneers like Antonie van Leeuwenhoek who studied among other things the flow of blood using his own single-lens microscopes (for a review, see Frischknecht et al. (2009)). With ongoing research and development a new interdisciplinary field was formed attracting physicist, biologists, chemists or mechanical engineers including famous identities like Robert Koch, Isaac Newton or Carl Zeiss (Dunn and Jones, 2004, Frischknecht et al., 2009). Only the collaboration among the disciplines lead to scientific breakthroughs like discovery of bacteria, tuberculosis or viruses (Frischknecht et al., 2009) and was honored by several Nobel prizes.

After photography has been established in the early 19th century, it soon was combined with microscopy, called cinematography (Talbot, 1913), recording the first cell image in 1891. In the late 20th century photography was revolutionized and simplified the process of generating images remarkably. Another important advancement in microscopy was the discovery of the green fluorescence protein (GFP) which emits green light after exposure to ultra violet light (Shimomura et al., 1962). Through fluorescence microscopy the determination of protein location and protein abundance became possible.

However, still images provide unconnected snapshot data and only by observations over time causality can be obtained. Furthermore, a higher magnification was needed to reach single-cell resolution to investigate dynamics of individual cells. The new requirements were initially pushed by embryologists and developmental biologist (for a review, see Coutu and Schroeder (2013)) leading to first single-cell tracking producing full lineage traces (Sulston and Horvitz, 1977). To fulfill these requirements, new hardware and software had to be developed to successfully accomplish these tasks.

The requirements and costs of generating such images has been greatly reduced leading to huge amounts of data. While in the early days, bioimages could be quickly analyzed by humans, today man-years may be needed to overview the data generated by a single experiment. Therefore, intelligent and efficient methods had to be developed to handle these needs. All these developments lead in the last decade to the new field of bioimage informatics, which applies common image processing approaches to microscopic images of biological samples (Myers, 2012, Peng et al., 2012).

In recent years, the quantification of cellular properties in such images has lead to important biological insights regarding cell-to-cell variability (Elowitz et al., 2002), drug response (Cohen et al., 2008), cell fate prediction (Cohen et al., 2010) and cell-cycle analysis (Neumann et al., 2010) (for a review, see Muzzey and van Oudenaarden (2009)). The study of molecular mechanisms is especially relevant for understanding stem cell differentiation in model systems.

1.1 Biological context

Nature developed manifold mechanisms allowing an organism to generate new tissue after injury or to constantly renew aging cells. Continuously as well as on demand new cells have to be produced keeping the cell type proportions in balance. This process is maintained by a cell type population called *stem cells* (Becker et al., 1963). Stem cells have the following properties: (i) they are pluripotent (i.e. they can become several, more specialized cell types), (ii) they can infinitely self-renew, keeping their actual pluripotent state after cell division. All over the adult body several types of pluripotent stem cells can be found e.g. in the blood, the bones, the brain or the epidermis (Campbell and Reece, 2001). Importantly, stem cells only come in very low cell numbers compared to the cell numbers of mature cell types. For example, there are only about 10^5 blood stem cells (Abkowitz et al., 2002) in an adult human keeping all mature blood cells in balance by renewing 10^{11} – 10^{12} cells per day (Kaushansky et al., 2010). Still, the mechanisms how, when and why a stem cell decides whether to differentiate into a certain lineage or to self-renew are poorly understood.

1.1.1 Embryonic stem cells

During the development of an embryo (embryogenesis) the fertilized single-celled egg undergoes cell division, cell differentiation and morphogenesis (Campbell and Reece, 2001). The so called zygote starts to divide generating cell stages containing 2, 4, 8 or 16 pluripotent cells, called morula. As the number of cells increases the cell cluster converts the morula into a hollow sphere, called blastocyst. An outer and inner morula is formed where cells already differentiate into the three germ layers endoderm, ectoderm

and mesoderm (Alberts et al., 2002). After 4 to 5 days, the outer morula becomes trophectoderm, the inner morula builds the inner cell mass (ICM) which will further differentiate into epiblast and primitive endoderm. At this stage the pluripotent ICM cells can be isolated from the blastocyst and are called embryonic stem cells (ESCs) (Bryja et al., 2006) (Figure 1.1). When cultured in appropriate medium conditions ESCs can be arrested in a pluripotent stage by self-renewal or can be differentiated into all three germ layers.

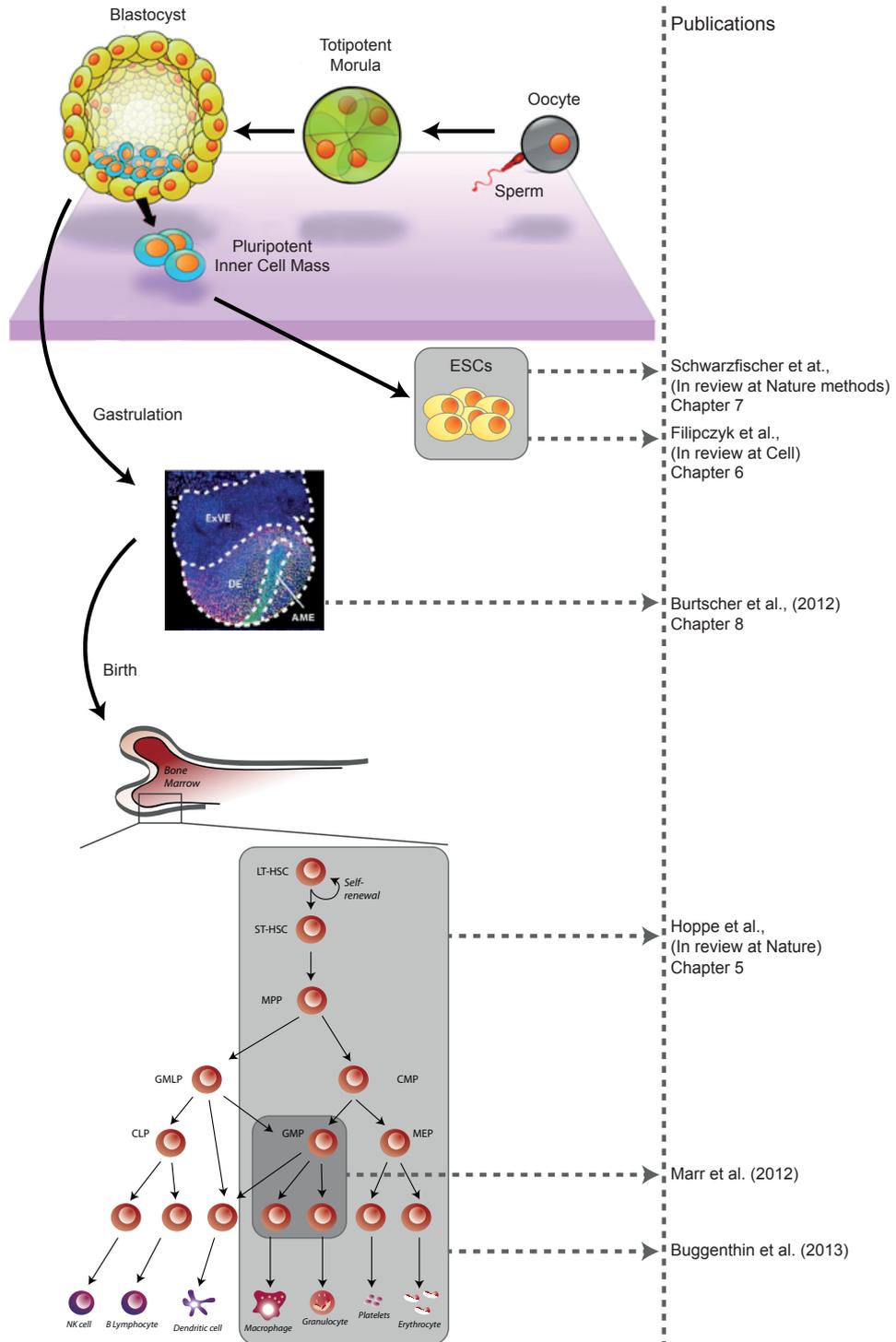
ESCs hold two main promises. First, given the correct treatment ESCs can be theoretically differentiated into all tissues and organs allowing regenerative medicine the repair and replacement of defective cells in a living organism (Wang and Orkin, 2008). Second, ESCs are one of the essential requirements to form chimaeras, the fundamental tool to form animals with genetically modified genomes (Alberts et al., 2002). Furthermore, ESCs can be cultured in large cell numbers and kept in pluripotency serving as an ideal model system. However, ESCs are controversial discussed, since the only way to derive them is to destroy an embryo (Robertson, 2010).

The development from a zygote to a multicellular organism is an astonishing process given the fact that all cells share the same DNA content (Campbell and Reece, 2001). A fine-tuned regulatory mechanism including transcriptional regulation and cell-to-cell signaling resulting in differential gene expression is the basis for the successful formation of an embryo. Also, to keep ESCs in their pluripotent state and to prevent them from differentiating an explicit state of transcriptional regulation has to be maintained. A variety of transcription factor proteins have been reported to play a major role in keeping ESCs pluripotent such as Octamer-binding transcription factor 4 (Oct4, also known as Pou5f1) (Niwa et al., 2000), SRY (sex determining region Y)-box 2 (Sox2) (Rizzino, 2009) or Nanog homeobox (Nanog) (Chambers et al., 2003, Mitsui et al., 2003). The well-defined expression of these factors is sustained by keeping the right stimuli in the medium conditions. The triumvirate Oct4, Sox2 and Nanog and their transcriptional regulation build the transcriptional core network for pluripotency (Chambers and Tomlinson, 2009, Chen et al., 2008, Kim et al., 2008, Loh et al., 2008).

Although pluripotent ESCs are a well-defined cell population, a heterogeneous expression profile of Nanog has been detected (compare Figure 1.2B) (Chambers et al., 2007). Only continuous single-cell observations can investigate how this heterogeneity is established and if it encodes a key function in the context of pluripotency and differentiation.

1.1.2 Hematopoiesis

Hematopoiesis is the process to constantly generate and replenish blood cells (Orkin and Zon, 2008). Hematopoietic stem cells (HSCs) are on top



of the classical hierarchical blood differentiation tree having the ability to self-renew and to give rise to all mature blood cell types (Figure 1.1). By differentiation cells become more and more specialized thereby losing their ability to self-renew. This differentiation process has to be carefully controlled to keep the abundance of mature cell types in balance. Lineage decisions are made either based on external stimuli or intrinsic factors (Rieger and Schroeder, 2007). It is assumed that transcription factors specific for their respective lineage are already expressed in low levels, so called lineage priming (Laslo et al., 2006). Biochemical analyses have shown that auto-activation as well as mutual inhibition exists for these lineage specific transcription factors. The transcription factors PU.1 and Gata1 are thought to play a major role in the lineage decision between granulocyte/macrophage progenitors (GMPs) and megakaryocyte/erythrocyte progenitors (MEPs) (Graf and Enver, 2009). Both factors show a transcriptional self-activation as well as they can inhibit each others activity (Zhang et al., 1999, 2000). Furthermore, both proteins are able to lock down the opposing downstream lineage-specific genes (Nerlov et al., 2000).

Theoretical models describing this toggle-switch proposed explanations to the observed biology. However, all data used was solely based on bulk analysis technologies and lacked single-cell resolution. Furthermore, contradictory data exists, with some reports even questioning the principle of HSCs, reporting self-renewing macrophages and myeloid progenitors (Sieweke and Allen, 2013, Yamamoto et al., 2013). Although research of more than half a century renders the hematopoietic differentiation system to one of the best studied systems, we still do not understand its molecular details (Foster et al., 2009). To get a better insight of the underlying processes it is indispensable to investigate hematopoiesis on a single-cell level from early stem cells up to differentiated cells. Only the continuous monitoring of the involved proteins during this process can clarify the discrepancies.

Figure 1.1 (*preceding page*): An overview of the relevant biology and publications discussed in this thesis. In mice, the fertilized egg gives rise to the blastocyst containing the inner cell mass, which can be used to derive embryonic stem cells (ESCs). We analyze the protein half-life of two important pluripotency transcription factors in Chapter 7 and their relevance for pluripotency maintenance in Chapter 6. At a later stage the blastocyst undergoes gastrulation where cell movement plays a major role investigated in Chapter 8. After birth, the adult mouse hematopoiesis is the process to maintain the blood cells in balance. We investigate the role of two transcription factors known to be involved in stem cell lineage decision in Chapter 5. Figures are adopted from Burtscher et al. (2012), Jones (2009) and Rossi et al. (2012).

1.2 Technical background

In the following, we will shortly reflect current technology discussed in this thesis. Although the following techniques already solved challenges regarding the biology outlined above, we will examine their limitations further motivating the need for a new technology to investigate molecular mechanisms on a single-cell level.

1.2.1 Western Blot analysis

Classically, Western Blotting is used to measure protein expression on a population scale. The method needs a lot of cell samples, which makes this kind of analysis almost impossible to work in systems where every cell is costly, such as primary HSCs. Furthermore, if cell content is numerously available Western Blot can only estimate population averages. Unfortunately, the accuracy and reproducibility of this method can be highly variable (Kreutz et al., 2007). Besides the laborious biological preparation of Western Blots the scanned images have to be carefully corrected for uneven illuminations before further analysis or quantification steps can be performed (Gassmann et al., 2009). Another disadvantage is the need for a cell lysate killing the cells of interest which does not allow to inspect a cell's future fate. Although this technology already gave insights into hematopoiesis by measuring PU.1 half-life (Nutt et al., 2005) or Gata1-PU.1 interaction (Zhang et al., 1999) and into ESCs by measuring Nanog stability (Muñoz Descalzo et al., 2013) or protein levels upon different treatments (Ramakrishna et al., 2011), it does not allow to investigate molecular mechanisms on a single-cell basis.

1.2.2 Fluorescent proteins

In the jellyfish *Aequorea victoria* a green fluorescent protein (GFP) was discovered by Shimomura et al. (1962). This protein can be used after excitation to visualize its location in biological specimen (Alberts et al., 2002). GFP can be used to monitor gene expression in living organisms by adding the promoter of the gene of interest in front of the GFP coding region. Whenever the promoter is activated GFP will be transcribed, translated and maturated within hours. Another application is to tag specific signal peptides to GFP which will transfer the protein to specific cell compartments which allows visualization in microscopy (Coutu and Schroeder, 2013). A further promising method fuses the GFP coding sequence to a gene locus of interest, either C- or N-terminally, resulting in a chimeric product of this particular encoded protein. After the normal function of the fusion protein has been validated, this creates a tool to easily visualize relative protein expression in single-cells. A lot of derivatives by directed genetic modifications using GFP have been created showing differences in half-life,

maturation rate, excitation or emission spectra to broaden the application of fluorescent proteins (Okita et al., 2004, Zhang et al., 2002).

1.2.3 Flow cytometry

Flow cytometry is a commonly used technology to achieve large-scale data with single-cell resolution. A cell sample solution is stained with several fluorescent dyes by the help of anti-bodies to characterize cell properties. A stream of small droplets, ideally containing only one cell at time, is scanned by several lasers measuring the emission of the different dyes. The overlapping emission spectra are unraveled by compensation measuring cell samples with only one dye (Tung et al., 2007). Together with the measured scattering of light, more than a dozen features per cell can be acquired. This technique quantifies complete distributions of several cell properties instead of population averages, but is only applicable if adequate anti-bodies exist. In contrast to Western Blotting, it is not as easy to visualize intracellular cell properties, since the technology is not based on a cell lysate but cells are kept intact and alive. On the other hand, cells can be used for further experiments and cells can be directly sorted into different tubes based on their individual properties.

To sort specific cell populations by certain cell properties (i.e. marker expressions) a user defines virtual gates for one dimension or more complex gates for two selected features (Figure 1.2A). If a feature does not show a clear separation but rather smooth gradient transitions between subpopulations a hard threshold might not be adequate. A few new concepts exist to take the higher dimensionality into account which try to cluster the data into naturally emerging sub-populations (Bendall et al., 2011, Qiu et al., 2011). The high dimensional data has to be carefully analyzed (Herzenberg et al., 2006), but common software only allows to visualize two dimensions at once resulting in scatter plots (Figure 1.2A).

However, flow cytometry is a well-established technology to quantify relative fluorescence of thousands of cells on a single-cell basis and thereby helped to gain insights into ESCs by identifying heterogeneity (Chambers et al., 2007) or investigating repopulation dynamics (Karwacki-Neisius et al., 2013) and into HSCs by identifying new subpopulations (Akashi et al., 2000, Iwasaki et al., 2005a), but still lacks the opportunity to keep the cell identity from one time point to the next.

1.2.4 Acquisition of microscopy images

After the first cell images were captured huge improvements were achieved resulting in today's digital image acquisition. Methods and technology to record images is a field on its own involving biologist, mechanical engineers or physicists (Coutu and Schroeder, 2013). A modern microscope consists

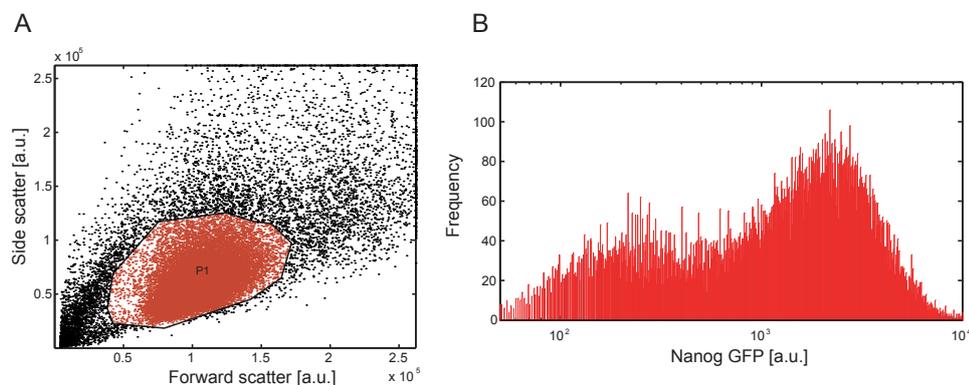


Figure 1.2: A typical flow cytometry gating scheme to sort ESCs. (A) A polygonal gate has been drawn by hand defining a subpopulation by forward scatter (FSC) and side scatter (SSC) properties. (B) ESCs of the TNGA cell line (compare Chambers et al. (2007)) show a heterogeneous NanogGFP expression profile.

of one or more light sources, several shutter systems, a stage (potentially motorized), several lenses, filtercubes and a digital camera. To tune the setup of a microscope for specific tasks one has to carefully choose between several trade-offs regarding the resulting image quality. On the one hand, a high signal to noise ratio for later image processing steps is desirable and can be achieved by using high doses of excitation light, but on the other hand biological cell samples are damageable and might react abnormally or even die due to phototoxicity.

One of the simplest microscopy images is obtained by using white light passing through a sample and by capturing all light which transmits through it resulting in so called bright field images (Campbell and Reece, 2001). Cell samples appear as dark objects since they absorb passing light, the background appears bright. These images can be taken within milliseconds and allow to easily observe tissue or cell outlines although they do not provide a good contrast. A slightly different light way is used to capture fluorescence material. The whole sample is illuminated by a small excitation spectrum and the emitted and reflected light is captured in a different wavelength resulting in fluorescence microscopy images. Ideally, these images show the fluorescence of interest only. However, mostly cell samples are kept in medium, which might react autofluorescently, resulting in additional background signal (Waters, 2009). Therefore it has been of great interest to untangle the cellular signal from the underlying background fluorescence which will be discussed later in this thesis in detail (Chapter 4). Bright field and this particular kind of fluorescence imaging are summarized in the term of wide field microscopy.

Different techniques or microscope setups exist based on different methodologies to capture images (Campbell and Reece, 2001). In contrast to wide-field microscopy, confocal microscopy only illuminates a small fraction of the specimen and scans the whole sample subsequently, resulting in higher spatial resolution and signal to noise ratio by keeping a low phototoxicity. It can also be used to acquire whole volumes of cell samples by capturing several narrow focal planes reconstructing the entire 3D structure of a specimen. Phase-contrast imaging can be used to achieve higher contrast compared to bright field microscopy but needs a more efforts in terms of microscope setup and control.

With this technology we are able to measure cell properties with single-cell resolution and to follow them continuously over time. However, new hard- and software is needed to acquire and analyze these images challenging current bioimage informatics.

1.2.5 Time-lapse microscopy challenges

For the biological questions in this thesis it is of great importance to continuously follow individual cells over a long time period to understand single-cell dynamics. Since this requirement can be achieved by time-lapse microscopy, we were faced by several technical and methodological challenges. Single-cell tracking and quantification has already been successfully applied to simple prokaryotic model systems such as *E. coli* (Elowitz et al., 2002, Taniguchi et al., 2010) or to eukaryotic model systems such as *S. cerevisiae* (Gordon et al., 2007). However, continuous imaging of mammalian cells is more challenging for the experimentalists since developmental or differentiation processes happen on larger timescales as compared to simple organisms. Besides challenging the biological handling and used hardware (such as the microscope or incubation system), increased data volumes will be generated per experiment which have to be efficiently stored, analyzed and quantified. For example, following a single HSC and keeping track of its whole pedigree over one week will give rise up to 16384 cells assuming a cell-cycle length of 12 hours. By imaging cells every two minutes in bright field and additional fluorescence images with a lower time interval in wells covered by 39 fields of view will lead to over 200,000 images and over 200 Gigabytes of data per experiment.

To the best of our knowledge (for a detailed discussion see Section 3.1), there exists no adequate software for single-cell tracking and quantification which is able to handle the requirements of a typical experiment involving mouse HSCs or ESCs: (i) long-term observation up to weeks, (ii) adequate magnification resulting in multiple fields of view, (iii) time-lapse acquisition with a time interval frequent enough to follow highly motile cells possibly changing their morphology (i.e. due to differentiation or cell-cycle) and (iv) several fluorescence channels.

1.3 Research questions

The main goal of this thesis is to quantify fluorescence of labeled proteins on a single-cell basis using time-lapse microscopy images and to investigate the role of these labeled proteins in their respective biological context. To this end, we need to measure individual cells over time requiring to follow the cells (i.e. cell tracking) and to identify correct cell outlines (i.e. cell segmentation). Furthermore, we were faced with different issues of fluorescence imaging challenging us to find adequate normalization methods. We required a method achieving quantification results that compare well to flow cytometry. In this work, we strongly collaborated with the Research Unit Stem Cell Dynamics of Timm Schroeder at the Helmholtz Zentrum München. In our experience, the biological experts did not have any adequate software tools at hand to continuously and reliably follow relative protein abundance of individual cells on a large time scale.

To answer biological research questions, we asked whether we could measure the total amount of protein instead of just relative amount of every single measurement. In particular, we were interested in the stoichiometric interplay of the two transcription factors PU.1 and Gata1 during hematopoietic stem cell differentiation. For this purpose, we developed a method to map protein abundances onto fluorescence quantification in our time-lapse movies. This technology enables us for the first time to investigate single-cell dynamics of the hypothesized stochastic interplay of the two proteins to understand their role in myeloid lineage differentiation.

Furthermore, we tackled long standing questions, how and why heterogeneity emerges within clonal embryonic stem cells. Recently the principles of heterogeneity in ESCs (Chambers et al., 2007) was questioned by Faddah et al. (2013) challenging us to investigate the protein dynamics of ESCs on a single-cell level. Several molecular details such as protein binding partners are already known in this field based on bulk analyses, but only by our imaging technology we were able to visualize co-expression of transcription factor abundances allowing a quantitative analysis of different pluripotency factors and their transcriptional interplay.

1.4 Overview of this thesis

In **Chapter 2**, we briefly outline the methods and algorithms used throughout the thesis. We introduce basic concepts of statistics and machine-learning, but focus on image processing methods. In particular, newly developed methods used for fluorescence image normalization are explained in detail.

After recapitulating existing methods, **Chapter 3** presents the developed software tools (background estimation tools, QTfY, sQTfY and CCT),

which incorporate the methods of **Chapter 2** in user-friendly GUIs. The tools provide an efficient interface to track cells and quantify fluorescence signal. To accurately normalize fluorescence images we present two different methods to estimate the underlying background signal. For all tools we optimized the graphical interface to allow efficient and reliable data generation.

In **Chapter 4** we first evaluate the tools and their reliability by comparing fluorescence quantification based on our toolbox to well-established flow cytometry measurements. Furthermore, we explore the technical detection limits of our quantification approach by simulating *in silico* background cells rendering additional control cells needless.

In **Chapter 5** we apply QTFy to blood stem cells incorporating two fluorescently labeled proteins known to play a role in myeloid lineage differentiation. We further present an approach to estimate real protein abundance and map these numbers onto single-cell trajectories. With this technique we are able to investigate the interplay between the two transcription factors on a single-cell basis on a long time scale. Our results reject current models of the literature and suggests that both proteins are rather lineage executing factors instead of lineage deciders.

After that, we apply our toolbox to mouse ESCs in **Chapter 6** and investigate the role of the pivotal pluripotency factor Nanog. We observe single-cell dynamics and clarify contradictory views in current literature. Additionally, we identify cell populations showing distinct differences in their transcription factor correlation networks.

We use QTFy to quantify protein half-life in **Chapter 7** based on single-cell trajectories. Our results are highly comparable to bulk analysis based methods and highlight that only by integrating the full time and single-cell resolution, we recover cell-to-cell variability in protein stability.

We discover two distinct cell populations emerging during gastrulation using CCT as tracking software in **Chapter 8**.

Chapter 9 summarizes all important insights of this thesis and briefly discusses future extension of the tools, methods and analyses. Furthermore, we outline future applications of our tools and suggest new experiments to confirm and to extend our findings.

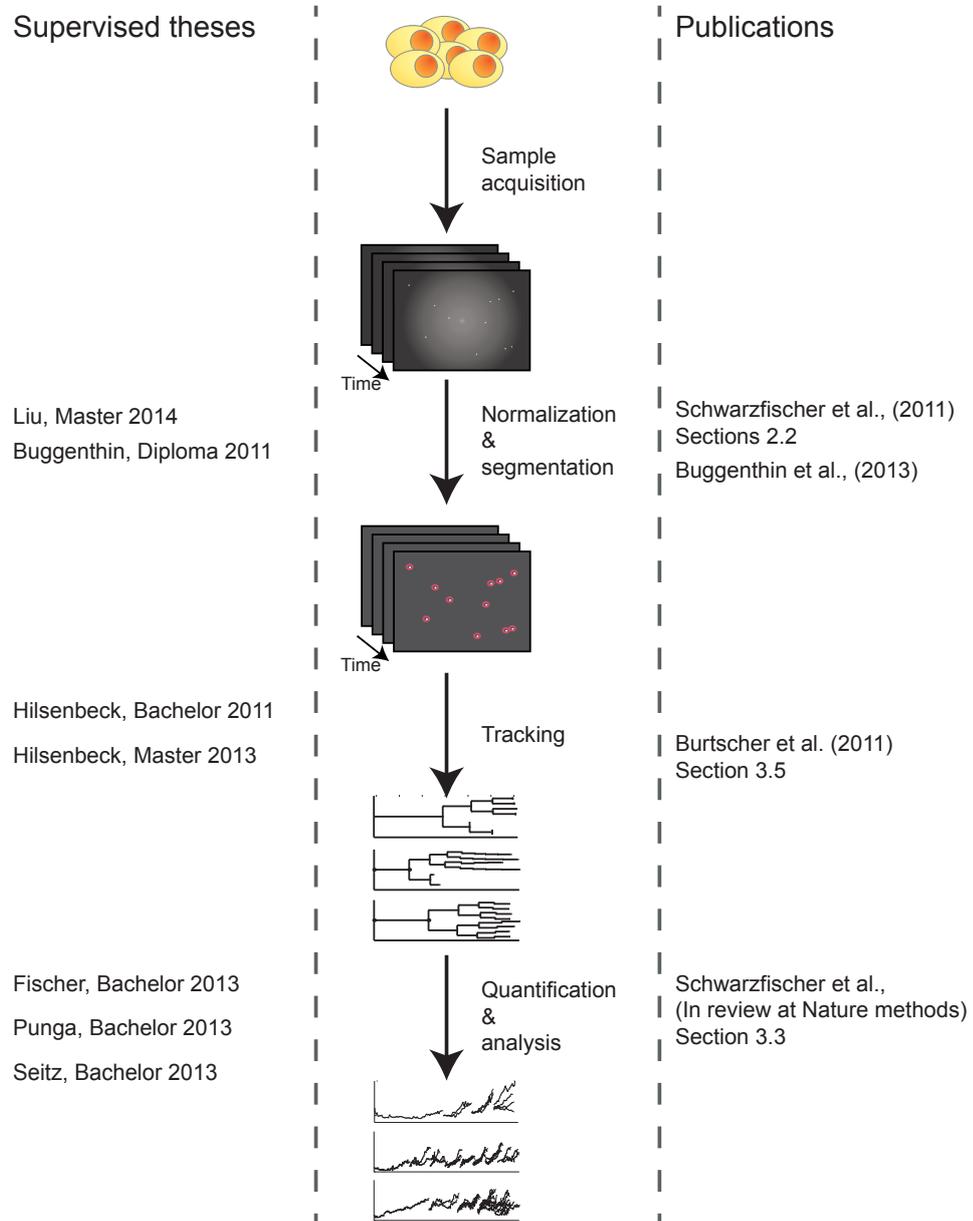


Figure 1.3: Summary of publications and supervised theses during this work from a technical point of view. We discuss general concepts of normalization and segmentation in Sections 2.2. We present a simple tracking tool in Section 3.5 to follow individual cells. The quantification of cellular fluorescence is performed by using the developed tool QTFy (Section 3.3).

Chapter 2

Methods

In this chapter we introduce basic machine-learning and image processing methods used to quantify cellular fluorescence. We discuss different ways how to correct fluorescence images to allow reliable cellular signal quantification. Finally, we introduce correlation analysis, the principle of maximum likelihood and model comparison.

2.1 Machine-learning

To filter out cellular signal in fluorescence images, we utilize the power of machine-learning algorithms. In order to apply these methods we collect f features for n data samples $x_i \in \mathbb{R}^f$ originating from an image to build up a feature matrix $M \in \mathbb{R}^{f \times n}$. A subdivision of M into distinct classes can be performed by a clustering algorithm, which is the process of finding similar subsets in a given data set based on some similarity score (unsupervised machine-learning) (Murphy, 2012). If, additional information, such as labels l_i for each sample x_i , exist, a classification algorithm can be applied trying to separate the data into the given class labels (supervised machine-learning). Classically, unsupervised machine-learning needs a set of parameters, on the other hand supervised machine-learning needs additional labels.

2.1.1 DBSCAN

A density based clustering, called DBSCAN (Density-based spatial clustering of applications with noise) (Ester et al., 1996), is an unsupervised machine-learning algorithm. It iterates over all data points in a multidimensional feature space and calculates the number m of neighboring points in a given radius ϵ . If m is below some threshold, MinPts, the data point is regarded as noise, otherwise the point is called density-reachable and the point is assigned to a cluster. DBSCAN automatically finds the number of clusters and only needs two input parameters: ϵ and MinPts.

2.1.2 Random forest classification

Decision trees, also called classification and regression trees (CART) (Murphy, 2012), need additional labels and are hence supervised machine-learning methods. They recursively partition the data space into sub regions by defining a local model on each region. The partitioning can be represented by a tree structure, with leaves of the tree encoding for the found classes (Breiman et al., 1984). To grow the tree, the data has to be split at each node according to a defined cost function evaluating the possible features. A very popular cost function for classification by decision trees is the Gini index (Gini, 1921). For a given feature variable t partitioning the data W into n classes W_1, \dots, W_n the Gini index is defined as

$$Q_t(W) = \sum_{k=1}^n p_{t_k}(1 - p_{t_k})$$

where p_{t_k} is the relative frequency of data points in the corresponding region of class k . This function favors class splitting where high proportions of the data are assigned to one class (Bishop, 2007). A decision tree is able to fully split the data into pure leaves, that is leaves containing one class only, if an indefinitely depth of the tree is allowed. To prevent this overfitting, the tree gets pruned until a preset depth is reached or according to an error rate (for further details see Breiman et al. (1984)).

The disadvantage of decision trees is that they are highly depending on the input data, even small changes can have large effects on the resulting tree. To get rid of this high variance estimator effect, random forests have been introduced (Breiman, 2001). For this method randomly sampled subsets of the data (this procedure is called bagging (Breiman, 1996)) with randomly chosen subsets of features train a whole ensemble G of decision trees. Every data sample x_i is classified according to every decision tree of G and the final class assignment is based on a majority voting.

Random forest classification has a very good predictive accuracy, is easily human interpretable and performs inherently a feature selection (Murphy, 2012).

2.1.3 Active batch learning

Since the gathering of labeled data can be very time consuming, Borisov et al. (2011) proposed a method to query for labels of data points, which have a huge benefit for the underlying classifier. Here, a random forest classifier is trained on training data including labels and used to evaluate every data point from a test set without labels. The algorithm proposes data samples of the test set providing a good improvement for the classifier. Furthermore, it has been shown that label querying in a whole batch instead of querying single labels for the next training round is very practical by nature, but

also improves effectiveness (Borisov et al., 2011). For our application it is very convenient to ask the user for a batch of labels which can be efficiently assigned in our application (e.g. by keyboard shortcuts, compare Section 3.2).

To perform the so called Stochastic Query-By-Forest (Borisov et al., 2011), we define a random forest classifier G with with R number of trees $T_i, i = 1, \dots, R$. The predicted class probability of class c is denoted by $p_{ic}(x)$ of the i th tree of data sample x with a committee disagreement $q(x) = sd(p_{ic}(x))$ being the standard deviation of these class probabilities. Additionally, a discard fraction α is introduced, which serves as a tradeoff between randomness and high utility. The algorithm performs the following steps:

1. Build an ensemble G as discussed in Section 2.1.2.
2. For each unlabeled data sample x , compute the committee disagreement $q(x)$ and sort remaining n_u unlabeled instances with respect to $q(x)$
3. Sample the next batch from $x_1, \dots, x_{\alpha n_u}$ using the sample probabilities with the following utility scores $L(x) = \frac{q(x)-q_0}{q(x_1)-q_0}$, where $q_0 = q(x_{\alpha n_u})$. By normalizing with the sum of L we get the sample probabilities $p_s(x) = \frac{L(x)}{\sum_x L(x)}$
4. Gather the labels of the newly sampled batch, rebuild the model G and return to step 2.

In the original publication, this iteration is performed until no unlabeled instances are left (Borisov et al., 2011). In our case we can directly visualize if the classifier is trained enough to meet the quality requirements of the user and let the user decide how much effort he wants to put into labeling (compare Section 3.2).

2.2 Image processing

Methods, results and figures of this section are partly based on Schwarzfischer et al. (2011) and Schwarzfischer et al., in review.

The methods of this section are used to perform accurate fluorescence normalization and cell segmentation to allow reliable cell quantification. We focused on methods, which are easily parameterizable for the tool QTfy and intuitively to understand (compare Section 3.3). Furthermore, we apply machine-learning methods previously introduced in the context of image processing (Section 2.1).

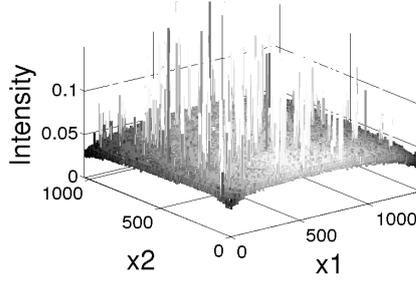


Figure 2.1: A typical fluorescence image with space coordinates (x_1, x_2) and fluorescence intensity plotted on the z-axis where peaks represent cellular signal. In long-term time-lapse microscopy, one has to deal with the following issues: (i) An inhomogeneous illumination due to the light source and camera lens, (ii) a non-zero background signal due to autofluorescence of the medium, and (iii) the effect of photo-bleaching. Figure adopted from Schwarzfischer et al. (2011).

2.2.1 Definitions

A raw fluorescence image $I(x, t) \in U$ in a set of gray scale intensity values $U = \{1, \dots, V\}$ with V being the highest possible intensity value, of w pixel in width and h pixel in height at time point t with space coordinates $x = (x_1, x_2) \in \mathbb{N}^{h \times w}$, can be decomposed into the following elements (compare Figure 2.1):

- A cellular signal $s(x, t)$, which changes over time.
- A homogeneous background signal $b(t)$ (e.g. autofluorescence of the culture medium), that decreases over time due to photo-bleaching.
- A coordinate-specific illumination function called *gain* $g(x)$ originating from the uneven illumination or the light source and the lens. The gain is defined as signal intensity per fluorescent molecule, which is assumed to scale linearly.
- A camera offset $o(x)$, which is constant over time.

Additional technical noise appears in equal measure at every position and time point. Since our normalization method uses robust fitting methods noise only has a marginal influence and is not discussed in the procedure presented here. Other means to infer the nature and intensity of the noise have to be applied separately.

We summarize all contributions in the following equation:

$$I(x, t) = s(x, t) \cdot g(x) + b(t) \cdot g(x) + o(x). \quad (2.1)$$

By rearranging the formula we get the cellular signal by

$$s(x, t) = \frac{I(x, t) - b(t) \cdot g(x) - o(x)}{g(x)}. \quad (2.2)$$

All contributions on the right hand side of Equation (2.2) are derived by the following approaches (Section 2.2.4).

Within the image, we define a subset c_i corresponding to the cell mask of cell i , $c_i \subset \{x\}_i$. The cell area a_i is defined as the cardinality of c_i , $a_i = |c_i|$.

2.2.2 Segmentation by thresholding

The process of segmentation can be seen as a labeling problem assigning one of two class labels (i.e. background or foreground) to every pixel of an image (Petrou and Petrou, 2010). After assignment, connected foreground pixels result in an object mask (i.e. cell mask c_i), which are used to measure cellular properties like intensity, size, shape or more sophisticated morphological features.

The simplest methods for gray level image segmentation is assigning an intensity threshold, dividing the image into the two classes. Every pixel above this threshold refers to foreground, everything below is regarded as background. To find the optimal cell mask in fluorescence images many methods have been proposed working on the intensity distribution of an image (Al-Kofahi et al., 2010, Arteta et al., 2012, Bergeest and Rohr, 2012, Dima et al., 2011, Sommer et al., 2011). A simple, popular method is Otsu's thresholding (Otsu, 1979). This method finds the optimal threshold τ_{opt} based on the assumption that the probability density function p_y of the intensities y is a mixture of two classes, which can be well separated by considering only the first and the second moments of the two classes. To find this optimum, we minimize the weighted sum of the intra-class variance

$$\sigma_W^2(\tau) = \theta(\tau)\sigma_1^2(\tau) + (1 - \theta(\tau))\sigma_2^2(\tau)$$

where $\sigma_i^2(\tau)$ is the variance of the two classes and $\theta(\tau)$ is the cumulative probability function of the threshold τ ,

$$\theta(\tau) = \sum_{y=1}^{\tau} p_y.$$

Otsu (1979) has shown that this is equal to maximizing the inter-class variance. The inter-class variance $\sigma_B^2(\tau)$ of a hypothesized threshold τ is defined as:

$$\sigma_B^2(\tau) = \frac{[\mu(\tau) - \mu\theta(\tau)]^2}{\theta(\tau)[1 - \theta(\tau)]}$$

where μ is the mean gray value of the image

$$\mu = (|p_y|^{-1}) \cdot \sum_{y=1}^{|p_y|} y$$

and $\mu(\tau)$ is the class mean of class 1,

$$\mu(\tau) = \sum_{y=1}^{\tau} x \cdot p_y$$

with p_y being the probability function of the intensity distribution (Petrou and Petrou, 2010). After exhaustively computing all possible thresholds τ , the value with the maximal inter-class variance is used as the optimal threshold τ_{opt} (compare Figure 2.2B and D).

The calculated threshold level can be intuitively interpreted. Decreasing the value will result in bigger objects (Figure 2.2C), increasing the value will shrink the object (Figure 2.2E). Cell masks c_i can thereby easily be manually adjusted. However, cell segmentation by thresholding probably fails if two equally bright cells are clumping together. Even an adjustment of the threshold might not resolve this issue.

2.2.3 Watershed segmentation

To resolve clumping objects, we consider the spatial proximity of pixels and adapted the watershed method. In general, this method performs two steps. First, a seed pixel for every object has to be found and second, neighboring pixels are assigned to each object by predefined rules until all pixels have an assignment (Petrou and Petrou, 2010). For the first part we introduce two different approaches, either working on the raw image intensities or based on a distance transformation of the image using the cell segmentation mask.

As seed points of the watershed method we want to utilize local maxima based on the raw image intensities and therefore introduce the principles of image dilation. For an image $I(x)$ with pixel coordinates $x = (x_1, x_2)$, we compute the dilation image $g(x)$ in dependence of a structuring element E (Petrou and Petrou, 2010). The shape and size of E defines the neighborhood of each pixel of $g(x)$. In our case we use a circular structure E_r with a radius r depending on the typical cell size as a free parameter. For every pixel of $I(x)$ we compute the largest value inside this structure and assign it to the dilation image at position x . This operation is denoted by $I(x) \oplus E_r$ and calculates the dilation image (compare Figure 2.3B):

$$g(x) = I(x) \oplus E_r$$

After subtracting the dilation from the raw image we are left with the local maxima of the image, which can be used as seed points (Figure 2.3B). For filtering we only regard local maxima, which have higher intensity values in the raw image than the calculated threshold.

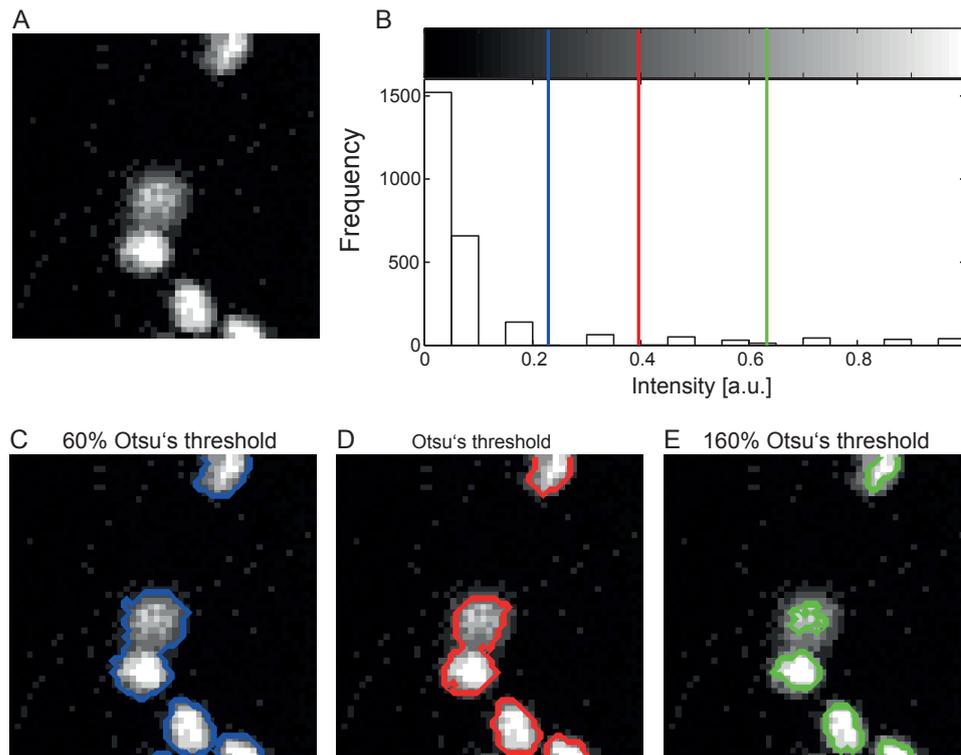


Figure 2.2: Thresholding allows to identify cell outlines. (A) Subimage of a raw fluorescence image $I(x)$ of blood cells. (B) The intensity histogram incorporates a high proportion of background pixel and a long tail originating from cellular signal. The red vertical line represents Otsu's threshold, the green line is 160% and blue 60% of this value. (C) Cell segmentation using 60% of Otsu's threshold results in big regions. (D) Otsu's thresholding results in good approximations of the cell outline, but cannot distinguish clumped cells. (E) A higher threshold can split clumping cells but results may not reflect accurate cell outlines.

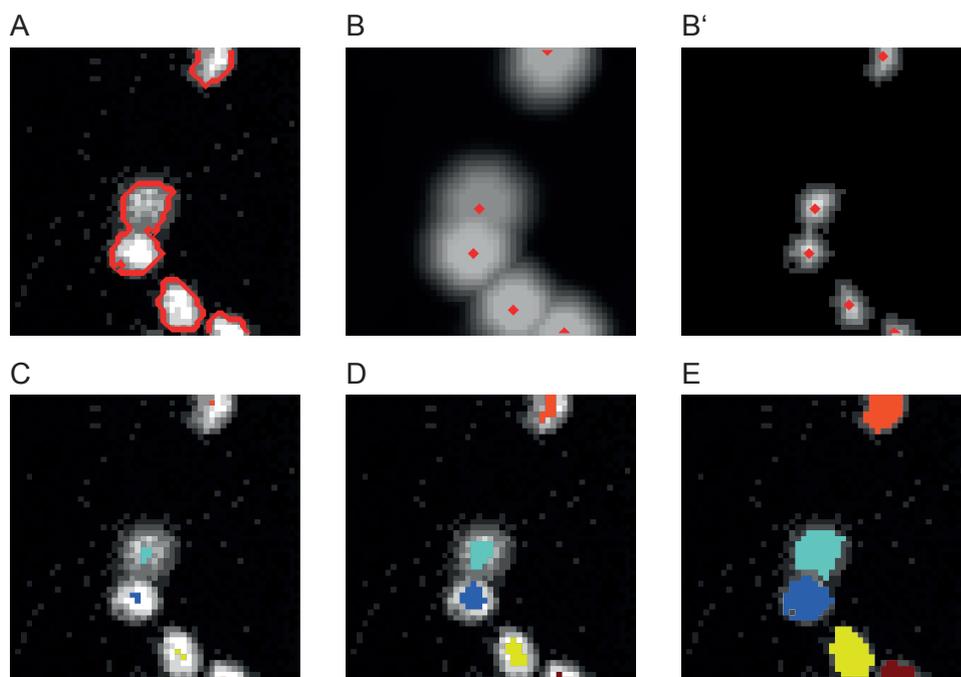


Figure 2.3: Watershed segmentation allows to distinguish clumped objects. (A) Subimage of an example raw fluorescence image $I(x)$ with an initial cell segmentation based on Otsu's threshold showing clumped objects. (B) The dilated image $g(x) = I(x) \oplus E_r$ with $r = 4$. Red dots highlight local maxima (i.e. seed points) derived by subtraction from $I(x)$. (B') The distance transformed image of the initial cell segmentation. Red dots highlight identified local maxima. (C) After initialization based on the seed point of (B') the first surrounding pixels are assigned to distinct objects (encoded by arbitrarily chosen colors) for all pixel greater than some τ . (D) After all pixel have been assigned, τ is decreased and unassigned pixel above this threshold are assigned to the objects. Pixels surrounded by two different classes are not assigned to a class but serve as watershed boundaries. (E) The final watershed segmentation overlaid with the initial cell segmentation correctly splits clumping object resulting in adequate cell outlines.

If the intensity profile does not serve as a good reference to find local maxima, alternatively, we can estimate local maxima based on a distance transformation based on the cell mask (Maurer and Raghavan, 2003). Every foreground pixel x gets replaced by its euclidean distance to the next background cell. After computing the dilation of the distance transformed image using the same E_r , the dilated image is subtracted from the distance transformation leading to the local maxima (i.e. seed points, Figure 2.3B').

To compute the watersheds we again apply two alternative methods, either based on real intensity values or the distance transformed image. In both scenarios the calculated seed points are used to initialize the watershed algorithm (Meyer, 1994), applying the following steps (compare Figure 2.3C to E):

1. Set the threshold τ to the maximum level of the raw image
2. All pixels of the image, which are higher or equal to this threshold, have not yet a label and are in adjacent neighborhood of the labeled pixels, are assigned to the label to which they are adjacent.
3. Repeat step 2, until no pixels can be assigned anymore.
4. Decrease the threshold τ and return to step 2, until no more unlabeled pixels exist.

Pixels having two or more differently labeled classes in their direct neighborhood are referred as the barriers (i.e. watersheds) (Petrou and Petrou, 2010). Since this algorithm assigns a label to every pixel within the image and does not detect the border of a cell, we only use labeled pixels within the cell mask identified by cell segmentation to divide clumped objects resulting in individual cell masks for every group (Figure 2.3E).

In our quantification tool QTFy (Section 3.3) the user can change the radius r of E_r as an intuitive parameter to tune the maxima suppression by dilation. A larger radius r induces a larger filter and results in less object splitting. By decreasing the radius r , local maxima will be more frequent leading to more seed points, and eventually more split objects.

2.2.4 Tiling method for background estimation

To estimate the background $B(x, t)$ of a given image $I(x, t)$, we developed the tiling method, which can be divided into the following two steps. First, we divide $I(x, t)$ into small overlapping parts of the image called *tiles* (compare Figure 2.4). We want to distinguish tiles containing cellular signal from tiles containing only background fluorescence, similarly to cell segmentation (Section 2.2.2). Here, we do not care about the correct cell outline, but rather focus on minimizing false positives, that is tiles containing cellular

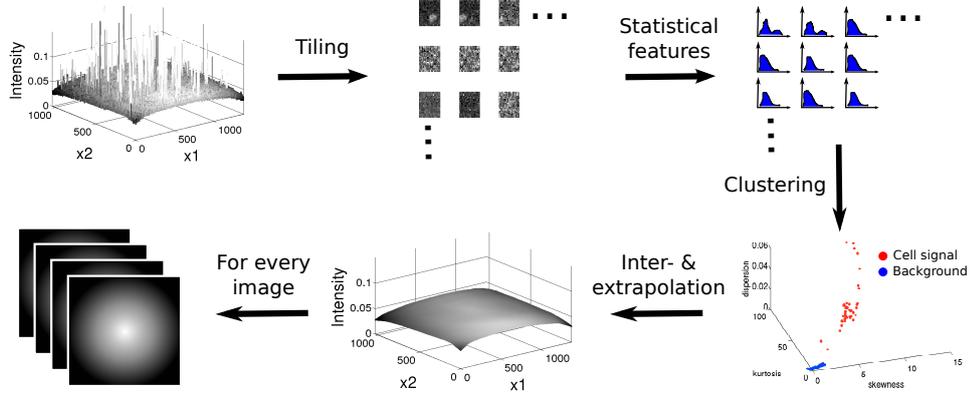


Figure 2.4: Time-dependent background estimation $\hat{B}(x, t)$. Each fluorescence image $I(x, t)$ is tiled into small overlapping sub-images. For each tile, several features are calculated. A machine-learning algorithm splits the tiles into two groups. The tiles containing only background are kept and their mean intensity is used to construct a grid which serves as a basis for a two dimensional inter- and extrapolation to estimate the full background $\hat{B}(x, t)$ of the image. This procedure is applied to every fluorescence image of a time-lapse movie.

signal regarded as background tiles resulting in underestimation of cellular signal after image normalization. On the downside we discard information of the background signal (i.e. true negatives), which typically does not severely impair the final result, unless the cell density becomes too dense. To distinguish the two classes of tiles, we compute a variety of features for each tile, which allows to apply machine-learning. We will discuss two alternative methods to perform this task. Either we manually label tiles of the image as only background or containing cellular signal allowing to perform supervised machine-learning (compare Section 2.1.2) or we apply unsupervised methods (Section 2.1.1) requiring a set of parameters. We evaluate these methods in Section 4.

After one of the machine-learning methods grouped the tiles, the mean intensity of each background tile is used to reconstruct an initial background grid. Finally, a two dimensional natural neighbor inter- and extrapolation is applied, which results in an estimation of $B(x, t)$, $\hat{B}(x, t)$. This procedure is applied to every image independently.

Image features

For fluorescence images it turned out, that statistical moments based on the intensity distribution of each tile serve as powerful features (Schwarzfischer et al., 2011). For each tile we calculate the first four central moments, namely

the mean μ , the variance σ^2 , the skewness γ_1 and the kurtosis γ_2 of the underlying intensity distribution. Additionally, we calculate the coefficient of variation, defined as the normalized standard deviation $\frac{\sigma}{\mu}$.

We compute textural features to also enable classification on bright field data (Haralick et al., 1973). To that end, we calculate a co-occurrence matrix P describing the frequency of pixel-pairs of specific gray levels $(i, j) \in 1, \dots, V$, in a set with V different gray values, in some defined distance d and some relative inclination Θ . For an image $I(x)$ with space coordinates $x = (x_1, x_2)$ the matrix is defined as

$$P(i, j) = \sum_{x_1=1}^h \sum_{x_2=1}^w (I(x) = i) \wedge (I(x') = j) \quad (2.3)$$

where the x-coordinate x' is the offset given by

$$x'_1 = x_1 + d' \cos(\Theta), \forall (d' \in \{1, \dots, \max(d)\}) \wedge (\Theta \in \{0, 2\pi\}),$$

and analogously

$$x'_2 = x_2 + d' \sin(\Theta), \forall (d' \in \{1, \dots, \max(d)\}) \wedge (\Theta \in \{0, 2\pi\})$$

(Nixon and Aguado, 2012).

Using a distance of one, $d = 1$, and angle of zero degree, $\Theta = 0^\circ$, will count the co-occurrence of pixel gray values adjacent to the right of another pixel (compare Figure 2.5). The resulting matrix P is a squared matrix whose dimensions equal V (Figure 2.5C and D). Based on this matrix we calculate 14 features proposed by Haralick et al. (1973), which describe the textural features of an image. For example, the first feature is defined as

$$f_1 = \sum_{i=1}^V \sum_{j=1}^V \left(\frac{P(i, j)}{R} \right)^2,$$

with R being a normalization constant. f_1 describes the angular second-moment measure indicating if a dominant gray tone transition occurs and is a measure for homogeneity within the image (compare Figure 2.5). Similarly, all other features are calculated based on the matrix and measure contrast, entropy, etc.

Furthermore, we calculate six further textural features, which have been proposed formerly (Tamura et al., 1978). This set of features describes the occurrence of defined, naturally observed pattern, namely coarseness, contrast, directionality, line-likeness, regularity and roughness. Based on human classification, it has been shown that these feature serve as a robust feature set for diverse kinds of images (Tamura et al., 1978).

Lastly we calculate Gabor filters, which are similar to wavelet functions applied to an image (Gabor, 1946). In general, the filter is a sine wave modulated by a Gaussian kernel function (Nixon and Aguado, 2012). It has been shown that the theory is very powerful by integrating spatial and spectral information in one feature.

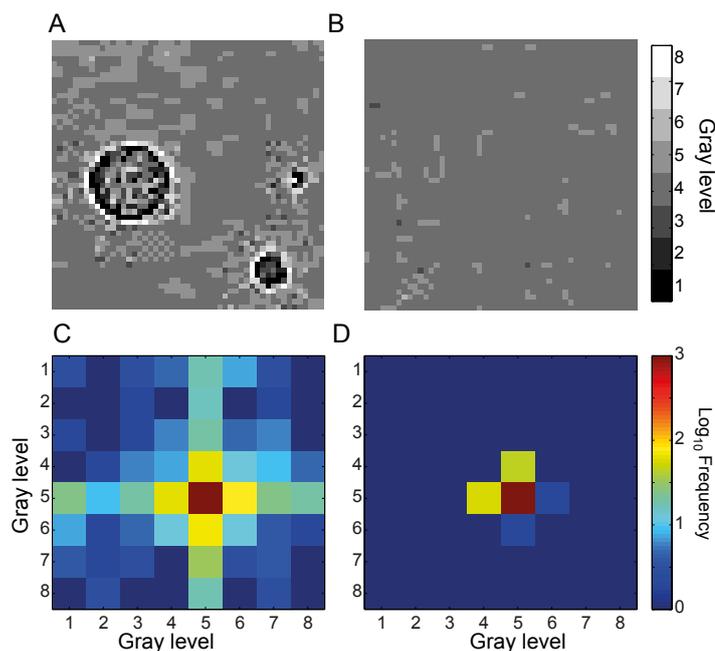


Figure 2.5: Haralick features provide a measure for texture properties. (A) A subimage of a bright field image containing HSCs has been discretized to 8 gray level values. (B) A subimage of a bright field image containing background pixels only has been discretized to 8 gray level values. (C) The co-occurrence matrix $P(i, j)$ has been computed for (A) using $d = 1$ and $\Theta = 0^\circ$ leading to a high homogeneity measure $f_1 = 0.5340$. (D) The co-occurrence matrix $P(i, j)$ of (B) shows less variable counts leading to a higher homogeneity measure $f_1 = 0.9118$ serving as a potential feature to distinguish background tiles (B) from tiles containing cells (A).

Unsupervised background estimation

We estimate the illuminated background signal including the offset, denoted as $B(x, t) = b(t) \cdot g(x) + o(x)$. By estimating $B(x, t)$ on each image separately we account for bleaching of the medium and image specific illumination. First, our method divides the image $I(x, t)$ into small overlapping sub-images, called *tiles* (see Figure 2.4). The distribution of tile intensities with cellular signal considerably differs from tiles without cellular signal and can be distinguished by the moments of the distribution. The appropriate selection of moments depends on the properties of the background image: for flat images, the first two moments (mean and variance) provide the most powerful features to discriminate background from cell signal containing tiles, for more complicated illumination shapes with varying gradients (as shown in Figure 2.4), additional higher moments will be more appropriate.

For the following steps, we represent each individual tile as a point in the multi-dimensional space of distribution moments. All background tiles have almost similar moments and will accumulate in a small dense volume (compare Figure 2.4). Therefore, we use a density-based clustering approach, called DBSCAN (Section 2.1.1), which returns two clusters (see Figure 2.4): One very dense cluster referring to background tiles, and a dispersed cluster from tiles containing cellular signal.

After the clustering, all background tiles are used for the second step of the tiling method (Section 2.2.4) to estimate all $\hat{B}(x, t)$.

Although this unsupervised machine-learning based method is robust for many applications (see Chapter 5 and Chapter 6), it heavily depends on the selection of features in dependence of the image characteristics.

Supervised background estimation

To facilitate an appropriate feature selection and an elaborate parameter tuning, we developed an alternative technique which allows the user to manually label image tiles from all over the movie. From the labeled data we collect 28 features including the statistical moments of the intensity distribution of each tile as well as textural features (Haralick et al., 1973, Tamura et al., 1978) and Gabor wavelet features (Gabor, 1946) (compare Section 2.2.4). These features seem to be essential for bright field images. By training a random forest classifier with 1000 trees we separate the feature matrix M according to their labels (compare Section 2.1.2). The method inherently performs a feature selection choosing the features to best distinguish the two classes based on the underlying image characteristics. Thereby, images of all kinds (fluorescence, bright field and phase-contrast) can be processed.

To reduce the hands-on time required to label tiles, we followed the work of Borisov et al. (2011) and implemented an active learning step (compare Section 2.1.3). We randomly sample 100 tiles of an image and calculate

all features and apply the trained classifier. According to the proposed algorithm we ask the user to label a batch of 10 elements which have a high impact on the prediction of the classifier. Again the selection of the tiles in the batch inherits a random process to avoid overfitting. This new labeled data is used to train and improve the random forest classifier.

Since false positives have a severe effect on the estimated background and subsequently on the cell quantification we lower the false positive rate by adjusting the prediction class rate. The well-trained classifier is applied to the whole set of images within a movie and the background classified tiles are used for the second step of the tiling method (compare Section 2.2.4) to estimate all $\hat{B}(x, t)$.

2.2.5 Fitting the time-independent gain

The bleaching of each pixel is position-dependent due to the uneven illumination of the experimental setup. A pixel in the center will experience a high irradiation and therefore bleach at a faster rate (Figure 2.6B). Plotting the absolute background pixel intensity against the mean background intensity for every time point reveals a linear correlation (see Figure 2.6C). Every pixel behaves linearly with respect to the overall mean intensity but with different slopes. A linear regression for every pixel x yields the slope, which represents the relative gain $g'(x) = c \cdot g(x)$ (see Figure 2.6D) while the ordinate-intercept represents the offset $o(x)$. The relative gain $g'(x)$ defined as intensity per mean background signal represents the real gain $g(x)$ multiplied by a factor c . Since we correlate every pixel with the mean background intensity instead of a known fluorescent molecule concentrations we can only determine this relative gain. The factor c , which is independent of time t and space x , cannot be further determined without additional experiments. However, since it is a constant factor, it does not alter the signal fold changes within the fluorescence images. To determine the factor c and finally infer protein abundances, a further calibration step must be applied. This can be done experimentally by comparing known concentrations in normalized time-lapse images or computationally with additional tracking methods as described in (Komorowski et al., 2010).

2.2.6 Normalization of fluorescence images

Depending on the application and the image characteristics, we propose the following normalization approaches:

1. Gain normalization

The final relative cellular signal can be derived by

$$s'(x, t) = \frac{I(x, t) - \hat{B}(x, t)}{g'(x)}, \quad (2.4)$$

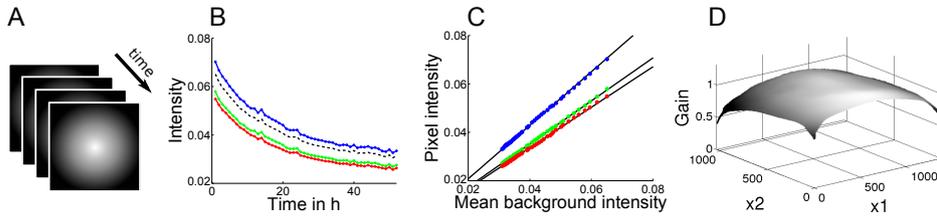


Figure 2.6: Fitting the time-independent gain $g(x)$. (A) Monitoring each pixel as well as the mean in every background image $\hat{B}(x, t)$ over time shows (B) the bleaching behavior of the medium. For demonstration we picked three pixels from the upper left corner (red), the left edge (green) and the center of the image (blue). The black dashed line indicates the mean background intensity. (C) A scatter plot of the mean vs. each pixel intensity reveals a linear dependence. From the ordinate intercept and the slope of the linear regression, we infer the offset $o(x)$ and the relative gain $g'(x)$ (D). Figure adopted from Schwarzfischer et al. (2011).

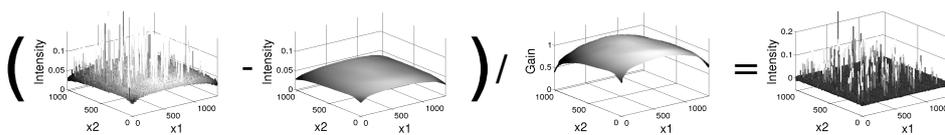


Figure 2.7: Final correction $(I(x, t) - B(x, t))/g'(x) = s'(x, t)$. The illuminated background and the offset, $B(x, t)$, derived by the tiling method described in Figure 2.4 are subtracted from the fluorescence image. The resulting image is then divided by the time-independent gain $g'(x)$, calculated from the linear regression as shown in Figure 2.6. The final image contains normalized cell signal and a homogeneous background around 0. Figure adopted from Schwarzfischer et al. (2011).

with

$$s'(x, t) = s(x, t) \cdot c^{-1}, \quad (2.5)$$

illustrated in Figure 2.7. After this normalization step all cellular signals are on a comparable intensity level. Ratios of the relative cellular signal are identical to ratios of real cell signal (compare Section 4).

Our proposed method estimating the gain function for fluorescence images only works with a series of images with varying overall background levels.

2. Background subtraction

A simple background subtraction $s'(x) = I(x) - \hat{B}(x)$ can be performed whenever no severe uneven illumination effect can be observed ($g'(x) \approx 1, \forall x \in I(x)$).

3. Background division

When the background $\hat{B}(x)$ is a good estimate for the underlying illumination ($\hat{B}(x) \approx g'(x)$), a division by the background corrects all background pixels to distribute around 1. A final subtraction of 1 leads to background pixels centering around zero $s'(x) = \frac{I(x)}{\hat{B}(x)} - 1$. Cellular signal gets corrected for uneven illumination and constant background signal, but cellular signals get not necessarily normalized to similar scales between images, if the overall background intensity varies.

4. Background division with varying background signal

To make normalized cellular signal comparable between different images, we scale the signal according to the mean signal of the background image $\bar{B}(x)$. In this scenario we assume, that the background level varies between images, but the cellular signal is not affected by this difference $s'(x) = \bar{B}(x) \cdot (\frac{I(x)}{\hat{B}(x)} - 1)$. This method is comparable to the gain normalization used for our time series data, but less robust.

5. Normalization of bright field images

To facilitate cell segmentation in bright field images, where cells typically appear as dark objects compared to surrounding background, we resolve uneven illumination by a simple division of the raw image by the estimated background $s'(x) = \frac{I(x)}{\hat{B}(x)}$. To eliminate bright halos around cells, we truncated intensities brighter than the background to 1. Thereby all cells became dark object on evenly illuminated bright background (Buggenthin et al., 2013).

2.2.7 Quantification

After we determined the cell mask c_i by cell segmentation (Section 2.2.2) and normalized the fluorescence image and derive $s'(x, t)$, we determine the absolute intensity $s_i(t)$ of a cell i by integrating all pixel intensity values within the boundaries of the segmentation:

$$s_i(t) = \sum_{x \in c_i} s'(x, t). \quad (2.6)$$

2.3 Correlation analysis

2.3.1 Pearson correlation coefficient

Given a p -dimensional vector of continuous random variables $X = (X_1, \dots, X_p)$, the covariance between two random variables X_i and X_j is defined as

$$\text{cov}(X_i, X_j) = \mathbb{E}[(X_i - \mathbb{E}[X_i]) \cdot (X_j - \mathbb{E}[X_j])]. \quad (2.7)$$

This single number describes the linear dependence of two variables and can be intuitively interpreted. By normalizing the covariance by the product of the standard deviations of both variables (Grimmett and Stirzaker, 2001), we get the Pearson correlation coefficient

$$\text{Corr}(X_i, X_j) = \rho_{X_i, X_j} = \frac{\text{cov}(X_i, X_j)}{\sigma_{X_i} \sigma_{X_j}} = \frac{\mathbb{E}[(X_i - \mathbb{E}[X_i]) \cdot (X_j - \mathbb{E}[X_j])]}{\sigma_{X_i} \sigma_{X_j}} \quad (2.8)$$

Obviously, the correlation is undefined if any $\sigma_{X_i} = 0$.

2.3.2 Partial correlation coefficient

Since a Pearson correlation coefficient cannot distinguish between direct and indirect effects, we use the concept of partial correlation, which calculates the correlation of two variables by conditioning on all others. To derive this measure, we first briefly introduce the theory of linear regression.

A general multiple linear regression model for the response variable Y is given by

$$Y(X) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \epsilon \quad (2.9)$$

with β_0 representing the intercept coefficient, β_i contains the regression coefficients and ϵ follows a normal distributed, $\epsilon \sim \mathcal{N}(\mu, \sigma^2)$, with mean $\mu = 0$ and variance σ^2 (Weisberg, 2005).

We calculate the coefficients by optimizing random variable realizations $x_i \in X_i$ with n samples:

$$(\hat{\beta}_0, \hat{\beta}_i) = \arg \min_{\beta_0, \beta_i} \sum_{k=1}^n \left(y_k - \beta_0 - \sum_{i=1}^p \beta_i \cdot x_{ik} \right)^2 \quad (2.10)$$

where $\hat{\beta}_0$ and $\hat{\beta}_i$ are the least square error estimates of β_0 and β_i , respectively, and x_{ik} represents the i th sample of the k th variable. The residuals are then defined as

$$\hat{\epsilon}_k^y = y_k - \hat{\beta}_0 - \sum_{i=1}^p \hat{\beta}_i \cdot x_{ik} \quad (2.11)$$

Regressing two variables x_m and x_n , with $m \neq n$ to the set $x_i, i \in \{1, \dots, p\} \setminus \{m, n\}$ yields the residual vectors $\hat{\epsilon}_k^m, \hat{\epsilon}_k^n$.

The partial correlation coefficient $r_{mn|x_i}$ of x_m and x_n given x_i is then defined as the correlation between the respective residuals $\hat{\epsilon}_k^m$ and $\hat{\epsilon}_k^n$ from these regressions:

$$r_{mn|x_i} = \text{Corr}(\hat{\epsilon}_k^m, \hat{\epsilon}_k^n). \quad (2.12)$$

We would like to note, that the slope of the linear regression β_1 of two variables X and Y can also be derived alternatively by normalizing the covariance of both variables by the variance of X (Toutenburg, 2013): the partial derivatives with respect to β_0 and β_1 of the argument of Equation 2.10 results in two equations with two unknown variables. After setting these equations to zero to find the minimum of the function, we are left with:

$$\sum_{k=1}^n (y_k - \hat{\beta}_0 - \hat{\beta}_1 x_k) = 0 \quad (2.13)$$

$$\sum_{k=1}^n (y_k - \hat{\beta}_0 - \hat{\beta}_1 x_k) x_k = 0 \quad (2.14)$$

After rearranging the equations we get

$$\hat{\beta}_1 = \frac{\sum_{k=1}^n x_k y_k - n \bar{x} \bar{y}}{\sum_{k=1}^n x_k^2 - n \bar{x}^2} = \frac{\text{cov}(X, Y)}{\sigma_X^2} \quad (2.15)$$

with $\bar{x} = \mathbb{E}[X]$ and by exploiting the displacement law (Toutenburg, 2013).

2.3.3 Coefficient of determination

To explain how much of the variance of one variable Y can be explained by the linear regression of variable X , the coefficient of determination is defined as

$$R^2 = \frac{\sum_k (\hat{y}_k - \bar{y})^2}{\sum_k (y_k - \bar{y})^2} = \frac{\sigma_{\hat{Y}}^2}{\sigma_Y^2} \quad (2.16)$$

with $\hat{y}_k = \hat{\beta}_0 + \hat{\beta}_1 \cdot x_k$ and $\sigma_{\hat{Y}}^2$ being the variance of the fitted regression points (Toutenburg, 2013), and furthermore

$$\sigma_{\hat{Y}}^2 = \sigma_Y^2 - \sigma_\epsilon^2. \quad (2.17)$$

By using Equation 2.15 we reformulate the variance of the residuals as

$$\begin{aligned} \sigma_\epsilon^2 &= \sum_{k=1}^n (y_k - \hat{\beta}_0 - \hat{\beta}_1 x_k)^2 & (2.18) \\ &= \sum_{k=1}^n ((y_k - \bar{y}) - \hat{\beta}_1 (x_k - \bar{x}))^2 \\ &= \sigma_Y^2 + \hat{\beta}_1^2 \sigma_X^2 - 2\hat{\beta}_1 \text{cov}(X, Y) \\ &= \sigma_Y^2 - \hat{\beta}_1^2 \sigma_X^2 \\ &= \sigma_Y^2 - \frac{\text{cov}(X, Y)^2}{\sigma_X^2} \end{aligned} \quad (2.19)$$

By combining Equations 2.17 and 2.19 we get

$$R^2 = \frac{\sigma_{\hat{Y}}^2}{\sigma_Y^2} = \frac{\sigma_Y^2 - \sigma_\epsilon^2}{\sigma_Y^2} = \frac{\text{cov}(X, Y)^2}{\sigma_X^2 \sigma_Y^2} = \rho_{X,Y}^2 \quad (2.20)$$

Therefore the squared Pearson correlation coefficient equals the R^2 value for a linear regression with an intercept (Toutenburg, 2013).

2.4 Maximum likelihood

An observed data set S can be explained by some defined statistical model depending on a set of parameters h (Murphy, 2012). A likelihood function $p(S|h)$ of a model evaluates h depending on the data S resulting in a measure how likely the data has been generated from the model with the given parameters. By the maximum likelihood approach we optimize the parameter set such that the model best explains the underlying data based on the likelihood function of the model and get an estimate for the best parameter set

$$\hat{h} = \arg \max_h p(S|h). \quad (2.21)$$

The maximized value of the likelihood function is derived by evaluating the model with the best parameter set

$$\hat{L} = p(S|\hat{h}). \quad (2.22)$$

2.5 Model comparison

2.5.1 Bayesian information criterion

Throughout this work, we proposed several models to explain our experimental observations. To evaluate, which model best describes the underlying data we perform model selection based on Bayesian information criterion (BIC) (Schwarz, 1978). The BIC validates the fitted model in comparison to the underlying data by the maximized value of the likelihood function \hat{L} and additionally penalizes the degrees of freedom D as well as the number of data points N :

$$\text{BIC} = -2\log(\hat{L}) + D\log(N) \quad (2.23)$$

Thereby each model gets a score how well it fits the data minus how complex the model is, allowing to select the better model by choosing the model with the lower BIC (Murphy, 2012).

Chapter 3

Software tools

In the following chapter, we first define the general requirements for our experimental setup and discuss published software in this context. Next, we present all software tools developed by our own, which are applied in Chapter 4 to Chapter 7. We present a tool to estimate the background signal in fluorescence images (Section 3.2), which includes all previously described methods (compare Section 2.2.4). Next, we introduce a toolbox, which includes tracking information, image normalization and cell segmentation to provide quantification of whole genealogies, called QTFy (Section 3.3). Furthermore, we present a tool which is capable to quantify fluorescent intensities of cells in single images (Section 3.4). It allows to reliably compare different image acquisition setups or quantify cells without tracking information in a high-throughput manner. Lastly, we developed a GUI to manually track individual cells including cell annotations (Section 3.5).

3.1 Existing tracking and segmentation software

The following section is partly based on Schwarzfischer et al., in review.

In the biological context of this thesis, we were interested in (i) following individual cells over long periods of time up to weeks (ii) using an adequate magnification to reach single-cell resolution resulting in multiple fields of view, (iii) incorporating adequate time-lapse intervals to follow highly motile cells during differentiation and (iv) simultaneously recording several fluorescence properties. To this end, we first have to apply a tracking tool to generate continuous cell trajectories, and second we need to accurately and reliably quantify the cellular fluorescence signal based on cell segmentation methods. Here, we use combination of fluorescent dyes with non-overlapping emission spectra. Finally, we get time-resolved measurements of single-cells allowing to investigate molecular properties.

3.1.1 Single-cell tracking

First of all, a cell tracking software should incorporate all the requirements listed above and furthermore should provide a intuitive and efficient user interface. The most crucial requirement of cell tracking software is that it must provide a high accuracy and robustness. Even one single ambiguous event can render the whole pedigree describing the cellular relationships useless (Meijering et al., 2009). Since this requirement can not (yet) be guaranteed a robust error detection and correction has to be offered.

A variety of tools already exist which allow to track single-cells in a fully-automatic manner, e.g. TimeLapseAnalyzer (Huth et al., 2011, 2010), TLM-Tracker (Klein et al., 2012), Celltracker (Scherf et al., 2012a). Since auto-tracking will possibly never be 100% correct, these must provide a robust and efficient error detection and potential manual correction. Furthermore, all of these methods will certainly have some bias in their tracking results. In order to facilitate their quality a good signal to noise ratio between cells and background is preferred, ideally achieved by strong fluorescence signals. However, even after the laborious process of generating transgenic cell lines or mouse models which express an appropriate fluorescence marker, the frequent fluorescence imaging challenges the cell's health due to phototoxicity. Trying to use the non-phototoxic bright field images will again fail for current available tools due to the low contrast and thus cell recognition.

All these tools cannot guarantee a perfectly correct tracking result and therefore a semi-automatic approach will be more appropriate by first trying to automatically track all cells and second asking for user input in ambiguous situations. Since in this scenario hours of work of human users cannot be compensated by computational power, the tool has to be as efficient as possible to minimize the human effort. A variety of tools already offer semi-automatic single-cell tracking, e.g. Trackmate (Jaqaman et al., 2008) or Tactics (Shimoni et al., 2013). These tools still potentially have bias in their tracking algorithms and cannot exclude human mistakes. Furthermore, manual cell tracking might also be erroneous but in contrast to current automated cell tracking methods these errors can be detected and corrected by the user. Unfortunately, all available tools do not fully achieve all required points discussed above. Especially, they lack efficient implementation of important steps such as intelligent storing of intermediate results and therefore do not come into consideration to be used on a daily basis, but rather thwart productivity.

3.1.2 Timm's Tracking Tool

In the Research Unit Stem Cell Dynamics the tracking software Timm's Tracking Tool (TTT) has been developed and published in a former version

(Eilken et al., 2009, Rieger et al., 2009). Currently, the tool is being further developed fulfilling all requirements of a typical experiment. It focuses on providing an efficient user interface for manual, computer-assisted cell tracking. It enables the user to load voluminous imaging data (Figure 3.1B) including multiple fields of view (Figure 3.1A) and to display several fluorescence channels simultaneously (Figure 3.1C). In addition, the emerging pedigrees is shown and the user can annotate tracked cells by various cell properties (Figure 3.1D and E).

3.1.3 Segmentation and quantification

Besides accurate single-cell tracking, the next crucial task is to accurately estimate the outline of a cell (i.e. cell segmentation) to measure subcellular properties like total fluorescence signal. A most accurate outline of a cell is of interest, especially, when quantifying total fluorescence intensity or shape features. For this process another own squadron of tools exists to provide cell segmentation for all kinds of scenarios. As already stated above, these tools also have to fulfill all requirements and have to provide efficient user interfaces to correct for potential false segmentation. Furthermore, imaging will be performed with conditions optimized for survival and fitness of the cells resulting in low contrast and thereby challenging segmentation algorithms. This is further complicated since minute differences have to be quantified such as endogenous protein levels. When quantifying cell types where every cell is costly (e.g. primary stem cells) it is of great interest to quantify the last bit of every experiment. Lastly, quantified fluorescence levels of individual cells should be highly comparable to well-established methods like flow cytometry.

Cell segmentation basically can be performed by either parametric methods (Bergeest and Rohr, 2012, Piersma et al., 2013) or machine-learning based methods (Arteta et al., 2012, Sommer et al., 2011). No matter how well their parameters are chosen or their methods have been trained there will be some degree of failure in segmentation. The difficulties in most setups are the changing cell characteristics due to cell-cycle, cell differentiation or even changing imaging conditions throughout a movie. Recently, we have shown that the popular tool CellProfiler (Carpenter et al., 2006) lacks of robustness and accuracy when cells should be segmented in bright field images on a large scale (≈ 315.000 images) with a standard pipeline (Buggenthin et al., 2013). Other existing tools like TLM-Quant (Piersma et al., 2013) try to overcome all requirements but completely fail in terms of usability. In particular, this tool is based on a series of source-code scripts which requires some degree of computational and programming experience which, on the one hand, is advantageous for developers, but on the other hand it excludes a majority of potential users. Further tools like Cell-ID (Gordon et al., 2007) or FARSIGHT (Bjornsson et al., 2008) can quantify a

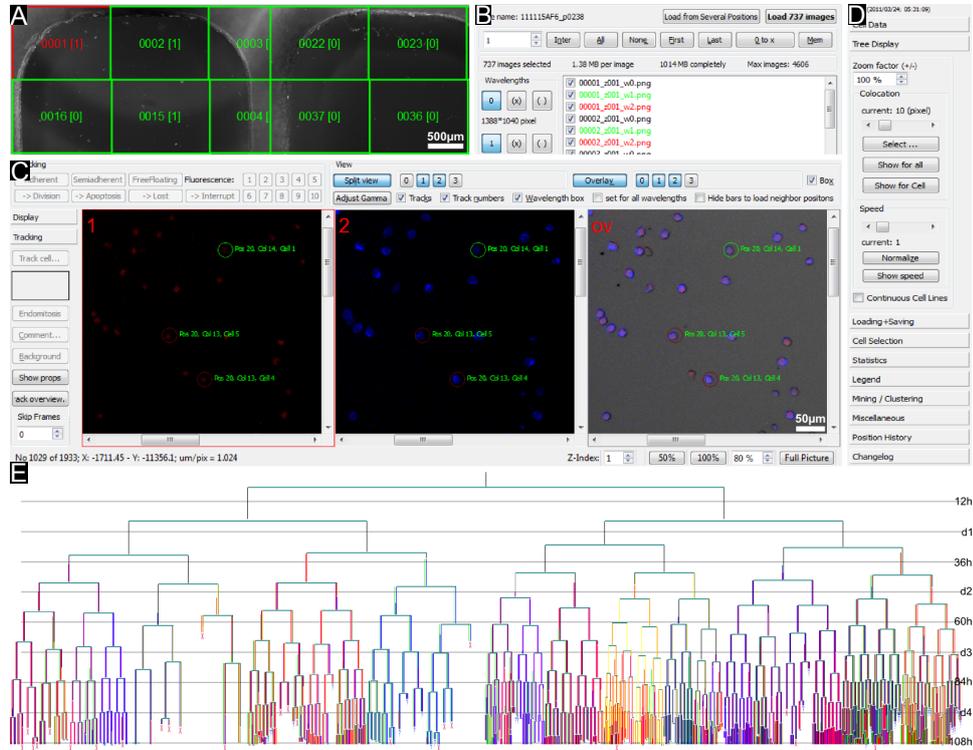


Figure 3.1: TTT graphical user interface and data visualization developed by the Research Unit Stem Cell Dynamics. (A) Part of experiment overview showing different overlapping fields of view (=positions), each with a resolution of 1388x1040 pixels. The selected position is highlighted in red and the number of tracked colonies starting in each position is shown in square brackets. (B) Selection of images for loading by time or imaging channel. Images of fluorescence channels are highlighted with colors. (C) Display of bright field image data (OV) including colored overlay of signals from two different fluorescence channels (1,2) for inspection and cell tracking. Circles indicate already tracked cells. Any set of imaging channels can be displayed in an arbitrary layout, and display settings such as color transformations can be set individually for each channel. (D) Controls for adjustment of image display, pedigree editing and statistics. (E) Example cellular pedigree. Colors qualitatively visualize cell adherence status and fluorescence in different channels for each time point. Cell death and loss of cell events are indicated by 'X' and '?' symbols respectively. A mESC pedigree is shown, but cell adherence and fluorescence properties were arbitrarily changed to illustrate multi-channel display possibilities of TTT. Figure taken from Schwarzfischer et al., in review.

variety of morphological cell features on a semi-automatic basis providing a nice graphical user interface, but it cannot quantify fluorescence signals. Astonishingly, up to our knowledge none of the tools incorporates an adequate normalization for fluorescence signals nor have they be evaluated against well-established technologies like flow cytometry. In current literature, fluorescence quantification of single-cells has been performed by custom-written unpublished software whose accuracy remains questionable due to the lack of proper evaluation (Bjornsson et al., 2008, Elowitz et al., 2002, Piersma et al., 2013, Taniguchi et al., 2010).

3.1.4 Conclusion

Hence, we claim that to date the required accuracy in cell tracking and quantification can only be achieved by a combination of automated methods followed by manual supervision and correction of their results by the user (see Section 3.3). For cell tracking we found TTT to achieve our requirements by a manual tracking approach. In our experience, we found that none of the existing tools are efficient, reliable and user-friendly enough to meet our requirements for long-term single-cell quantification experiments.

3.2 Background estimation tools

In order to accurately quantify fluorescence intensities, images have to be corrected for background signal (compare Section 2.2.1 and Schwarzfischer et al. (2011)). Since the estimation of one background image based on the tiling method (Section 2.2.4) takes more than 10 seconds on a standard computer we precalculate every background with the following tools and store them as 16bit pngs for later quantification of cellular intensities. Here, we present two tools, which perform all necessary image processing steps to accurately estimate the underlying background signal based on either unsupervised or supervised machine-learning methods (compare Section 2.1).

3.2.1 Unsupervised background estimation

We developed a tool based on unsupervised machine-learning tiling method (Section 2.2.4) automatically performing all computational processes on a distributed computing system. Although parameters of these methods have been tweaked to be applicable on a broad range of different images the performance is highly depended on the image characteristics. We optimized the feature set and its parameters to perform best on fluorescence images, which we are using on a daily basis (compare Chapter 5 and Chapter 6).

We provide a small interface (Figure 3.2) allowing a biologist to select certain wavelengths, which should be corrected for background. The tool submits queue jobs in our SUN grid engine processing all selected images in a highly parallelized manner and notifies the user by email when done.

3.2.2 Supervised background estimation

To widen the applicability and usability of our tiling method to images with different characteristics (e.g. bright field images), we developed a tool, which also incorporates the tiling method, but based on supervised machine-learning methods (compare Section 2.2.4). To successfully estimate background images for a whole movie, the procedure is subdivided into the following steps. To get a first set of labels, the tool asks to select tiles containing cellular signal or containing background only (Figure 3.3). In this mode ('Select/Train') 28 features for every selected tile are computed (see Section 2.2.4) and a random forest classifier is trained (compare Section 2.1.2). Next, the user continues to the 'Active Learning' mode. In this mode random tiles from the current image will be sampled and classified by the random forest classifier and the tool queries for further labels according to the active learning routine (compare Section 2.1.3). This method ensures, that new labels are acquired with a low cost for the user (i.e. as less labels as possible) with a high benefit for the classifier. After additional tiles have been labeled the third mode ('Results') will estimate the final background

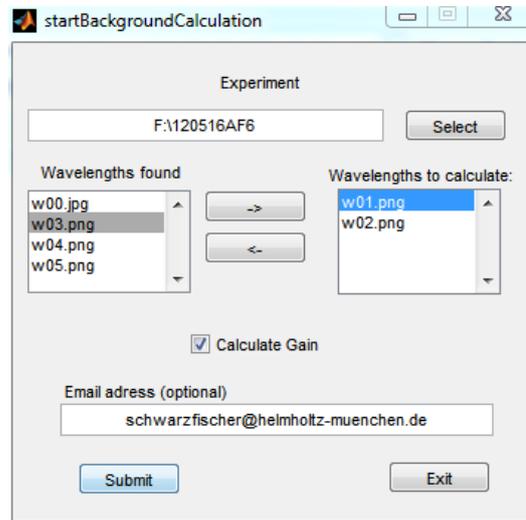


Figure 3.2: The background estimation tool lists all available wavelengths for a given experiment. After the user made a selection, for which wavelengths background images should be estimated, the computational jobs will be distributed into a SUN grid engine. If the user specified an email address, he will get a notice whenever the calculations are done.

signal of the whole image by classifying all tiles of the image. Finally, the program allows to store the trained classifier and to run the background estimation of the whole movie on the grid engine or locally.

At every time, the user has the possibility to check whether the classifier is already sufficiently trained enough or if further labeling is needed to improve the background estimation. Thereby the user can control the quality of the resulting background images by himself.

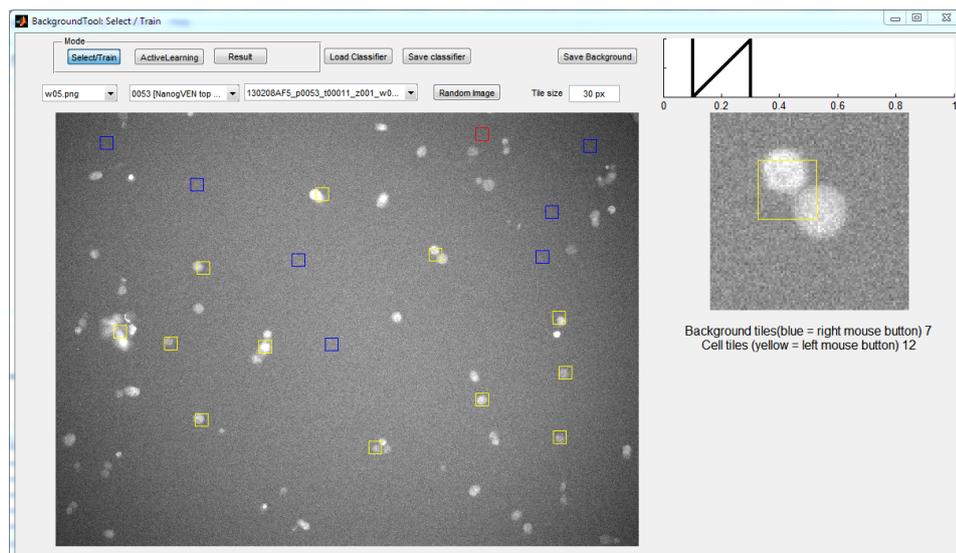


Figure 3.3: Background estimation tool incorporating supervised machine-learning methods. A selected image will be displayed with a close-up on the upper right showing the last selected tile. The graph on the top right can be used to adjust optical contrast for every wavelength. The GUI provides three modes on the upper left:

Select/Train: The user is able to select tiles and give them labels by using the left or right mouse button. Yellow tiles correspond to tiles containing cellular signal. Blue tiles contain background only. After the initial tiles have been labeled, textural and statistical features are calculated to train a random forest classification.

ActiveLearning: Using active batch learning allows to create an easy user dialogue to get new tile labels of randomly sampled tiles at a low cost of user input and a great benefit for the classification.

Result: Finally, the GUI allows to estimate all background images based on the trained classifier locally or using distributed computing.

3.3 Quantification of tracked fluorescence (QTFy)

The tool QTFy, methods, results and figures in this section are partly based on Schwarzfischer et al., in review, Hoppe et al., in review and Filipczyk et al., in review.

To quantify cellular fluorescence of tracked cells in long-term time-lapse microscopy experiments, we developed the tool QTFy (Quantify Tracked Fluorescence, [kju: ti: fai]). The tool incorporates the background normalization (compare Section 2.2.6) using precalculated background images (see Section 3.2) as well as segmentation algorithms (See section 2.2.2) to quantify the intensity of individual cells (Section 2.2.7). Furthermore, it allows to visualize single-cell time courses with different cell-cycle normalizations and provides an efficient GUI to inspect every individual measurement.

3.3.1 QTFy workflow

In a first step, single-cell tracking needs to be performed, by e.g. using Timm’s Tracking Tool (TTT) as described (Figure 3.4A, Section 3.1.2).

After tracking, we apply QTFy to quantify all tracked cells (Figure 3.5). To handle the vast amount of data we chose a semi-automatic approach, which allows to robustly quantify cells on a large scale. First, QTFy automatically segments all tracked cells in the fluorescence channel using a predefined set of parameters (Figure 3.4B). Image normalization improves auto-segmentation (Figure 3.4C), but mis-segmentation and artifacts may still compromise the fluorescence quantification. To remedy this, QTFy’s GUI allows the user to efficiently inspect each single-cell trace and, if necessary, to tune segmentation parameters or even draw cell boundaries by free-hand for every time point individually (Figure 3.4D). Furthermore, QTFy supports segmentation based on one fluorescence channel (e.g. containing a nuclear membrane marker) and quantification of any other channels (e.g. containing the labeled proteins of interest). QTFy is written in MATLAB using the Image Processing Toolbox and the Statistics Toolbox with about 15,600 lines of code and 10,200 lines of comments.

3.3.2 Quantifying tracked cells

In a first step, the user selects an experiment, one or several trees containing tracked cells and which channel(s) should be used for segmentation and which channel(s) should be used for quantification (Figure 3.6A).

Next, the user specifies the segmentation algorithm and can choose between thresholding methods and a simple ellipse method which would place an ellipse with a given radius around the track points (Figure 3.6B). For every algorithm a set of robust default parameters can be adjusted. The expected minimal and maximal area in pixel should be tuned to meet the

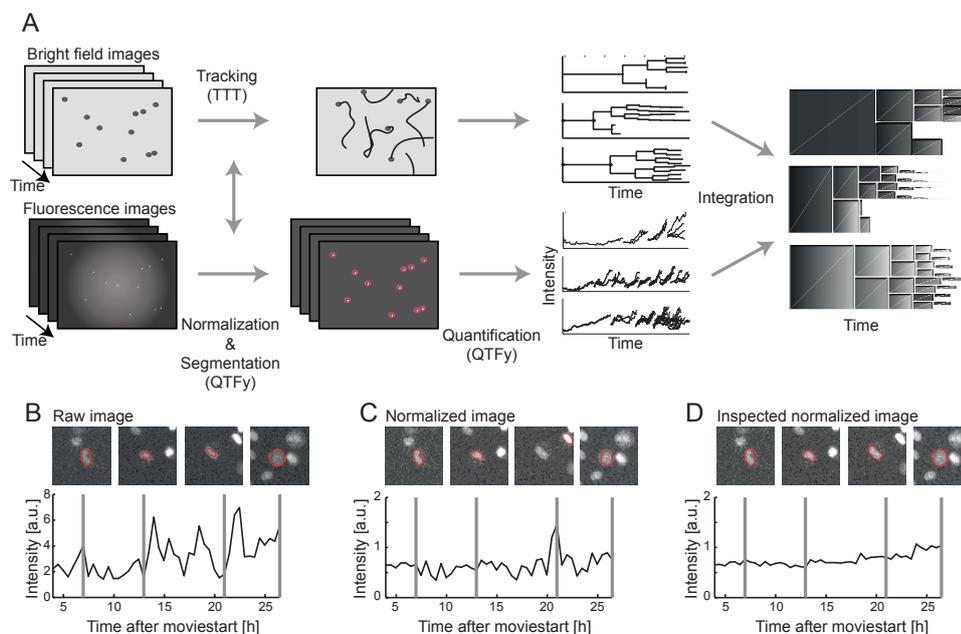


Figure 3.4: QTFy allows to conveniently quantify and visualize cellular fluorescence of tracked cells. (A) QTFy uses the tracking information of individual cells (e.g. derived by using TTT) and performs fluorescence image normalization (Section 2.2.6) and cell segmentation (Section 2.2.2). Combined with the tracked tree and single-cell quantification time-courses emerge visualizing the cellular expression over time. To get a better overview of the data QTFy allows to integrate all information to achieve intuitive visualization using heat trees (Section 3.3.5). (B) to (D) Large-scale cell quantification demands for efficient segmentation and computer assisted error correction. (B) Cell quantification based on unnormalized data will fail because of frequent miss-segmentation. (C) After normalization, an automatic segmentation approach might still fail (e.g. segmentation of the wrong object at 22h) but general trends will be visible. (D) After manual inspection of the data, accurate protein dynamics are observed. Figure adopted from Schwarzfischer et al., in review.



Figure 3.5: The main window of QTFy shows all opened trees and provides the core functionality to quantify newly tracked trees, to save or open quantified trees and to call further visualization and inspection tools.

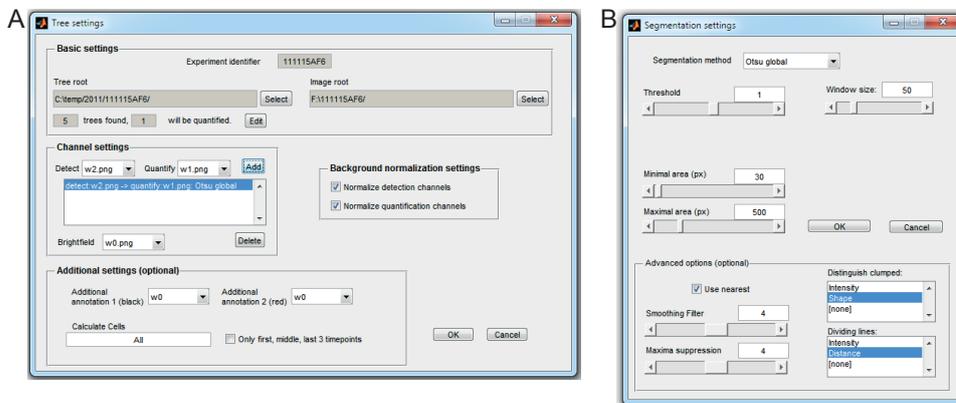


Figure 3.6: QTFy provides an efficient GUI to quantify newly tracked cells. (A) To quantify newly tracked trees of an experiment basic settings have to be entered. (B) Cell segmentation settings are needed to perform an initial automatic quantification (compare Section 2.2.2 and 2.2.3). QTFy already suggests a set of robust default parameters to reduce user input.

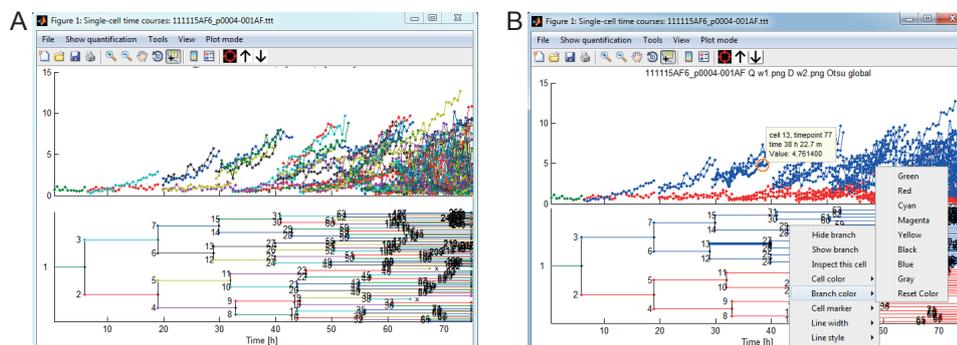


Figure 3.7: The time course viewer allows to display and customize visualization of quantified trajectories on the top with the corresponding pedigree at the bottom. (A) An exemplary tracked tree has been quantified using QTFy. (B) An additional context menu allows to change the appearance of every trajectory allowing to highlight specific cells of interest. Here, all cells without fluorescence expression have been colored in red, cells switching on fluorescence are represented in blue.

experiment settings. The segmentation based on Otsu’s thresholding (compare Section 2.2.2) and watershed segmentation (compare Section 2.2.3) is adapted from an implementation of CellProfiler (Carpenter et al., 2006).

After setting all initial parameters QTFy automatically calculates all segmentations and runs a simple optimizer to fit the resulting cell area in the preset range. Afterward, all quantifications based on the segmentations are assessed considering the correct fluorescence normalization (Section 2.2.6).

Once a tree has been quantified QTFy goes hand in hand with TTT. Whenever an tracking update in TTT has been done (i.e. further tracking, re-tracking of miss-tracked cells, deleting time points. etc.), QTFy can check for updates on demand or it will automatically check for updated whenever a saved tree is loaded and will automatically apply the changes if needed.

3.3.3 Visualization of time courses

To inspect the quantification results, QTFy allows a variety of visualizations. The time course viewer shows all quantified trajectories of individual cells (Figure 3.7A). As an additional feature the color, marker style and line thickness of specific cells or whole branches can be changed (see Figure 3.7B).

QTFy allows in this time course viewer to adjust x-axis, y-axis, all labels over intuitive context menus to create publishable figures. Besides showing the absolute quantified intensity (Figure 3.8A) a simple moving average filter with a window size of 5 time points can be applied and visualized (Figure

3.8B). Furthermore, the GUI offers to calculate and to directly visualize the following different representations of the data:

- Linear regression

To hide noisy data, every trajectory can be fitted by a linear regression (compare Section 2.3.1) resulting in straight lines (Figure 3.8D).

- Cell area and mean intensity

In the 2D projection of a cell into the image we observe the cell area as number of pixels determined by the cell segmentation step (Figure 3.8C). If this measure is a good proxy for the growing cell volume the area also increases its initial value by two-fold. Dividing the absolute intensity for every measurement by the obtained area results in a measure proportional to concentration (Figure 3.8E).

- Concentration

Single-cell trajectories are normalized by an abstracted nuclear volume V ranging from 1 to 2 over cell-cycle (compare Schwarzfischer (2009)). Assuming a spherical nucleus volume with a linear increase of area leads to the following volume growth over cell-cycle progress

$$t \in [0; 1] : V(t) = (1 + (2^{\frac{2}{3}} - 1)t)^{\frac{3}{2}} \quad (3.1)$$

Here, we assume that every cells has the same volume V_0 at cell birth throughout the whole experiment. The resulting volume of a cell is a slightly bend line which grows two-fold from its initial value over the whole cell cycle.

When we normalize the absolute intensity by this hypothetical volume we get an intensity concentration per standardized cell birth volume (Figure 3.8F).

- Net production

By looking at the first derivative differences in cell cycle can be easily divided into subgroups (Schwarzfischer, 2009). Cells below zero have a decrease in the net production which includes protein production and protein decay, cells above increase their net production. Using this method, experiments with different conditions, like varying light sources, become comparable since the absolute amount of intensity is canceled out (Figure 3.8G).

- Doubling hypothesis

If the dynamics can be estimated by a straight line doubling its initial intensity during lifetime, we highlight differences in cell cycle with

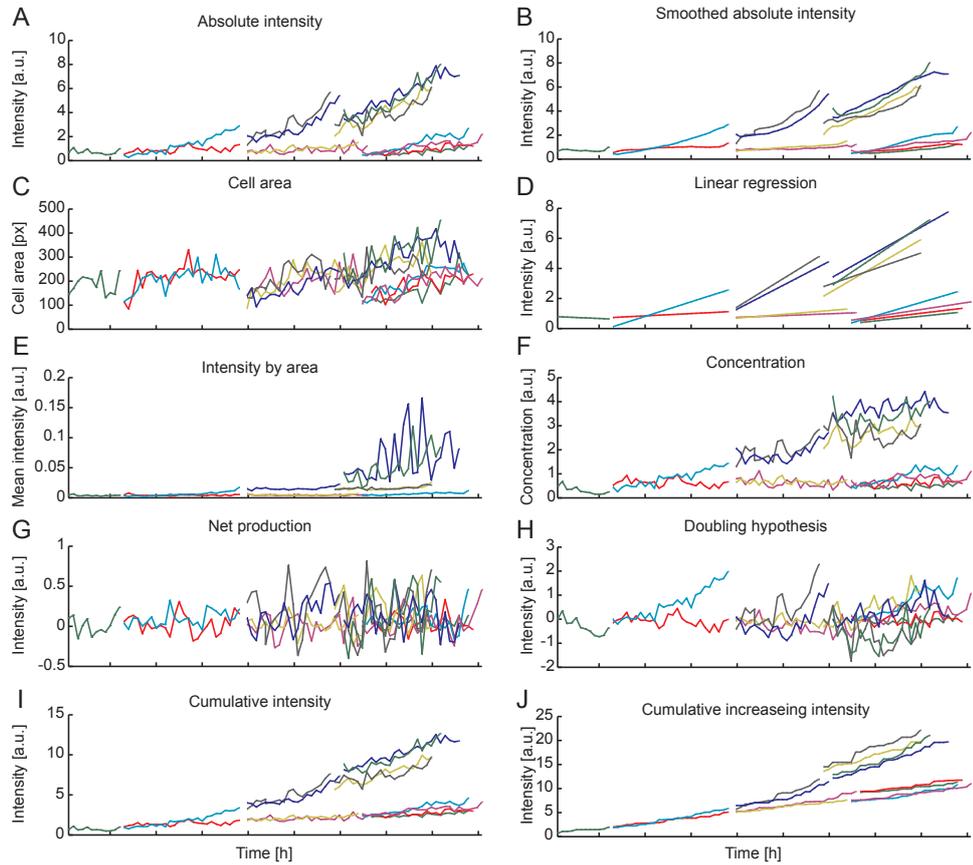


Figure 3.8: An overview of all provided visualization methods of the QTFy time course viewer. The different methods visualize every trajectory based on (A) total fluorescence intensity, (B) a smoothed representation of the absolute intensity using a moving average with a window size of 5 time points, (C) the area of a cell in pixel, (D) a fitted linear regression, (E) total intensity divided by nuclear area, (F) a concentration estimation based on a hypothetical standard volume doubling over each cell-cycle, (G) the first derivative of (A), (H) the total intensity subtracted by a straight line doubling a cell's initial intensity over the cell-cycle, (I) a cumulative representation over generations adding half of the intensity of each respective mother cell or (J) a cumulative representation ignoring decreasing consecutive time points.

this method. Every cell trajectory gets subtracted by a straight line starting at the initial value increasing to the double amount of this value. Cells maintaining concentration start with zero intensity and stay at this level; cells increase their intensity if overproduction can be observed and decrease below zero if more protein decays than protein gets produced (Schwarzfischer, 2009) (Figure 3.8H). We evaluate this hypothesis in a steady-state population in Section 6.5.

- Cumulative production

In order to emphasize certain branches or complete subtrees which have a different overall protein production we add to every cell half of the mother's last measurement. As an optional argument we allow to ignore decreasing intensity between consecutive time points, which further accentuates the overall protein production on a branch basis (Figure 3.8I and J).

The trajectory smoothing by a moving average filter can optionally be applied to every representation.

3.3.4 Tree browsing

A full lineage tree over many generations results in many cells living in parallel and visualization of many time traces in a well-arranged manner becomes a challenging task (compare Figure 3.7). For a user it is hard to judge if segmentation errors may have occurred by looking at all time traces at once. Even general trends will be completely hidden in the mass. For these reasons we developed several different views on the data and introduced the following mechanisms.

The most intuitive way enables the user to show just certain parts of the data, by selecting and deselecting cells or whole branches. Every cell number or branch point of the tree can be interactively clicked and will display or hide all descendents. We also allow the user to enter a sequence of cell numbers which should be displayed (Figure 3.5). Combined with a zoom function browsing through huge trees becomes convenient. All visualization updates are highly optimized to operate in real time allowing an efficient everyday usability.

3.3.5 Heat tree

As the time traces become unreadable to a human eye in the later generations we developed a different perspective of the data. We keep the lineage tree and plot time on the x axis. Additionally, we plot on top of the lines the intensity by color coding. We stretch every generation to take as much space as possible in the y direction which finally ends up in *Heat trees* (Figure 3.9). In QTFY we present a GUI to further modify a heat tree in order to

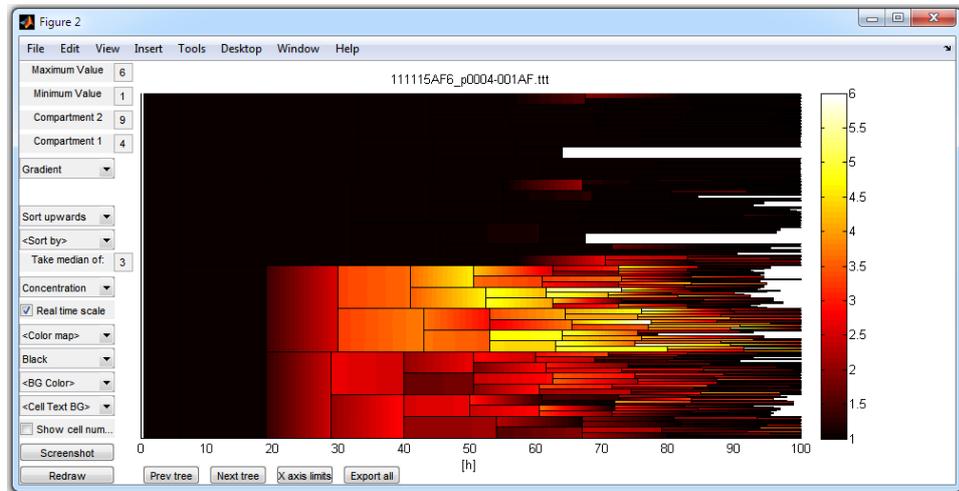


Figure 3.9: The heat tree GUI of QTFy provides numerous possibilities to customize the visualization of the quantified data. Here, a sorting is chosen, such that sisters with a higher intensity at the beginning of the cell-cycle are at the bottom. This visualization allows to get an immediate overview of the overall tree behavior.

prepare publishable figures. The user can choose between a generation view or real time scale. All the different representation methods mentioned above (Section 3.3.3) can be used. Additional smoothing of the data is resulting in a smooth gradient for every cell.

Furthermore, we can sort the tree in different ways. Since the numbering of cells is arbitrarily chosen by the tracking procedure we are allowed to flip sisters at every cell division. This allows us to sort cells to appear at the top or bottom depending on their inherent states. We allow the user to sort sister cells according to their initial or their final intensity values. This method allows to get an immediate overview of the overall tree behavior (compare Figure 3.9).

3.3.6 Cell inspector

With the automatic cell segmentation we already get reasonable time courses (Figure 3.4C) but one of the main feature of QTFy is the possibility to manually inspect every single measurements. For every time point the user can optimize individual parameters to get the most accurate cell segmentation and thereby the most reliable cell quantification. The user can click at every time point in the time course viewer, which will highlight the cell itself in all channels and show a tooltip containing detailed information (compare Figure 3.7B). By opening the Cell Inspector on this particular time point a GUI is presented showing the result of the initial segmentation method with

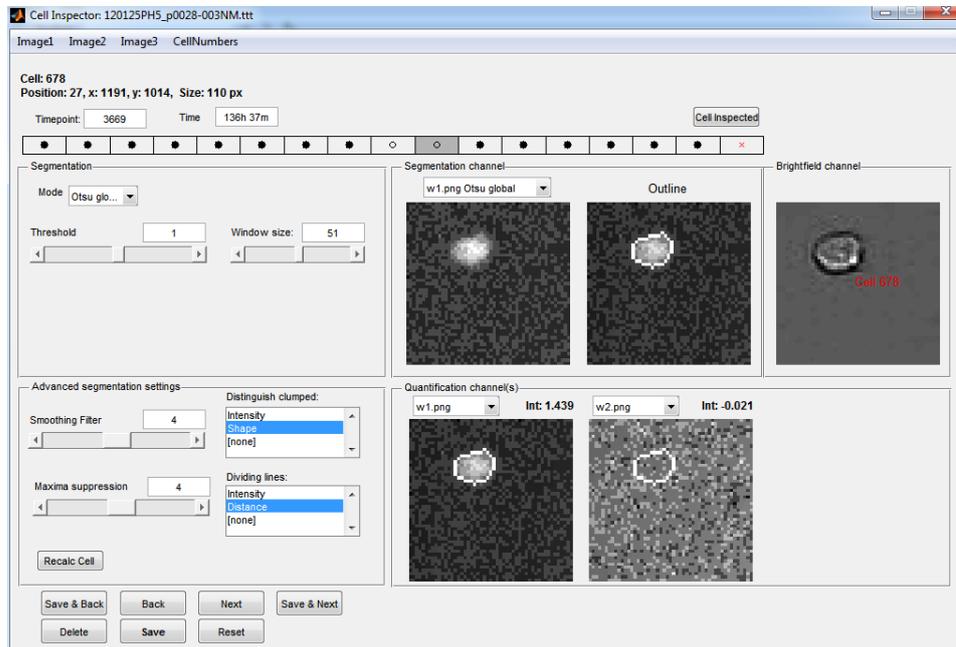


Figure 3.10: The Cell Inspector of QTFy allows to inspect and to adjust the segmentation with all available parameters. The time bar on the top represents the status of time points of a particular cell, whereas filled circles indicate manually inspected data points, open circles indicate only automatic segmentation and red crosses are shown whenever a time point has been deleted.

all its parameters. Additionally images of all quantification wavelengths affected by this segmentation are displayed (compare Figure 3.10). Using keyboard shortcuts the user can quickly go through all the measurements and inspect and, if necessary, adjust the segmentation parameters. Whenever a segmentation of a cell fails even after manual adjustment, the user is able to draw an ellipsoid or even by freehand an individual cell mask. Every time a measurement got manually inspected by pressing the *Save* or *Save & Next* button a flag will be recorded allowing later analysis to filter on inspected data (i.e. manually curated data).

3.3.7 Summary

QTFy has been optimized to achieve the needs of efficient and reliable quantification of cellular fluorescence in a semi-automated manner. We evaluate the normalization, segmentation and quantification pipeline of the tool and compare it with well-established quantification methods in Chapter 4. We apply QTFy to quantify fluorescence of ESCs in Chapter 6 and Chapter 7 and of HSCs in Chapter 5.

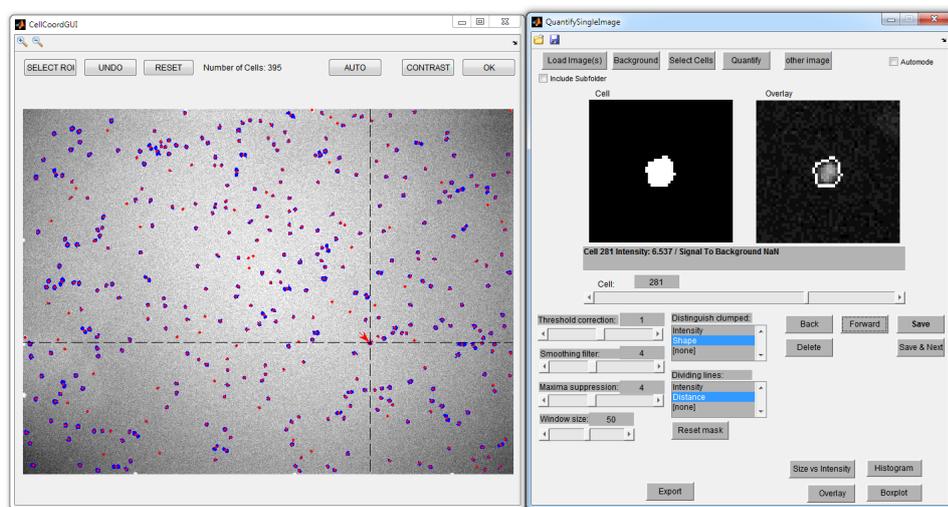


Figure 3.11: sQTFy: A tool to quantify fluorescence in snapshot image data. After detecting cell locations with a rough segmentation algorithm (blue dots in left window, with red dots representing cell center), the user can apply the QTFy steps to perform semi-automatic cell segmentation. The user can adjust the parameters of every individual cell (right window). The segmentation can be used to quantify cellular fluorescence of the image itself or based on a different image of another wavelength.

3.4 Quantification of fluorescence snapshot images (sQTFy)

To quantify cellular fluorescence of a single image without tracking information, we developed a derivative of QTFy, called sQTFy ('s' means snapshot, Figure 3.11). The tool basically performs similar steps as QTFy, but also incorporates methods to identify cell locations. The general workflow comprises the following steps:

1. Load a single image (i.e. snapshot data).
2. Perform background correction, either based on the gain normalization method or otherwise by using background subtraction and background division (Section 2.2.6). Background images can be calculated directly in the tool with a semi-automatic approach or precalculated background images can be used.
3. Detect cell locations by using a blob detection algorithm called MSER (Matas et al., 2002) (Figure 3.11 left). We use a linear time implementation (Hilsenbeck, 2011, Nistér and Stewénus, 2008) to optimize cell segmentation performance on the whole image based on given param-

eters. Use either the segmentation performed by MSER, or perform the manual adjustment segmentation pipeline used in QTFy (compare Section 2.2.2 and Section 2.2.3) for each identified object individually.

4. After that, the user can adjust segmentation parameters for each cell individually, similarly to QTFy (Figure 3.11 right).
5. Use the (manually optimized) cell masks to quantify cellular fluorescence intensities in the image itself, or overlay them onto a different image (e.g. a different wavelength) and quantify cellular fluorescence in another image.
6. The tool can export quantification and segmentation features into simple text files or Excel spreadsheet for further analysis.

In contrast to QTFy, the output contains two measures benchmarking the underlying image quality: *signal to background* κ_{1i} and *signal over background* κ_{2i} of a given signal s_i of cell i and its underlying background signal $b_i = \sum_{x \in c_i} \hat{B}(x)$ (compare Equation 2.6) defined as

$$\kappa_{1i} = \frac{s_i}{b_i} \quad (3.2)$$

$$\kappa_{2i} = \frac{s_i - b_i}{b_i} \quad (3.3)$$

These criteria have often been used in the Schroeder lab to benchmark new cameras and experimental setups, and to optimize our image quantification approach. Furthermore, we heavily used the tool to segment on DAPI signal and to quantify endpoint staining fluorescence to investigate transcription factor activities in ESCs (Section 6.7). The additional feature to draw a particular region of interest for every image allowed us to efficiently segment cells of clonal colonies.

The current version of sQTFy is written in MATLAB. Recently, the requirements of this tool changed from benchmarking single images to quantifying cells in an automated manner in the context of high-content screening. Simple batch processing algorithms are already implemented, but for a more efficient work flow large parts of the tools have to be adapted. Furthermore, to extend the group of potential users of the tool, we tried to completely transfer and extend sQTFy as an ImageJ plug-in in the bachelor thesis of André Seitz (Seitz, 2013). However, it turned out that ImageJ is not the optimal platform for our application. We want explore the approach further and e.g. provide the tool as a CellProfiler (Kamentsky et al., 2011) or stand-alone application.

3.5 Annotated single-cell tracking

The following section is based on the publication of Burtscher et al. (2012).

Ingo Burtscher of the Institute of Stem Cell Research (today Institute of Diabetes and Regeneration Research, Helmholtz Zentrum München) investigated if cell populations with different motility properties exist and analyzed single-cell time-lapse microscopy experiments monitoring developing embryos. Although a variety of tools exist that allows tracking of single-cell trajectories, only a few of them provide additional annotations for each tracked cell which was a crucial requirement for the this application. It turned out that automated methods are unfeasible due to high cell density and insufficient temporal resolution. The only tool available at this point to address our needs was TTT (Eilken et al., 2009, Rieger et al., 2009), which provides a framework to manually track cells in an efficient way. However, it required special formatted data as input, was only runnable on Linux based system and not yet conveniently practicable.

Therefore, we developed a custom written tracking software by our own called CCT, Colored Cell Tracker (Figure 3.12). We optimized the user interface to allow for convenient and efficient manual tracking of individual single-cell trajectories with different annotations or cell labels. Furthermore, it incorporates the following properties:

- Track cells forward/backward in time in an efficiently and intuitive manner
- Provide individual cell labels for different cell types, visualized in different colors
- Track only individual cells, no cell division necessary
- Provide individual zoom levels
- Visualize all tracked cells at once
- Play time-lapse movie with given tracks forward/backward with custom speed
- Save/Resume tracking sessions
- Export cell properties like coordinates and labels for further analysis
- Export videos including all tracked cells
- Implemented in MATLAB

CCT provides all necessary tools to efficiently track and annotate individual cells resulting in coordinate specific outputs and exported videos. We

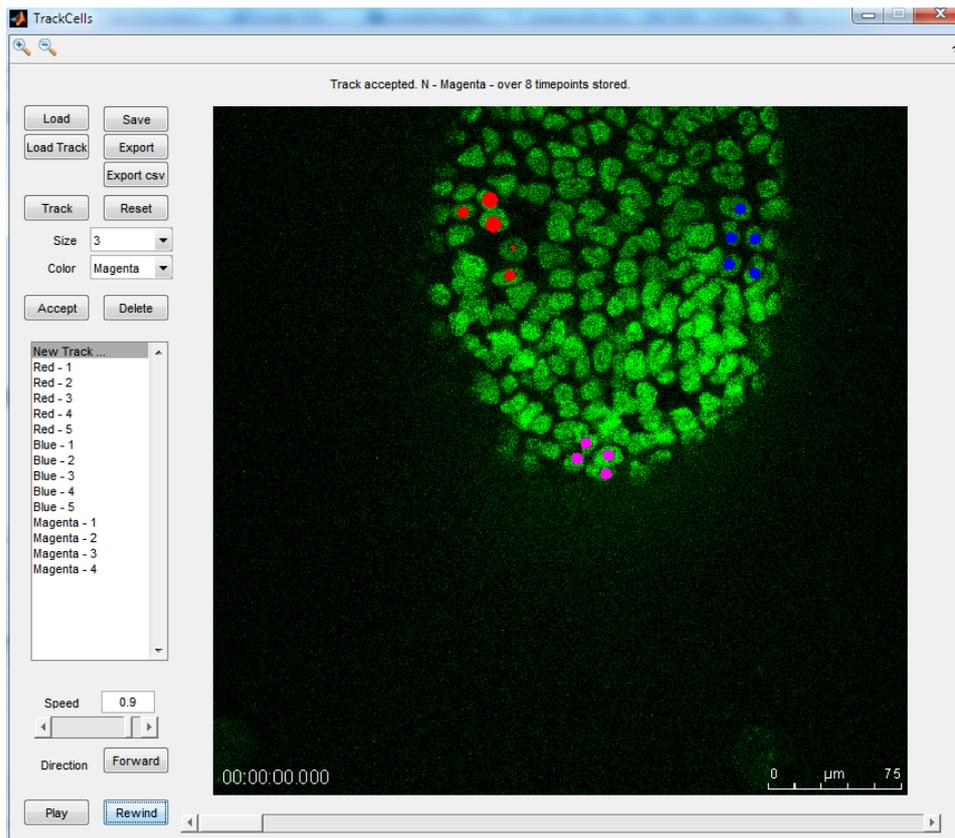


Figure 3.12: Colored Cell Tracker(CCT): an efficient manual single-cell tracking tool. The program allows to track forward and backward in time, simply by clicking and holding the mouse over the cell of interest. It is possible to use individual cell labels for different cell populations as annotations visualized by different colors (cyan and red dots in this example). An arrow highlights the currently tracked cell. Tracked cells can be exported into *csv* files for further analysis or into movies with overlaid trajectories. Saving and resuming the actual session is possible at all times.

could show that using our software allows identifying distinct cell types with different properties of cell movement (compare Chapter 8). In the meantime, due to the fast development and progression in the field of bioimage informatics, more powerful tools emerged which rendered the further development of CCT redundant. We stopped the development of CCT and switched to the current version of TTT as tracking software which has made huge progress in handling compared to its former release (Schwarzfischer et al., in review).

Chapter 4

Benchmarking and evaluation of QTFy

After we discussed the usability and efficiency of our tool QTFy (Section 3.3) we next evaluate the resulting cellular quantifications by comparing them to fluorescence measurements derived by flow cytometry. Furthermore, we compare our normalization method to existing methods currently used in fluorescence quantification. We present a method to determine the detection limit of fluorescence image quantification.

4.1 Fluorescence normalization

In order to accurately quantify cellular signal in fluorescence images several problems have to be handled. In a typical fluorescence image a certain degree of autofluorescence originating from the medium or the plastic well might be observed, which results in a background signal (compare Figure 4.1A). In addition the whole image is typically unevenly illuminated such that intensity values in the corner of the image can deviate up to 40% compared to the center of illumination, which is not necessarily in the center of the image. Since the shape and the center of this illumination pattern can be different between different fields of view within one experiment and between experiments there is no stable reference, which fits to every image. Finally, by observing fluorescence pictures over time, we noticed that the background signal itself bleaches over time and decreases in intensity (Figure 4.1B). This behavior strongly depends on the experimental setup (e.g. exposure time, time interval between images, etc.) and will thereby also vary between experiments.

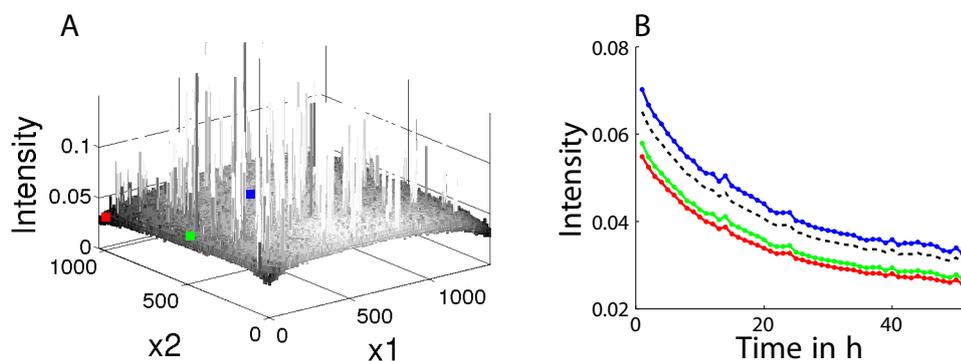


Figure 4.1: Properties of time-lapse fluorescence images. (A) A typical fluorescence image with space coordinates (x_1, x_2) and fluorescence intensity plotted on the z-axis where peaks represent cellular signal. In long-term time-lapse microscopy, one has to deal with the following issues: (i) An inhomogeneous illumination due to the light source and camera lens, (ii) a non-zero background signal due to autofluorescence of the medium, and (iii) the effect of photo-bleaching (B). During an experiment the setup is sequentially illuminated with light. This leads to a bleaching effect in the medium where cells are cultured in and decreases the image intensity. The black dashed line represents the mean image intensity while the colored lines correspond to the colored dots in (A) (red: pixel in the upper left corner, green: pixel in the middle left, blue: pixel in the image center). Figure taken from Schwarzfischer et al. (2011).

4.1.1 Existing methods

There exist already a variety of retrospective methods, which estimate the background signal and the gain function from the cellular image itself (Babaloukas et al., 2011). The most common methods used in the field is a method called 'Rolling Ball' (Sternberg, 1983), which estimates the local background in a given radius by virtually rolling a ball through the 2D image. It basically applies the principles of dilution (Section 2.2.3) with a structuring element similar to the upper half of a sphere. This local filtering eliminates peaks originating from cellular signal and estimates the underlying background signal, but only works for images without large clumping cell colonies. Alternatively, median filtering (Lim, 1990) has been used leading to comparable results and having similar issues with clumping cells. A more sophisticated method fits a B-spline surface into raw images and iteratively filters out cellular signal (Lindblad and Bengtsson, 2001).

To experimentally correct for fluorescence image issues, one proposed method is commonly used, called flat-field correction (Wolf et al., 2007). First, the illumination shape is monitored by measuring a fluorescent dilution. Either the dye is recorded in a different wavelength but in the same well as the cells or the illumination is recorded in the same wavelength in a physically separated well without cells (Schwarzfischer, 2009). Additionally, an empty well is imaged over time to estimate the underlying background signal. The raw image is then normalized by subtracting the background image from the sample image, which is then divided by the background subtracted illumination image (Wolf et al., 2007). For cases where the illumination and background images are representative for the actual image containing cells this procedure was successfully applied previously (Halter et al., 2007, Taniguchi et al., 2010). However, this method implies additional experimental work and requires either additional wells or one additional wavelength.

4.1.2 Our method

Our process of normalizing fluorescence images can be divided into the following main steps. First of all, the background signal of the autofluorescence is estimated by the tiling method (Section 2.2.4). This first step allows to subtract background fluorescence signal, but it still cannot correct for the uneven illumination; cellular signal in the center of illumination is still brighter compared to cellular signal in less illuminated regions. Therefore, the illumination function, called gain (Section 2.2.5), has to be estimated by using the estimated background images. We normalize our fluorescence images by subtracting the background followed by a division by the gain function (compare Section 2.2.6).

4.1.3 Evaluation based on bead fold-changes

We applied our method to a time-lapse microscopy experiment of fluorescent beads, which are commonly used for calibration in flow cytometry. Here, we used unbleached and bleached beads as a representative for two different intensities of fluorophores in a cell matrix. The bleaching itself is of no importance to our test. Thus, we were able to validate our normalization method in a maximally clean and controllable setup.

FITC-Beads were quantified and sorted by flow cytometry. We included only single-bead events according to the FSC/SSC Plot and further gated the beads for high FITC fluorescence. Beads were thereby quantified and subsequently sorted. Furthermore, they were plated on a μ -slide in medium together with medium or a fluorescein-solution in different wells. Tiff-Images of fluorescence were acquired at 1388x1040 pixel resolution over 17 hours at a 5 min interval. Afterwards, the FITC-Beads were harvested, washed and reanalyzed on a flow cytometer using the exact same FSC/SSC gates as before.

For cell segmentation, we used the published tool *ilastik* (Sommer et al., 2011) to detect bead outlines on the bright field images, resulting in binary images. We did not apply the segmentation on fluorescence images using QTFy to exclude any bias. The same bead detections were used throughout all following quantifications with different normalization methods.

For our method, we used tiles of 30x30 pixels overlapping by 15 pixel and clustered on skewness, kurtosis and the fano factor (i.e. the normalized variance). For DBSCAN we set minimal number of objects considered as a cluster to 4 (number of dimensions+1) and used an adhoc value for the neighborhood radius. Following the protocol described of flat-field correction, we subtracted the background from the original image and divided it by the background-subtracted fluorescein images.

To test the applicability of this approach, we compared the intensity fold change between unbleached (at the start of the movie) and bleached (at the end of the movie) beads with flow cytometry analysis and time-lapse fluorescence microscopy imaging (see Figure 4.2). From flow cytometry analysis, we inferred a 8.5 fold change between unbleached and bleached beads. This is in accordance with a 8.0 fold change, derived with our normalization method. In contrast, the raw, unnormalized data yields a fold change of 4.0, while the fluorescein normalized method described in (Taniguchi et al., 2010, Wolf et al., 2007) yields negative intensity values for the bleached beads (see Figure 4.2).

4.1.4 Discussion

Our approach to estimate fluorescence background images can be well compared to already published methods. The simplest thing to solely estimate

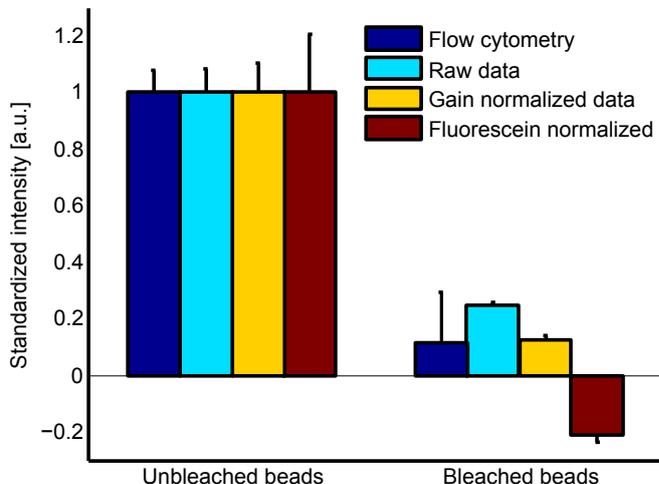


Figure 4.2: We compare the intensity fold change between unbleached and bleached beads derived from flow cytometry analysis and time-lapse microscopy imaging. Our normalization method (yellow) yields ratios comparable to the flow cytometry data (blue). Additionally, the null hypothesis of equal distribution medians cannot be rejected by a Wilcoxon rank-sum test (p-value = 0.5531). The raw data (cyan) and the flat-field corrected (red) data deviate substantially (p-value of Wilcoxon rank-sum test $< 4 \cdot 10^{-8}$). Figure adopted from Schwarzfischer et al. (2011).

the background in real data images is to take the original image and apply a 2D median filtering step with a large window size (Lim, 1990). This method is a quick and easy way as long as there are no colony forming cells and the background does not have a complex illumination pattern. Moreover, it is only appropriate for images where the median of a window always corresponds to the background. Similarly, a rolling ball filter (Sternberg, 1983) also fails for large cell colonies. A more sophisticated way of classifying intensities has been described in (Lindblad and Bengtsson, 2001). The authors iteratively fit a two dimensional cubic spline surface to a grid which is based on estimated background pixels. With this fit they are able to redefine the first poor estimation of pixels corresponding to background and achieve an approximation for the true background after the algorithm reaches the convergence criteria. It turned out that the difference for both methods is minute regarding accuracy as well as computational performance.

Previous approaches to estimate both background and gain commonly used a calibration step to infer the space-dependent gain by imaging fluorescent dilutions like fluorescein (Taniguchi et al., 2010, Wolf et al., 2007). The background or camera offset was determined by imaging empty bins or non-fluorescent dilutions. These images were taken at different spots of the experiment and did thus not reflect the exact illumination and background

at positions where the actual cell images were taken. To countervail this spatial inconsistency, it is also possible to directly include fluorescein in the cell culture and capture the illumination image in a different wavelength than the signal of interest. However, the gain detection is then flawed by cells and other contaminants in the fluid which will lead to deviations from the exact illumination pattern. Our approach uses the bleaching medium and does not rely on additional dilutions. First, we computationally estimate the time-dependent illuminated background from each cell image itself. The gain is calculated by a linear regression for each pixel against the estimated mean background intensities. This is similar to the approach described in Sigal et al. (2006), where different dilutions of GFP are used to infer the pixel gain. In this respect, the different levels of bleached medium resemble different fluorescein or GFP concentrations. Second, we have no need for further chemicals like fluorescein in the cell culture which reduces the experimental work.

Since the availability of background tiles can indeed become crucial for our method we tested the performance of our tiling method by randomly adding bright spots of 10x10 pixel, representing cellular signal within a perfect background. For each iteration we added one hundred of these spots, estimated the background with our method using 30x30 pixel tiles and calculated the root mean square difference. It turned out that our method is applicable for up to 1500 bright spots, which means that about 13% of the image is full of cells. After setting the tile size to 20x20 pixel we could still get accurate background estimations for higher cell densities. We require about 10% of all possible grid points to get an accurate background estimation. However, this threshold also depends on the performance of the used inter- and extrapolation step. Especially the extrapolation step can introduce crucial deviations from the real background.

Concerning computational needs, a two core processor, each 2.80GHz, takes about 15 seconds per single image background estimation and about one hour per gain calculation with MATLAB written software. If computational power is a limiting factor we propose a more parsimonious version of the method: It would suffice to apply the computationally expensive linear regression only to data points of the initial background grid instead of all points including interpolated pixel. After that, the inter- and extrapolation is used for the sparse gain-grid.

With fluorescent beads as a representative for cellular signal, we here validated our normalization method against an accurate flow cytometry analysis. We showed that correct fold changes are only preserved if an accurate normalization method is applied.

In summary, we provide a powerful toolbox to accurately normalize long-term time-lapse fluorescence microscopy movies and to reliably quantify cellular fluorescence.

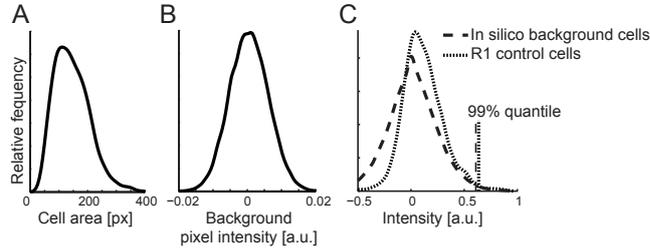


Figure 4.3: Autofluorescence of control cells can be well estimated by *in silico* simulated background cells. (A) Distribution of inspected cell areas based on mCHERRYnucmem segmentation of R1 control cells. (B) Regions without any cellular signal in normalized fluorescence images have been manually selected resulting in a distribution of normalized background pixel intensities ($n=100,000$) (C) Based on a selected cell area a of (A) we are able to simulate *in silico* background cells by summing up a randomly selected background pixel intensities of (B) ($n=100,000$). We compare *in silico* background cells against real control cells, which have been tracked and quantified ($n=3,920$). Figure adopted from Schwarzfischer et al., in review.

4.2 Determination and evaluation of a negative gate

After normalizing an image, technical noise remains in equal measure at every pixel. To estimate a potential positive intensity level of negative cells reached by summing up noisy background pixels, we determine a threshold (i.e. negative gate) to judge if a cell intensity is real or just noise. Motivated by the gating procedure used by flow cytometry, we refer cells with intensities below the gate as negative cells, and everything above the gate as positive cellular signal.

4.2.1 Background level of *in silico* cells

From an experiment we manually select regions in the background normalized fluorescence images where no cell can be found in the bright field channel from several positions and time points. All pixel intensities create a distribution of background intensities centering on zero (Figure 4.3B). From the segmentation of real R1 control cells we get a distribution of cell sizes (Figure 4.3A). By randomly choosing one cell of size a and randomly selecting a pixel from the background intensity distribution we can simulate *in silico* background cells. Finally, we end up with a distribution of *in silico* background cell intensities and define the negative gate as the 99 % quantile of this distribution (see Figure 4.3C).

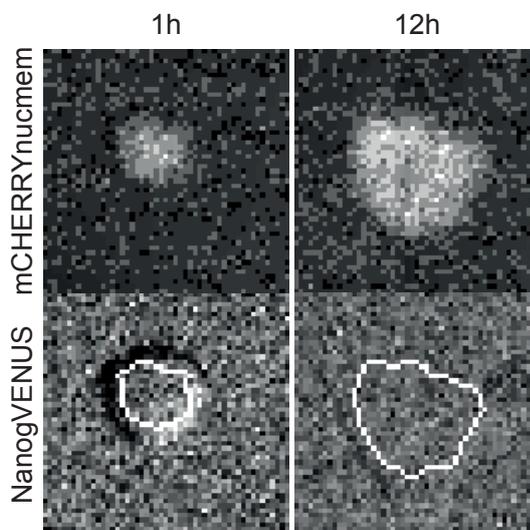


Figure 4.4: Autofluorescence of R1 control cells vanishes after 12 hours. mCHERRYnucmem signal is used to perform cell segmentation. NanogVENUS artifacts vanish after 12 hours after movie start.

4.2.2 Background level of control cells

To evaluate the approach of *in silico* generated control cells, we used freshly sorted R1 control ESCs which have been infected with mCHERRYnucmem and imaged for 24 hours. We applied our complete pipeline and tracked, segmented every cell in mCHERRYnucmem and quantified NanogVENUS background normalized fluorescence images. We end up with a distribution of control cell signals. We define the negative gate in imaging as the 99% quantile of this distribution (compare Figure 4.3C). The negative gate of *in silico* background cells was slightly lower than the 99% quantile of real negative cells.

4.2.3 Autofluorescence of cells vanishes after 12 hours

We found that some cells form artificial shadows and become visible in fluorescence images although the actual cell does not incorporate any fluorescent marker in this wavelength (compare Figure 4.4A). Quantifying the signal of these cells resulted in positive values above background. Furthermore, we investigated if the background level varies over time and found that after approximately 12 hours these shadows disappear. We quantified 98 trees over 20 hours and illustrate that the 99% quantile of half an hour bins of measurements becomes comparable to the negative gate of *in silico* background cells after about 12 hours (see Figure 4.5). Taking the 99% quantile of all measurements after 12 hours results in an almost overlapping negative gate compared to *in silico* background cells.

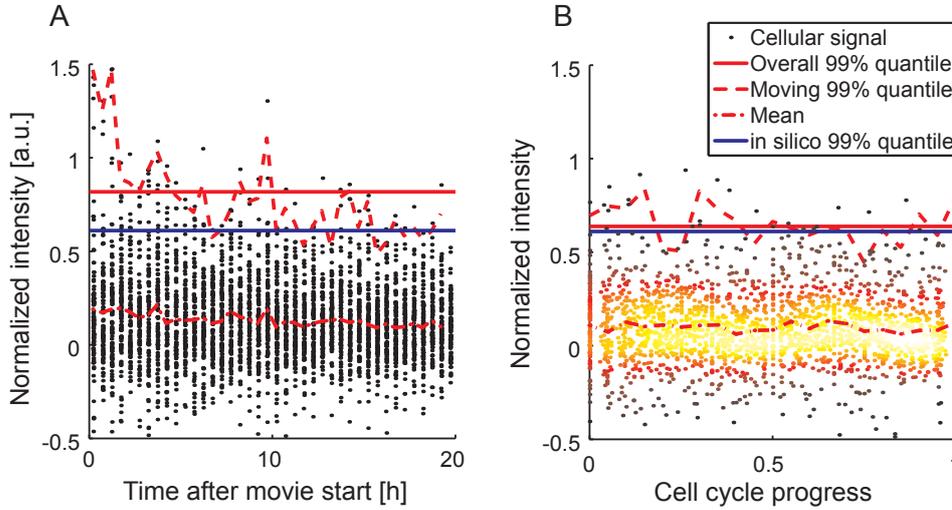


Figure 4.5: The negative gate of real control cells decreases over time and is independent of cell-cycle progression. (A) The 99% quantile of 98 tracked and quantified R1 control cells converges to the 99% quantile of *in silico* background cells. (B) The 99% quantile of quantified cells is highly comparable to the 99% quantile of *in silico* background cells throughout the cell-cycle.

Next, we asked if these artificial shadows form because cells have not yet flattened out on the glass bottom of the well. Interestingly, we also observe an effect of cells forming spheres right before cell division. If this causes any arbitrary optical fluorescence signal real cellular signals of ESCs are affected falsifying the quantification. We took all cells which have been tracked over one generation and normalized their time courses to standardized cell-cycle times (see Figure 4.5B). The moving 99% quantile and the mean over the cell-cycle phases does not show a trend towards cell division. Also, the overall 99% quantile of the quantified cells is highly comparable to the 99% quantile of *in silico* background cells validating our method to estimate the negative gate. We finally conclude, that the signal of auto fluorescence from cells originates from fluorescent molecules, which bleach after approximately 12 hours in the current setup, and originates not from optical effects.

Note, that in a second experiment we used freshly sorted C67Bl/6 MEP control cells from bone marrow and imaged them right after sorting. We quantified randomly selected cells at the first time point over many different positions and compared the resulting distribution of control cell intensities to *in silico* background cell intensities. For these blood cells we did not detect any shadows or autofluorescence. The 99% quantiles of both distributions were highly comparable (data not shown).

By using a density based approach we estimate the empirical distribution of control cells by a kernel density estimator. In a mixture model including the estimated control distribution together with one Gaussian distribution, we fit the distribution of sample cells assessing for every measured cell a probability if originating from the control distribution or the fitted Gaussian. However, for the data available at this point the maximum likelihood approach estimated the weight of the control distribution close to zero (data not shown). Therefore, the sample distribution was already well-explained by the Gaussian. To make a fair comparison of this approach to our negative gate further data is needed. However, this approach needs measurements of control cells in every movie and assumptions about the sample distribution have to be made.

Summarizing, we are able to estimate the detection limit of our fluorescence image technology by a negative gate for every movie independently by simulating *in silico* background cells from selected normalized background image regions and a distribution of expected cell areas. Additional real control cells are not needed for our experimental setup which decreases the biological work and redundantizes the quantification of control cells.

Chapter 5

Single-cell kinetics refute the stochastic PU.1 / Gata1 switch as the basis of myeloid lineage choice

Methods, results and figure in this section are partly based on Hoppe et al., in review. All biological experiments have been performed by Philipp Hoppe and are not part of this thesis. My contributions are:

- Development and application of QTFy in the context of HSC time-lapse experiments
- Statistical validation of normal PU.1eYFP/Gata1mCHERRY mouse line behavior
- Comparison of developed quantification technology to flow cytometry
- Estimation of protein abundance from given Western Blots
- Data handling, analysis and statistics of single-cell quantifications

In the following section we will discuss the biological relevance of the two transcription factors PU.1 and GATA-1 during myeloid differentiation. Philipp Hoppe constructed a hemi-/homozygous PU.1eYFP/Gata1mCHERRY double knock-in mouse strain to visualize protein abundance during *in vitro* cell differentiation on a single-cell basis. First, we show that this mouse line does not have any severe phenotype that would effect the differentiation process. Next, we show that time-lapse microscopy quantification compares well to flow cytometry analysis in terms of fluorescence fold-changes, as well as dynamic range and sensitivity. Furthermore, we map estimated protein molecule numbers to image intensities. Finally, we can conclude that the

current opinion about the PU.1/GATA-1 paradigm is not a key mechanism in myeloid lineage differentiation. Therefore, existing models have to be carefully revised.

5.1 Biological background: hematopoiesis

Hematopoiesis is the continuous process of generating mature blood cells and of maintaining a healthy blood system (Orkin, 2000). Hematopoietic stem cells (HSCs) can self-renew and have to give rise to all mature blood cell types in the right proportion. Early hematopoiesis evolves from the ICM and more specifically in the yolk sac and fetal liver derived from the mesoderm (Orkin and Zon, 2008). After birth the adult hematopoiesis is transferred to the bone marrow (Figure 5.1) (Mikkola and Orkin, 2006). HSCs can differentiate into more specialized cells and undergo different multipotent progenitor cell (MPP) stages and lose their ability to self-renew. However, subpopulations of more specialized cells have been recently identified which can also indefinitely self-renew (Sieweke and Allen, 2013, Yamamoto et al., 2013). In literature MPPs are assumed to choose between differentiation into common myeloid progenitors (CMPs) or into common lymphoid progenitors (CLPs) (Rossi et al., 2012). CMPs can further give rise to granulocyte/macrophage progenitors (GMPs) or megakaryocyte/erythrocyte progenitors (MEPs) (Akashi et al., 2000). All these progenitor cells constantly divide and rebuild the whole pool of mature blood cell types including erythrocytes, platelets, granulocytes, macrophages, dendritic cells, NK cells, B and T lymphocyte (Figure 5.1, (Rossi et al., 2012)). Throughout all differentiation steps shortcuts and crosslinks have been reported (Adolfsson et al., 2005).

To regulate this highly dynamic system a careful fine tuning is necessary that constantly maintains normal and healthy hematopoiesis. Malfunctions disturbing the balance will lead to severe phenotypes such as leukemia or anemia motivating to understand its underlying processes for future therapies (Wolff and Humeniuk, 2013). The influence of extrinsic and intrinsic mechanisms is highly debated and after decades of research not fully understood. In current literature, it is widely assumed that hematopoietic stem and progenitor cells (HSPCs) differentiate based on an intrinsic stochastic system e.g. due to transcriptional noise (Cantor and Orkin, 2001, Cross and Enver, 1997, Enver et al., 1998, Graf, 2002, Graf and Enver, 2009, Orkin, 2000, Orkin and Zon, 2008). While the outcome of an individual HSPCs is only determined by random fluctuations, the overall probabilities for all cell populations are deterministically determined by the wiring of the underlying transcription factor networks. Hence the balance between blood cell types is maintained.

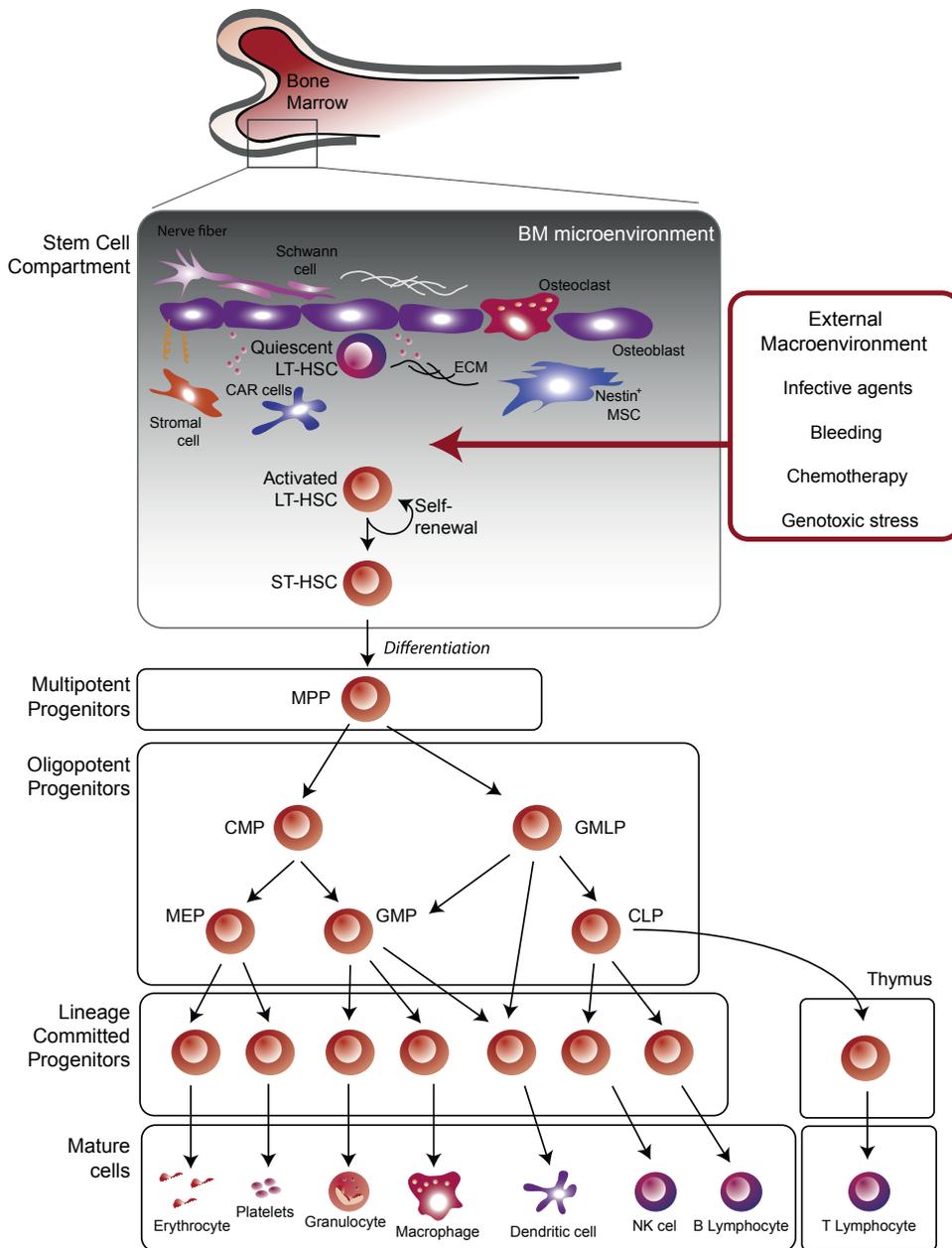


Figure 5.1: Hematopoiesis is the lifelong replenishment of mature blood cells originating from long-term hematopoietic stem cell (LT-HSCs) taking place in the bone marrow. HSCs have the ability to self-renew and to give rise to all mature blood cell types by differentiating into more specialized multipotent (MPP) and oligopotent progenitor cells and thereby losing their self-renewal capability. Figure adopted from Rossi et al. (2012).

The regulatory network involved in hematopoiesis consists of different biochemical motifs, such as auto-activation (Chen et al., 1995, Li et al., 2001, Okuno et al., 2005, Tsai et al., 1991) or mutual inhibition (Nerlov et al., 2000, Zhang et al., 1999, 2000). These two mechanisms can describe a robust bistable switch behavior where either one or the other transcription factor is on. The effect of the mutual inhibition assures that the antagonistic player is suppressed and a lineage decision is established. It is assumed that this motif is one of the key regulators in hematopoietic lineage decision. For the myeloid lineage decision the transcription factors PU.1 and Gata1 form such a particular motive (Wolff and Humeniuk, 2013). Furthermore, it has been demonstrated that transcription factors involved in bistable switches also repress downstream targets of their antagonists which allows to further lock-down differentiation programs of rivaling lineages (Wolff and Humeniuk, 2013). However, these hypotheses are based on previous gain- and loss-of-function experiments measuring mRNA data. Early expression analysis discovered specific transcription factors which are only expressed in a certain lineage (Evans et al., 1988). Further loss-of-function studies recovered more and more transcription factors proven to be crucial for specific cell types (McKercher et al., 1996, Pevny et al., 1991, Scott et al., 1994). On the other hand, by gain-of function experiments or induced overexpression it has been shown that committed progenitor cells can be driven into one lineage direction or even be reprogrammed to their antagonistic lineage (Kulesa et al., 1995, Nerlov and Graf, 1998). Further studies have shown that candidates involved in such a bistable motif are expressed in low levels in progenitor cells before commitment which would make this scenario plausible (Delassus et al., 1999, Hu et al., 1997, Miyamoto et al., 2002). The resulting consequence from this hypothesis is that every multipotent cell should express all lineage specific transcription factors (Cantor and Orkin, 2001, Cross and Enver, 1997, Enver et al., 1998, Orkin, 2000).

Based on these experimental results theoretical models have been developed trying to explain the differentiation process of HSPCs. In particular, the bistable switch has been extensively investigated. The simple models quantitatively describe the interaction of the two involved factors using ordinary differential equations (ODEs). By investigating the parameter space a bistable and a tristable behavior has been described which could explain a multipotent meta-stable state with its two downstream lineage-restricted stable-states (Chickarmane et al., 2009, Huang et al., 2007, Roeder and Glauche, 2006). By adding additional noise terms, it has been recently shown that a full stochastic two-stage model reveals not only one but two meta-stable states which have a bias for only one specific lineage respectively (Marr et al., 2012, Strasser et al., 2012). Although this meta-stable multipotent state is already primed for a lineage they can easily be driven towards their antagonistic regime. In a different approach a more complete view onto HSPC differentiation was implemented by including 11 core transcrip-

tion factors (Krumstiek et al., 2011). By Boolean modeling all four mature myeloid cell types could be reproduced and observed biological mRNA data could be very well fitted.

Although these theoretical work can well-explain the observed measurements, they all lack of experimental validation with continuous single-cell resolution and protein level. So far, mostly population average based techniques such as Western Blot, microarray or snapshot data based techniques such as flow cytometry were available. With these techniques it is impossible to unravel the underlying biological processes since they cannot measure time-resolved data of individual cells. Furthermore, a non-invasive method is needed to quantify the expression of transcription factors continuously. For a deeper understanding of HSPC differentiation only continuous long-term single-cell observations can give new insights.

5.2 The PU.1eYFP/Gata1mCHERRY double knock-in mouse

For this purpose Philipp Hoppe created in the lab of Timm Schroeder (Research unit Stem Cell Dynamics, Helmholtz Zentrum München) a hemi-/homozygous PU.1eYFP/Gata1mCHERRY double knock-in mouse strain (PG) which allows to quantify PU.1 and Gata1 abundance in dependence of their respective fluorescent reporter (Hoppe et al., in review). The homozygous PU.1eYFP fusion knock-in mouse has been previously described to serve as a faithful reporter for PU.1 function without any hematopoietic effect compared to wild type mice (Kirstetter et al., 2006). The fluorescence protein mCHERRY was directly fused to the C-terminus of Gata1 gene sequence which is located on the X-chromosome (Figure 5.2A). The resulting Gata1mCHERRY mice were mated with PU.1eYFP mice to create double knock-in PU.1eYFP/Gata1mCHERRY mice (PG) which are homozygous for both transcription factors (male mice are hemizygous) (Hoppe et al., in review).

5.3 Phenotypical comparison of wild type and double knock-in mice

In the following section, we show that PG mice show no severe phenotype and undergo normal hematopoiesis compared to a wild type C67Bl/6 mouse (WT) in terms of blood counts, composition of bone marrow, protein half-life and composition of colony assays.

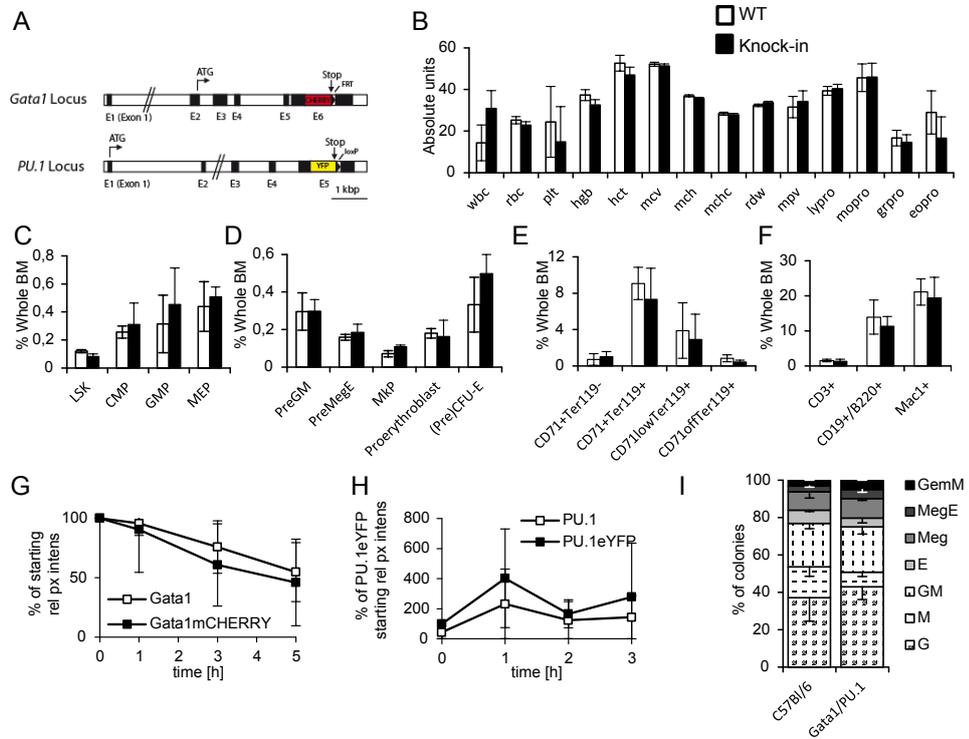


Figure 5.2: Homo-/Hemizygous PU.1eYFP/Gata1mCHERRY double knock-in mice exhibit normal hematopoiesis. (A) Structure of endogenous loci of Gata1 and PU.1 after knock-in of mCHERRY and eYFP, respectively. (B) Peripheral blood counts of adult male wild type C57BL/6 (WT, n=6) and PU.1/Gata1 (PG, n=9) mice. wbc=white blood cells ($200/\text{mm}^3$), rbc=red blood cells ($4 \cdot 10^5/\text{mm}^3$), plt=platelets ($4 \cdot 10^2/\text{mm}^3$), hgb=hemoglobin ($0,4 \cdot \text{g/dl}$), hct=hematocrit (%), mcv=mean corpuscular volume (μm^3), mch=mean corpuscular hemoglobin ($0,4 \cdot \text{pg}$), mchc=mean corpuscular hemoglobin concentration (g/dl), rdw=red cell distribution width ($0,4 \cdot \%$), mpv=mean platelet volume ($0,2 \cdot \mu\text{m}^3$), lypro=($2 \cdot \%$ of wbc), mopro=($0,1 \cdot \%$ of wbc), grpro=(% of wbc), eopro=($0,2 \cdot \%$ of wbc). A one-way multivariate analysis of variance (MANOVA) did not detect any significant difference between WT and PG mice (p-value > 0.09). (C)-(F), Composition of adult bone marrow (BM) of WT and PG mice (n = 4). A MANOVA for each panel independently detected one significant difference between WT and PG mice (D) (p-value = 0.03). For (D), (E) and (F) a MANOVA did not detect a significant difference between WT and PG mice (p-value > 0.16). (G) Western Blot quantification of Gata1(mCHERRY) pixel intensities after Cycloheximid treatment of E14.5 fetal liver (FL) cells (50% wild type Gata1, 50% Gata1mCHERRY). (H) Western Blot quantification of PU.1(eYFP) pixel intensities after Cycloheximid treatment of PU.1^{wt/eYFP} progenitor cells. (I) Colony assay of whole BM WT and PG mice (n = 3 each). A Wilcoxon rank sum test did not detect any significant difference between WT and PG mice for every population independently (p-value > 0.2). All data are mean \pm s.d. Figure adopted from Hoppe et al., in review.

5.3.1 Blood counts

We measured the peripheral blood counts of six adult male WT mice and nine adult male PG mice of the same age (12 to 16 weeks). The complete measurement data consists of 18 different features for each mouse (see figure 5.2B). Some features are redundant such as rbc (red blood count) and hct (hematocrit) and therefore have a high Pearson correlation ρ (compare Section 2.3.1). Since highly correlated variables can be explained by one or the other we filtered the 18 features to make statistical tests more effective. We detected the following high correlation pairs: grpro and lypro ($\rho = -0.99$), wbc and lnum ($\rho = 0.99$), hct and rbc ($\rho = 0.99$), hct and hgb ($\rho = 0.98$), rbc and hgb ($\rho = 0.96$), eonum and grnum ($\rho = 0.96$). We filtered out lypro, lnum, hct, hgb and eonum. Using the remaining 13 features a one-way multivariate analysis of variance (MANOVA) did not reject the null hypothesis that the means of each group are the same and variances are only due to random fluctuations (p-value = 0.094). With this analysis we did not detect a significant difference between the two groups. With any other filtering combination (e.g. wbc instead of lnum, etc) we could not detect a significant difference as well (all p-values > 0.09).

5.3.2 Composition of bone marrow

We applied different staining schemata to whole bone marrow of four WT and four PG mice which allowed us to investigate the composition of 16 well-defined cell populations (compare Figure 5.2C-E). For each staining (i.e. each panel of Figure 5.2) we independently applied a MANOVA. We detected one significant difference between WT and PG mice for the Bryder staining (Pronk et al., 2007) with a p-value of 0.03 (compare Figure 5.2D) which could be explained by the non-overlapping error-bars for MkP (Megakaryocyte progenitors), which is the only group showing a significant difference when applying individual Wilcoxon rank sum tests to every cell population (p-value < 0.03). Every other staining did not show a significant difference between WT and PG mice (all p-values > 0.16). In summary, although a small difference could be detected, the composition of 16 bone marrow cell populations is highly comparable between WT and PG mice.

5.3.3 Protein half-life

We compared the stability of the fused protein and its wild type reference for PU.1 and Gata1 using Western Blot analysis. Protein synthesis was inhibited using cycloheximide treatment. In the experiment investigating Gata1 and Gata1mCHERRY, E14.5 fetal liver cells of 50% WT and 50% PG mice were loaded onto the Western Blot resulting in two bands per line, separating the wild type and the fusion protein. We monitored protein dynamics over 5 hours in three independent biological replicates. Signal intensities have

been background normalized and measured using ImageJ (Schneider et al., 2012). We normalized each time course to its initial intensity and illustrate the mean with its standard deviation for each measurement in Figure 5.2G.

For the comparison of wild type PU.1 and PU.1eYFP we used progenitor cells of heterozygous PU.1eYFP mice. In these experiments we assume that 50% of wild type and 50% knock-in protein are present. Finally, we applied the same western blot and quantification analysis (compare Figure 5.2H).

A statistical test for differences between paired time-resolved observations (Brand et al., in preparation) did not detect any significant difference between WT and PG mice for all proteins, respectively (p-value > 0.16).

5.3.4 Colony assay contribution

For WT and PG mice we created colony assays by seeding cells of the whole bone marrow in three independent biological replicates, respectively. For every colony assay we counted the frequency of possible outcome colonies (compare Figure 5.2I). To test for differences for this kind of data between the two groups a Hotelling's T^2 test is most adequate, but it requires more samples than features (Hotelling, 1931). Since we only have three samples for each group and seven different cell populations a Hotelling's T^2 test is not applicable to this kind of data. Therefore, we applied a Wilcoxon rank sum test for every cell population independently. With this statistical test we did not detect any significant difference between WT and PG mice (all p-values > 0.2).

Summarizing, we conclude that myeloid lineage decision of hematopoietic stem and progenitor cells is not significantly affected by our knock-in strategy. The PG mouse strain provides two faithful reporters to investigate the role of PU.1 and Gata1 during hematopoiesis in a noninvasive manner.

5.4 Simultaneous detection of PU.1eYFP and Gata1-mCHERRY

The hypothesis that PU.1 and Gata1 are already expressed in HSPCs was based on mRNA expression data (Akashi et al., 2000, Iwasaki et al., 2005b). Furthermore, it was shown that GMPs solely express PU.1 and no Gata1 whereas for MEPs the opposite is the case. The double knock-in mouse allows to visualize these two transcription factors on protein level for different live blood cell populations by flow cytometry. First, we can confirm that MEPs express Gata1mCHERRY but are low for PU.1eYFP and GMPs show negative expression for Gata1mCHERRY but have high levels of PU.1eYFP

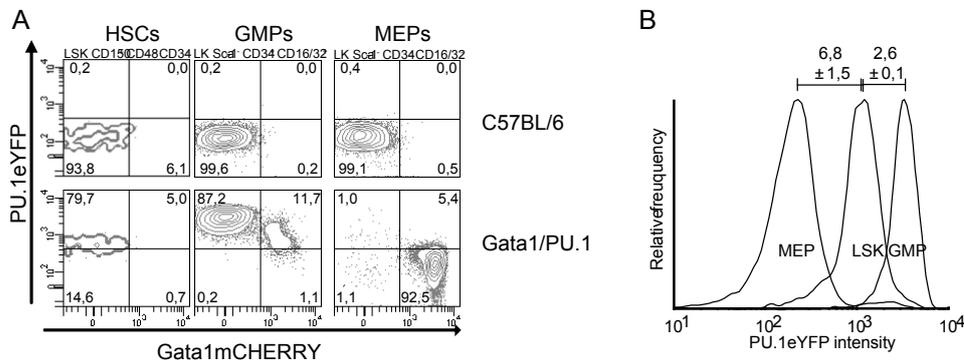


Figure 5.3: The PU.1eYFP/Gata1mCHERRY mouse model allows simultaneous quantification of both transcription factors. (A) Flow cytometry analysis of wild type C57BL/6 and PU.1eYFP/Gata1mCHERRY stem and progenitor cells. Representative example. Data are mean (%) ($n = 4$ each). (B) Fold-changes of PU.1eYFP fluorescence intensity between MEPs, LSKs and GMPs ($n=4$). Data are mean \pm s.d. Figure adopted from Hoppe et al., in review.

(Figure 5.3A). Interestingly, GMPs do have a small fraction of cells expressing both, PU.1eYFP and Gata1mCHERRY. Strikingly, in HSCs we do not see any Gata1mCHERRY expression at all. On the other hand PU.1eYFP is expressed at intermediate levels as expected. Compared to current models HSPCs should express both players at comparably means before and during lineage decision.

The fact that HSCs do not express Gata1 at all does not necessarily exclude that a transient subpopulation exists that has real myeloid progenitor potential which expresses both, PU.1 and Gata1, at the same time. Since flow cytometry data can only show snapshot data of cell populations with known marker expressions we might miss a myeloid progenitor cell population which explains the transition of HSPCs to GM vs MegE cells. With the data shown in Figure 5.3A we are not able to tell how PU.1eYFP and Gata1mCHERRY behave during differentiation. Furthermore, fold-changes in PU.1eYFP between MEPs, LSK, which are downstream of HSCs, and GMPs are rather minute compared to a fold-change of 2 which a cell encounters due to cell-cycle (Figure 5.3B). Looking solely at PU.1 is not enough to judge if a cell is in a late cell phase of a LSK population or already differentiated into GMP lineage after cell division.

Hence, monitoring HSCs during differentiation and quantification of its transcription factors PU.1eYFP and Gata1mCHERRY over a long period of time on a single-cell basis is inevitable.

5.5 Comparison to flow cytometry

We investigated how our fluorescence quantification by imaging performs compared to the well-established method flow cytometry. We sorted three distinct subpopulations in flow cytometry, namely PUhigh (PU.1eYFP⁺ Gata1mCHERRY⁻), GATAmid (PU.1eYFP^{mid} Gata1mCHERRY^{mid}) and GATAhigh (PU.1eYFP⁻ Gata1mCHERRY⁺), from PG progenitor cells (Lin⁻ Sca-1⁻ cKit⁺, compare Figure 5.4A) and put them into physically separated wells. We imaged each well with multiple fields of views for about 30 time points (representative examples are given in Figure 5.4B). For image quantification we use all visible cells of the first time point. However, we need all other time points to well-estimate the gain function of each field of view (compare Section 2.2.6). We used three independent microscopy experiments and quantified every cell using sQTFy (Section 3.4). One representative distribution is given in Figure 5.4C, lower row, with all fold-changes between all cell populations. For comparison we used seven independent flow cytometry experiments to determine the fold-changes between the three populations with its standard deviation in absolute scale for PU.1eYFP and Gata1mCHERRY (see Figure 5.4C, upper row).

We standardized the intensity such that the negative gate of both techniques equals one. For imaging we used *in silico* background cells (compare Section 4.2) whereas for flow cytometry we used WT control cells. From every control distribution we took the 99% quantile as negative gate g . Although imaging and flow cytometry have differences in the low signals, especially the high ranges of both wavelengths are highly comparable in terms of fold-changes between cell populations. The fold-change of PUhigh and GATAmid in eYFP is 8.3 ± 2.4 for flow cytometry and 8.8 ± 1.3 for imaging. Similarly, the fold-change for mCHERRY between GATAhigh and GATAmid is comparable between flow cytometry (2.2 ± 0.3) and imaging (2.2 ± 0.2).

To compare flow cytometry and imaging more quantitatively, we took the negative gate, g , for independent flow cytometry and imaging experiments and calculated the dynamic range d by

$$d = 10 \cdot \log_{10} \left(\frac{Q^{99}}{g} \right) \quad (5.1)$$

with Q^{99} being the 99% quantile of the cell population of interest.

Secondly, we calculate the sensitivity s , defined as the percentage of cells above the negative gate. Note that we calculate d and s for every cell population and both techniques independently.

By integrating four flow cytometry and three imaging replicates we estimate the average dynamic range and sensitivity for every cell population (compare Figure 5.4D and E). Sensitivity trivially reaches 100% for cell

populations which are completely positive (e.g eYFP sensitivity for PUhigh cells). For other cell populations imaging always outperforms flow cytometry in terms of sensitivity and dynamic range. The negative gate of imaging is rather low in relation to the intensity of cell populations which is the major reason why imaging performs better in terms of sensitivity and dynamic range. This holds true for both wavelengths, PU.1eYFP and Gata1-mCHERRY.

5.6 Inference of protein abundance

A current standard to determine protein abundance in cells is to use Western Blot analysis. For this method a high number of cells ($> 50,000$) is needed which makes it unfeasible for primary stem cells as there are roughly 1,000 HSCs per mouse (Warr et al., 2011). To gain enough cell material we used highly abundant cell populations and applied additional mapping steps to assign protein numbers to the cell populations of interest. The variables used throughout this section are independent of Chapter 2 but refer to the manuscript of Hoppe et al., in review.

5.6.1 Gata1mCHERRY

We use fetal liver cells from E14.5 embryos which have a high number of Gata1 positive cells, resulting in about $c \approx 10^6$ cells per western blot. We loaded the sample band with these fetal liver cells. Additionally, we loaded each western blot with up to seven different known concentrations of recombinant mCHERRY. After applying the standard operation procedure for Western Blots, the whole process results in a gray scaled image showing the sample band and all dilution bands (Figure 5.5A). Every dilution band corresponds to one of the recombinant mCHERRY concentrations x . We quantify the absolute intensity of each band, $B(x)$, by selecting an area $A(x)$ that contains the band, and integrate the signal within $A(x)$. We use the same $A(x)$ to quantify the background intensity above and below each band ($a(x)$ and $b(x)$). We determine the background corrected intensity $I(x)$ of each band by subtracting the mean estimated background around that band via

$$I(x) = B(x) - \frac{(a(x) + b(x))}{2} \quad (5.2)$$

To average out any human bias in selecting the regions $A(x)$, we repeat these steps three times and take the average intensity $I(x)$ for further calculations. The intensity of our sample band, S , is calculated analogously three times. From the three measurements we get the average intensity \bar{S} . For the number of cells c we assume an uncertainty of about $\Delta c = 10\% \cdot c$ originating from the counting procedure.

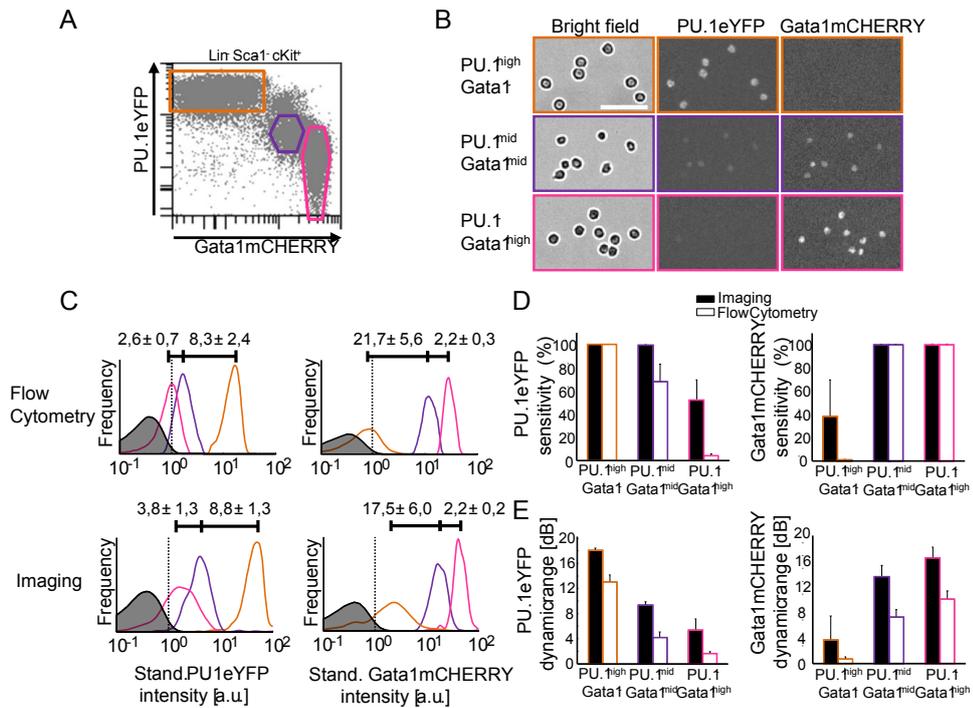


Figure 5.4: Live cell imaging allows reliable quantification of PU.1eYFP and Gata1mCHERRY compared to flow cytometry. (A) Three different Lin⁻ Sca-1⁻ c-Kit⁺ hematopoietic progenitor cell populations were sorted according to their PU.1eYFP and Gata1mCHERRY expression. (B) Right after sorting each population got subsequently imaged and quantified. (C) Fold-changes of PU.1eYFP and Gata1mCHERRY within flow cytometry and imaging aligned to the detection threshold. One representative example is shown for illustration, numbers are mean \pm s.d. (D) Sensitivity of PU.1eYFP and Gata1mCHERRY in flow cytometry and imaging. Data are mean percentages \pm s.d. of cells that are above the detection threshold ($n \geq 3$). (E) Dynamic ranges of PU.1eYFP and Gata1mCHERRY in flow cytometry and imaging. Data are mean \pm s.d. ($n \geq 3$). Figure adopted from Hoppe et al., in review.

We found a sigmoidal function to best describe the relationship between the concentrations x and the background corrected intensity (compare Figure 5.5B):

$$I(x) = \frac{\lambda x^n}{K^n + x^n} \cdot \epsilon(x) \quad (5.3)$$

Here, the exponent n determines the steepness of the sigmoidal function, K sets the inflection point, λ is the maximum of the curve and ϵ is a log-normally distributed error term with expectation 1 and standard deviation σ as suggested for Western Blot data (Kreutz et al., 2007). This model outperformed linear models with and without intercept as well as a sigmoidal function with offset according to BIC and coefficient of variation between replicates.

After solving and normalizing the equation from above for x , we determined the average number of proteins P_j per cell from the sample intensity \bar{S}_j independently for every replicate j (i.e. Western Blot) as

$$P_j = \frac{K_j}{\left(\frac{\lambda_j \epsilon_j}{\bar{S}_j} - 1\right)^{\frac{1}{n_j}}} \cdot \frac{1}{c \cdot w} \quad (5.4)$$

where the second term relates mCHERRY dilution in nanogram to absolute protein numbers per cell ($w = 69992$ Da is the molecular weight for Gata1mCHERRY). The model parameter set $h_j = \{\lambda_j, K_j, n_j, \sigma_j\}$ was obtained by local optimization of the likelihood function $L(h_j|x)$ depending on the data x using multiple restarts initialized according to Latin-hypercube sampling (McKay et al., 1979):

$$L(h_j|x) = \frac{1}{x\sigma_j\sqrt{2\pi}} \cdot \exp\left(\frac{\ln(x) + \frac{\sigma_j^2}{2} - \ln\left(\frac{\lambda_j x^{n_j}}{K_j^{n_j} + x^{n_j}}\right)}{2\sigma_j^2}\right) \quad (5.5)$$

As P_j is a combination of uncertain variables, we obtained error bars for each P_j individually by applying standard error propagation to account for uncertainties in the number of cells c and uncertainties in the model (estimated via the standard deviation σ_j of our noise model ϵ_j) (Sivia, 1996):

$$\Delta P_j = \sqrt{\left(\Delta c \cdot \frac{K_j}{\left(\frac{\lambda_j}{\bar{S}_j} - 1\right)^{\frac{1}{n_j}}} \cdot \frac{1}{c^2 w}\right)^2 + \left(\sigma_j \cdot \frac{K_j}{\left(\frac{\lambda_j}{\bar{S}_j} - 1\right)^{\frac{1}{n_j}}} \cdot \frac{1}{c w}\right)^2} \quad (5.6)$$

However, we find that the uncertainties ΔP_j for each individual replicate j are always at least a factor 3 smaller than the across-replicate standard deviations $\Delta \bar{P}$ ($\Delta P_j < 0.3 \cdot \Delta \bar{P}$). Therefore, we only consider the standard deviation across replicates, as this is the dominant source of uncertainty in

Western Blot j	Number of proteins per cell P_j	Uncertainty ΔP_j
1	$11.8 \cdot 10^3$	$1.9 \cdot 10^3$
2	$27.0 \cdot 10^3$	$2.7 \cdot 10^3$
3	$30.1 \cdot 10^3$	$3.5 \cdot 10^3$
Combined	$\bar{P} = 23.0 \cdot 10^3$	$\Delta \bar{P} = 9.8 \cdot 10^3$

Table 5.1: Gata1mCHERRY protein numbers with uncertainty for three independent Western Blots. Since uncertainties of individual replicates are at least 3 times smaller than the inter-replicate deviations these three replicates are combined and result in average and only standard deviations across replicates.

our procedure and end up with a final number of Gata1mCHERRY proteins per cell for Gata1mCHERRY positive fetal liver cells of $\bar{P} = 23.0 \pm 9.8 \cdot 10^3$ (compare table 5.1).

To estimate the number of proteins in different cell populations we sorted bone marrow of a PG mouse together with fetal liver cells by flow cytometry. From flow cytometry analysis we can calculate the fold changes of mCHERRY intensity between fetal liver cells (used for the western blot) and all other HSPC populations $k = \{\text{PUhigh, GATAmid and GATAhigh (see Chapter 4), HSC, LSK, GMP, MEP, negative gate(MEP), negative gate(GMP)}\}$. Every population k has a mean fluorescence intensity M_k and a standard deviation ΔM_k over different independent experiments. From six experiments for fetal liver cells (FL) we derive M_{FL} and ΔM_{FL} . From seven experiments containing all other populations we derive all M_k and ΔM_k . We calculate the fold change and its uncertainty between two populations by

$$f_k = \frac{M_k}{M_{\text{FL}}} \quad (5.7)$$

and

$$\Delta f_k = \sqrt{\left(\Delta M_k \cdot \frac{1}{M_{\text{FL}}}\right)^2 + \left(\Delta M_{\text{FL}} \cdot \frac{M_k}{M_{\text{FL}}^2}\right)^2} \quad (5.8)$$

By combining all steps in a final propagation of uncertainty we determine the average number of proteins \bar{P}_k for every population k by

$$\bar{P}_k = \bar{P}_{\text{FL}} \cdot f_k = \bar{P}_{\text{FL}} \cdot \frac{M_k}{M_{\text{FL}}} \quad (5.9)$$

and its uncertainty by

$$\Delta \bar{P}_k = \sqrt{\left(\Delta \bar{P}_{\text{FL}} \cdot \frac{M_k}{M_{\text{FL}}}\right)^2 + \left(\Delta M_k \cdot \bar{P}_{\text{FL}} \frac{1}{M_{\text{FL}}}\right)^2 + \left(\Delta M_{\text{FL}} \cdot \bar{P}_{\text{FL}} \frac{M_k}{M_{\text{FL}}^2}\right)^2} \quad (5.10)$$

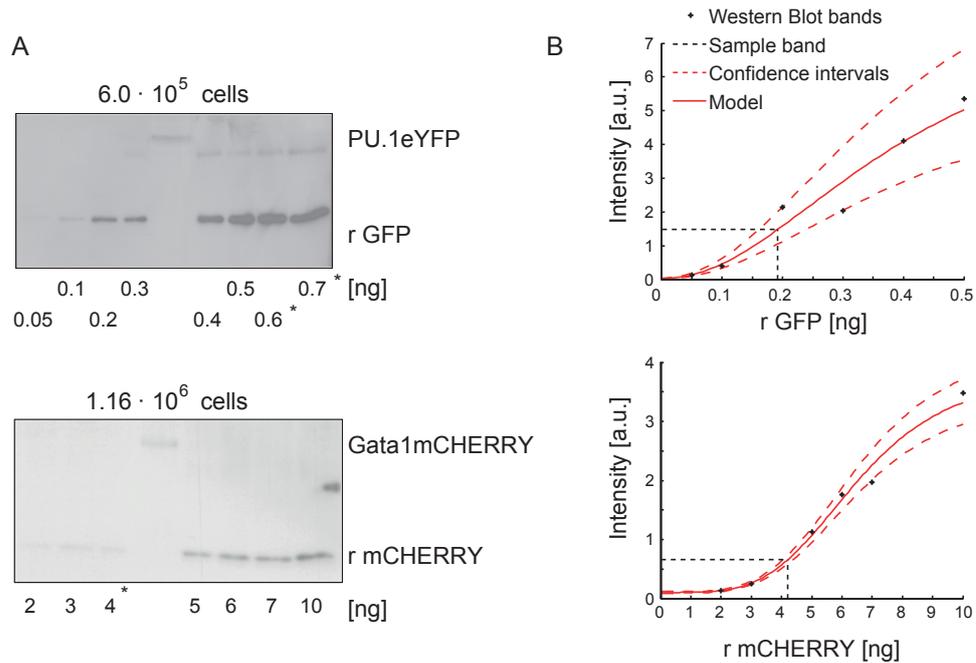


Figure 5.5: Protein abundance determined by Western Blot analysis. (A) Quantification of molecule numbers per sorted PU.1eYFP⁺ LK progenitor cells (top) and E14.5 Gata1mCHERRY⁺ FL cells (bottom) with defined amounts of recombinant (r) eGFP and mCHERRY, respectively. One representative example is shown ($n = 3$). Bands which are marked by an asterisk are excluded from quantification due to overexposure or false pipetting. (B) A sigmoidal fit incorporating a lognormally distributed noise model best explains the measured intensity bands. The sample band lies between measured Western Blot bands. Red dashed lines show the 5 and 95% confidence intervals.

Cell population k	MFI M_k	Fold change f_k	Proteins per cell \bar{P}_k
Fetal liver (FL)	1630.0 ± 40	1	$23.0 \pm 9.8 \cdot 10^3$
PUhigh	—*	—	—
GATAmid	1810 ± 409	1.11 ± 0.25	$25.5 \pm 12.3 \cdot 10^3$
GATAhigh	3875 ± 366	2.38 ± 0.23	$54.6 \pm 23.8 \cdot 10^3$
HSC	—*	—	—
LSK	—*	—	—
GMP	—*	—	—
MEP	3502 ± 287	2.15 ± 0.18	$49.4 \pm 21.4 \cdot 10^3$
Negative gate (GMP)	461 ± 184	0.28 ± 0.11	$6.5 \pm 3.8 \cdot 10^3$
Negative gate (MEP)	596 ± 202	0.37 ± 0.12	$8.4 \pm 4.6 \cdot 10^3$

Table 5.2: Gata1mCHERRY protein numbers have been determined for fetal liver (FL) cells with three independent Western Blot analyses. Using the mean fluorescence intensity (MFI) over several independent flow cytometry experiments allows to map protein numbers from FL cells to every other population. Error bars are derived using error propagation and incorporate deviations FL cells and of standard deviations of MFIs. * = MFI is below negative gate which does not allow to reliably determine the number of proteins.

The final protein numbers of Gata1mCHERRY in 10 HSPC populations are summarized in Table 5.2 and partly visualized in Figure 5.6. Some populations have a mean fluorescence intensity which is below the negative gate for flow cytometry. Protein abundance calculation have not been applied to these populations since their fluorescence quantification is below the detection threshold and thereby not reliable.

In summary, we estimated the number of proteins by widely used Western Blot technology and additionally determined the uncertainty which incorporate not only imprecision of Western Blot analysis, but also uncertainties of the applied mapping procedure to estimate protein numbers in low abundant cell populations. Compared to studies which estimated the number of molecules for various proteins on a large scale (Schwanhäusser et al., 2011), our estimates fit well into the reported scale having a median of 16,000 proteins per cell.

5.6.2 PU.1eYFP

To get PU.1eYFP protein abundance in certain cell populations we applied the exact same technique as in the previously described section. We loaded three independent Western Blots with recombinant GFP since Anti-GFP should recognize both, wild-type GFP and mutant-forms of GFP (i.e. eYFP, Figure 5.5A). Additionally, we loaded PU.1eYFP positive cells (i.e. PU⁺ LK progenitor cells) in the sample lane. For PU.1eYFP we use a molec-

Western Blot j	Number of proteins per cell P_j	Uncertainty ΔP_j
1	$42.8 \cdot 10^3$	$5.4 \cdot 10^3$
2	$53.8 \cdot 10^3$	$8.4 \cdot 10^3$
3	$32.6 \cdot 10^3$	$5.2 \cdot 10^3$
Combined	$\bar{P} = 43.1 \cdot 10^3$	$\Delta \bar{P} = 10.6 \cdot 10^3$

Table 5.3: PU.1eYFP protein numbers with uncertainty for three independent Western Blots. Since uncertainties of individual replicates are at least 3 times smaller than the inter-replicate deviations these three replicates are combined and result in average and only standard deviations across replicates.

ular weight of $w = 60102$ Da in Equation 5.4. We fitted all Western Blot quantification with the same sigmoidal function (Equation 5.3, Figure 5.5B) for every replicate and after applying the Equations 5.4 and 5.6 derived a total amount of PU.1eYFP protein of $43.1 \pm 10.6 \cdot 10^3$ in PU⁺ LK progenitor cells (compare Table 5.3).

To determine the protein abundance in 9 other HSPC populations we use PU⁺ LK progenitor cells as reference distribution. We apply the equations 5.7 to 5.10 and end up with protein numbers for all HSPCs (compare table 5.4, Figure 5.6).

5.6.3 Mapping protein numbers to imaging intensity

Identifying the same cell population in time-lapse microscopy movies as in flow cytometry allows calculating the amount of protein for every individual cell for every time point. The mean fluorescence intensity of the first time points corresponding to freshly sorted HSCs was used to calibrate PU.1 protein abundance independently for each time-lapse experiment. Since we did not observe any Gata1mCHERRY expression in HSCs, we had to find a different reference population to calibrate Gata1mCHERRY molecule numbers. For Gata1mCHERRY we determined the detection limit by *in silico* background cells (compare Section 4.2). Whenever a movie cell exceeds twice the detection limit in the mCHERRY channel for more than 5 consecutive time points, the cell itself and all its descendants were annotated as Gata1 positive. Mean protein abundance of these Gata1 positive imaging cells has been calibrated to the mean protein abundance of PU.1^{mid}Gata1^{mid} in flow cytometry. By assuming a linear behavior between protein numbers and imaging intensity, protein levels are then interpolated linearly for every single measurement.

We used the fold-changes of PU^{high} to GATAmid, GATA^{high} and the negative gate quantified by imaging (compare Section 5.5) to map protein numbers between these populations (Figure 5.6). For PU.1eYFP we used protein numbers from PU^{high} cells as reference (Table 5.4) and for Gata1-

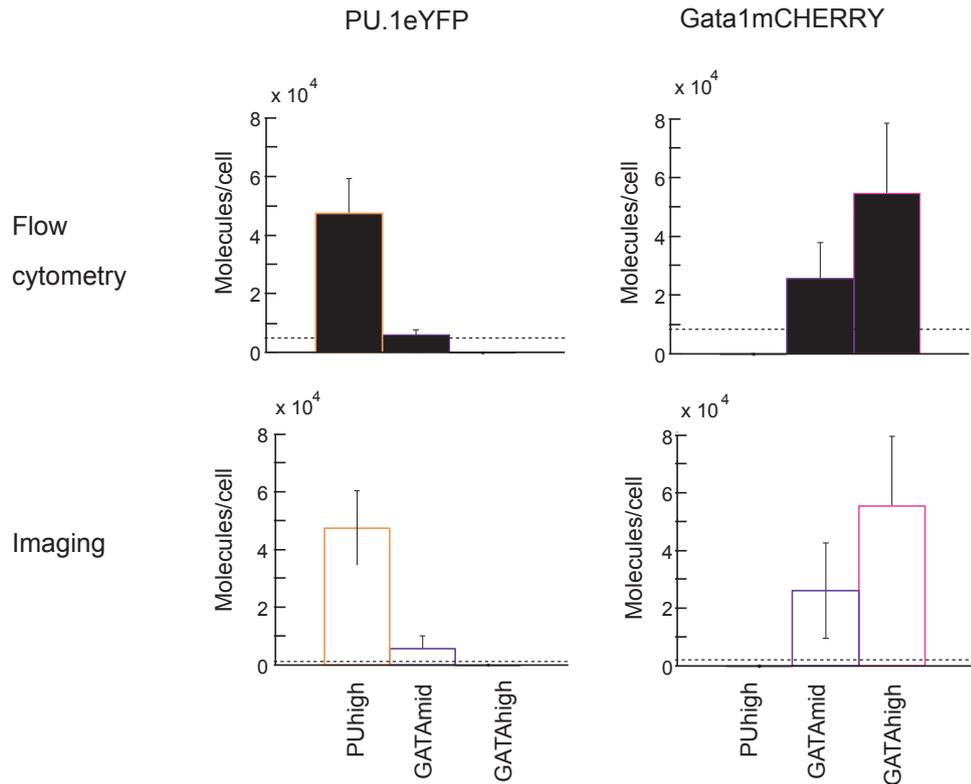


Figure 5.6: Estimation of molecule numbers per sorted cell of the designated population. Bars represent mean protein abundance per cell and error bars include uncertainty from western blot quantification and fold-changes of flow cytometry. Dashed lines represent negative gates estimated by control or *in silico* background cells. Imaging and flow cytometry are highly comparable in their estimated protein numbers. The negative gate in imaging is much lower compared to flow cytometry. Figure adopted from Hoppe et al., in review.

Cell population k	MFI M_k	Fold change f_k	Proteins per cell \bar{P}_k
PU.1 western (PW)	3077 ± 228	1	$43.1 \pm 10.6 \cdot 10^3$
PUhigh	3376 ± 320	1.10 ± 0.13	$47.3 \pm 12.9 \cdot 10^3$
GATAmid	425 ± 66	0.14 ± 0.02	$5.9 \pm 1.8 \cdot 10^3$
GATAhigh	—*	—	—
HSC	575 ± 19	0.19 ± 0.02	$8.1 \pm 2.1 \cdot 10^3$
LSK	1171 ± 54	0.38 ± 0.03	$16.4 \pm 4.3 \cdot 10^3$
GMP	3051 ± 279	0.99 ± 0.12	$42.7 \pm 11.7 \cdot 10^3$
MEP	—*	—	—
Negative gate (GMP)	318 ± 32	0.10 ± 0.01	$4.4 \pm 1.2 \cdot 10^3$
Negative gate (MEP)	337 ± 68	0.11 ± 0.02	$4.7 \pm 1.5 \cdot 10^3$

Table 5.4: PU.1eYFP protein numbers have been determined for PU⁺ LK progenitor cells (PW) with three independent Western Blots. Using the mean fluorescence intensity (MFI) over several independent flow cytometry experiments allows to map protein numbers from PW cells to every other population. Error bars are derived using error propagation and incorporate deviations PW cells and of standard deviations of MFIs. * = MFI is below negative gate which does not allow to reliably determine the number of proteins.

mCHERRY we used GATAhigh cells as reference (Table 5.2). Finally, we mapped HSCs protein numbers from flow cytometry to GMPs quantified by imaging (compare Table 5.5).

Importantly, the number of molecules determined for PU.1eYFP for GMPs in imaging ($40.1 \pm 4.7 \cdot 10^3$) compares well to the number of molecules in flow cytometry ($42.7 \pm 11.7 \cdot 10^3$). In general, fold-changes between flow cytometry and fluorescence imaging are comparable throughout all measured cell populations which indicates that a reliable quantification of protein molecules on a single-cell level can be well achieved using our technology.

5.6.4 Detection limit in protein numbers

We determined the detection limit of every movie by calculating *in silico* background cells (compare Chapter 4). Given this intensity value allows to calibrate the detection limit to protein numbers as we did before for individual cells (Table 5.5). Above this threshold we are sure that we detect real protein signal. Below this number we cannot distinguish if a cell has no protein or if some minute abundant proteins are still expressed. Although we are using different wavelengths (eYFP and mCHERRY) the detection limit is rather similar between PU.1eYFP ($1.1 \pm 2.0 \cdot 10^3$) and Gata1-mCHERRY ($1.9 \pm 4.4 \cdot 10^3$, Table 5.5). The negative gate is much lower for imaging for both measured fusion proteins which allows to detect positive protein expression much more sensitive compared to flow cytometry (nega-

Cell population	# PU.1eYFP	# Gata1mCHERRY
PUhigh	$47.3 \pm 12.9 \cdot 10^3$	—*
GATAmid	$5.5 \pm 4.4 \cdot 10^3$	$25.6 \pm 16.3 \cdot 10^3$
GATAhigh	—*	$54.6 \pm 23.8 \cdot 10^3$
Negative gate	$1.1 \pm 2.0 \cdot 10^3$	$1.9 \pm 4.4 \cdot 10^3$
HSC	$8.1 \pm 2.1 \cdot 10^3$	—*
GMP	$40.1 \pm 4.7 \cdot 10^3$	—*

Table 5.5: Protein numbers have been determined by flow cytometry (bold numbers) and mapped to different cell populations using their fold changes quantified by fluorescence imaging. * = average fluorescence signal is below negative gate which does not allow to reliably determine the number of proteins.

tive gate(eYFP)= $4.7 \pm 1.5 \cdot 10^3$, negative gate(mCHERRY)= $8.4 \pm 4.6 \cdot 10^3$), which will be beneficial to investigate potential co-expression in low abundances.

5.7 Tracking and quantification of differentiating HSCs until lineage marker expression

We performed experiments using true multipotent HSCs to identify a cell population which can differentiate into granulocyte/macrophage (GM) or megakaryocyte/erythrocyte (MegE) lineage (compare Figure 5.7B). Although about 50% of cells do not show any committed phenotype after 6 days counted by manual investigation, the development of GM as well as MegE committed cells could be observed in our culture conditions (Figure 5.7A). In addition to the two labeled transcription factors PU.1eYFP and Gata1mCHERRY we monitored another fluorescent label, which is bound to an antibody against CD16/32, which is a surface marker only expressed in GMPs (Akashi et al., 2000). The labeled antibody is added to the medium and accumulates at specific GMP receptor proteins at the cellular surface and thereby induces a fluorescent signal which is much brighter than the background (Eilken et al., 2011). We could show that Gata1mCHERRY is never coexpressed together with CD16/32 (Figure 5.7C) and Gata1mCHERRY positive cells give mainly rise to the MegE lineage (Figure 5.7D), which allows to use Gata1mCHERRY expression as a lineage readout. Finally, the morphology can support the determination of megakaryocytic cells (Figure 5.7E).

By applying TTT (Section 3.1.2) to long-term time-lapse microscopy movies of differentiating HSCs we were able to follow individual cells resulting in lineage trees up to 11 consecutive generations (Figure 5.7F). Next, we applied our toolbox QTFy (compare Section 3.3) and quantified PU.1-

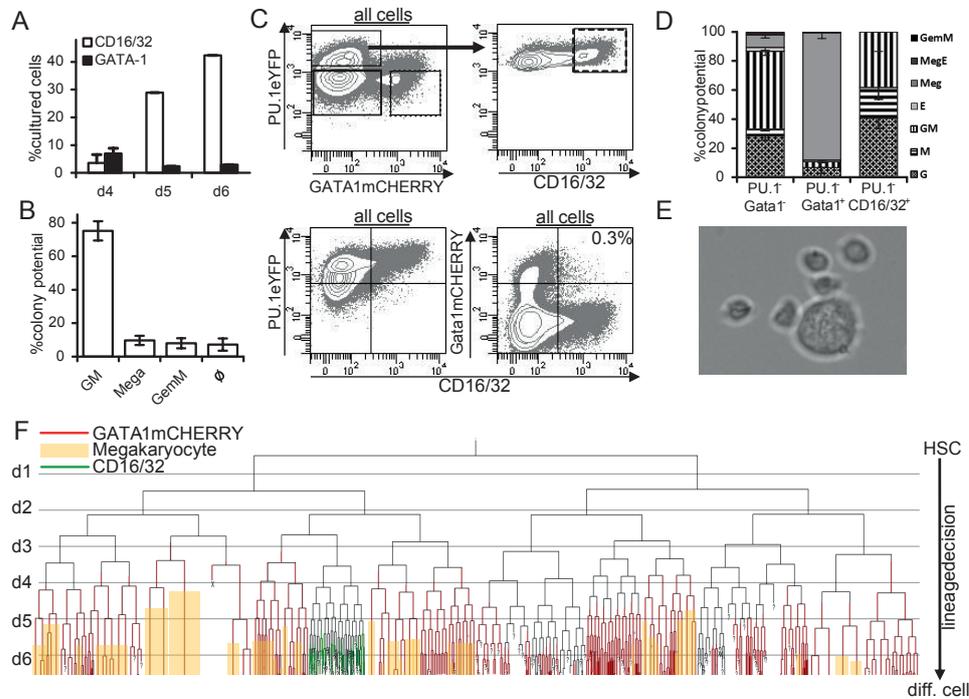


Figure 5.7: CD16/32 and Gata1 onset as well as morphology serve as reliable lineage output marker. (A) Percentage of cells showing expression of CD16/32 and Gata1mCHERRY of differentiating HSCs on days 4-6 of culture. Mean \pm s.d. with $n = 3$ (day 4), $n = 1$ (day 5,6). (B) Colony potential of differentiating HSCs on day 5-7 of culture. GM = colony containing CD16/32 expressing cells, Mega = colony containing megakaryocytes, GemM = colony containing CD16/32 expressing cells and megakaryocytes, \emptyset = colony containing neither CD16/32 expressing cells nor megakaryocytes. Note that most colonies still contained cells that did not express CD16/32 or Gata1mCHERRY. Data are mean \pm s.d ($n = 3$). (C) Sort of HSC progeny on day 4 of culture. (D) Colony potential of sorted cells (c). Data are mean \pm s.d ($n = 3$). (E) Morphology of megakaryocytes (big round cell) allows to reliably determine cell fate compared to other HSPC types (small round cells). (F) Single-cell genealogy of a differentiating HSC over 11 consecutive generations. Annotated are Gata1mCHERRY and CD16/32 expression read out by live-antibody staining. Additionally macrophages could be determined by morphology. Figure adopted from Hoppe et al., in review.

Experiment	1	2	3	Total
Track points	1,481,757	1,312,419	699,442	3,393,618
QTFy data points	83,462	72,604	54,668	210,734
Final data set	50,667	57,559	37,177	145,403

Table 5.6: Three independent long-term time-lapse microscopy experiments with differentiating HSCs have been tracked on bright field using TTT and quantified and manually inspected on fluorescence images using QTFy. For further analysis only cells which have at least 70% manually inspected data points are used. Every QTFied data point consists of at least two fluorescence quantifications.

eYFP and Gata1mCHERRY protein numbers from undifferentiated HSCs up to committed cells, determined by CD16/32 signal (Figure 5.8A), morphology (Figure 5.8B) or Gata1mCHERRY onset (Figure 5.8C). Whenever a cell exceeded twice the detection limit in the mCHERRY channel for more than 5 consecutive time points, the cell itself was annotated to have a Gata1mCHERRY onset and all its descendants were annotated as Gata1mCHERRY positive. We annotated all observed cell fates and continuously followed cells resulting in over three million data points arising from over 100 HSCs over three independent biological replicates which will be used for all further analysis (for a detailed data description see Table 5.6).

From all three million manually tracked data points, which are based on the frequently acquired bright field images, there remain about 210,000 data points which have fluorescence images which allows to quantify PU.1eYFP and Gata1mCHERRY expression. Since QTFy records which time-point has been inspected manually, we filter for cells which have at least 70% of all time points per cell manually inspected. We also manually exclude cells which are in regions within the well which show artifacts or are close to the boarder. We also manually investigated for every position if backgrounds and gain calculation depict any artifacts and recalculated them with optimized parameters, if necessary. For most of the following analysis we ignore the first cell of every tree since its time point of cell birth is unknown and use only cells which have been tracked until cell division. All in all, we end up in a carefully surveyed data set incorporating over 140,000 segmentation masks. Every data point consists of two fluorescence quantifications (PU.1-eYFP and Gata1mCHERRY), which are based on a segmentation performed on either one of both fluorescence channels, whichever shows a better signal. Furthermore, for every data point of each cell there exist additional features: cell area, position and time within the movie and a qualitative status about the morphology and CD16/32 expression.

With the carefully curated data set we are able to answer the following biological questions:

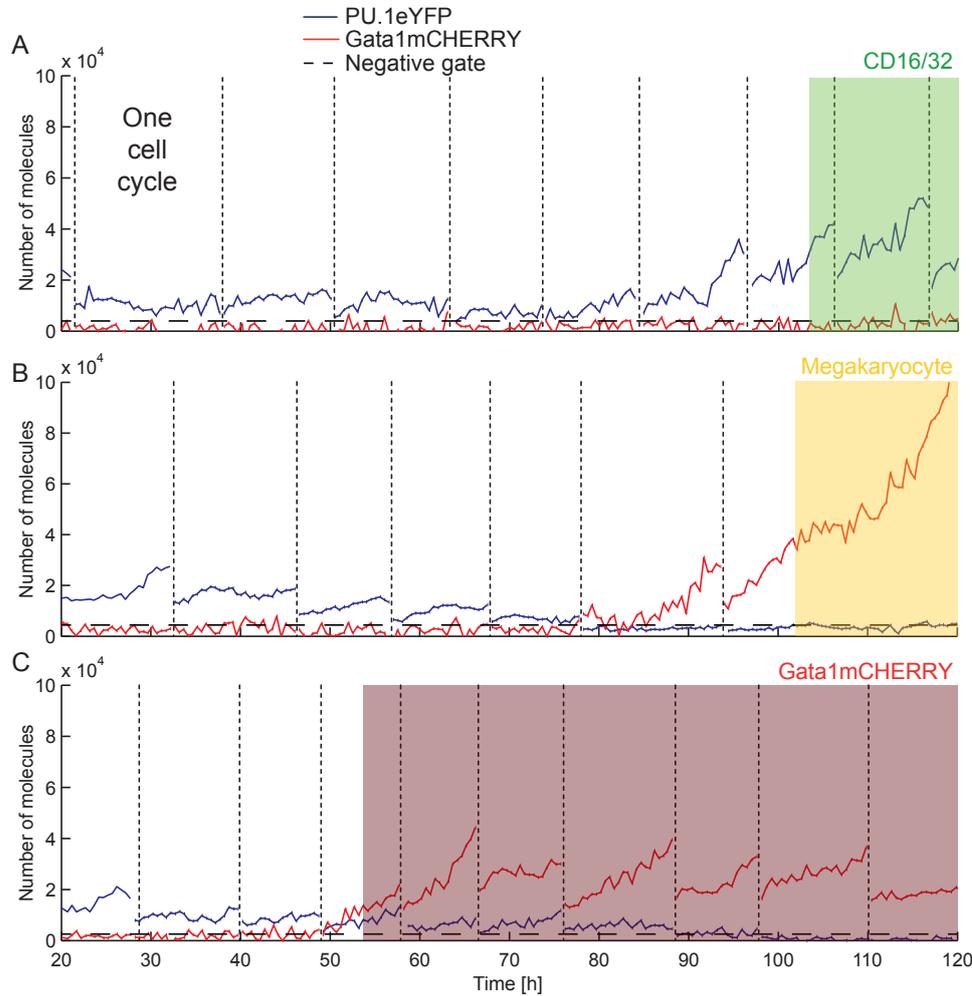


Figure 5.8: Exemplary single-cell time courses of PU.1eYFP and Gata1mCHERRY starting from freshly sorted HSCs up to lineage committed cell types. (A) Exemplary traces showing PU.1eYFP and Gata1mCHERRY of a branch that differentiates into GMP, determined by positive CD16/32 expression (green box). (B) Exemplary traces of a HSC giving rise to megakaryocytes, determined by morphology (yellow box). PU.1eYFP expression constantly decreases without any Gata1mCHERRY expression above the negative gate (horizontal dashed line) (C) Exemplary trace showing an early onset of Gata1mCHERRY, which stays positive throughout the whole branch (red box). PU.1eYFP is still expressed after onset, but decreases slowly and constantly. Figure adopted from Hoppe et al., in review.

- Is there an observable interplay between PU.1 and Gata1 during differentiation HSC to GM or MegE lineage?
- Is there a distinguishable subpopulation within differentiated GM or MegE lineages?
- How is the differentiation program for every lineage established? What dynamics can be observed on average?
- At what time point can GM versus MegE differentiating cells be distinguished?
- Do specific cell properties change during differentiation already before lineage marker onset?

5.8 Analysis of GM and MegE differentiating branches

By investigating PU.1eYFP protein abundance during differentiation into GM lineage, we observe increasing expression, as expected, but even before the onset of CD16/32 (Figure 5.8A). Importantly, by investigating 369 CD16/32 onset branches we observed almost exclusively only PU.1eYFP expression and only very rarely ($8 \pm 9\%$, $n=3$) Gata1mCHERRY expression slightly above the negative gate. By further manual investigation we could exclude every transient Gata1mCHERRY signal to origin from real cellular expression and prove them as artifacts. With this first results, we cannot fully exclude any transient double positive state but it indicates that a double positive state does not reflect the major GM lineage differentiation process.

Interestingly, the PU.1eYFP levels at CD16/32 onset was very heterogeneous as already indicated by flow cytometry (Figure 5.3A) and spans in imaging a dynamic range of 16.44 ± 0.65 dB ($n=3$, Figure 5.9A).

By investigating the GM differentiation in more detail we found that $45 \pm 8\%$ ($n=3$) of all branches never drop below their initial PU.1eYFP level over generations before CD16/32 onset (Figure 5.9B, black lines). The other half of differentiating branches shows a decrease of PU.1eYFP in various pattern but increase again before CD16/32 can be observed (Figure 5.9B, red lines). Cells which do not drop below initial PU.1eYFP levels seem to be already decided and show a straight differentiation process where Gata1mCHERRY is not involved at all. Given a detection limit of about 1,900 Gata1mCHERRY molecules (compare Table 5.5) brings cells which drop below the initial HSC level of about 8,100 PU.1eYFP molecules into a regime where a potential equilibrium could exist (Figure 5.9C). However, out of all investigated GM branches only $31 \pm 8\%$ ($n=3$) reach this state. This renders Gata1mCHERRY not to be a major antagonist of PU.1eYFP during the GMP differentiation.

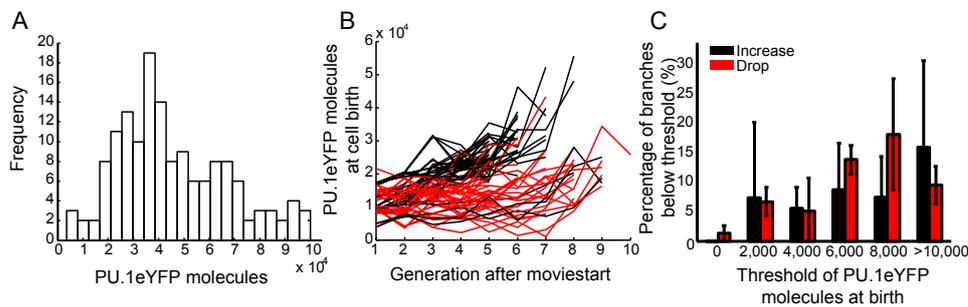


Figure 5.9: GM differentiation shows heterogeneous behaviors of PU.1eYFP expression. (A) PU.1eYFP signal at CD16/32 onset is very heterogeneous and spans a dynamic range of 16.44 ± 0.65 dB ($n=3$). One representative example shown including 156 onset cells. (B) $45 \pm 8\%$ ($n=3$) of 369 traces show that they never drop below their initial starting level of PU.1eYFP at cell birth over generations. Only a minority of these cells start with an initial level of PU.1eYFP below 8,000 molecules. The other half of all traces shows a decrease of initial PU.1eYFP followed by an increase before CD16/32 onset (one representative experiment shown). (C) Only a few branches drop to PU.1eYFP levels below initial HSCs level of 8,100 PU.1eYFP molecules where a potential equilibrium of PU.1eYFP and Gata1-mCHERRY could exist, given the detection threshold of $1.9 \pm 4.4 \cdot 10^3$ Gata1-mCHERRY molecules.

Next, we investigate cells which differentiate into the MegE lineage. The corresponding branches also show diverse behaviors (compare examples in Figure 5.8B and C) but have one major astonishing property in common: Whenever a Gata1mCHERRY onset has been detected in a cell (367 investigated branches), all its progeny never shows a CD16/32 onset, which is in accordance with the previous colony assay analysis (Figure 5.7D). This also confirms Gata1mCHERRY to be a good readout for MegE lineage decision.

By further investigating MegE differentiation trees, we selected branches which reach Gata1mCHERRY abundance of at least 13,200 molecules for more than five consecutive time points. This threshold refers to the lower Gata1mCHERRY bound of GATAmid cells in flow cytometry (mean minus one standard deviation, compare Table 5.2, see Figure 5.10A). Additionally, we defined a gate specific for HSCs, which is the 25 and 75% quantile of the PU.1eYFP and Gata1mCHERRY distribution of cells of the second and third generation, respectively. For every branch we calculate the transition time t between the last time point within the HSC gate and the first time point entering the GATAmid gate (Figure 5.10B). We detected that a broad distribution of transition times arises (Figure 5.10C), which is comparable between all three replicates. We fitted Gaussian mixture models with one, two or three components (see Figure 5.10C) and calculated the Bayesian information criteria (BIC, compare Section 2.5.1) (Schwarz, 1978) for every model. A BIC validates the fitted model in comparison to the underlying data by the log likelihood function and additionally penalizes the degrees of freedom as well as the number of data points. For the transition times the BIC always favored a mixture model with two components for all three replicates. We calculated the split point ($s = 20 \pm 6 \text{ h} \approx 1 - 2$ generations, mean \pm s.d., $n=3$) of each experiment by the intersection of the two Gaussian distributions weighted by their mixture properties (green vertical line in Figure 5.10C).

By filtering on all cells which transit faster than the split point ($s < t$) we see that these cells simultaneously downregulate PU.1eYFP and upregulate Gata1mCHERRY (Figure 5.10D). On the other hand out of all branches $37 \pm 3\%$ ($n=3$) reach the GATAmid gate slower than their respective split point ($s > t$) and first downregulate PU.1eYFP and subsequently upregulate Gata1mCHERRY (Figure 5.10E). As we have shown previously GATAmid cells will differentiate into MegE lineage. We investigated further features of the two distinct differentiation programs and found that in two out of three replicates, slow differentiating cells became primarily Megakaryocytes determined by manual inspection of representative samples. Furthermore, the lifetime of the slow differentiating cells which enter the GATAmid gate was significantly higher (Wilcoxon rank sum test p-value < 0.001 , $n=2$), which is typical for Megakaryocytes (compare Figure 5.7F and Figure 5.8B). However, the third experiment did not show any difference in lifetime (p-value = 0.69) or cell fate differences between the two groups and needs further inves-

tigation. Notably, this particular experiment includes the large tree shown in Figure 5.8F, which inherits half of all investigated Gata1mCHERRY onsets in this experiment, which might introduce a bias. Additionally, this experiment shows only few quantified Megakaryocytes. All investigated differentiating branches show no clear influence of PU.1eYFP, but rather follow a given program instead of suggesting an interplay between both transcription factors.

5.9 Population analysis of GM and MegE differentiation

Due to the highly heterochronous lineage marker onsets (Figure 5.11A) the average transcription factor dynamics cannot be observed. By synchronizing these onsets, we can investigate the dynamics before, during and after the onset using an absolute time scale. Second, by the synchronization investigation of a specific time point before a lineage marker onset showing a clear distinction between GM and MegE differentiation becomes possible.

Here, we utilized the tracking information to shift individual branches in time. Due to the complex tree structure and the fact that measurements become more frequent in the later movies due to cell division it is not trivial to cut trees into individual branches and to shift them in time, especially if they have multiple non-synchronized onsets even of different lineages (Figure 5.11A). The naive approach takes complete branches from the HSC until marker onset. However, this allows cells to appear multiple times in the data set, which especially holds true for early cells in the movie introducing statistical artifacts. We therefore developed a method which randomly cuts a tree into branches, each containing one marker onset and thereby excluding complete subtrees, which do not show any marker onset at all ($40 \pm 3\%$, $n=3$, compare Figure 5.7A). We shift all branches such that the marker onset is always at time $t = 0$ (Figure 5.11B). Multiple randomization of the branch cutting procedure showed that the average behavior of PU.1eYFP and Gata1mCHERRY in the following analyses did not show any obvious differences.

For the shifted and separated GM and MegE differentiating sub branches we calculated a moving density estimate on PU.1eYFP and Gata1mCHERRY concentration (compare Section 3.3.3) to visualize the average behavior in hours before lineage marker onset (Figure 5.12A and B). The general differentiation program of GM differentiating cells seems to split from 60 to 40 hours before CD16/32 onset which is in accordance with our previous findings (compare Figure 5.9) referring to either a constant increase of PU.1eYFP or slight dropping behavior which finally increases to reach GMP levels. The heterogeneous distribution at marker onset ($t=0$) can be well compared to previous findings (Figure 5.9A). Note, that a typical PU.1eYFP

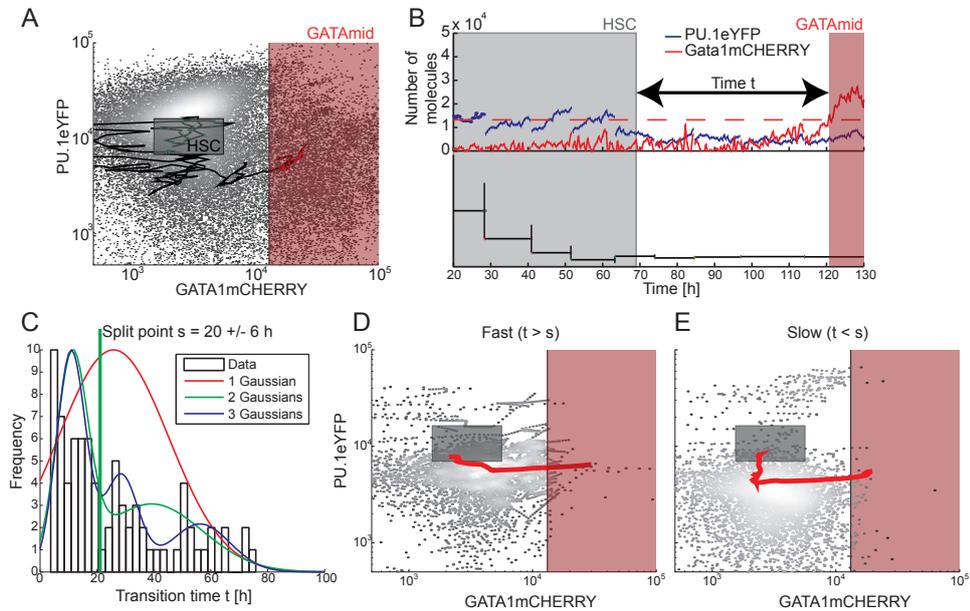


Figure 5.10: MEP differentiation shows two distinct behaviors of transcription factor dynamics. (A) A typical HSC gate (gray box) has been drawn into a scatter cloud showing all measured time points of one exemplary experiment. The density is encoded in the brightness of the dots. We selected branches which reach the $\text{Gata1mCHERRY}^{\text{mid}}$ gate (red box) and followed their time courses between the two gates (black lines shows one exemplary trace). (B) One exemplary branch shows an upregulation of Gata1mCHERRY , which reaches the threshold of 13,200 Gata1mCHERRY molecules (red dashed line). We measure the time t between the first contact with the $\text{Gata1mCHERRY}^{\text{mid}}$ gate and the last time point within the HSC gate. (C) Transition times t between the gates are broadly distributed. Based on BIC ($n=3$), a Gaussian mixture model with two components outperforms the other models (1 or 3 components). The vertical green line illustrates the split point s of the two normal distributions which separates the data into two groups. (D) The fast fraction ($t < s$) of cells simultaneously downregulates PU.1eYFP and upregulates Gata1mCHERRY leaving the HSC gate and entering the $\text{Gata1mCHERRY}^{\text{mid}}$ gate. (E) Cells which take longer ($t > s$) from the HSC to the $\text{Gata1mCHERRY}^{\text{mid}}$ gate show an initial downregulation of PU.1eYFP followed by Gata1mCHERRY upregulation.

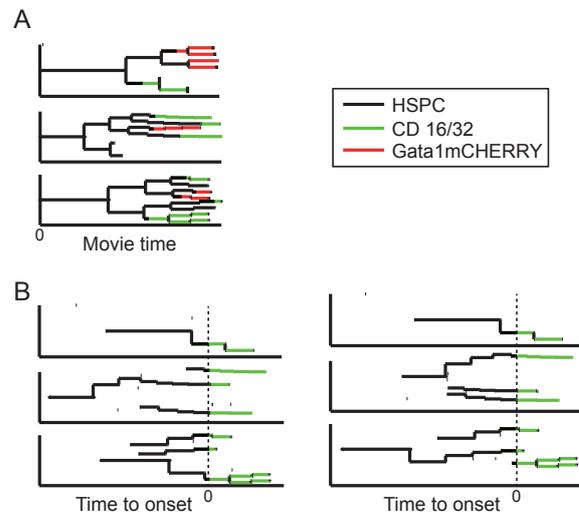


Figure 5.11: The branch cutting method allows shifting individual sub-branches in time. (A) Three example trees showing non-synchronized multiple CD16/32 and Gata1mCHERRY onsets. There are various possibilities to cut the tree in individual branches since onset cells might share the same progenitor cell. (B) We excluded all cells which do not show CD16/32 onset and randomly cut the tree such that every cell can occur only once in the data set (two realizations shown). Shifting all branches to the same onset time allows to further investigate the general protein dynamics for this specific lineage.

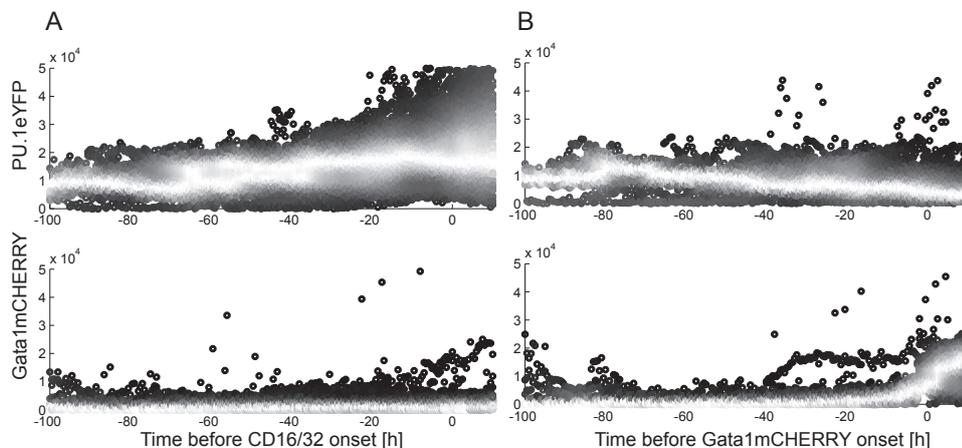


Figure 5.12: Average transcription factor dynamics before lineage marker onset illustrates the detachedness of Gata1mCHERRY and PU.1eYFP (one representative example shown). A moving local density estimate is encoded in the brightness of the spots. (A) PU.1eYFP signal before CD16/32 onset seems to drop before increasing to final GMP levels. (B) PU.1eYFP signal already slowly decreases over time until final Gata1mCHERRY onset can be observed.

abundance of $40.1 \pm 4.7 \cdot 10^3$ (compare Table 5.5) is not reached on average since we recalculated PU.1eYFP time courses in concentration (compare Section 3.3.3). Gata1mCHERRY signal is does not show any clear function in these average time courses besides some outliers.

The average behavior before Gata1mCHERRY onset shows a clear down-regulation of PU.1eYFP (Figure 5.12B). By fitting a linear regression into PU.1eYFP abundances after cell birth over generations, we found a negative slope in the majority of all branches ($78 \pm 20\%$, $n=3$) indicating that the lineage decision has already been made and that PU.1eYFP acts independently of Gata1mCHERRY. The onset of Gata1mCHERRY shows a very steep and fast onset on average, but becomes very heterogeneous within a few hours, originating from the two different differentiation program we described earlier (compare Figure 5.10). Without the synchronization of the onsets these dynamics would be hidden in the data.

To investigate when a difference between the average GM and MegE can be observed, we combined GM and MegE branches anchored to the marker onset ($t = 0$ h), which showed an initial cloud of multipotent cells 100 hours before marker onset ($t = -100$ h) centering around the typical gate for HSCs (one realization given in Figure 5.13). Over time, the cloud hardly shows any change from about 100 hours to 60 hours before lineage marker onset (i.e. CD16/32 and Gata1mCHERRY). About $t = 60$ to $t = 40$ hours before marker onset the cloud broadens and slowly two sub-populations arise

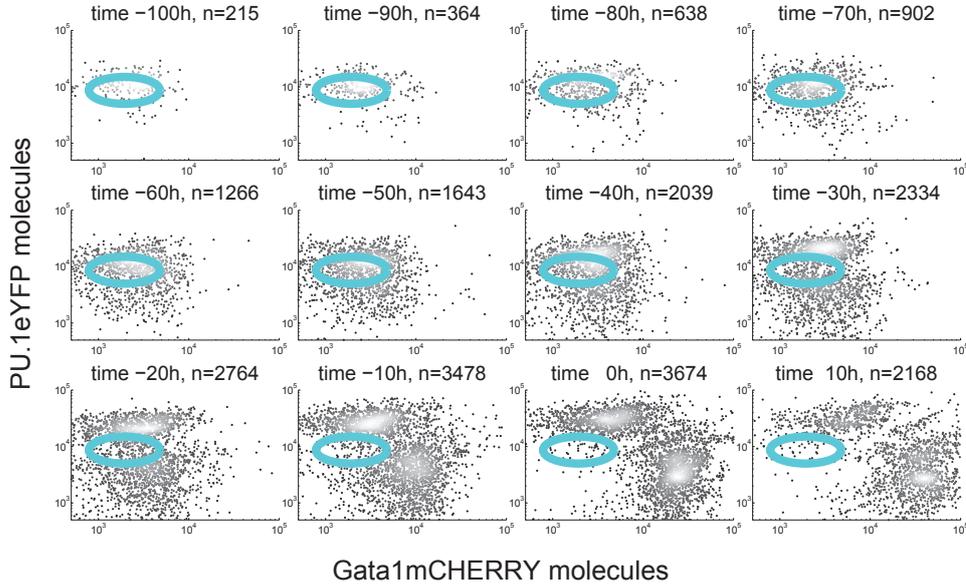


Figure 5.13: Time resolved density scatter plot of PU.1eYFP and Gata1mCHERRY of all tracked cells that had a defined fate (CD16/32 or Gata1mCHERRY expression) within a 10hour moving window, normalized to the onset ($t = 0$ h) of CD16/32 or Gata1mCHERRY, respectively. The cell density is encoded in the brightness of each dot. The blue circle marks the typical initial HSC gate. Figure adopted from Hoppe et al., submitted.

slowly approaching regions, which are typical for GM or MegE differentiated cells. GM differentiating cells just increase their PU.1eYFP level, whereas MegE cells move towards their PU.1eYFP^{low}Gata1mCHERRY^{high} state. Compared to the time scale of the whole movie (≈ 6 days) the transition happens quite fast (20 hours, that is, 1 to 2 generations). Once more, no transient double positive state can be observed and only very rare events show up in the upper right quadrant, which are probable artifacts. In the visualization here, we used absolute abundances and not concentration to be more comparable with flow cytometry resulting in a close approximation of initial flow cytometry plots (compare Figure 5.3). This results was highly unexpected since flow cytometry was based on freshly sorted HSPCs and our imaging results are based on *in vitro* differentiation of HSCs in time-lapse experiments. By the synchronization of trajectories we identified that differentiation programs can be distinguished 40 hours before lineage marker onset.

To investigate the time point to distinguish differentiation programs from a different perspective we represent both transcription factors simultaneously in an alternative visualization inspired by the Waddington landscape

(Goldberg et al., 2007, Waddington, 1957). To describe the landscape in the classical way, either static data (Figure 5.3) or the projection of all data within one time-lapse experiment (Figure 5.10A) should be used to visualize the parameter space (i.e. PU.1eYFP and Gata1mCHERRY). However here, we use time on one axis, which drives the development, illustrating a trajectory space or time space. So far, landscapes have been based on simulated dynamics (Wang et al., 2011, Zhou and Huang, 2011), but also in terms of stem cell differentiation (Graf and Enver, 2009, Heinäniemi et al., 2013). Here, we are able to describe a landscape of CMP differentiation based on real data (Figure 5.14). To reduce the number of dimensions, we show the ratio of Gata1mCHERRY and PU.1eYFP over time before lineage marker onset. In accordance with previous findings all cells seem to behave equally, up to ≈ 30 hours before marker onset. Then the valley splits into two valleys representing a high and a low ratio of PU.1eYFP to Gata1mCHERRY (i.e. GMPs and MEPs). Additionally one can see that the GMP valley is deeper and broader with the higher abundance of GMPs compared to MEPs (compare Figure 5.7A). When looking at specific ratios over time (i.e. we investigate a cut along the time axis), we see that a ratio typical for HSCs shows a high frequency of cells long before the lineage marker onset (Figure 5.14B), which gets depleted starting ≈ 20 hours before marker onset. A population with a transient high abundance of cells can be observed for an equilibrium between Gata1mCHERRY and PU.1eYFP in the time from 20 hours before onset until the lineage marker onset, originating from MEP committing cell types (Figure 5.14C). Cells which have a clear tendency toward Gata1mCHERRY can only be observed at time shortly before and with positive marker expression, representing MEP differentiated cells (Figure 5.14D). In accordance with the previous analysis, only by synchronizing the branches to the marker onset the distinct differentiation programs can be characterized before during and after the onset.

5.10 PU.1 dynamics before lineage marker onsets

To investigate more precisely if a differentiation program already starts before any cell fate marker onset can be observed, we looked at PU.1eYFP within individual cells more closely. We calculated the fold-change of PU.1eYFP within one cell-cycle, that is the fold-change of the median of the last three time points divided by the median of the first three time points. The average fold change should center around two to maintain protein concentrations constant. However, since we use the median of the last and first three time points we will always slightly underestimate the fold-change.

We investigate the PU.1eYFP fold-change within one cell-cycle for many cells over generations before Gata1mCHERRY onset (Figure 5.15A). In general, the fold-change is always clearly below 2 indicating an overall decrease

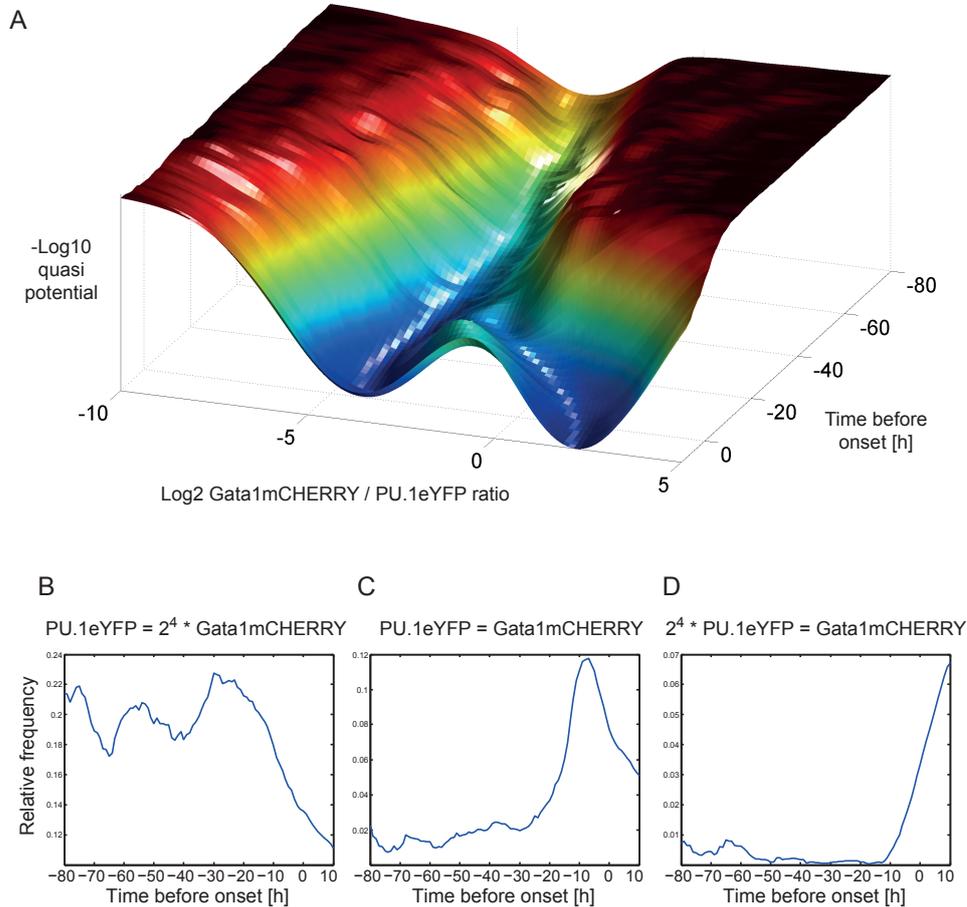


Figure 5.14: Illustration of the ratio between \log_2 Gata1mCHERRY and PU.1eYFP over time before lineage marker onset. (A) The time linearly influences the quasi-potential of the time-shifted branches and deepens the valleys illustrating their nonreversible differentiation process. The left valley refers to GMP differentiation showing a high PU.1eYFP expression. The second valley which emerges as subbranch illustrates MEP differentiation. (B) The relative frequency over time for one specific ratio, typical for HSCs, shows a decrease of this pool 20 hours before lineage marker onset. (C) The equilibrium of PU.1eYFP and Gata1mCHERRY is only transiently maintained by cell differentiating into MegE lineage. (D) A clear over representation of Gata1mCHERRY is only accomplished by cell which become MEPs.

of PU.1eYFP already starting 9 generations before Gata1mCHERRY onset. We applied a two sample t-test to analyze whether one of the generations shows significant differences to any other generation (Figure 5.15C). After correcting for multiple testing by Bonferroni correction we did not detect any significant differences between every generation combination.

Similarly, we looked at PU.1eYFP fold-change within one cell-cycle before CD16/32 onset (Figure 5.15B). 9 to 4 generations before CD16/32 onset an overall decrease of PU.1eYFP concentration can be observed, which switches for generation 3 to 1 before and at CD16/32 onset to an increase of PU.1eYFP concentration. When testing for significant differences between the generations, we detected that two blocks emerge (Figure 5.15D). The first block including early generations up to generation 3 did not show any significant difference within the generations, but showed significant differences to generations -2 and -1. Furthermore the block of generation 0 to -3 is not significantly different within it. This hints to a marker delay of 2 to 3 generations after the lineage decision has been made.

Next, we tested if significant differences in PU.1eYFP fold-change within one cell-cycle exist between generations of differentiating MEPs and GMPs (Figure 5.15E). Interestingly, we did not detect any significant differences in the early generations, indicating that cells seem to behave comparably, independent of their future cell fate. Between 4 and 5 generations before the CD16/32 onset, a certain change in PU.1eYFP production can be observed leading to significant changes for PU.1eYFP fold-change. Similarly, the fold-change before Gata1mCHERRY onset changes allowing to detect significant changes to earlier generations. The generations right before lineage marker onset shows highly significant changes between both differentiating lineages as expected.

These results illustrate that early generations behave similarly and are not yet lineage decided. Independent of CD16/32 onset PU.1eYFP seems to be upregulated at least 3 generations before a definite cell fate, hinting at a lineage marker delay (compare Figure 5.8A). For MEP differentiation, we also observe a change of PU.1eYFP fold-change which happens 3 generations before any Gata1mCHERRY expression can be observed demonstrating once more the independence of both transcription factors.

In a recent study of Kueh et al. (2013) the mechanistic principles of PU.1 increase of macrophage differentiation was investigated. The authors asked, if this increase originates from an increased protein production or if an elongation of the cell-cycle suffice to accumulate higher protein abundances. Our experiments lack of an appropriate marker for macrophages, but we are able to investigate cell-cycle times from HSCs until GMPs. In contrast to the published study we find the opposite resulting in a significant constant decrease of GM differentiating cells compared to HSCs (Figure 5.16, after Bonferroni correction all p-values 7 generations before and 1 generation after CD16/32 onset < 0.01). Since we observe a indistinguishable mixture of

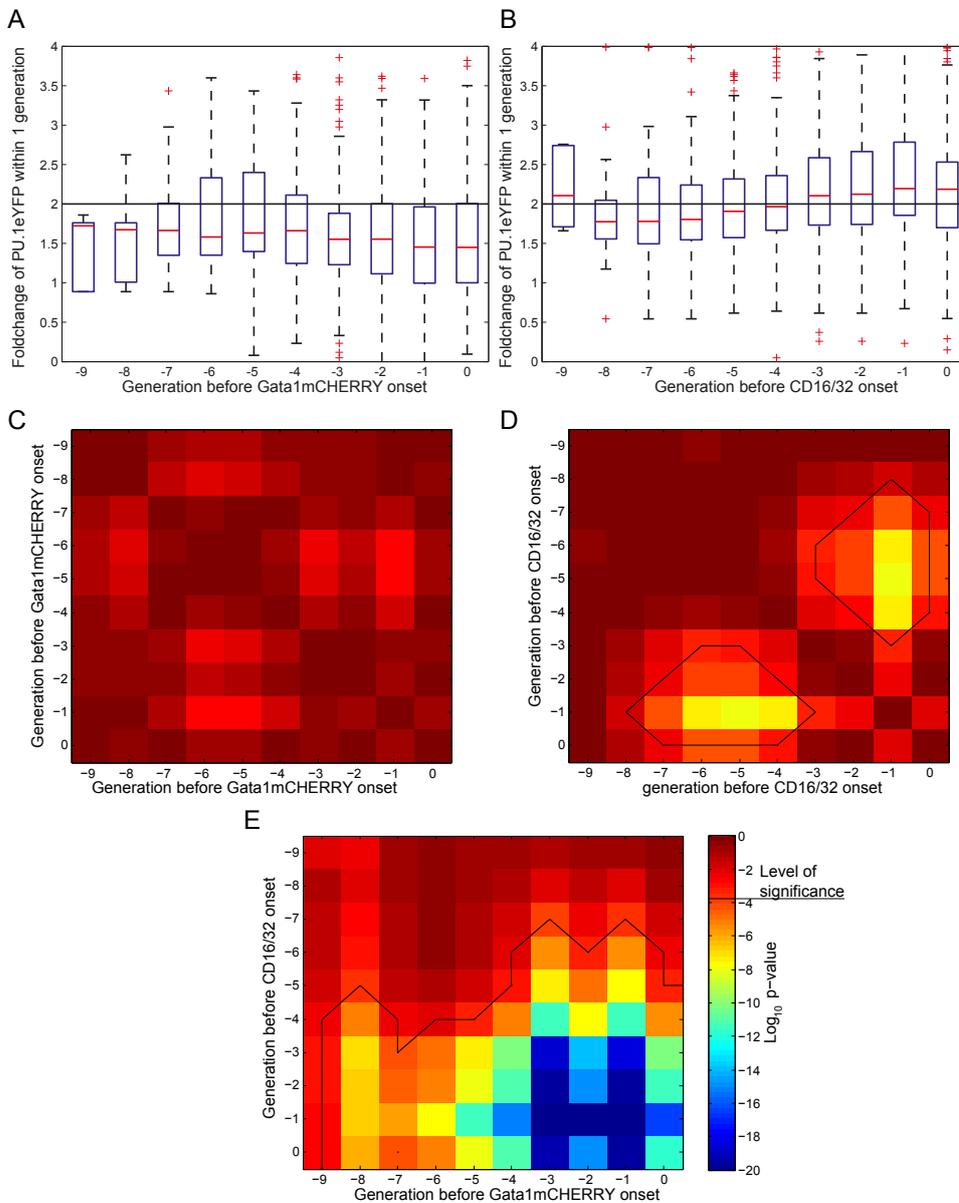


Figure 5.15: The fold change of PU.1eYFP within one cell-cycle is comparable in early generations independently of their later cell fate until five generations before lineage marker onset. (A) The fold-change of PU.1eYFP within one cell-cycle for cells becoming Gata1mCHERRY positive is constantly below 2 resulting in an overall downregulation over generations. (B) The median fold-change of PU.1eYFP of cells which become CD16/32 positive is below 2 up to four generations before marker onset indicating an overall drop in PU.1eYFP abundance. Three generations before CD16/32 onset PU.1eYFP gets upregulated to reach common GMP protein levels. (C) A two-sampled t-test did not detect any significant difference between every generation from (A). (D) A two-sampled t-test discovered two emerging blocks. Note that (C) and (D) are symmetric. (E) PU.1eYFP fold-change within one cell-cycle is not significantly different for generations up to 5 generations before lineage marker onset (upper left part). The difference in PU.1eYFP fold-change in generations near the marker onset is highly significant. Significant differences are surrounded by black polygons.

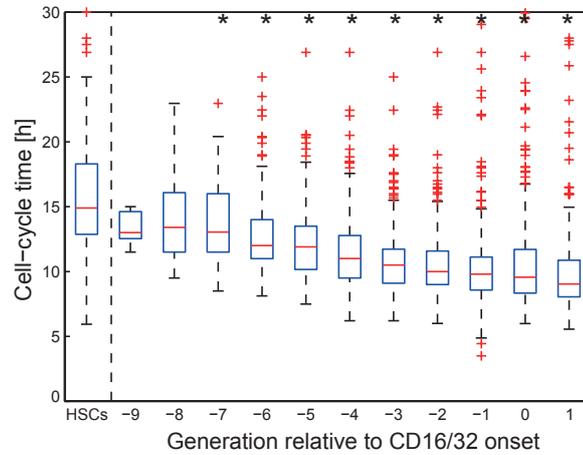


Figure 5.16: Cell-cycle lifetime of GM differentiating cells decreases constantly, which is in clear contrast to Kueh et al. (2013). Asterisks indicate a significant difference to the reference cell-cycle lifetime distribution of HSCs determined by a Wilcoxon rank sum test (p-value <0.01).

granulocytes and macrophages, further investigations are necessary if cells having a longer cell-cycle time mature to macrophages.

5.11 GemM trees show no interdependence of PU.1 and Gata1

To investigate PU.1eYFP and Gata1mCHERRY interplay on a tree basis we analyze multipotent trees, giving rise to GM as well as MegE cells. These trees show no discernible difference in the PU.1eYFP and Gata1mCHERRY protein dynamics in cells where the lineage decision potentially could take place (Figure 5.17). We find trees where two clear sub trees emerge each of them giving rise to one or the other lineage, but we also find trees with more mixed outcomes. This illustrates the heterogeneity within the HSC starting population as well as the heterogeneity of the differentiation programs. The first cell exclusively gives rise to a lineage must not necessarily make the lineage decision, also its two daughter cells potentially can independently make the decision. Based on our data we cannot exactly pinpoint the time point of decision, but independent decisions become more and more unlikely in later generations. However, also later generations show no clear PU.1eYFP or Gata1mCHERRY expression pattern which could be linked to lineage outcome.

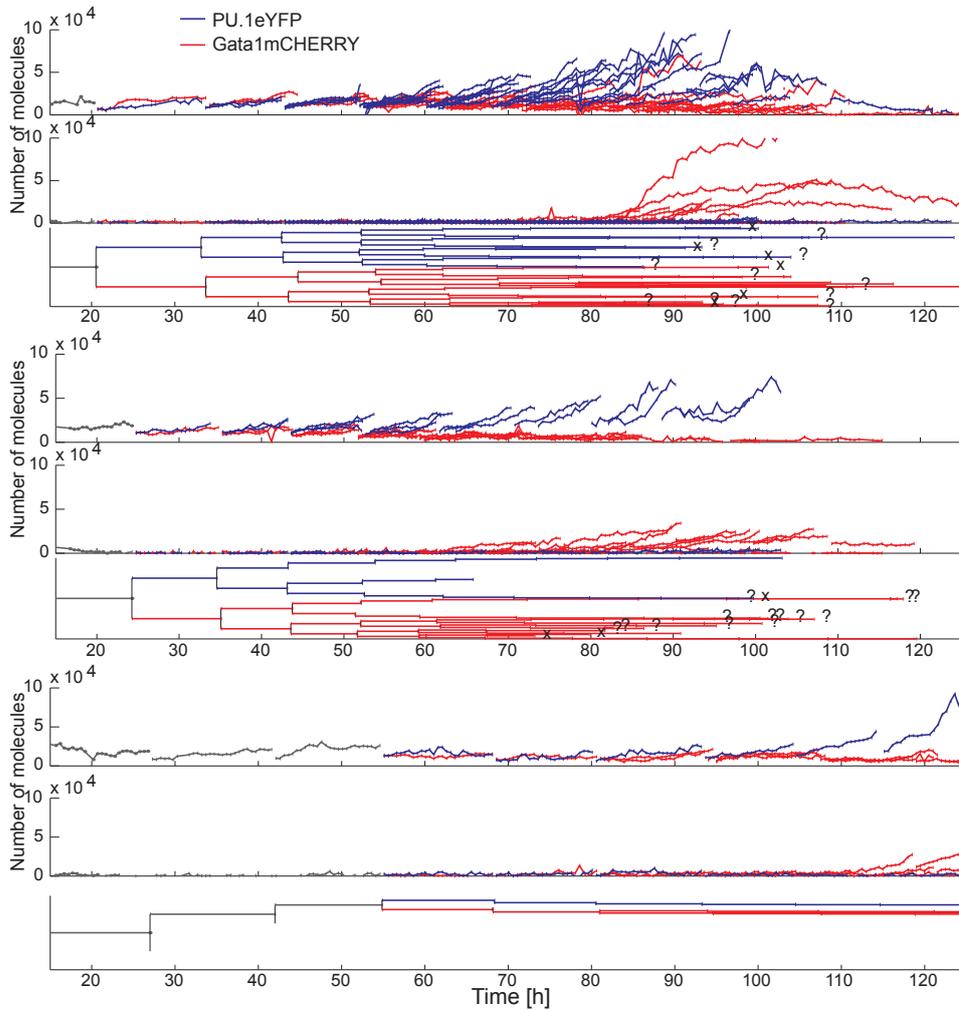


Figure 5.17: Single HSCs give rise to both lineages, GM and MegE. The cells which are the most probable decision makers are highlighted and do not show any differences in PU.1eYFP or Gata1mCHERRY expression pattern in this generation. Also later generations show no clear behavior which could be linked to later cell fate. Cell death and tracking loss events are marked by 'x' and '?', respectively.

5.12 Discussion

We investigated a double-knock-in mouse model comprising two reliable and faithful reporters for PU.1 and Gata1 which does not show any altered phenotype. Together with our newly developed hard- and software it allows for the first time to follow individual HSCs during myeloid differentiation *in vitro* on a long time scale in a non-invasive manner. We compared image quantification to well-established flow cytometry data and achieved a higher dynamic range and sensitivity with our technology. Previously published experiments showed a dynamic range of about 7 dB for PU.1 and Gata1 in mRNA experiments throughout all blood cell types (Seita et al., 2012) whereas we show a dynamic range of about ≈ 18 dB for both transcription factors on protein levels. This higher variability at the protein level, which is probably ignored by the mRNA level, highlights the relevance to investigate protein dynamics.

To investigate the stoichiometric interplay of both transcription factors simultaneously, we developed a method to reliably determine the number of molecules for both factors based on Western Blot analysis. Additionally, we mapped these numbers to flow cytometry and imaging fluorescence intensities, which allowed to follow protein abundances over time. To confirm our estimated abundances, different experimental (Heinrich et al., 2013) or theoretical (Komorowski et al., 2010, Rosenfeld et al., 2005) approaches to estimate the number of molecules should be applied.

We investigated how freshly isolated HSCs develop *in vitro* with time-lapse microscopy experiments. Interestingly, we found that Gata1mCHERRY always determines the lineage decision into MegE direction whenever Gata1mCHERRY exceeded the negative gate. On the other hand, every GM committed branch showed no Gata1mCHERRY expression at all. A potential state, expressing both lineage transcription factors in comparable means could exist in cells which decreased the PU.1eYFP level to levels equal to the Gata1mCHERRY detection threshold. Only this state would be applicable to describe the common model of a bi-stable switch. Since this state was only reached very rarely we exclude that the stochastic interplay between the two transcription factors plays a major role in GMP differentiation. Therefore, all current models and literature based on this theory have to be carefully revised.

The two distinct GMP differentiation patterns could emerge from different processes which differentiate into either one of their mature cell blood types, but are indistinguishable here, due to the lack of additional markers. A further investigation based on more sophisticated features such as morphology quantification could unravel if such differences exist.

We showed that two well-defined differentiation programs exist for the MegE lineage. Interestingly, we found a fast differentiating cell type, which could be potentially associated with the reported shortcut of HSCs towards

MEP differentiation (Adolfsson et al., 2005). In two out of three experiments we could well identify that slowly differentiating cells have a Megakaryocytic cell fate. This process has to be investigated further by mapping the two found groups onto the tree and imaging data to validate these findings by morphology or further tracking.

By integrating the whole tracking information, we shifted differentiating cells with known outcome in time and visualized their general transcription factor dynamics. Interestingly, this *in vitro* differentiation based on time-lapse imaging perfectly fits to flow cytometry data of *ex vivo* sorted HSPCs of bone marrow.

In a recent publication, the authors asked the question whether the increase in PU.1 during GMP differentiation originates from an increased protein production or an elongation of the cell-cycle times (Kueh et al., 2013). Here, we found that the PU.1eYFP fold-changes within one cell-cycle increase towards GM lineage, but Kueh et al. (2013) found an elongation of the cell-cycle times of differentiating macrophage progenitors. However, in our analysis we observed the opposite resulting in significantly decreasing cell lifetimes of GM differentiating cells (Figure 5.16). Together with the increased PU.1 fold-change within one cell-cycle we claim that the increased protein production is leading the accumulation of PU.1 abundances. A potential third scenario hypothesizing an increased PU.1 protein stability during GM differentiation to reach higher PU.1 abundances is discussed later in Chapter 7.

Chapter 6

Network heterogeneity of pluripotency transcription factors in embryonic stem cells

Methods, results and figures in this section are partly based on Filipczyk et al., in review and Schwarzfischer et al., in review. All biological experiments have been performed by Adam Filipczyk and are not part of this thesis.

My contributions are:

- Development and application of QTFy to mESC data
- Validation of cell lines and validation of quantification technology
- Analysis of quantified time-lapse data
- Analysis and statistics of endpoint staining data
- Spatial analysis

In the following section we apply QTFy (compare Section 3.3) on the data of murine ESCs which incorporate VENUS as fusion proteins to visualize key regulator proteins (Nanog and Oct4) and their single-cell dynamics. After demonstrating the normal behavior of the knock-in cell lines, we illustrate the power of our toolbox which can be used to track and quantify complete genealogies for up to eight generations on a single-cell basis. We evaluate our quantification technique by comparing intensity distributions to flow cytometry results. We discovered two distinct subpopulation emerging from NanogVENUS low expressing cells. First, a subpopulation which stably represses NanogVENUS expression and second, a subpopulation which activates NanogVENUS. However, both subpopulations are still positive for

pluripotency factors Oct4, Sox2 and Rex1. On the other hand, these two populations can be well separated by their correlation structure between measured transcription factors. Finally, we show that these networks look constant between subpopulations on a population scale, but may vary on a colony basis.

6.1 Biological background: embryonic stem cells

Classically, it was assumed that in conditions sustaining pluripotency the three core network players, Oct4, Sox2 and Nanog, were expressed homogeneously (Chambers et al., 2003, Mitsui et al., 2003). However, Nanog has been shown to inherit a high heterogeneity and there exist cells which barely express any Nanog protein or show high amounts of them. Yet these cells stay pluripotent and are defined as ESCs (Chambers et al., 2007). It was suggested that cells can travel between Nanog high and low states where a transient low state is referred to a window of opportunity to differentiate whereas the high state is referred to the true pluripotent state supporting self-renewal (Chambers et al., 2007). Importantly, it has been shown that the characteristic expression distribution will be reproduced from low Nanog cells as well as high Nanog cells (Kalmar et al., 2009). Since then, Nanog has gained great interest and biologists tried to understand the heterogeneity and its role in the pluripotent state. One main question arose whether the the heterogeneity is just irrelevant molecular noise or if these fluctuations are important regulatory mechanisms (Garcia-Ojalvo and Martinez Arias, 2012, MacArthur et al., 2009). Furthermore, there exist contradictory data on the molecular details it self. On the one hand, a self activating auto-regulatory loop has been affirmed for all three core network players (Boyer et al., 2005), but on the other hand auto-repression has been reported for Nanog and Oct4 (Fidalgo et al., 2012, Navarro et al., 2012, Pan et al., 2006).

The tripartite core network has been further investigated by associated transcriptions factors (Figure 6.1) leading to whole databases of interactions based on literature mining or manual annotations (Som et al., 2010, Xu et al., 2013). Zink finger protein 42 (Rex-1, also known as Zfp-42) (Shi et al., 2006) as well as Kruppel-like factor 4 (Klf4) (Zhang et al., 2010) are also tightly connected and associated to all other pluripotency factors (Chen et al., 2008, Pardo et al., 2010) and are assumed to play a major role in ESC pluripotency (Lu et al., 2009).

Medium conditions are a crucial part to keep cells in their artificial pluripotent status. The most important ingredient is the cytokine leukemia inhibitory factor (LIF), which activates signaling pathways such as Jak-Stat3, which is thereby linked to the core network and promotes pluripotency (Niwa et al., 2009). In combination with serum or bone morphogenic protein (BMP), this resembles a widely used medium condition where only

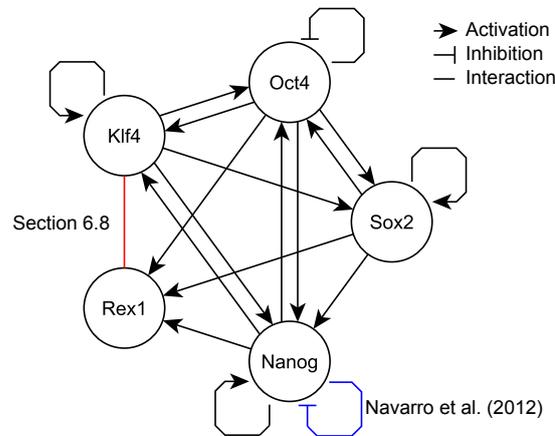


Figure 6.1: The core pluripotency network model consists of Nanog, Oct4 and Sox2. Additional transcription factors Rex1 and Klf4 are tightly connected to the network. Every black edge is based on manually curated literature mining and extracted from the PluriNetWork (Som et al., 2010). The blue self-inhibition edge is based on Navarro et al. (2012). The red edge is a link between Rex1 and Klf4 missing in current literature to be able to explain the correlation structure analyzed in Section 6.8.

a minority of cells differentiate and ESCs can be expanded over multiple passages. LIF withdrawal can be used to initiate the differentiation process (Smith, 1991). Eventually, a small fraction of cells might also differentiate in the presence of LIF. Two small-molecule inhibitors have been introduced (SU5402 and PD184352) which inhibit the FGF receptor tyrosine kinases and the ERK cascade (Ying et al., 2008). By using this so called *2i* condition, a high apoptosis rate could be observed (Ying et al., 2008). Therefore, by adding a more selective inhibitor, CHIR99021, to the medium conditions (*3i* medium) could further improve pluripotency and self-renewal of ESCs (Ying et al., 2008).

Based on previous bulk studies a variety of mathematical models have been developed trying to explain mechanisms underlying Nanog heterogeneity. The reported heterogeneous Nanog distribution has been described by a noise-driven excitable system (Kalmar et al., 2009). In this minimalistic model Oct4 acts as activator for Nanog given a low abundance of Oct4, but it acts as repressor if Oct4 levels get higher. Nanog also activates Oct4 in this model. After adding a descent amount of noise to the system, the parameters can be tuned to fit to the data described by Chambers et al. (2007). The transient fluctuation of low Nanog molecules are quickly counterbalanced which drives Nanog back up to its high steady-state level. Two more models have been proposed, which could both successfully fit the observed data (Glauche et al., 2010). These models include Nanog auto-activation and external ac-

tivation by the Oct4-Sox2 complex, which is also self auto-activating. The model explains a bistable system where every compound (i.e. Nanog and Oct4-Sox2 complex) is either highly expressed or only at basal levels. Additional noise terms acting on Nanog expression construct a fluctuation scenario or by alternatively including another factor establishing a negative feedback with Nanog results in an oscillatory scenario. Both models are highly generic allowing to fine-tune all model parameters to fit the flow cytometry measured static Nanog expression profile. More recently, another model has been proposed to detect the distinct, inherent Nanog states (Luo et al., 2012). By fitting Gaussian distributions to stationary Nanog distributions of different medium conditions the authors inferred that three, not two, stable steady-states could explain the data best.

These models provide very diverse explanations for the underlying molecular mechanisms although they all fit the observed Nanog distribution. None of them could be proven or rejected so far due to the lack of time-resolved data showing the real Nanog protein expression in continuous single-cell resolution. For this reason adequate ESC lines have to be created, which allows monitoring absolute Nanog protein amount in contrast to formerly used transcriptional reporter lines. Furthermore, non-invasive methods are required to follow single-cells throughout many generations on a large time scale. Only by single-cell long-term Nanog quantification, important points like existence, frequency and timescales of transitions between low and high states can be answered.

6.2 Novel pluripotency transcription factor protein reporter ESC lines

Murine ESCs from two different parental lines were used for fusion knock-in experiments. First, a GFP was fused to one of the endogenous Nanog and Oct4 gene loci in E14 ESCs, respectively. Similarly, NanogVENUS and Oct4VENUS fusion knock-in cell lines were generated based on R1 ESCs (Figure 6.2A and B) (Filipczyk et al., 2013). The fluorescent proteins reflect the true protein abundance due to their direct fusion and in contrast to former studies (Chambers et al., 2007), our constructs are not transcriptional reporters. NanogVENUS shares similar protein half-life compared to wild type protein (see Chapter 7). The NanogVENUS cell line also shows normal functionality and pluripotency which validates our approach (Filipczyk et al., 2013). Although a significant difference can be observed between Oct4VENUS and Oct4 wild type in terms of protein half-life (Welch's t-test p-value <0.047 , compare Chapter 7), the resulting protein half-lives are in a comparable range (Table 7.1). All cell lines share similar proliferation and differentiation characteristics (Filipczyk et al., in review). In accordance

with previous findings, NanogVENUS and Oct4VENUS are heterogeneously and homogeneously expressed in the nucleus, respectively (Figure 6.2C).

Recently, it has been shown that both Nanog and Oct4 share comparable expression profiles (Faddah et al., 2013). After quantifying protein expression of our knock-in cell lines in LIF conditions by flow cytometry we observed indeed a very broad distribution of NanogVENUS (Figure 6.2D). The expression profile includes a very low regime, which overlaps with the auto fluorescence of R1 control cells, as well as very high levels. In direct comparison, Oct4VENUS shows a more narrower expression profile, which does not include negative expression levels. Importantly, when quantifying NanogVENUS and total Nanog expression simultaneously by immunofluorescence imaging based on cell segmentation of DAPI signals, both expression profiles highly overlap (Figure 6.3B) and a strong correlation between both could be observed (Figure 6.3A, $R^2=0.74$, 492 cells, compare Section 2.3). The quantified profile of endogenous Nanog is also highly comparable to flow cytometry data of NanogVENUS (Figure 6.2E). This again indicates that our cell lines show normal regulation of the protein fusion reporters. Finally, we investigated if a strain specific difference can be observed. By comparing NanogVENUS (based on R1 background) to our NanogGFP lines (based on E14 background), we cannot observe big differences in the expression profiles and NanogGFP still extends into the negative gates (Figure 6.2F). Similarly, Oct4VENUS (R1 background) is highly comparable to Oct4GFP (E14 background) in their protein expression profiles (Figure 6.3C).

Previous studies, based on GFP reporter lines instead of using direct fusions, showed almost a bimodal distribution of Nanog GFP (TNGA) (Chambers et al., 2007). By comparing the expression profiles of this reporter to our cell lines we do not see these two modes, but still NanogVENUS and NanogGFP share a very broad distribution. Furthermore, the GFP profile of TNGA cells shows a higher dynamic range, possibly caused by the longer half-life of this reporter protein which is estimated to be around 26 hours (Corish and Tyler-Smith, 1999). In comparison, the half-life of Nanog wild type is about 3.69 hours (Chapter 7).

Although differences exist in the protein expression profiles between our cell lines and the cell line previously used, we still observe similar distributions. Furthermore, we argue that our cell line better resembles endogenous wild type Nanog expression in ESCs.

6.3 Live continuous long-term single-cell quantification of NanogVENUS protein expression in ESCs

To investigate NanogVENUS dynamics in individual cells, we performed movies of NanogVENUS knock-in cells in LIF+Serum conditions for up to

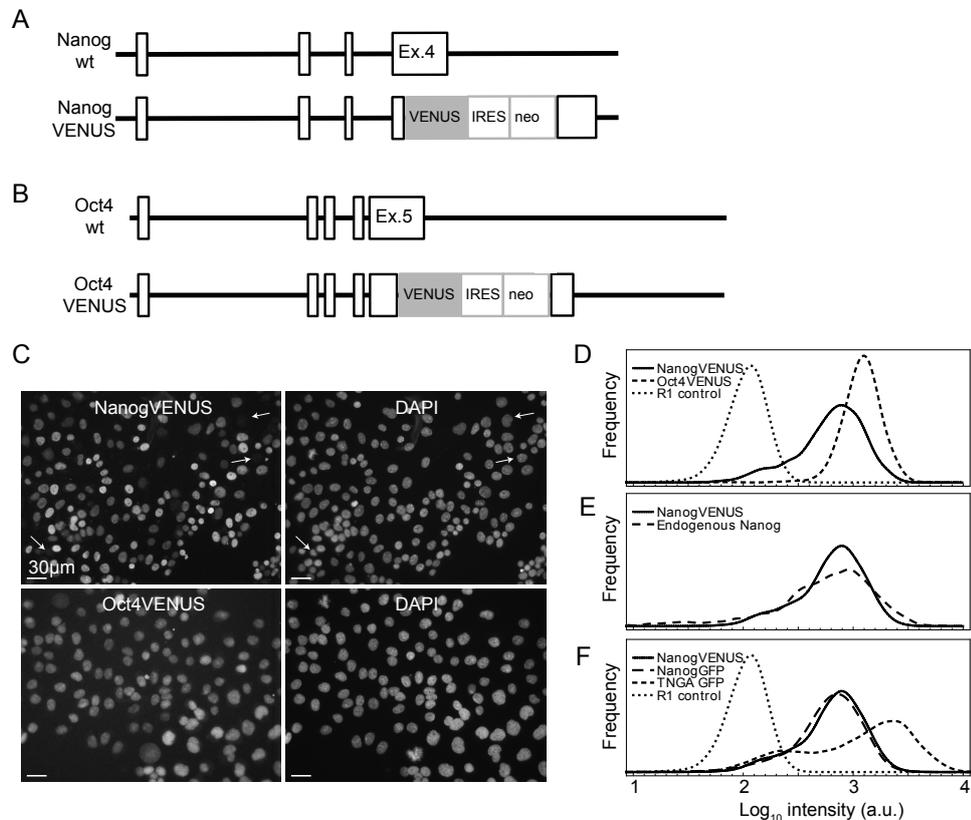


Figure 6.2: Nanog and Oct4 fluorescent protein fusion ESC lines faithfully report endogenous Nanog and Oct4 protein expression. (A) Targeting strategy for generation of allele-specific fluorescent Nanog protein reporter. White boxes denote Nanog exons. (B) Targeting strategy for generation of allele-specific fluorescent Oct4 protein reporter. White boxes denote Oct4 Exons. (C) Nanog VENU S has a more heterogeneous expression range compared to Oct4 VENU S as assessed by immunofluorescence microscopy. Scale bar is 30 μm . Arrows indicate instances of Nanog VENU S negative cells. (D) The Nanog VENU S reporter has a much broader dynamic range of expression than the Oct4 VENU S fusion reporter as shown by flow cytometry analysis. (E) The Nanog protein fusion distribution matches that of endogenous Nanog, as detected by an antibody against endogenous Nanog protein. Nanog signals from immunofluorescence imaging of ESCs were electronically gated and quantified using sQTFy (Section 3.4). (F) Nanog knock-in fusion reporters Nanog GFP (in E14 ESCs) and Nanog VENU S (in R1 ESCs) retain similar protein expression profiles. However, they differ in their expression distributions when compared to the GFP transcriptional reporter (TNGA), especially in the high end of the Nanog distribution. Figure adopted from Filipczyk et al., in review.

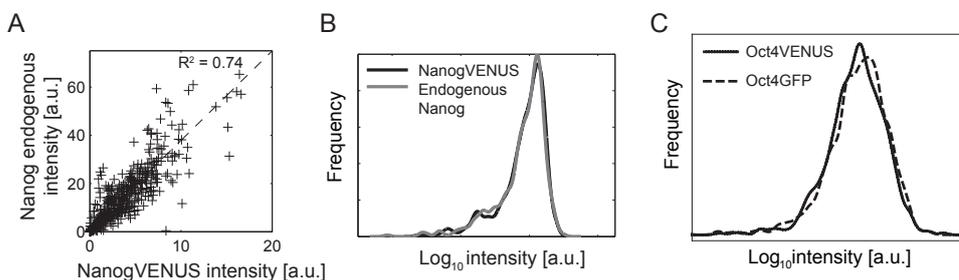
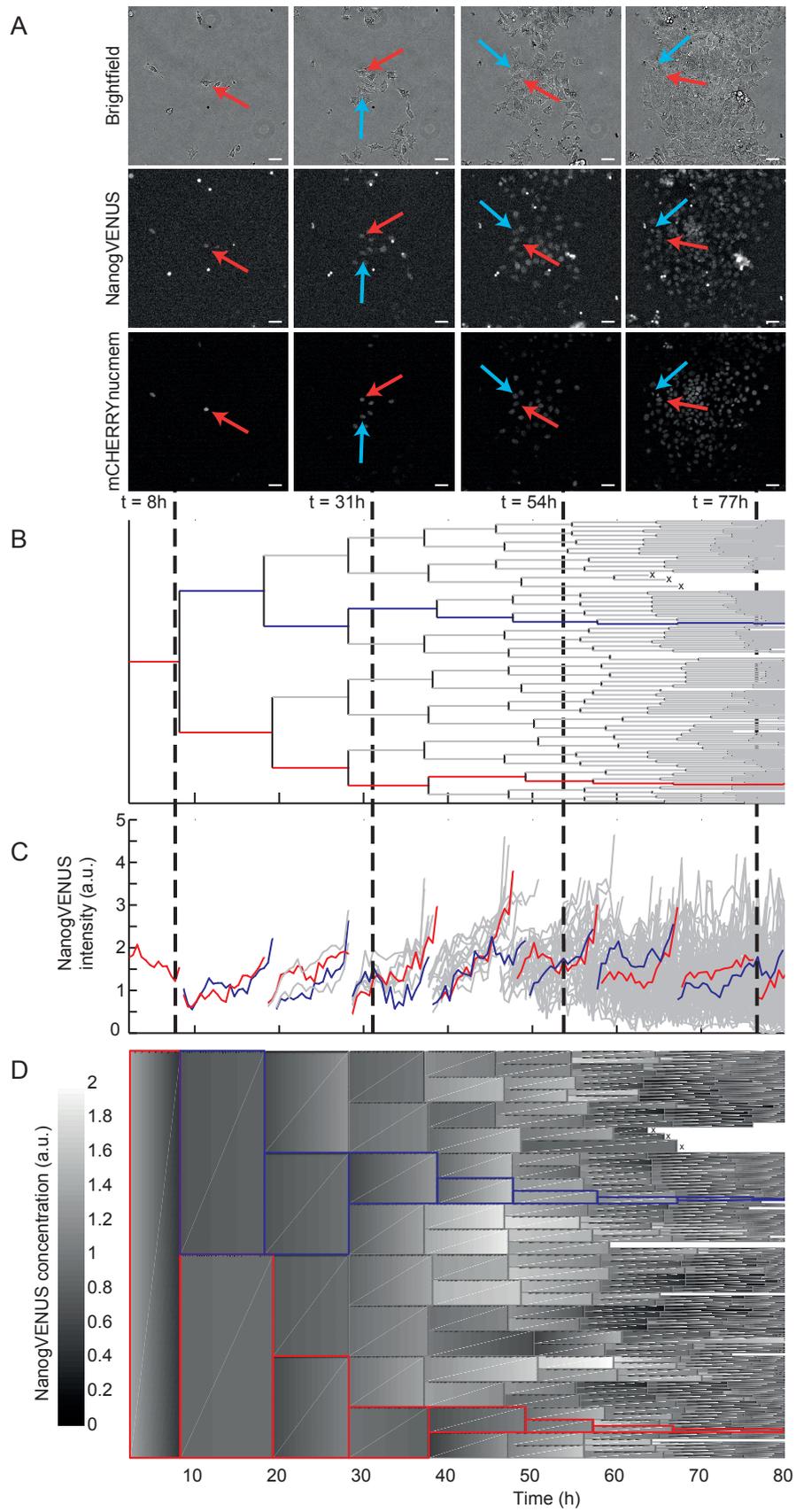


Figure 6.3: Knock-in fluorescence quantifications serve as reliable reporter for endogenous protein levels. (A) Endogenous Nanog levels strongly correlate to NanogVENUS levels ($R^2=0.74$, 492 cells). NanogVENUS ESCs were immunostained with an antibody against Nanog, imaged, background corrected, segmented, and quantified in both fluorescence channels. (B) The intensity distribution of the antibody stained endogenous Nanog matches closely to that of the NanogVENUS intensity distribution. (C) Oct4GFP and Oct4VENUS are expressed similarly, regardless of genetic background and fluorescent protein (R1 in VENUS line and E14 in GFP line). Figure adopted from Filipczyk et al., in review.

7 days. Additionally, ESCs were engineered to express a fluorescent nuclear marker (mCHERRYnucmem) in the form of the mCHERRY fluorescent protein N-terminally fused to a nuclear membrane protein which allows to easily identify the nucleus of each cell in the fluorescent channel. We tracked colonies with the current version of TTT (Section 3.1.2) using the bright field images (Figure 6.4A), which were acquired every 30 minutes. We were able to follow individual cells with all their progeny up to 8 generations or up to 80 hours after movie start resulting in pedigrees with up to 512 cells (Figure 6.4B). Additional fluorescence channels captured mCHERRYnucmem and NanogVENUS or Oct4VENUS expression simultaneously with the bright field images (Figure 6.4A). Using QTFy (Section 3.3) we normalized all fluorescence images (Section 2.2.6), segmented the nucleus of tracked cells based on mCHERRYnucmem signal and quantified the corresponding NanogVENUS or Oct4VENUS intensity (Figure 6.4C). The obtained time-courses demonstrate the large heterogeneity of NanogVENUS expression ranging from cells with almost no visible signal to very high intensities. Even within one cell-cycle we observe diverse expression pattern emphasizing the need for continuous single-cell measurements. After correcting for cell-cycle progression and recalculating intensity into concentration (Section 3.3.3), we visualize the whole genealogy as heat tree (Section 3.3.5), which still shows large heterogeneity in NanogVENUS expression, even between siblings (Figure 6.4D).

For all experiments, only cells with unequivocal identity that could clearly be identified when evaluating the movie were used (i.e. cells growing



in monolayer colonies, compare Figure 6.5A) and all cells with questionable identity were excluded from relevant analyses (Figure 6.3B). Cells which grow in 3D ($\approx 10\%$ of all colonies) flatten out again in $\approx 50\%$ of these cases enabling the quantification of their NanogVENUS intensities. A comparison of cell intensity distributions of flattened out 3D colonies to cell intensity distributions of monolayer colonies does not show any significant differences indicating that the exclusion of questionable cells does not affect analysis results ($p > 0.05$ two sample t-test, Figure 6.5C).

For all following analyses we performed 4 long-term experiments, 3 4-day experiments and 3 1-day experiments (Table 6.1). All in all, we have more than 200,000 manually inspected data points with multiple features such as intensity levels for NanogVENUS or Oct4VENUS and mCHERRYnucmem, cell area in pixel and the position within the bright field images. In almost every movie we have additional endpoint staining data, an immunofluorescence staining directly after the movie. For this purpose ESCs were fixed and one last image of NanogVENUS was taken, which shows a slight shift of the plate due to the fixation step. After that, immunofluorescence analysis has been applied, which again results in a slight shift in the image series due to plate removal and back insertion. Immunofluorescence allows to measure additional cellular properties, here we stained DAPI, Klf4, Sox2 or Rex1 and Oct4 in the case of NanogVENUS cells. For single-cell quantification one had to carefully select only those cells, which do not show any mCHERRYnucmem expression since their emission spectra highly overlap with antibody dyes. All images were background corrected, segmented on DAPI signal and quantified in all other channels as described before using

Figure 6.4 (*preceding page*): Long-term imaging, single-cell tracking at and quantification generates genealogies of NanogVENUS expression for up to 8 generations (A) bright field, VENUS and mCHERRY fluorescence microscopy images obtained every 30 minutes permit us to observe the development of clonal colonies for up to 80 hours and up to 8 generations. Scale bars are $40\mu\text{m}$. (B) Computer-assisted manual cell tracking creates cellular genealogies. We highlight the lineage of two arbitrarily chosen single-cells (red and blue arrows in (A), lines in (B-D)) over 8 generations. Cell apoptosis is denoted by 'x'. Any cell that could not be tracked or quantified is omitted from the tree. (C) NanogVENUS fluorescence intensity quantification of single-cells reveals generation of a wide intensity range. Identification and segmentation of individual cellular nuclei is based on mCHERRY fluorescence imaging. (D) Cellular NanogVENUS quantifications are combined with genealogies to produce heat trees, where NanogVENUS concentration over lifetime is represented by a smooth gradient. Figure adopted from Filipczyk et al., in review.

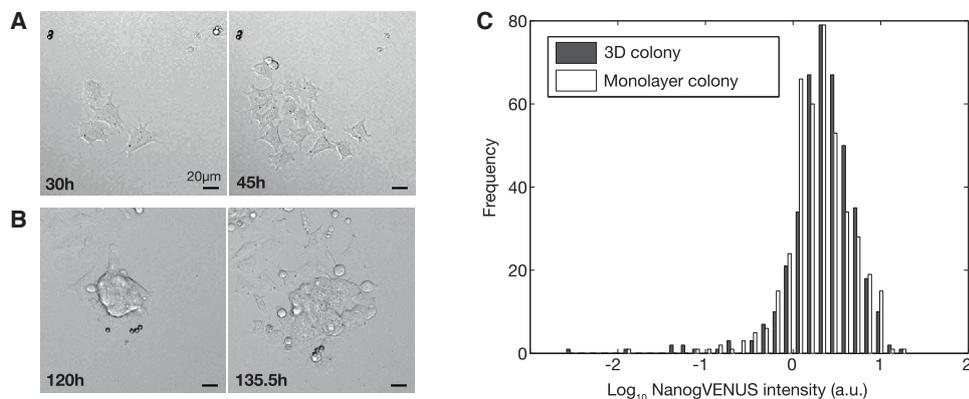


Figure 6.5: Colonies with 3D growth can be excluded from analyses without introducing bias. (A) Monolayer colonies (shown at 30h and 45h after movie start) can be reliably tracked and quantified throughout the whole movie. Scale bars are $20\mu\text{m}$. (B) Approximately 10% of ESC colonies grow in 3D (example colony shown at 120 h) and cannot be tracked reliably. In $\approx 50\%$ of these cases, colonies flatten out later (flattened out colony shown at 135.5 h) and can be quantified. (C) NanogVENUS intensity does not differ significantly ($p > 0.05$, two sample t-test) between the reference NanogVENUS distribution and cells grown in 3D colonies ($n=429$). Figure adopted from Filipczyk et al., in review.

sQTFy (Section 3.4). Since the gain normalization could not be applied due to lack of a time series, we normalized by dividing the raw images by their estimated background (compare Section 2.2.6) and subtracting 1. Potential spatial shifts between DAPI images and NanogVENUS images were corrected by rigid image registration using the program Elastix v4.500 (Klein et al., 2008). The registered DAPI cell masks were then used to quantify corresponding NanogVENUS intensity ensuring correct cell identities between images. Thus we obtained the background-corrected total signal intensity for individual cells for each factor (i.e. DAPI, Oct4, Sox2 or Rex1, Klf4 and NanogVENUS).

6.4 Comparison to flow cytometry

To compare a NanogVENUS distributions from imaging with flow cytometry, we segmented and quantified individual randomly selected NanogVENUS cells based on mCHERRYnucmem signal in an unbiased way using sQTFy, (Figure 6.6A and B). The sensitivity (percentage of cells above the negative gate, Figure 6.6C, compare Section 5.5) was higher for imaging ($94.9 \pm 3.8\%$, mean \pm SD, $n=3$) and less variable ($\text{CV}=4.0\%$, $n=3$) compared to flow cytometry (sensitivity $86.8 \pm 10.1\%$, $\text{CV}=11.6\%$, $n=3$).

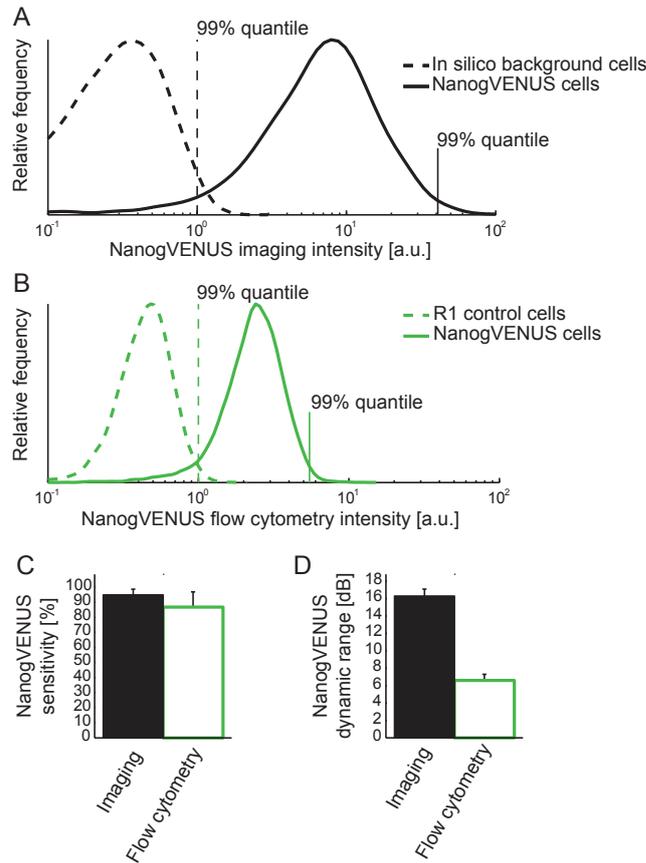


Figure 6.6: Improved fluorescence quantification by imaging versus flow cytometry. (A) VENUS quantification of NanogVENUS expressing ESCs by imaging ($n=3$, one representative shown). Negative gate was determined by *in silico* background cells. (B) VENUS quantification by flow cytometry of same cell population ($n=3$, one representative shown). Negative gate determined by VENUS negative control ESCs (C) Sensitivity (% of detected cells above the negative gate) is comparable between approaches ($n=3$ each). (D) The dynamic range (fold change between the 99% quantile over the negative gate in decibel) is higher in imaging than flow cytometry ($n=3$ each). Figure adopted from Schwarzfischer et al., in review.

Experiment	Trees	Cells	QTFied points	Internal ID
Long-term 1	404	12,844	25,362	110613AF6
Long-term 2	124	1,312	10,194	110907AF6
Long-term 3	206	16,864	31,398	110930AF6
Long-term 4	194	30,210	138,890	111115AF6
4-day 1	12	1,152	5,432	120410AF6
4-day 2	2	230	3224	120516AF6
4-day 3	36	2,552	3,830	130503AF6
1-day 1	98	594	7,968	120509AF6
1-day 2	436	1,620	1,732	130204AF5
1-day 3	1,550	2,552	1,258	130208AF5
Total	3,062	68,980	223,856	

Table 6.1: All experiments and data points used in this study. We performed movies containing NanogVENUS and/or Oct4VENUS knock-in cell lines in three different manners. Long-term movie are typically at least 6 days long and best visualize single-cell dynamics over a long period of time. 4-day movies are used to facilitate comparison of single-cell trajectories with endpoint staining data since colonies are not as dense compared to long-term experiments. 1-day movies are used to correlate endpoint staining data with single-cell trajectories in a shorter time scale.

The dynamic range was calculated as the 99% quantile, Q_{99} , of the NanogVENUS distribution, divided by the negative gate. We give the dynamic range in decibel, defined as $10 \cdot \log_{10} \frac{Q_{99}}{\text{negative gate}}$. The dynamic range was significantly higher for our imaging technique (16.3 ± 0.8 dB, $n=3$) than for flow cytometry (6.6 ± 0.7 dB, $n=3$, Figure 6.6D), probably due to the fluorescence image normalization, in particular its estimated background subtraction step (compare Section 2.2.6).

To test this hypothesis, we normalized flow cytometry data by subtracting the mean fluorescence signal of control cells. The dynamic range of flow cytometry slightly improves (8.7 ± 0.9 dB, $n=3$), but is still significantly lower compared to our imaging technique. This could probably be caused by the more precise segmentation, which only quantifies nucleus expression and excludes potential auto fluorescence in the cytoplasm. Furthermore, our normalization accurately normalizes every pixel below a cell, thereby considering the actual nuclear size.

Next, we segmented individual ESCs based on normalized DAPI fluorescence intensity (compare Chapter 6). Segmentations were used to quantify normalized NanogVENUS fluorescence and, as further comparison, non-normalized NanogVENUS fluorescence resulting in two single-cell expression distributions. After standardizing by multiplication every distribution to the same mean we calculated the Kolmogorov-Smirnov distance to a stan-

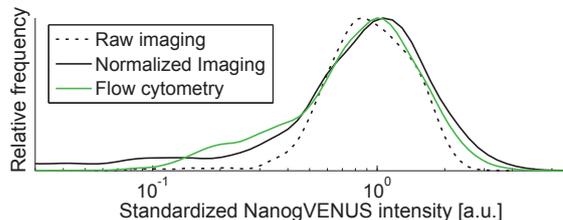


Figure 6.7: Normalized fluorescence intensity distribution quantified by imaging shows significantly increased overlap to a reference distribution derived by flow cytometry compared to non-normalized cell intensities (paired t-test p-value < 0.02 , $n=3$). Every distribution has been standardized to have the same mean equal to one. Figure adopted from Schwarzfischer et al., in review.

standardized NanogVENUS reference distribution obtained by flow cytometry of normalized and unnormalized data (Figure 6.7). Normalized intensity distributions derived by our imaging technique compares well to the reference distribution and its overlap is significantly higher compared to non-normalized intensity distributions (paired t-test p-value < 0.02 , $n=3$).

6.5 ESCs maintain Nanog concentration on average

We integrate three long-term experiments containing unsorted and unbiased tracked and quantified cells and calculate the following two general statistics. First, we check if the intensity is conserved during cell division. We select all cells, which give rise to two daughter cells and calculate the difference of the intensity of the mother cell right before cell division and the sum of intensities of both daughter cells after cell division (-0.16 ± 0.62 , $n=522$, Figure 6.8A). With a Wilcoxon signed rank test we can reject that this distribution comes from a continuous, symmetric distribution with median 0 (p-value $< 10^{-7}$). Since intensity measurements have half an hour intervals, daughters cells have time activate protein production leading to a higher sum of intensity.

Since ESCs are a steady-state population, we investigate if cells double their intensity during one cell-cycle (compare doubling hypothesis, Section 3.3.3). For every cell, we calculate the fold-change of the mean of the last two measurements to the mean of the first two measurements (2.12 ± 0.88 , $n=604$, Figure 6.8B). Indeed, ESCs double their amount of intensity on average. Furthermore, a Wilcoxon signed rank test cannot reject the hypothesis that this distribution comes from a continuous, symmetric distribution with median 2 (p-value = 0.1149).

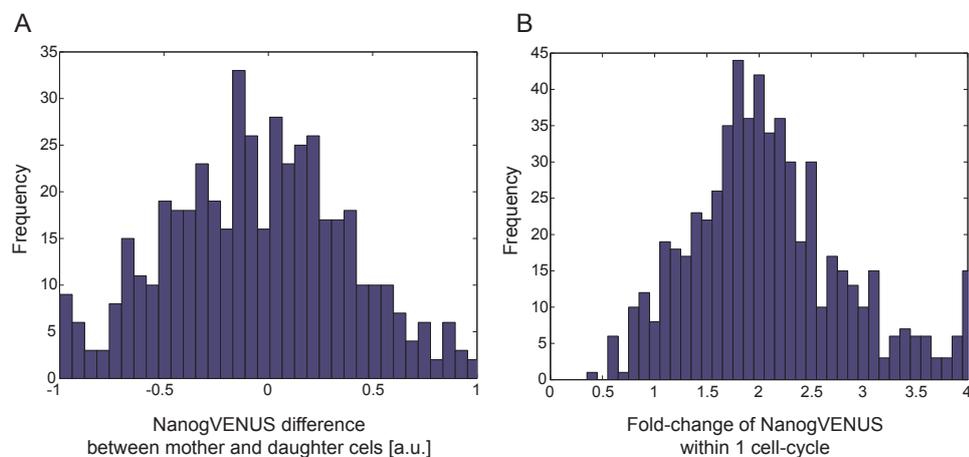


Figure 6.8: NanogVENUS intensity is maintained during cell-cycle on population average. (A) The difference in NanogVENUS between the intensity before division of a mother cell to the sum of the intensity after cell birth of the daughter cells is slightly below zero (-0.16 ± 0.62 , $n=522$). (B) NanogVENUS intensity is conserved on average within one cell-cycle resulting in the double amount of intensity before cell division compared to the initial intensity (2.12 ± 0.88 , $n=604$).

6.6 Nanog negative ESCs are a heterogeneous cell population

To investigate if the doubling hypothesis holds true for individual trajectories, we performed additional experiments. ESCs have been sorted by flow cytometry and only cells without any NanogVENUS expression are imaged, tracked and quantified over four days (Figure 6.9A). This Nanog negative state is also highly debated in literature to be a potential exit from pluripotency (Chambers et al., 2007, Dietrich and Hiiragi, 2007, Mitsui et al., 2003).

Unexpectedly, we found whole pedigrees which stayed NanogVENUS low throughout the whole movie (Figure 6.9C). From all negative sorted cells we found $40 \pm 5\%$ ($n=3$) clonal colonies which did not show any NanogVENUS onset at all. This finding clearly shows a memory of cells over many generations. Furthermore, we observed emerging trees which upregulate NanogVENUS during the movie time giving rise to colonies with mosaic NanogVENUS expression (Figure 6.9B) or which are asymmetrically fated trees, where one sub-branch of cells stayed negative for the whole time and the other half of the tree repopulated the whole intensity range (Figure 6.9D) to rebuild a mosaic colony. To exclude that cells with low NanogVENUS expression are differentiated, we quantified other pluripotency factors at the end of the movie by immunofluorescence staining. Importantly, Sox2 and

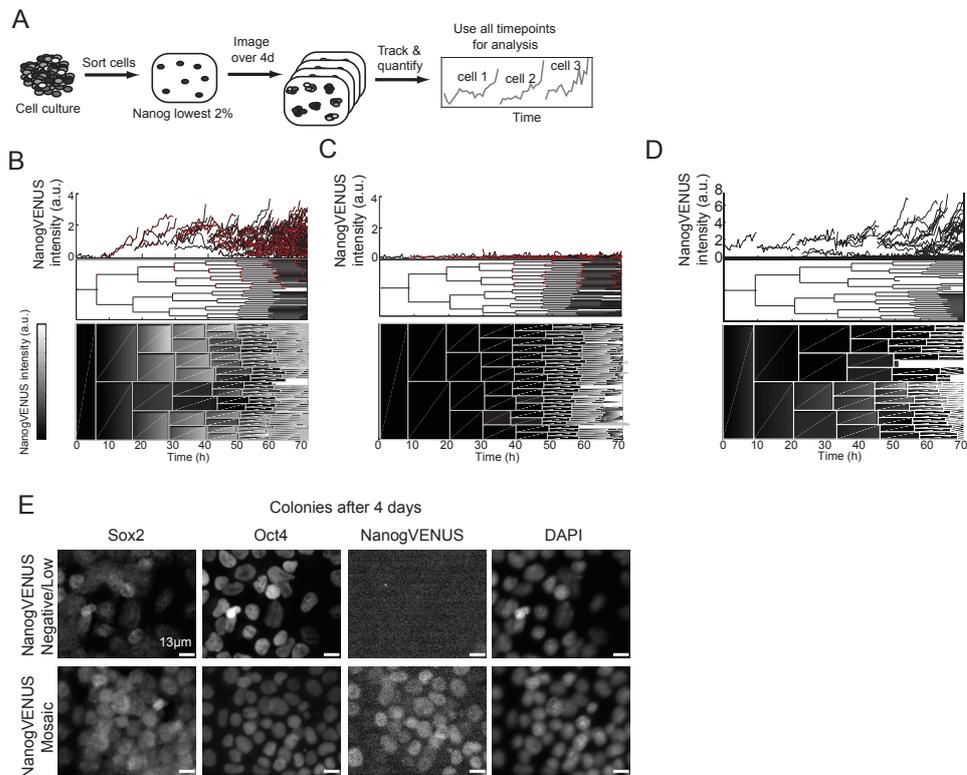


Figure 6.9: Nanog production can be biased over generations, and a unique subpopulation of ESCs stably maintains a Nanog negative/low state. (A) Experimental approach. (B) Nanog negative cells can upregulate Nanog and can produce progeny with Nanog intensities spanning all compartments. (C) Nanog negative cells can yield perpetually negative/low colonies. (D) Nanog negative cells can also give rise to colonies with divergent subtrees. (E) Cells in Nanog negative colonies express pluripotency factors Oct4, Sox2 and Rex1. Figure adopted from Filipczyk et al., in review.

Oct4 are positively expressed (Figure 6.9E) for mosaic colonies as well as for low colonies. Combined with the comparable proliferation rate of cells in low colonies, we exclude that these cells are differentiated.

6.7 Pluripotency factor correlation networks reveal different Oct4 and Klf4 wiring in Nanog negative versus re-expressing ESC clones

As described above we quantified all stained transcription factor levels at the end of a movie for each cell and used the full movie information to group cells into clonal colonies. For the following analysis we used cells, which

are sorted by flow cytometry to have no NanogVENUS expression at the beginning of the movie. Based on the NanogVENUS expression within a colony at the endpoint we divided the colonies into low or mosaic colonies. low colonies showed positive pluripotency markers Sox2 and Oct4, but were low to negative for Klf4 expression (Figure 6.9E, Figure 6.10A), which is a factor known to be upstream of Nanog (Chan et al., 2009, Zhang et al., 2010) (Figure 6.1). Mosaic colonies showed positive expression for all measured transcription factors, also with a higher overall level (Figure 6.10B).

In the following, we used Pearson and partial correlation analysis (Lauritzen, 1996) to investigate potential causalities between all measured quantities (compare Section 2.3). Since this kind of analysis is independent of absolute levels we could easily apply it to relative intensity levels. Previously, Pearson correlation was used to identify associations between Nanog and Oct4 depending on the cell medium (Muñoz Descalzo et al., 2012).

Pearson correlations showed that Oct4 and Sox2 are tightly associated with a high correlation coefficient for mosaic colonies ($R^2 = 0.81 \pm 0.07$, $n=3$) and negative colonies ($R^2 = 0.82 \pm 0.07$, Figure 6.10B). This finding is in accordance with previous reports and implies the co-regulation of both transcription factors (Muñoz Descalzo et al., 2012, Thomson et al., 2011). The correlation coefficient between Klf4 and NanogVENUS was also considerably high for mosaic colonies ($R^2 = 0.52 \pm 0.09$) and considerably high for negative colonies ($R^2 = 0.31 \pm 0.09$). This close co-expression can be well-explained by the upstream regulation of Nanog by Klf4 (Chan et al., 2009, Zhang et al., 2010). Interestingly, we observe even for negative colonies, which express NanogVENUS almost at the noise level correlations to other transcription factors although weaker when compared to mosaic colonies (Figure 6.10C).

To get a better insight into transcription factor interactions (which have already been investigated in high-throughput manner (Chen et al., 2008, Lu et al., 2009)), we next applied partial correlations (Lauritzen, 1996). Partial correlations are able to reveal direct interactions by simultaneously removing all indirect interaction between all factors. We applied partial correlation not only to the intensity levels of Nanog, Oct4, Sox2 and Klf4, but also included DAPI intensity as well as nuclear area in pixel. This allowed us to correct for cell-cycle induced effects which are encoded in the growing area or increasing DAPI intensity. We visualize partial correlations as graph networks and analyzed five different subsets of the data (Figure 6.10C).

We found different networks depending on the colony properties and tested for significantly different correlations using a partial correlation population test (Levy and Narula, 1978). The partial correlation between Klf4 and Oct4 was negative only for NanogVENUS low colonies and also significantly different compared to all other ESC population indicating that transcription factor interactions change depending on the overall colony NanogVENUS expression (p-value <0.01 , $n=3$, Figure 6.10C).

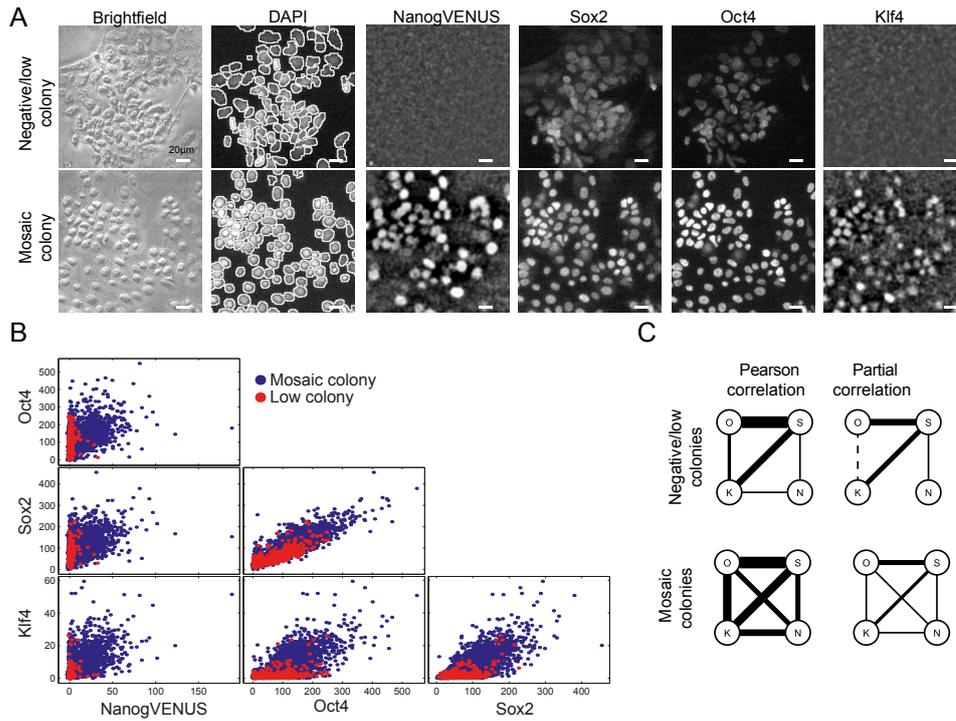


Figure 6.10: Pluripotency factor expression correlation reveals different Oct4 and Klf4 wiring in NanogVENUS negative/low vs. mosaic colonies. (A) ESC colonies can remain NanogVENUS negative/low for up to 4 days following sorting negative Nanog sorting, while still expressing other pluripotency factors. Images are quantified using nuclear segmentation masks obtained from the DAPI signal. (B) Oct4 and Sox2 show strong expression correlation in both types of colonies. Overall, correlation and relative levels are reduced in negative/low colonies. (C) Partial correlation networks computed from mosaic and negative/low colonies differ significantly. Dashed lines indicate negative correlation, solid lines indicate positive correlation. Figure adopted from Filipczyk et al., in review.

6.8 Rex1 influences the interplay between Klf4 and Nanog

In an additional data set (which was not part of Filipczyk et al., in review), we investigated ESCs which have been stained for Oct4, Klf4 and Rex1 instead of Sox2 after four days.

Rex1 is also tightly connected and associated to all other pluripotency factors and is known to be a marker for undifferentiated ESCs (Shi et al., 2006) (Figure 6.1). Similarly to Nanog, Rex1 is heterogeneously expressed in pluripotent ESCs (Tanaka, 2009).

As described previously we normalize all fluorescence images, segment clonal colonies (mosaic in terms of NanogVENUS expression) based on DAPI intensity and quantify all stained transcription factors. We quantify the fluorescence of NanogVENUS by registering the segmentation onto the NanogVENUS image.

First, we investigate if the data set including Rex1 staining is comparable to the data set including Sox2. Since we want to apply correlation analysis later, which is independent of absolute scales, we standardized the intensity distribution I_k for each factor k individually to have the same mean: $\bar{I}_k = 1$ (Figure 6.11). Both data sets have similar histograms on a qualitative scale. Since the cells have been low sorted for NanogVENUS, initially NanogVENUS shows a bimodal expression profile with a much more pronounced low peak compared to the unbiased data set (compare 6.2D). Oct4, Sox2 and Klf4 show a quite narrow distribution compared to NanogVENUS. Rex1 in comparison to Sox2 has a very broad distribution, showing a low and a high mode, as expected. Although some differences in DAPI and area can be observed the overall impression persists that both data sets can be well compared for further correlation analysis.

As in the previous section, we apply Pearson and partial correlation analysis to investigate the interactions of the measured transcription factors including DAPI intensity and cell area. As previously shown, Pearson correlations induce fully connected networks (Figure 6.12A and B). Edges, which could be measured in both data sets show very similar correlation coefficients, indicating that both data sets can be well compared. In particular the Pearson correlation coefficient ρ of NanogVENUS and Klf4 is well comparable between the Rex1 data set ($\rho = 0.46$) and the Sox2 data set ($\rho = 0.51$). The Pearson and partial correlation network of the Sox2 data set shows the same correlation structure as previously (compare Figure 6.10C and Figure 6.12B and D). However, a comparison of the partial correlation networks show some unexpected results such as the partial correlation between NanogVENUS and Klf4, which shows a significant negative partial correlation for the Rex1 data set (Figure 6.12C, partial correlation coefficient $\rho = -0.14$) and a significant positive partial correlation for the

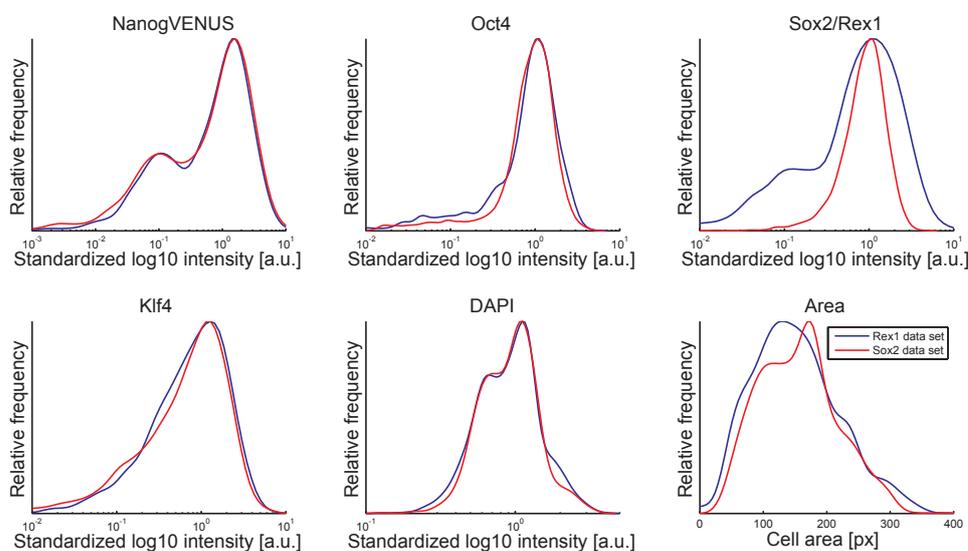


Figure 6.11: Two data sets with different staining procedures have been quantified ($n(\text{Rex1})=1776$, $n(\text{Sox2})=2337$). Every intensity distribution can be well-compared between both data sets.

Sox2 data set (Figure 6.12D, $\rho = 0.24$). Excluding the Rex1 measurements from the Rex1 data set and calculating the partial correlation results in a positive partial correlation coefficient for NanogVENUS and Klf4 ($\rho = 0.09$). Furthermore, the partial correlation between Rex1 and NanogVENUS and Rex1 and Klf4 are very high compared to other edges ($\rho > 0.63$, for both). These findings indicate that only by correcting for the influence of Rex1 the correlation between NanogVENUS and Klf4 turns into an anti-correlation, placing Rex1 rather between the two other transcription factors. This is in contrast to previous studies which assume that Klf4 is upstream of Nanog, which is again upstream of Rex1 (Shi et al., 2006, Zhang et al., 2010) (see Figure 6.1).

Note, that this analysis is based on one experiment only. However, in the bachelor thesis of André Seitz (Seitz, 2013), we quantified these two data sets similarly, but did not have the information of clonal colonies. Instead, we segmented all possible cells within the experiments and did not select especially for mosaic NanogVENUS colonies resulting in a higher number of cells per data set ($> 15,000$ cells). However, similar results could be achieved showing that the partial correlation of NanogVENUS and Klf4 is reduced by including the measurements of Rex1.

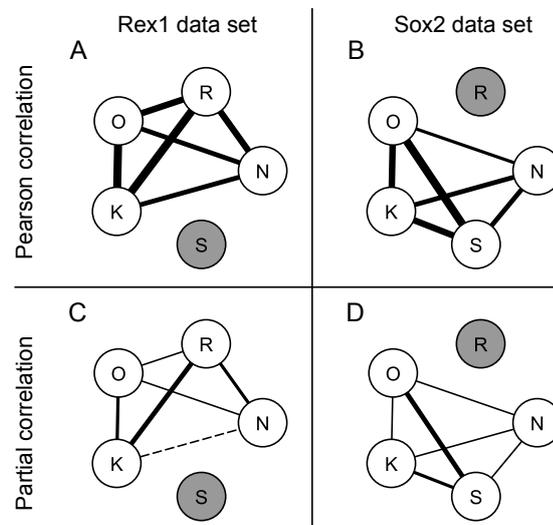


Figure 6.12: Rex1 influences partial correlation structure between NanogVENUS and Klf4. Two data sets have been measured including Rex1 or Sox2. (A,B) Pearson correlation of both data sets show a fully connected and highly comparable network. (C) The partial correlation between Klf4 and NanogVENUS is significantly negative (dashed line) due to Rex1 inclusion. (D) The partial correlation between NanogVENUS and Klf4 is significantly positive (solid lines). The thickness of the lines encodes the strength of the underlying correlation coefficient. $n(\text{Rex1})=1776$, $n(\text{Sox2})=2337$, N=NanogVENUS, O=Oct4, S=Sox2, R=Rex1, K=Klf4

6.9 Nanog expression is not coordinated with the pluripotency network

Next, we asked if we can monitor partial correlation results in dependence of time. Here, we fully tracked and quantified normalized fluorescence of NanogVENUS and all stained transcription factors instead of just using endpoint staining data. Based on their NanogVENUS behavior we grouped the trees into mosaic and low colonies as previously introduced. We subdivide each tree into four equally sized subtrees and perform partial correlation analysis.

Although all subtrees share the same founder cell they differ in absolute transcription factor levels as well as the network structure (Figure 6.13 A). Similarly, low sub trees show very heterogeneous networks within one tree (Figure 6.13B). A general trend of transcription factor levels compared to the overall NanogVENUS expression can be observed, but the differences of partial correlation wiring stays inconsistent. This means that the interplay between core network players is not as coordinated as presumed by previous publications reporting tightly connected gene regulations. Note, that partial correlations can break down if the sample sizes gets too low. Therefore we decided not to further split the tree into more sub sets.

In the following, we investigate if a connection of the pluripotency network can be observed on a shorter timescale. We performed movies of 24 hours of low sorted NanogVENUS ESCs and again tracked and quantified NanogVENUS and quantified stained Oct4, Sox2, Klf4 and DAPI intensity at the endpoint. As before, we found asymmetric behavior of NanogVENUS dynamics (Figure 6.13C) and unexpectedly found that Klf4 expression does not always follow accordingly (Figure 6.13D). The majority of all cells showed positive Klf4 expression and negative NanogVENUS expression ($54\pm 3\%$, mean \pm SD, $n=3$) while the minority showed the opposite negative Klf4 and positive NanogVENUS expression ($2.2\pm 0.8\%$, $n=3$, Figure 6.13E). Only $29\pm 6\%$ ($n=3$) of all cells showed positive co-expression of Klf4 and NanogVENUS. Interestingly, we also found NanogVENUS mosaic colonies after 4 days, which show no Klf4 expression at all (Figure 6.13F).

The discrepancy between Klf4 and NanogVENUS (Figure 6.13D) inspired us to investigate this effect between all transcription factors. We called sisters divergent if the expression of one factor of interest in one sister cell is at least two fold higher for NanogVENUS or 1.2 fold for all other transcription factors compared to its sister cell. We further count sister cells as divergent if one cell stays above the negative gate and the sister stays below the negative gate. In trees where we observe more than just one cell division we sum up the intensity of both daughter cells and count them as one virtual cell which is compared to the the progeny of their aunt cell, i.e.

the sum of their two cousins. After identifying all divergent sisters in one factor of interest, we ask if

- another factor shares the same divergence (=congruent divergent sisters in this factor)
- another factor shows the opposite divergence (=incongruent divergent sisters in this factor),
- the other factor is not divergent (=non-divergent sisters for this factor).

We test for all other factors if we find a congruent divergence more often than we would expect from random. For this means, we apply a paired Wilcoxon signed-rank test.

We identified 136 sisters, which are divergent in NanogVENUS, and found that most of them are also significantly congruent divergent for Oct4 (Figure 6.13G). Similarly, but less frequently, Klf4 is significantly congruent divergent with NanogVENUS. For Sox2 we cannot see a significant congruent divergence, indicating that the onset of NanogVENUS has no impact on Sox2 on this timescale. When identifying divergent sisters in Klf4, we find significant congruent divergence for Sox2 and Oct4, but not for NanogVENUS (Figure 6.13H). This indicates that the impact of NanogVENUS on Klf4 as seen before (Figure 6.13G) is stronger than in the other direction. If we assume that both factors were negative in the beginning of the movie, we might conclude that if NanogVENUS goes up, Klf4 follows more often congruently than the other way around. By looking at divergent Sox2 sisters, we find that Oct4 and Klf4 are significantly congruent divergent (Figure 6.13I), but NanogVENUS seems to be independent of Sox2 divergence, similarly as for divergent NanogVENUS cells in respect to Sox2 (Figure 6.13G). Finally, when identifying Oct4 congruent sister pairs, all other transcription factors significantly congruently follow the divergence (Figure 6.13J).

Although this data set is very small, it clearly puts Oct4 in the center of the core pluripotency network. Every other factor seems to adapt the expression profile of divergent sisters. Nanog itself is disconnected from Sox2 and also partially from Klf4, which seems to follow NanogVENUS and not the other way around on this short time scale. These results are in strong contradiction to previous studies which assume that Nanog is a tightly connected core regulator determined experimentally (Chambers et al., 2003, MacArthur et al., 2012) or based on modeling studies (Glauche et al., 2010, Kalmar et al., 2009, Luo et al., 2012). Another architecture which has been recently proposed (Navarro et al., 2012) where Nanog has no regulatory effect on Oct4 and Sox2 seems to fit our observations better.

6.10 NanogVENUS intensity is independent of spatial position within ESC colonies

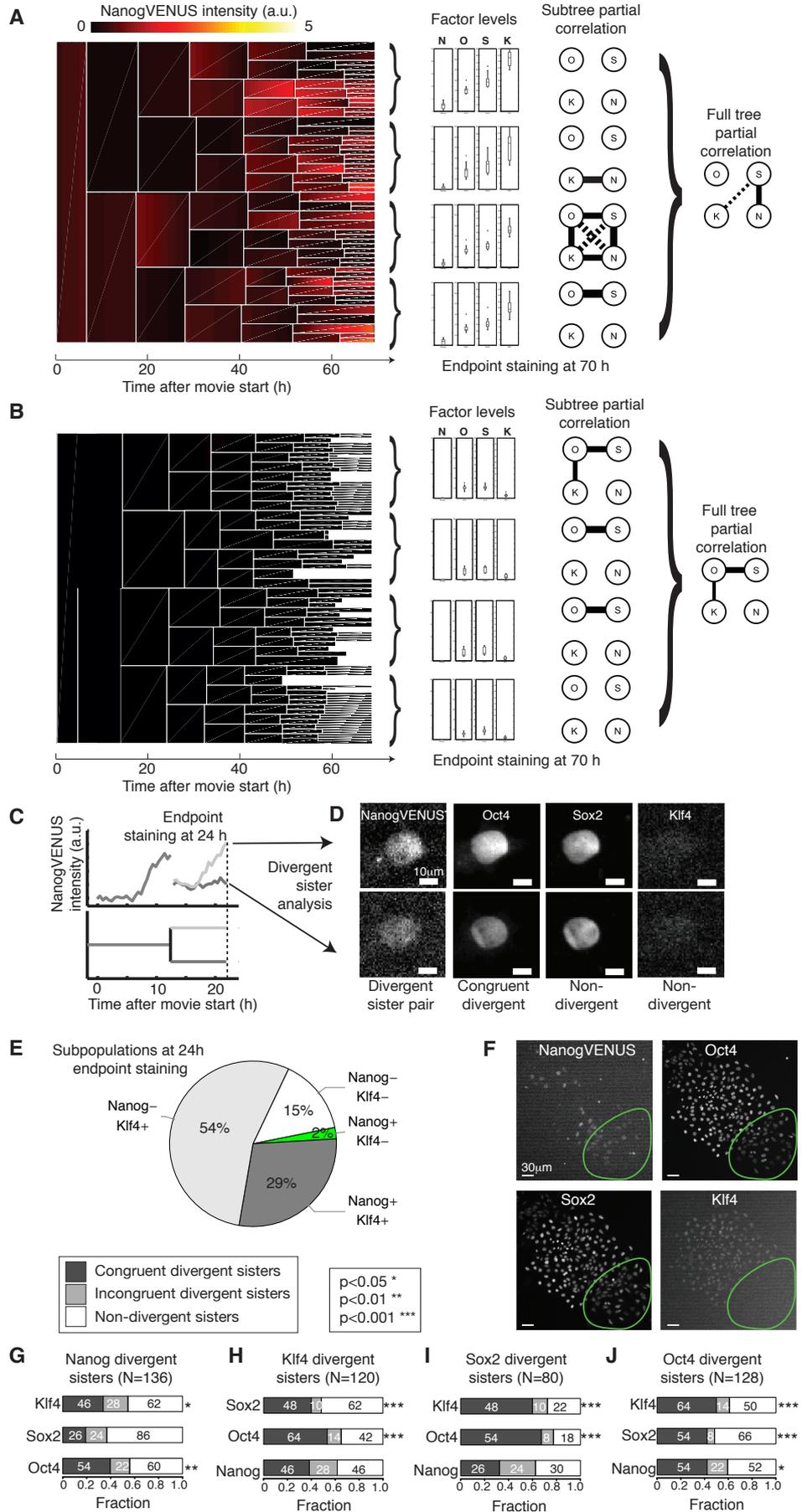
Since ESCs are derived from cells during embryogenesis where spatial interactions and morphogenesis play a major role, we investigate potential spatial dependencies in the densely growing ESC colonies. We correlate single-cell features to the position within a colony and analyze if potential differences exist between the border of a colony compared to the center of a colony. We use the same data set as described before, containing colonies which have been tracked for at least 5 generations.

We utilize the NanogVENUS intensity, nuclear cell area and lifetime of a cell and calculate the additional features: (i) the Euclidean distance to the center of the colony in pixel and (ii) a local density estimate using equally sized track points of all relevant cells based on a bivariate kernel density estimator, which inherently chooses an optimal bandwidth (Botev et al., 2010). The local density serves as an inverse proxy for the actual cytoplasmic cell size (Figure 6.14A).

First, we select one exemplary colony, calculate all features for every time point having more than four simultaneously tracked and quantified cells and determine the Pearson correlation coefficients (Section 2.3) of each feature to the distance to the colony center (Figure 6.15). Although some variability over time can be observed, NanogVENUS intensity is weakly correlated to the distance to the colony center ($\rho = 0.29 \pm 0.24$, Figure 6.15A). Similarly, the nuclear area is weakly correlated to the distance to the colony center for this particular colony ($\rho = 0.26 \pm 0.20$, Figure 6.15B). However, the cell lifetime shows no correlation ($\rho = -0.04 \pm 0.17$, Figure 6.15C), but a fluctuating pattern appears probably originating from the cell division events. As expected, the local density of a cell is highly anti-correlated to the distance to the colony center ($\rho = -0.83 \pm 0.28$, Figure 6.15D).

The movement of ESCs within the colony during the movie is rather low since cells are not very motile and grow in dense colonies (Figure 6.16A). Here, we selected all cells which live at the time point 50 hours after movie start and show their spatial trajectories colored by movie time. The different colors at the end of each trace highlight nonsynchronized cell division events of a clonal colony. By standardizing the movie time to percentage of cell-cycle progression and thereby synchronizing the cells, we cannot observe a clear pattern of coordinated cell movement.

To increase the power of correlation analysis, we want to include all available colonies. Since Pearson correlation is scale invariant, we can integrate three independent experiments of unbiased and unsorted NanogVENUS trees. To reduce every colony to one correlation coefficient, we select one individual time point for each colony showing a maximum number of simultaneously living cells. Since one specific time point incorporates



cells in different cell-cycle stages and cell division events are not synchronized, we take cell features calculated at cell birth for each cell, map them to the position of the selected time point and correlate them to the distance to the colony center (Figure 6.14B). Thereby, we correct for cell-cycle effects, but use the spatial information of one particular time point.

We calculate the correlation coefficients for 156 colonies (Figure 6.17). On the global scale, the nuclear area at cell birth is not correlated to the distance to the colony center ($\rho = -0.06 \pm 0.26$), indicating that the nuclear area is independent of the position within the colony. Similarly, the NanogVENUS intensity is not affected by the distance to the colony center ($\rho = -0.02 \pm 0.28$) validating that all analyses on NanogVENUS do not

Figure 6.13 (*preceding page*): Differences of Nanog expression in related cells show little correlation with other components of the pluripotency network. (A) Low-sorted NanogVENUS fusion ESCs were imaged continuously for 70h and then fixed before antibody staining of other TFs. Partial correlation networks of subtrees generated from a single-cell were compared. We find no apparent connection between Nanog expression and that of other TFs (Pearson or partial correlations) in the subtrees. Indeed, subtrees with similar NanogVENUS expression show variegated correlation structures. (B) Partial correlation analysis of subtrees from a Nanog negative/low colony also shows variability between subtrees. Only significant ($p < 0.05$, partial correlation population test) edges are shown in (A) and (B). (C) To detect changes in TF expression shortly after Nanog upregulation, ESCs were sorted for low NanogVENUS, imaged continuously for 24 hours and fixed, stained and others TFs quantified. (D) A sister pair divergent in NanogVENUS shows congruent divergence in Oct4, and non-divergent levels for Sox2 and Klf4. Notably, Klf4 is only marginally expressed. (E) Klf4+/NanogVENUS- and Klf4-/NanogVENUS+ subpopulations were detected 24h after low sorting for NanogVENUS ($nE=3$, $n\rho = 783$). (F) Example of a NanogVENUS+/Klf4- colony (green outline) next to a Klf4+ colony. Interestingly, NanogVENUS expression can be mosaic in the absence of Klf4. (G) Sister cells with divergent NanogVENUS expression show significant congruent divergence of Oct4 and Klf4, but not for Sox2 (Wilcoxon rank sum test, $nE=3$, $n\rho = 136$, combination of experiments shown). The asterisks indicate significant, pairwise difference between transcription factor expression of the sister cells. Bar plots show absolute number of sister pairs which are congruently divergent, incongruently divergent or non-divergent (cf. Figure S6). (H,I,J) Sister cells which show divergent expression of Oct4, Sox2 or Klf4 levels also show divergence for the other factors, with the exception of Nanog (H,I) ($nE=3$, combination of experiments shown). Figure adopted from Filipczyk et al., in review.

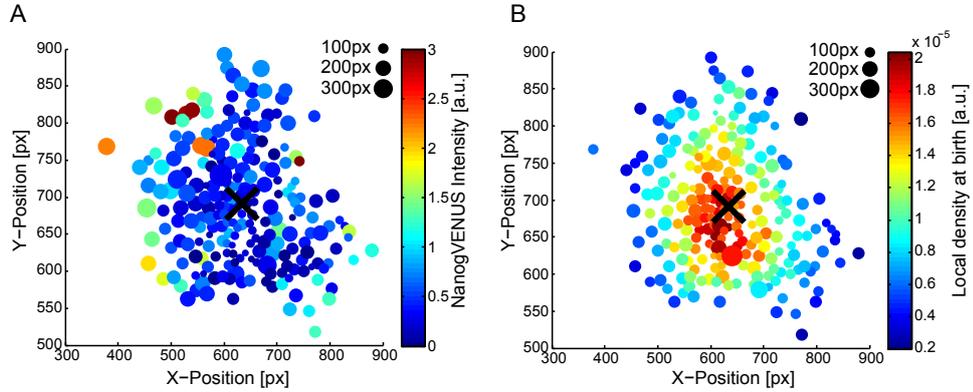


Figure 6.14: Quantification of cellular features allows correlation analysis of spatial dependencies. (A) One exemplary colony has been tracked and quantified. Every dot represents a cell at its specific location at a specific time point in the movie. The size of a circle encodes for the cell area and the intensity of NanogVENUS is color-coded. The 'x' marks the euclidean center of the colony. (B) Tracking information allows to map the cell-specific nuclear area and the local density at cell birth to the current time point. A strong anti-correlation of density at cell birth to the distance to colony center can be observed ($\rho = -0.95$). Number of cells $n = 223$.

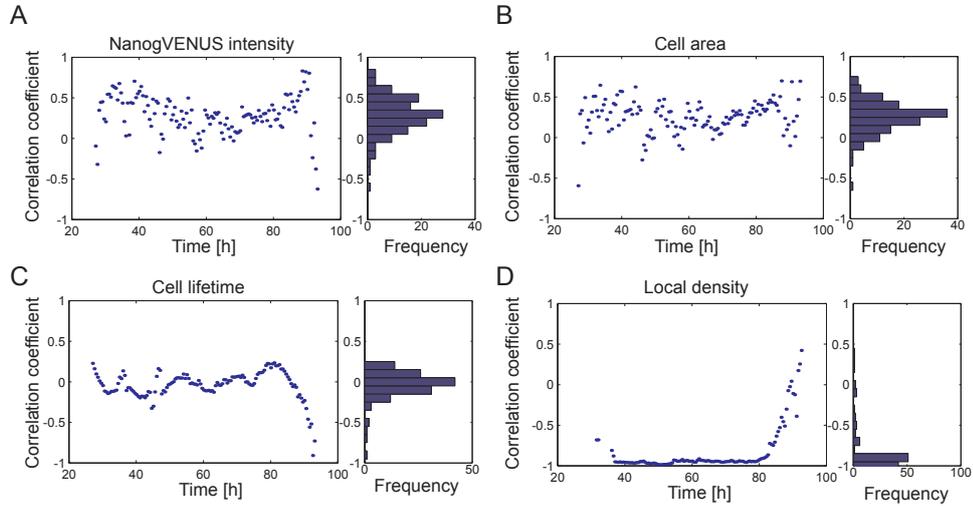


Figure 6.15: Pearson correlation coefficients to the distance of the colony center are independent of time. (A) NanogVENUS intensity shows a weak correlation to the distance of the colony center ($\rho = 0.29 \pm 0.24$). (B) The nuclear area is weakly correlated to the distance to the center of the colony ($\rho = 0.26 \pm 0.20$) (C) Lifetime is independent of the distance to the colony center ($\rho = -0.04 \pm 0.17$) (D) The local density is highly anti-correlated to the center of the colony ($\rho = -0.83 \pm 0.28$)

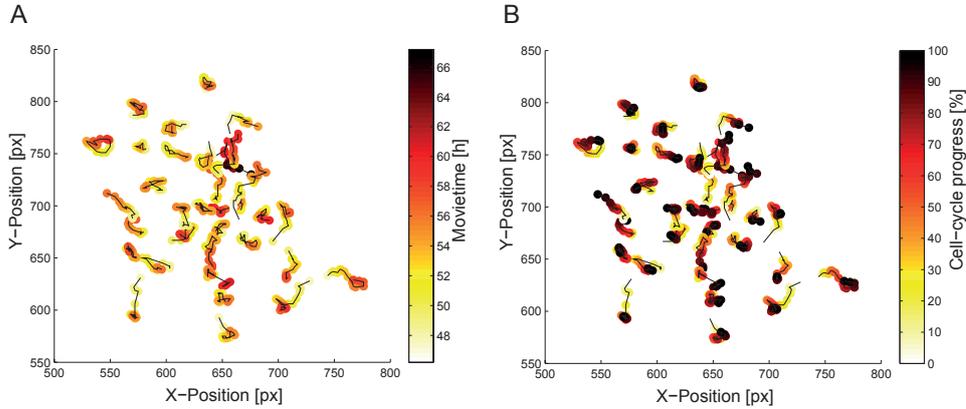


Figure 6.16: ESCs have a slow movement within a colony. (A) All cells living at the time point 50 hours after movie start are plotted. The color-coded movie time highlights the nonsynchronized cell division events. (B) Cell trajectories have been normalized to cell-cycle progression.

include any spatial bias. The lifetime of a cell is also independent of the position within the colony ($\rho = 0.07 \pm 0.25$). Finally, a strong negative correlation can be observed throughout all colonies between the distance to the colony center and the local density ($\rho = -0.90 \pm 0.09$). This result indicates that cells do not show a high diffusion property resulting in a very dense colony center and a sparse periphery resulting in an increased cytoplasmic cell area for cells at the periphery. However, the nuclear area is independent of the colony position probably due to its rigid structure.

For this analysis we always selected an individual time point with a maximum number of cells, but also a fixed time point for all colonies did not change the observed results (data not shown). Similarly, when using the median cell as colony center instead of the euclidean center of the colony results were not affected (data not shown).

6.11 Discussion

The understanding of pluripotency and its regulation is of great importance in stem cell research. ESCs are one well-defined cell population, yet they can differ amongst each other largely and show a variety of transcription factors heterogeneously expressed. Therefore, ESCs also serve as a perfect model system to investigate well-defined heterogeneous systems. Up to now, most studies are build one static analyses, some already incorporating single-cell resolution. Kinetics of transcription factors involved in the core pluripotency network are poorly understood due to lack of continuous time series data.

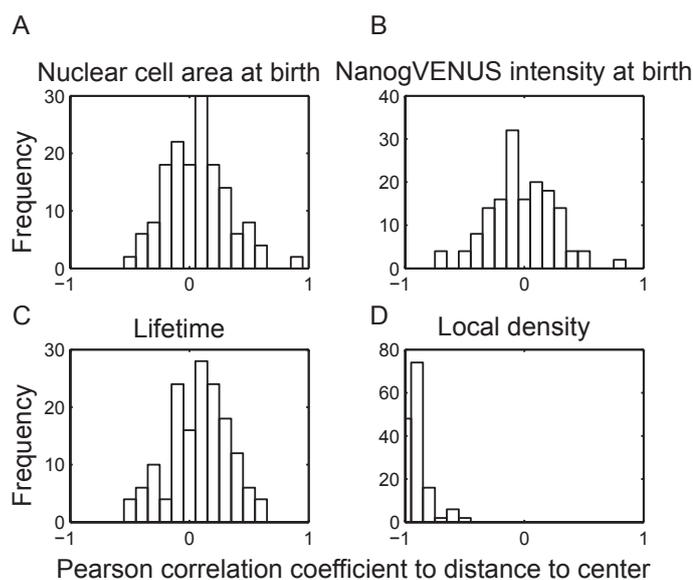


Figure 6.17: Three independent experiments have been combined to summarize correlations within colonies to the distance to their respective colony center ($n=156$). Overall, cell area at cell birth is not correlated to the distance of the center indicating a constant nucleus size ($\rho = -0.06 \pm 0.26$). Similarly, NanogVENUS intensity at cell birth is unaffected by the distance to the colony center ($\rho = -0.02 \pm 0.28$). The lifetime of a cell does not show a correlation to the distance of the colony center ($\rho = 0.07 \pm 0.25$). A strong negative correlation can be observed in almost every individual colony between the local density and the distance to the colony center ($\rho = -0.90 \pm 0.09$). All correlation coefficients are mean \pm s.d.

In our work we used a knock-in fusion line which in contrast to former used reporter lines shows the actual underlying protein amount in individual cells.

Using this line, we could show that NanogVENUS indeed shows a much broader distribution compared to Oct4VENUS, which also includes cells with negative NanogVENUS expression. Heterogeneous expression has been recently questioned in a different study (Faddah et al., 2013). In this artificial environment of ESCs, culture conditions play a crucial role, which could potentially explain some of the diverse results which have been reported (Smith, 2013).

On a population scale we showed that NanogVENUS maintains its steady-state distribution and that NanogVENUS is not lost during cell division.

We observed whole genealogies which stayed NanogVENUS negative for many generations. We exclude the previously reported mono-allelic behavior (Miyanari and Torres-Padilla, 2012), since our cell line has been shown to express Nanog biallelically (Filipczyk et al., 2013). On the other hand, we observed NanogVENUS negative cells, which repopulated the whole intensity distribution resulting in NanogVENUS mosaic colonies. Importantly, when quantifying other pluripotency factors at the endpoint of the movie, different wiring of the underlying correlation structure between both subtypes could be identified. The partial correlation of Oct4 and Klf4 shows a significant difference between negative colonies and mosaic colonies. Furthermore, the correlation analysis could show that Rex1 is rather placed between Klf4 and Nanog instead of downstream of both factors.

In current literature it has been widely discussed if heterogeneously expressed transcription factors either share one common upstream master regulator (Young, 2011), are the result on biophysical constraints (Chalut et al., 2012) or originate from cell states leading to specific protein stability (compare Chapter 7). We investigated the wiring of Oct4, Sox2, Klf4 and Nanog in different scenarios, where NanogVENUS low sorted ESCs have been cultured for one or four days. We observed no clear wiring, but only heterogeneous correlations between all measured factors. No clear pattern could be observed which could be linked to any cell or colony fate. This was highly unexpected and much more complex than previously assumed.

After investigating spatial dependencies within colonies we found that NanogVENUS intensity, nuclear area as well as cell lifetime is not affected by the position within a colony. On the other hand, the local density of a cell decreases with the distance to the colony center indicating that cells spread out on the periphery.

The results of our work supports the idea of micro-heterogeneity, which has been introduced recently (MacArthur et al., 2012, Trott et al., 2012). In such an environment, the antagonistic lineage transcription factors Oct4 and Sox2 (Thomson et al., 2011) have the possibility to maintain pluripotency while at the same time it allows them to orchestrate their roles as lineage

priming regulators. So far, we did not get a clear insight into how this chaotic and non-intuitive behavior is coordinated, which will be of great interest for further experimental and modeling studies (MacArthur and Lemischka, 2013).

Our work highlights the importance of continuous single-cell quantification which is crucial for the understanding of kinetics in heterogeneous systems. Without them, predictive model of the pluripotency network structure would be unfeasible, which may ultimately lead to controlled differentiation programs or routed cell reprogramming supporting future regenerative medicine applications.

Chapter 7

Single-cell quantification of protein stability

Methods, results and figures in this section are partly based on Schwarzfischer et al., in review. All biological experiments have been performed by Adam Filipczyk and Philipp Hoppe and are not part of this thesis. My contributions are:

- Development and application of QTFy to mESC data
- Tracking and quantification of single-cell data
- Analysis of Western Blots
- Implementation and analysis of decay models
- Statistical analysis of model results

In the following section we investigate the protein half-life of individually tracked and quantified single-cells using QTFy (Section 3.3). We used movies with NanogVENUS knock-in ESCs where protein synthesis is blocked by cycloheximide treatment allowing to study the decay of proteins on a single-cell basis. We compare our approach to typical Western Blot experiments and validate our approach. We introduce a simple exponential model which estimates protein half-life for Western Blot data, snapshot population averages and single-cell trajectories. Only by integrating the full time and single-cell resolution we unravel cell-to-cell variability in protein half-lives and also identify a stable fraction of proteins. Finally, we apply our pipeline to rare cell populations to investigate protein half-lives of PU.1eYFP and Gata1mCHERRY in HSPCs of various differentiation stages.

7.1 Biological background

The regulation of protein abundance is of central importance for controlling cell fates. For example, the pluripotency transcription factor protein Nanog is known to play a pivotal role in differentiation and the maintenance of pluripotency in ESCs (compare Chapter 6) (Chambers and Tomlinson, 2009). Nanog is heterogeneously expressed in ESCs (Chambers et al., 2007, Navarro et al., 2012), however, the functional relevance for pluripotency as well as the regulatory mechanism underlying this heterogeneity are not well understood. Protein abundances are regulated in a highly complex manner, involving not just transcriptional and translational control, but also post-translational modification including the fine-tuned degradation of specific proteins (Muñoz Descalzo et al., 2013). Due to current technical limitations, protein half-life is typically investigated using population averages (Schwanhäusser et al., 2011) and only rarely at the single-cell level (Eden et al., 2011, Halter et al., 2007). However, ESCs can easily be cultured in large numbers, permitting the comparison of single-cell results with established biochemical analyses requiring many cells.

7.2 Population average analysis

Conventionally, Western Blot analysis is used to determine protein half-life after blocking protein translation by cycloheximide treatment (Figure 7.1A). We measured the relative population average of protein abundance, $y(t)$, using ImageJ software with its plug-in Gel Analyzer which also corrects for background intensity (Schneider et al., 2012). The decreasing protein abundance is fitted by an exponential decay model with decay rate β , unstable amount α and a potential steady-state level (stable amount) γ (Figure 7.2A):

$$y(t) = \alpha e^{-\beta t} + \gamma \quad (7.1)$$

The initial protein amount is given by $\alpha + \gamma$. The protein half-life $t_{1/2}$ can be calculated from the decay rate by $t_{1/2} = \frac{\log(2)}{\beta}$.

We assume that each Western Blot replicate j follows the same decay rate β and steady state level γ but a varying initial protein amount. We modify the model to include this by replacing α with a scaling factor α_j^* for each replicate j . This scaling factor is both responsible for correctly fitting the individual initial concentrations as well as ensuring that all replicates reach the same steady state level. In Figure 7.2A, we show a common time course for the three Western Blot replicates by using the average of the α_j^* 's. For every observed data point we expect some measurement error $\epsilon_j(t)$. In accordance with the literature (Kreutz et al., 2007), we assume this noise to be multiplicative lognormally distributed with expectation one.

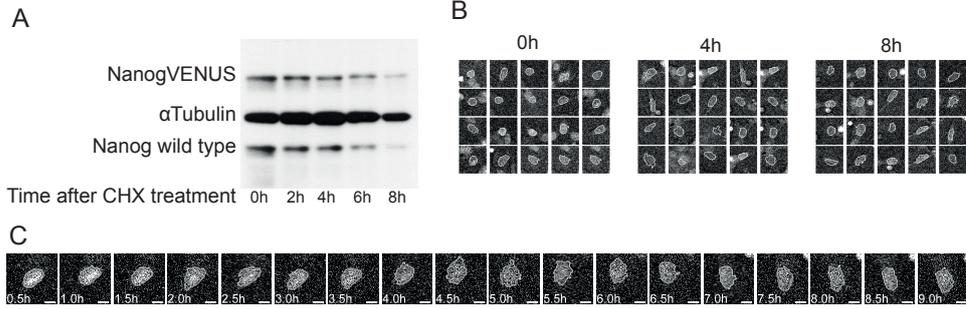


Figure 7.1: Three different experimental approaches allow to measure protein half-life. (A) Classical Western Blot analysis showing NanogVENUS and wild type Nanog protein decay after cycloheximide treatment. (B) Single-cell NanogVENUS levels have been quantified after cycloheximide treatment in a time series of images. (C) A single NanogVENUS knock-in ES cell has been continuously tracked and quantified. The cell intensity is displayed as s , scale bars are $10\mu\text{m}$. Figure adopted from Schwarzfischer et al., in review.

Furthermore, the noise parameter σ for this lognormal distribution should not depend on time. This yields the complete model for the j -th replicate:

$$y_j(t) = \alpha_j^* \cdot (e^{-\beta t} + \gamma) \cdot \epsilon_j(t) \quad (7.2)$$

$$\epsilon_j(t) \sim LN\left(\frac{-\sigma^2}{2}, \sigma\right)$$

We calculate the maximum likelihood estimate by optimizing the corresponding loglikelihood function with respect to the parameters using a local optimization routine in MATLAB. We achieve global convergence by restarting the optimization 10,000 times at different initial parameter values according to a Latin-hypercube sampling scheme (McKay et al., 1979). The uncertainty in the maximum likelihood estimate for the parameters can be assessed from the Hessian matrix H of the loglikelihood, which is in this case called the Fisher information matrix. Thus, the variances of the parameters can be directly found on the diagonals of the inverse of the corresponding Hessian:

$$H^{-1} = \begin{pmatrix} \Delta\alpha & \dots & \dots & \dots \\ \dots & \Delta\beta & \dots & \dots \\ \dots & \dots & \Delta\gamma & \dots \\ \dots & \dots & \dots & \sigma \end{pmatrix} \quad (7.3)$$

The uncertainty $\Delta t_{1/2}$ of protein half-life $t_{1/2}$ is computed by propagation of uncertainty (Sivia, 1996): $\Delta t_{1/2} = \frac{\log(2)}{\beta^2} \cdot \sqrt{\Delta\beta}$.

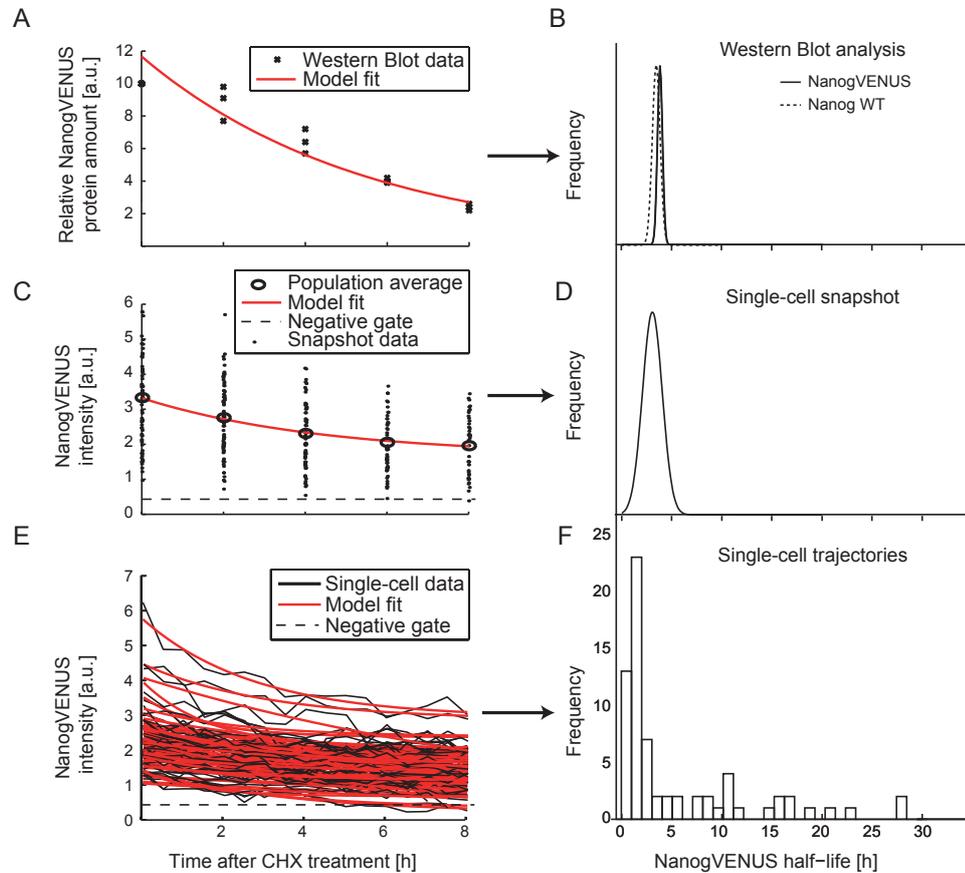


Figure 7.2: Continuous single-cell microscopy allows accurate measurement of heterogeneous protein decay. (A) Western Blot quantification of three independent biological replicates (black dots) was fitted by individual models with shared parameters for the exponential decay. (B) Half-lives of NanogVENUS estimated by Western Blot analysis including uncertainty over three independent biological replicates. Nanog wild type (WT) protein half-life is not significantly different to its respective VENUS fusion knock-in half-life (two sided Welch's t-test p -value > 0.69 , $n=3$). (C) Single-cell population averages over 5 consecutive time points were fitted by the same model as in (A) and revealed that NanogVENUS does not fully decay, but a fraction of stable NanogVENUS remained above the negative gate ($n=3$, one representative shown). (D) Half-life of NanogVENUS estimated by single-cell snapshot data including uncertainty over three independent biological replicates. (E) A cohort of NanogVENUS cells has individually been fitted by the decay model ($n=7$, one representative shown). (F) Half-life distribution of NanogVENUS of single-cell trajectories illustrating the heterogeneity within the population (one representative shown, $n=7$). Figure adopted from Schwarzfischer et al., in review.

We have previously shown that Nanog wild type (wt) and a Nanog-VENUS (NV) fusion protein exhibit comparable protein half-lives (Filipczyk et al., 2013). Western Blot analysis after cycloheximide treatment yields very similar stabilities for wild type ($t_{1/2}^w = 3.69 \pm 0.81$ h) and VENUS fusion proteins ($t_{1/2}^w = 3.81 \pm 0.50$ h, mean \pm 2s.d., n=3 each, Figure 7.2B), with no significant difference between wt and NV (p-value > 0.69 , two-sided Welch's t-test, Figure 7.2B).

If the cycloheximide treatment worked perfectly and every protein is degradable, the steady state level should be zero ($\gamma = 0$). Since we extended the simple exponential decay model by allowing the possibility of incomplete protein degradation ($\gamma \neq 0$), corresponding to a stable protein compartment we compared both models. The preferred model can be theoretically assessed with the likelihood ratio test (Vuong, 1989). However, we found that both wt and NV Western Blot data did not reject the more complex model with a non-decaying fraction (likelihood-ratio test p-value > 0.05). Thus, with Western Blot data no distinction between a complete or a partial decay can be obtained.

7.3 Single-cell snapshot analysis

To study the degradation dynamics in more detail and to unravel potential cell-to-cell variability, we performed single-cell time-lapse experiments with ESCs expressing NanogVENUS fusion transcription factor for 8 hours from the start of cycloheximide treatment. After quantifying single-cells in snapshot images every two hours by segmenting all single-cells in the image using sQTFy (Section 3.4), ending up with fluorescence quantifications leading to a population average over time (Figure 7.1B). The snapshot data thus consists of single-cell fluorescence measurements of NanogVENUS in approximately 60 to 100 cells at discrete time points for each of three replicates.

We use the same exponential decay model (Equation 7.1) to fit the resulting population averages ignoring temporal correlations (Figure 7.2C). In contrast to Western Blot data, but in accordance with literature (Harper et al., 2011), we now assume multiplicative Γ -noise with expectation one for fluorescence measurements:

$$y_j(t) = \alpha_j^* \cdot (e^{-\beta t} + \gamma) \cdot \epsilon_j(t) \quad (7.4)$$

$$\epsilon_j(t) \sim \Gamma(k, \frac{1}{k})$$

Here $\Gamma(k, \theta)$ denotes a Gamma-distribution with shape parameter k and scale parameter θ . Since the expectation should be one, we set $\theta = \frac{1}{k}$. As for the Western Blot data, we optimize the loglikelihood of the model to find the maximum likelihood estimate for the parameters, using the same restart scheme for the local optimizer.

The resulting protein half-life $t_{1/2}^s = 3.06 \pm 1.98$ h (mean \pm 2s.d., $n=3$, Figure 7.2D) is not significantly different to NV Western Blot results (p-value > 0.32 , two-sided Welch's t test). However, in contrast to Western Blot analysis, the NV single-cell image data were best modeled using a stable protein fraction (34.4 ± 13.8 %, mean \pm 2s.d., likelihood-ratio test p-value < 0.03 , $n=3$, Figure 7.2C). This result highlights the importance of performing single-cell investigations followed by manual data curation, as conclusions drawn only upon bulk analysis measurements may be wrong.

We would like to note that using a noise model with log normally distributed multiplicative noise (as used for Western Blot analysis) does not significantly change the results for snapshot data (p-value > 0.93 , determined by Welch's t-test).

7.4 Single-cell trajectory analysis

Finally, we quantified the decay process and the effective protein half-life of NanogVENUS in single-cell trajectories (Figure 7.1C). We tracked and quantified 50 to 100 cells per replicate using the standard procedure with TTT (Section 3.3.1) and QTFy (Section 3.3), ignoring cells which lived less than 5 time points.

We fitted the exponential decay model to each single-cell trajectory, obtaining cell-specific sets of parameters (Figure 7.2E). The model parameters now include an index j for the replicate and i for the cell within the replicate. As for the snapshot data, we assume Gamma-distributed noise of expectation one:

$$y_{ij}(t) = (\alpha_{ij}e^{-\beta_{ij}t} + \gamma_{ij}) \cdot \epsilon_{ij}(t) \quad (7.5)$$

$$\epsilon_{ij}(t) \sim \Gamma(k_{ij}, \frac{1}{k_{ij}})$$

For every cell i in each replicate j we again optimize the log-likelihood of the model to find the maximum likelihood estimate for the parameters, using the same restart scheme for the local optimizer as for the Western Blot and snapshot data.

As expected and described previously (Chambers et al., 2007), total NanogVENUS amount at the initial time point per cell was heterogeneous (Figure 7.2E). In addition, NanogVENUS half-lives were also remarkably heterogeneous ranging from 0.61h to 68.16h (1% and 99% quantile over replicates of the single-cell half-life distribution, $n=5$, Figure 7.2F). Since these distributions are not normally distributed and rather heterogeneous within replicates, we combine them by taking the average over respective medians of all replicates ($t_{1/2}^c = 2.34 \pm 0.64$ h, average over replicate medians \pm 2s.d., $n=7$). Although similar to previous methods, this finding is nonetheless significantly different from Western Blot analysis (p-value < 0.0006 , two-sided

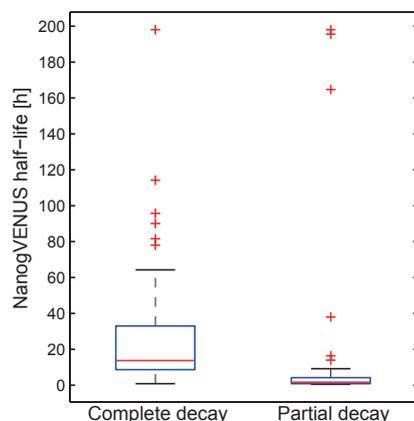


Figure 7.3: NanogVENUS half-life is significantly higher in cells that decay completely compared to cells which only partly decay (two sided Welch’s t test p-value <0.0127 , $n=5$, one representative shown). Figure adopted from Schwarzfischer et al., in review.

Welch’s t test). Importantly, the heterogeneous total NanogVENUS protein amount at the initial time point is not correlated with the heterogeneous NanogVENUS protein half-life (Figure 7.4A, $R^2 = 0.02 \pm 0.03$, $n=5$) in individual cells. We may thus conclude that differential protein stability is not the reason for heterogeneous Nanog expression.

Our approach also reveals two distinct categories of cells: cells with complete NanogVENUS decay ($30.7 \pm 7.0\%$, $n=5$) and cells with only partial decay and a stable amount of NanogVENUS (Figure 7.4B and C). When we further analyze differences between the two groups, we observed that, counter-intuitively, the half-life of NanogVENUS protein is significantly higher ($t_{1/2}^c = 9.16 \pm 8.08$ h, average median ± 2 s.d., $n=5$) in cells with complete decay than in cells with only partial decay ($t_{1/2}^c = 1.56 \pm 1.32$ h, average median ± 2 s.d., $n=5$, p-value <0.02 , two sided Welch’s t test, Figure 7.3).

In summary, our framework allows reliable protein half-life estimation which is highly comparable with conventional methods (Table 7.1). In addition, continuous single-cell trajectories unravels decoupled heterogeneous cellular protein half-lives and stable protein fractions, whereas population-level data cannot.

7.5 Cell-cycle dependence

To further investigate the sources of heterogeneity of NanogVENUS and a possible cell-cycle dependence of protein half-life, we performed additional movies where we added cycloheximide at day 2 of a time-lapse experiment.

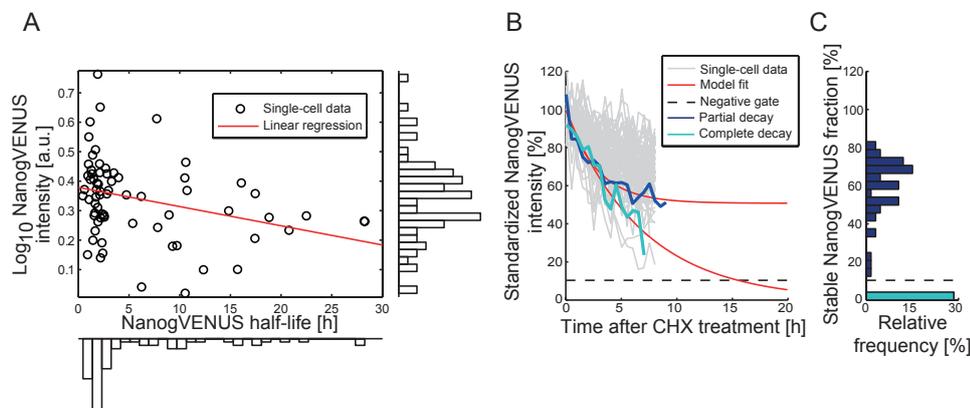


Figure 7.4: Single-cell analysis reveals a heterogeneous stable protein fraction. (A) The heterogeneity in NanogVENUS intensity cannot be explained by the emerging heterogeneity in NanogVENUS half-life ($R^2 = 0.02 \pm 0.03$, $n=5$, one representative shown). (B) On a single-cell level, we observe two types of dynamics upon cycloheximide treatment in ESCs: cells with partial protein decay and a stable amount of protein above the negative gate and cells with complete decay. For each group one representative time course and its model fit is shown. Gray lines illustrate all other cell trajectories of this representative example. All cell trajectories have been standardized to the same initial value after fitting. (C) The bimodal distribution of stable NanogVENUS fraction further demonstrates the two distinct groups. Figure adopted from Schwarzfischer et al., in review.

Method	NanogVENUS half-life [h]	Nanog wild type half-life [h]
Western Blot analysis ($t_{1/2}^w$)	3.81 ± 0.50	3.69 ± 0.81
Single-cell snapshot ($t_{1/2}^s$)	3.06 ± 1.98	NA
Single-cell trajectories ($t_{1/2}^c$)	2.34 ± 0.64	NA
	Oct4VENUS half-life [h]	Oct4 wild type half-life [h]
Western Blot analysis ($t_{1/2}^w$)	9.96 ± 1.23	7.32 ± 1.00

Table 7.1: Comparison of NanogVENUS with Nanog wild type and Oct4VENUS with Oct4 wild type half-lives determined by different techniques represented as mean \pm 2s.d. over replicates.

Since the divisional history of all individual cells was thus known at the time point of cycloheximide addition, protein decay could be quantified in dependence of cell-cycle progression (Figure 7.5A). A cell has to approximately double its protein amount during one cell-cycle in order to maintain a steady protein concentration. We therefore analyzed the correlation of total protein amount at cycloheximide addition and time in cell-cycle as a plausibility check. Indeed, on average, total amount of protein increased with cell-cycle progression for NanogVENUS (Pearson correlation coefficient $\rho = 0.41 \pm 0.26$, $n=5$, Figure 7.5B). Interestingly, the heterogeneous protein half-life showed no significant correlation with cell-cycle stage for NanogVENUS (Pearson correlation coefficient $\rho = 0.06 \pm 0.04$, $n=5$, compare Figure 7.5C). However, we found that the amount of stable NanogVENUS protein is correlated with cell-cycle (Pearson correlation coefficient $\rho = 0.39 \pm 0.18$, $n=5$, Figure 7.5D), indicating that NanogVENUS proteins get stabilized during cell-cycle. The later a cell was in cell-cycle at cycloheximide addition, the more non-decaying protein it had accumulated. Intriguingly, when looking at stable fractions this effect diminishes and the fraction of stable proteins is only weakly correlated with cell-cycle (Pearson correlation coefficient $\rho = 0.25 \pm 0.15$, $n=5$, data not shown). This indicates, that the percentage of stable NanogVENUS proteins slowly increases with cell-cycle, but also these proteins have to be destabilized right after cell division.

7.6 Protein half-life in siblings

Time-lapse experiments with cycloheximide treatment after 2 days additionally allowed us to investigate correlations between sister cells. A simple correlation test is not applicable to this kind of data since the parameters of two sisters can be arbitrarily sorted and do not have any order. Therefore, we took the model parameters $p \in \{\alpha, \beta, \gamma, \sigma, \dots\}$ of the individual fits of sister $s_1(p)$ and $s_2(p)$ and calculated the relative difference $d(p)$ for all sister pairs, defined as

$$d(p) = \frac{|s_1(p) - s_2(p)|}{\frac{1}{2}(s_1(p) + s_2(p))}. \quad (7.6)$$

We compare the resulting distribution with a distribution derived by calculating the relative difference of 10,000 randomly chosen cells. As a plausibility check of this method we calculate the real and a random distribution for the sister intensity difference after cell division which shows small difference for real sister cells (< 1) due to the symmetric segregation of NanogVENUS after division (Figure 7.6). The real distribution significantly differs from the distribution of random sister cells, as expected (Kolmogorov-Smirnov test p -value < 0.005 for all experiments, $n=5$). Secondly, real sister pairs have a significantly lower relative difference determined by a Wilcoxon rank sum test (p -value < 0.005 for all experiments, $n=5$). These three facts, the

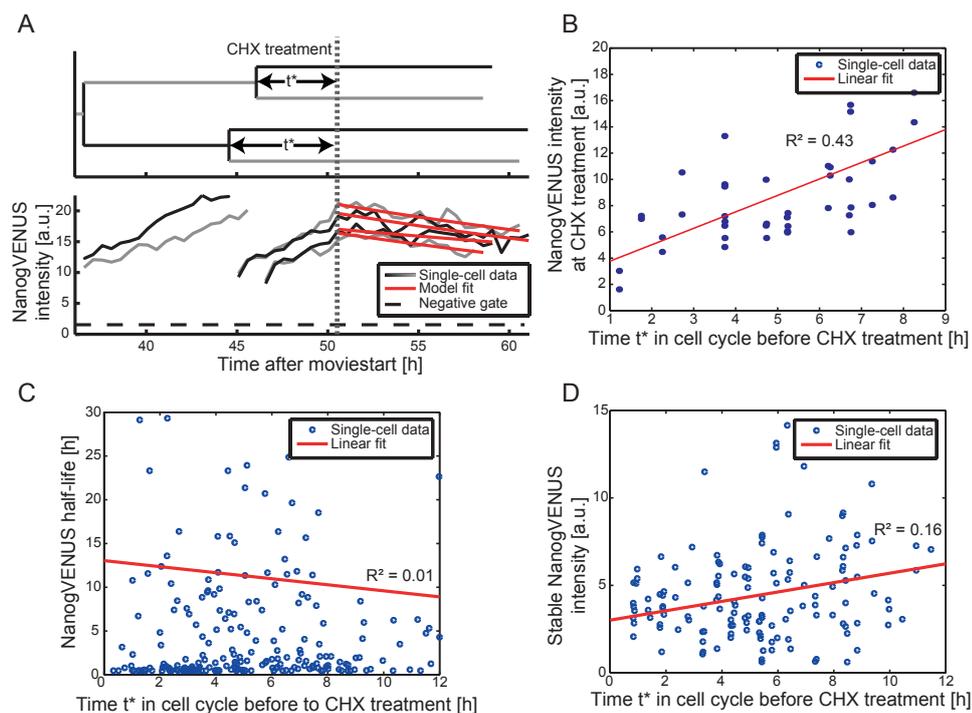


Figure 7.5: Continuous single-cell half-life quantification after cycloheximide treatment reveals cell-cycle dependence of model parameters. (A) Cells have been treated by Cycloheximide (CHX) after 2 days of normal movie time which allows to analyze model parameters in dependence of cell-cycle progression (t^*). (B) NanogVENUS intensity at cycloheximide treatment positively correlates with cell-cycle time (Pearson correlation coefficient $\rho = 0.41 \pm 0.26$, $n=5$, one representative shown). (C) The heterogeneously measured NanogVENUS protein half-life shows only weak correlation with cell cycle time (Pearson correlation coefficient $\rho = 0.06 \pm 0.04$, $n=5$, one representative shown). (D) The cell-cycle time t^* shows a positive correlation with the stable NanogVENUS intensity (Pearson correlation coefficient $\rho = 0.39 \pm 0.18$, $n=5$, one representative shown).

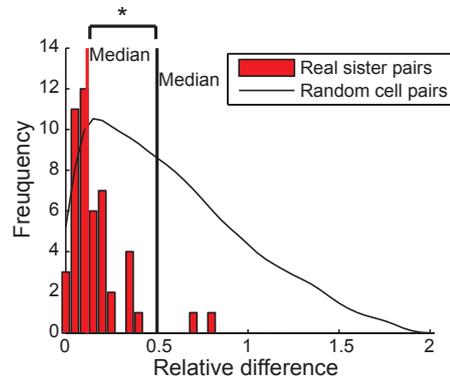


Figure 7.6: Sister cells are more similar in their model parameters compared to randomly picked cells. The relative absolute difference in intensity after birth shows a significantly lower median compared to a relative absolute difference distribution of 10,000 randomly picked cells (Wilcoxon rank sum test p-value < 0.005 for all 5 experiments, one representative shown).

difference in the distribution, the significant difference in the median and the lower median, demonstrate that real sister pairs are more equal to each other than randomly picked cells.

By investigating the other modeling parameters, we found that the median relative difference in NanogVENUS half-life is always smaller between real sisters compared to randomly picked cells. However, with our test of significance we could only show in three out of five experiments that real sister pairs are significantly more similar to each other compared to our null model.

7.7 Protein half-lives in rare cell types

single-cell analysis has the great advantage of being possible using only very few cells. This allowed us to investigate protein half-life of transcription factors in primary blood cells where cells are extremely rare and costly to purify (only about 1000 cells from one mouse) (Warr et al., 2011). Using the same approach as for ESCs, we analyzed the stability of PU.1eYFP and Gata1mCHERRY in hematopoietic stem and progenitor cells (HSPCs). HSPCs were isolated from a knock-in mouse line expressing these fusion proteins from the endogenous gene loci of PU.1 and Gata1, respectively (compare Chapter 5), two transcription factors that are widely thought to determine an important hematopoietic lineage decision (Graf and Enver, 2009, Nerlov and Graf, 1998).

Protein half-lives could be estimated for several HSPC populations and both transcription factors (see Table 7.2). Interestingly, we find that the

Protein (cell population)	Number of experiments	Median half-life [h]	Unstable percentage [%]
PU.1eYFP (HSC)	5	7.3 ± 2.0	67 ± 7
PU.1eYFP (MPP)	4	7.9 ± 4.7	69 ± 22
PU.1eYFP (GMP)	4	9.2 ± 7.7	65 ± 12
PU.1eYFP (Gatamid ery- throid progenitors)	3	7.0 ± 2.4	69 ± 3
Gata1mCHERRY (Gatamid erythroid progenitors)	3	11.9 ± 5.0	65 ± 8
Gata1mCHERRY (Gatahigh erythroid progenitors)	3	2.1 ± 2.3	42 ± 17
NanogVENUS (ESC)	7	2.3 ± 0.3	59 ± 5
Oct4VENUS (ESC)	8	4.3 ± 2.7	49 ± 15

Table 7.2: Our approach allows determining a variety of single-cell protein features even in rare cell populations like primary cells. Mean protein abundance has been determined by western blot dilution assays and subsequent flow cytometry intensity mapping. Error estimates are derived by propagation of uncertainty. The average and standard deviation of median half-life is shown for every cell population. The average amount and fraction of stable and unstable protein has been computed by the model (see Online Methods for details).

median PU.1eYFP half-life and single-cell variability remains unchanged throughout early hematopoietic differentiation and in all HSPC populations analyzed. No significant difference could be found for Gata1mCHERRY half-lives between different HSPC populations (p-value > 0.05 , determined by Welch’s t-test). However, by further investigating HSPCs expressing high amounts of Gata1, continuous single-cell microscopy revealed that these cells die earlier and more frequently upon cycloheximide treatment than other HSPCs. Furthermore, we could detect sudden drops in fluorescence intensity, which commonly only occurs at cell death, in cell trajectories even in cells maintaining normal morphology (data not shown). Conclusions regarding measured protein half-lives therefore have to be taken with care. Importantly, all these effects would be hidden and would lead to inaccurate half-life estimates if using Western Blot analysis. As a last application of our toolbox in this context, we also performed experiments using Oct4 fused to VENUS (compare Chapter 6) and determined its half-life in ESCs (Table 7.2).

7.8 Discussion

Reliable determination of molecular parameters like protein half-life is of paramount importance for a quantitative understanding of transcription factor dynamics (Eden et al., 2011, Schwanhäusser et al., 2011). Here, we showed that continuous single-cell quantification reveals a large cell-to-cell variability in the protein half-life of NanogVENUS, a feature that is not detectable with conventional methods. Single-cell data further showed that this heterogeneous protein stability is not correlated to the observed heterogeneous protein abundance.

For the ESC system, our results on protein stability can be well compared to the conventional method of Western Blot analysis. We measure Nanog stability in the same range as previous approaches (Abranches et al., 2013, Muñoz Descalzo et al., 2013, Navarro et al., 2012, Ramakrishna et al., 2011). However, we also find deviations from the results of these population average based methods, since a stable fraction of proteins could not be identified in Western Blot analysis. Most likely, this is due to the fact that protein half-life estimates determined by Western Blot data are based on a cell lysate using all available protein, including that of dead cells. This problem is alleviated by single-cell analyses and manual data curation excluding dead cells or debris from the analysis. Furthermore, quantification of fluorescence signal of fusion proteins, that reflect the underlying wild type protein amount, does not suffer from potential problems with antibody specificity or linearity of staining, which may confound Western Blots analysis. In contrast to Western Blot analysis, our single-cell quantification of protein kinetics revealed fractions of stable, non-decaying proteins in a subpopulation of ESCs. This finding indicates that the heterogeneous expression profile of absolute Nanog protein levels is a mixture of different, uncoupled and cell-specific mechanisms including protein decay rates, stabilized proteins or protein synthesis. Since ESCs with a complete Nanog decay can also be observed, we exclude that the observed heterogeneous NanogVENUS stability is induced by clustering of the fluorescence protein as previously suggested (Landgraf et al., 2012). Instead, we hypothesize that another regulatory mechanism like protein complex formation or DNA-binding underlies the variable stability of Nanog in different cells. In the same spirit, a model has been recently proposed incorporating complex forming of Nanog and Oct4 building the molecular basis of pluripotency (Muñoz Descalzo et al., 2013).

By exploiting the full potential of our toolbox, we performed movies where cycloheximide was added after 2 days of normal movie conditions which allows to track full genealogies. These movies provide information of the cell-cycle stage of every single-cell at the time of treatment as well as their related kinship. We could show that protein half-life is not affected by cell-cycle progression which rules out another potential source for the Nanog heterogeneity. Interestingly, the stable amount of protein was posi-

tively correlated to the cell-cycle time indicating that Nanog gets stabilized towards cell division.

To investigate if two related sister cells share similar protein half-lives, we developed a null model based on randomly picked cell pairs and compared it to real sister pairs. With our method we only found in three out of five experiments that real sister pair are significantly more equal in protein half-life than randomly picked cells, but the impression exists that sister behave quite similarly. Although every data point in this analysis has been inspected manually, another manual investigation might be necessary to unravel these discrepancies. As far as we investigated the data we do not see any obvious difference between the experiments.

Importantly, our pipeline for quantification of protein half-lives at the single-cell level also works for very small cell numbers, which is not guaranteed for bulk analysis methods (Schroeder, 2008, 2011). This is of great importance for both research with rare cells, like adult primary stem cells, and quantification of heterogeneous cell populations. A protein half-life of 5.5 hours for PU.1 had previously been determined in splenocyte populations (Nutt et al., 2005), which is comparable to our single-cell measurements throughout all investigated HSPC populations. In the work presented here, PU.1 half-life has not been directly measured for macrophages, but it has been reported to be about 50 hours (Kueh et al., 2013). This would indicate that stability gets massively increased of about 10fold during differentiation from GMPs (showing a half-life of 9.2 ± 7.7) into these macrophages. For Gata1 we determined a varying protein half-life ranging from 2 to 11 hours depending on the cell type (Table 7.2). Together with an 1 hour half-life in fibroblasts (Hernandez-Hernandez et al., 2006) this indicates that Gata1 protein degradation is regulated in different cell types. Measuring protein half-life in primary stem and progenitor cells has not been feasible up to now.

To further demonstrate the power of our toolbox, experiments with two differently labeled proteins would be of interest, which allows investigating if heterogeneity of protein half-life is rather a result of protein specificity or cell-to-cell variability. Furthermore, it would be of great interest to see how the heterogeneous protein half-life behaves when exposing ESCs to other conditions like $2i$, which is know to force Nanog to be more homogenous. In a recent study, Muñoz Descalzo et al. (2013) reported that Nanog and Oct4 protein stability may vary depending on culture conditions measured by Western Blot technology.

Chapter 8

Identification of two distinguishable cell populations during gastrulation based on cell movement

The following section is based on the publication of Burtscher et al. (2012). All experiments have been conducted in the Institute of Stem Cell Research by Ingo Burtscher. My contributions are:

- Development of CCT (Section 3.5) and its optimization to provide an efficient single-cell tracking tool with additional cell labels
- Visualization of tracked cells in still images and videos (Figure 8.1)
- Analysis and statistics of cell movement between different cell populations

8.1 Biological background: embryogenesis

The development of an embryo and the establishment of its different germ layers is a complex process which is poorly understood. In the early stages only a few cells exist already showing diverse characteristics, however hidden by applying bulk analysis techniques (Burtscher and Lickert, 2009). Only a high temporal and spatial resolution good enough to monitor individual cells will help to understand the underlying processes of embryogenesis. Therefore, continuous live time-lapse imaging and single-cell tracking is an appropriate technique to follow individual cells with fluorescently labeled proteins of interest.

The epiblast of an embryo is surrounded by the visceral endoderm (VE), a layer of cells which contributes to the development of the yolk sac. The definitive endoderm (DE) is a population of cells arising also from the epiblast, which enters the primitive streak and gives rise to the gut tube. The VE and the DE are known to disperse during development.

Two transcription factors are known to be the earliest markers to identify endoderm development; the forkhead transcription factor *Foxa2* (Monaghan et al., 1993) and the SRY-related HMG box transcription factor, *Sox17* (Kanai-Azuma et al., 2002). So far it is unknown if these markers are co-expressed in the very same cell populations or if they could be used to distinguish the DE and VE cell populations.

8.2 Application and results

Ingo Bartscher in the Institute of Stem Cell Research (today Institute of Diabetes and Regeneration Research, Helmholtz Zentrum München) developed a knock-in mouse line which heterozygously expresses a *Sox17-mCherry* fusion (SCF) fusion protein from its endogenous locus. This protein allows to investigate the protein abundance of *Sox17* by its fluorescence reporter on a single-cell level. Ingo Bartscher performed time-lapse experiments during gastrulation using confocal microscopy. We used a maximum intensity projections for the subsequent tracking of individual cells using CCT (Section 3.5, Figure 8.1E₁).

We identified and selected DE and extra-embryonic VE (ExVE) cells in the early movie (Figure 8.1E₁, green and yellow dots, respectively) and tracked them over the whole movie (Figure 8.1E₂). It turned out that DE cells are highly migratory compared to ExVE cells which were rather stationary (Figure 8.1E₃). The average displacement between two subsequent time points (5 min) per cell of ExVE cells ($0.68 \pm 0.22 \mu\text{m}$, $n=16$) was significantly lower than the displacement of DE cells ($3.76 \pm 0.92 \mu\text{m}$, $n=35$, mean \pm s.d., Welch's t-test p-value $< 8 \cdot 10^{-22}$).

The newly generated mouse line which expresses SCF can be utilized to monitor *Sox17* protein expression during endoderm development using fluorescence microscopy or flow cytometry. We could identify two distinguishable populations of cells separating into DE and ExVE based on cell movement. Here, SCF was only used to identify and track the cells, but further investigations could quantify the expression to see if also differences on a molecular scale exist. In this work, we established a framework to get a better understanding of embryo development on a single-cell resolution. Together with other recently developed mouse strains which express fluorescently labeled transcription factors *Foxa2* or *T*, they provide all necessary tools to investigate and even to quantify labeled protein expression during differentiation from an early embryo stage up to the formation of the germ

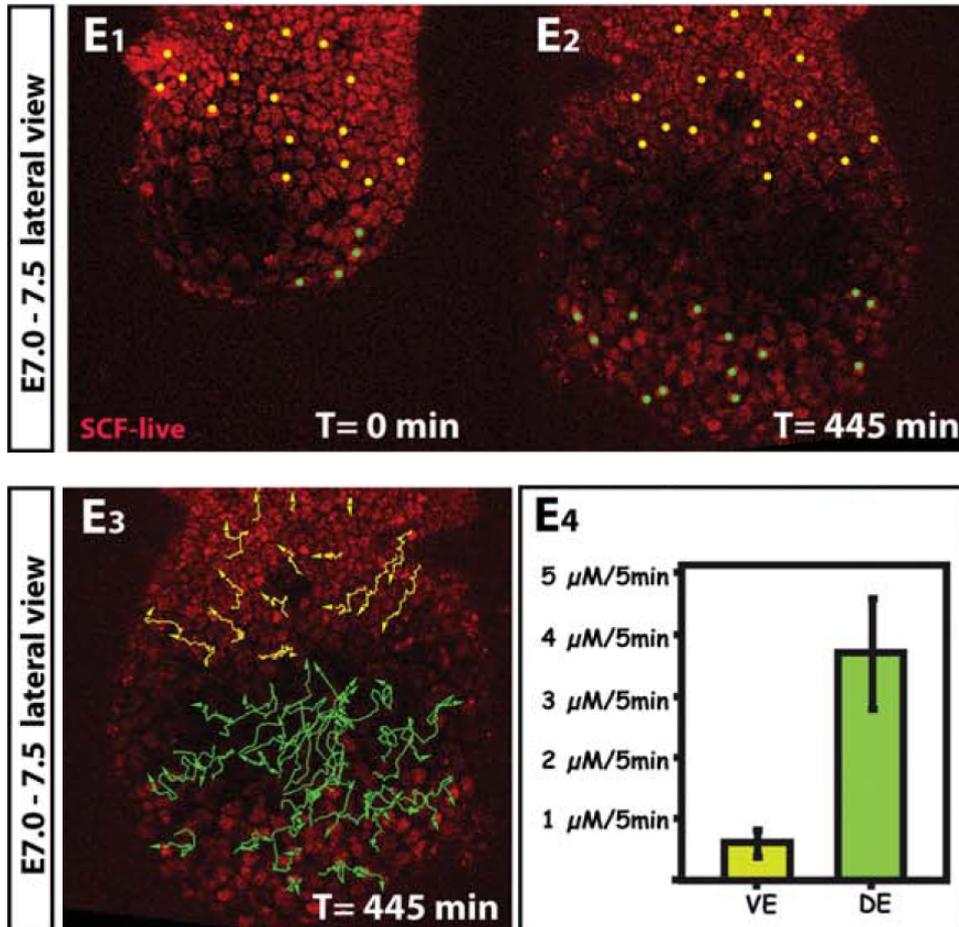


Figure 8.1: Spatial and temporal SCF reporter activity during gastrulation. (E₁₋₄) Single-cell tracking of DE (green dots) and ExVE (yellow dots) reveals distinct cellular behaviors during gastrulation. (E₁₋₂) Still images from a maximum intensity projections at (E₁; T = 0 min) and (E₂, T = 445 min) taken from a time-lapse movie of cultured mid- to late-streak stage SCF embryos. (E₃) Overview of the migratory paths of VE (yellow arrows, top) and DE cells (green arrows, bottom). (E₄) Quantification of average displacement for each cell for ExVE ($0.68 \pm 0.22 \mu\text{m}$) or DE ($3.76 \pm 0.92 \mu\text{m}$) per time frame of 5 min. VE = visceral endoderm, DE = definitive endoderm, ExVE = extra-embryonic VE. Figure adopted from Burtscher et al. (2012)

layers. Especially, the quantification of fluorescence in 3D growing colonies challenges current 2D software like QTFy (Section 3.3) and requires new methods and algorithms to work with 3D time-resolved image stacks.

Chapter 9

Summary and outlook

In this work, we were interested in quantifying cellular fluorescence signals over a long time in individual cells. In the last years, single-cell time-lapse microscopy is an emerging field in biology still lacking of reliable and efficient quantification methods and tools. Incorporation of continuous monitoring of individual cells allows to investigate the development of protein dynamics to answer the biological questions asked in this thesis. In particular, we asked how PU.1 and Gata1 act on the single-cell level to establish myeloid lineage decision. Similarly we asked, how a heterogeneous expression profile can be established in pluripotent ESCs and how involved pluripotency factors orchestrate their roles. Furthermore, we were interested if we can measure protein half-life by fluorescence imaging on a single-cell level to reveal potential cell-to-cell differences.

9.1 Summary

To reliably quantify cellular signals in fluorescence images, we developed and applied a variety of image processing methods. We adopted basic cell segmentation methods, which allow an easy parametrization to facilitate parameter tuning in a later step. We proposed two methods to estimate underlying background signal in fluorescence images and a variety of normalization procedures based on these estimated background images (Section 2.2.4 to Section 2.2.7). For time-lapse series data we proposed a gain normalization method, which exploits the varying background levels to estimate a calibration function allowing to bring fluorescence signal to a comparable level throughout the whole image series. We evaluated our pipeline from several perspectives. First, we showed that fold-changes of fluorescent beads are comparable to flow cytometry measurements (Chapter 4). We showed in two different systems (HSCs as well as ESCs) that a potential higher dynamic range can be reached using our imaging technique due to the pixel specific background correction (Section 5.5 and Section 6.4). The sensitivity

was on a comparable scale between the two technologies. Lastly, we quantified *in vitro* differentiating blood stem cells in a time-lapse movie leading to a similar expression profile of PU.1eYFP and Gata1mCHERRY compared to freshly sorted bone marrow measurements by flow cytometry.

During this work, we developed several software tools incorporating cell segmentation methods as well as our fluorescence image normalization (Chapter 3). Here, we focused on cell segmentation methods, which allow an intuitive parametrization scheme facilitating the adjustment of segmentation results for the everyday user. These methods are incorporated in an efficient tool (Section 3.3) optimized to meet the requirements of modern bioimage software (Carpenter et al., 2012). Furthermore, the tool incorporates a variety of visualizations of the quantified data.

With this tool, we investigated a newly developed mouse model allowing to visualize the two key factors PU.1 and Gata1, involved in myeloid lineage differentiation (Chapter 5). Additionally, we developed a procedure to estimate protein abundances of reference cell distributions and were able to map protein abundances to fluorescence measurements. In general, our results are in accordance to literature and show high levels of PU.1eYFP and low levels of Gata1mCHERRY for GMPs and the other way around for MEPs. Interestingly, Gata1mCHERRY is not expressed in undifferentiated HSCs emphasizing the need to carefully revise current hematopoiesis differentiation models.

Using single-cell tracks we were able to *in silico* filter on GM differentiating cell branches only. This procedure discovered two main differentiation programs either showing a constant increase of initial PU.1 or a slight drop before high PU.1 levels are reached. In either case Gata1 was never involved in GMP differentiation. Again this finding is in clear conflict with current literature. In contrast to Kueh et al. (2013), we detected that GM differentiating cells do not elongate but reduce their cell-cycle time. Furthermore, these cells increase their PU.1 production to reach high PU.1 levels.

For MegE lineage committed cells, we found two distinct dynamics of the two transcription factors. In two out of three experiments we linked cells having a fast differentiation program to the megakaryocytic cell fate indicating that for these cells not only the lineage decision between GM and MegE was made but already between Meg and E. However, slow differentiating cells could not be linked to a mature cell fate due to the lack of additional markers.

In another application of our toolbox, we quantified transcription factors involved in pluripotency of ESCs (Chapter 6). Similarly to the quantification of HSCs, we were able to quantify fluorescence in individual cells for up to 8 generations although these cells are less motile and grow in colonies. We validated former reported heterogeneity of absolute transcription factor abundances. We found instances of emerging pedigrees staying negative in NanogVENUS over many generations. Still, this subset of cells was posi-

tive for other pluripotency factors defining them as true pluripotent ESCs. By measuring immunofluorescence and using partial correlation analysis, we found further differences between these negative colonies and colonies, which showed NanogVENUS activity after low sorting. In particular, we found that the relationship of Oct4 and Klf4 significantly changes from anti-correlation to positive correlation. By similar analysis we found indications that Rex1 is rather placed between Klf4 and Nanog instead of downstream of both transcription factors. Furthermore, by basic spatial analyses, we demonstrated that cellular areas are larger at the periphery of a colony, but the nuclear area shows no spatial correlations at all - implying that the nucleus is a very rigid structure.

Finally, we applied our quantification pipeline to investigate potential variability of protein half-life between individual cells (Chapter 7). First, we used classic Western Blot analysis to quantify protein half-life of Nanog on a population level as a reference. Next, we reconstructed a similar approach based on fluorescence imaging to validate our technology. By incorporating the full time resolved single-cell information we measured individual protein half-lives. Besides the already known heterogeneous protein distribution of Nanog we found heterogeneous distributed half-lives between individual cells. Interestingly, these two sources of heterogeneity are not correlated.

Since we validated that protein half-life measurements are possible using single-cell data, we were able to measure protein half-life in rare cell types. Thus, we measured this molecular parameter in primary blood stem cells for the transcription factors PU.1 and Gata1 and excluded changes in protein stability to be involved during differentiation.

Furthermore, our technology allowed to explore potential cell-cycle dependencies of protein stability, which turned out not to be the case. Lastly, we found a tendency of two related sister cells to behave more similar in terms of protein half-life.

9.2 Outlook

For the work presented in this thesis, several extension for the developed methods and future applications for the developed tools are possible investigating follow-up experiments:

Methodological extensions

- Our background estimation procedure lacks of an automatic quality measure and manual quality assessment is necessary. Since the ground truth of the underlying background signal is unknown for our experiments, this measure cannot be derived trivially. Since we already process each pixel of a set of given background images to estimate the gain function by a linear fit, a goodness of fit for every pixel assess

an overall quality for every position. Large deviations of the linear fit highlight probably miss-classified tiles containing cellular signal. Furthermore, an in depth evaluation and comparison to other methods is missing either based on synthetic data using SimuCell (Rajaram et al., 2012), SimCep (Lehmussola et al., 2007) or annotated image sets (Ljosa et al., 2012).

- To estimate the quality of the used trackings a set of trees has to be independently tracked by several persons. The average of the tracked data then allows to assess the tracking quality of every individual person as well as of automated tracking methods.
- As the currently used cell segmentation methods are very basic but still show a good tradeoff between usability and performance, we could extend these methods by more sophisticated approaches. The methods described in Buggenthin et al. (2013) provide a robust and fast automatic segmentation based on bright field images. However, a validation that segmentation of the whole cell shows a good overlay to segmentation of the cell nucleus is needed.
- Methodologically, it would be a great benefit to jointly link all the methods presented here as already several image platforms suggested by Endrov (Henriksson et al., 2013), Icy (de Chaumont et al., 2012), BioImageXD (Kankaanpää et al., 2012) or Omero (Allan et al., 2012) (for a review, see Eliceiri et al. (2012)). For example, the background estimation routine could improve its quality by incorporating tracking and segmentation information. The exact location of a tracked cell would already tell, which tiles contain cellular signal can thereby excluded automatically. In the same spirit, the supervised background estimation profits from features calculated of every wavelength at the same time point. By incorporating textural features calculated on bright field images combined with histogram based features based on underlying fluorescence images better separations could be reached.
- Finally, we want to extend our experience in segmentation and quantification to 3D image stacks. To perform fluorescence quantification on 3D fluorescence data, a variety of validation steps are necessary to benchmark fluorescence normalization and quantification.

Software extensions

- To widen the application of our tools and to reach a larger community of potential users, our implementation has to be ported into widely used frameworks, such as ImageJ/Fiji (Schindelin et al., 2012, Schneider et al., 2012) or CellProfiler (Carpenter et al., 2006, Kamensky

et al., 2011). Especially, the need for high-content screening increased in the recent years (Bray et al., 2012), asking for a efficient and reliable quantification tool, such as sQTFy. However, in the current version we still have to improve the usability to successfully allow large-scale screening tasks.

- Since we already successfully applied active learning (Section 3.2), the adaption of the concept to manually adjusted cell segmentations improves overall quality of automatic segmentation. Furthermore, it would be of great interest to incorporate time information even in the segmentation procedure.

Experimental follow-up studies for HSCs

- The average behavior of GM differentiating cells can be addressed with several different models. In our study we already ruled out that an elongation of the cell-cycle length is causing the increase of PU.1 abundance. Furthermore, two different models can explain the observed data equally well. Either an increase of protein production or an increase of protein stability during differentiation is possible. Since we measured protein half-life in these rare cell populations we rule out that the protein stability of PU.1 changes within these cells and conclude that the increase of protein production is the major molecular function. Experimental proofs are needed measuring the translation rate, transcription rate or mRNA numbers to verify this hypothesis.
- C/EBP α (Hasemann et al., 2012) and FOG-1 (Mancini et al., 2012) are known to play a role in myeloid differentiation serving as ideal next candidates to investigate their dynamics on a continuous single-cell level.
- Protein abundances are a crucial point of the interplay between PU.1 and Gata1 during myeloid lineage decision. In this work, we developed a method to infer protein numbers, but still a validation of these numbers using different technologies (e.g. fluorescence correlation spectroscopy (Heinrich et al., 2013)) would be preferable.

Experimental follow-up studies for ESCs

- For the ESC system, we suggest experiments, monitoring Nanog, Klf4 and/or Oct4 simultaneously in different wavelengths. Combined with endpoint staining of even more factors this can provide new insights how the involved factors synchronizes their expressions to maintain their pluripotent status. Furthermore, by incorporating more and more factors it becomes methodologically challenging how to unmix the resulting overlaying emission spectra (Zimmermann, 2005).

- We already investigated basic spatial dependencies in ESC colonies, but did not find relevant correlations of the measured features with the distance to the center. Already distinguishable structural measures on a colony base have been reported between different conditions (Scherf et al., 2012b) asking for a redo of our analysis of colonies cultured in different conditions and with different seeding densities. Furthermore, additional features like the cell movement or the angle of movement should be incorporated. Furthermore, co-location analysis should be applied as suggested by Shivanandan et al. (2013) to analyze potential cell-to-cell signaling.
- The investigation of protein stability of individual cells unraveled unexpected cell-to-cell variability. In future experiments monitoring more than just one labeled protein can clarify if this variability stems from protein specificity or is generally conserved within a cell originating from the protein degradation machinery.
- Furthermore, experiments with ESCs monitoring Nanog half-life on a single-cell level in different conditions (e.g. 2i) would be of great interest, since already differences in half-life on the population average could be found (Muñoz Descalzo et al., 2013). A condition forcing Nanog to be more homogeneously expressed allows investigating the response of the heterogeneous half-life between individual cells.

One crucial experiment we are lacking for all application of QTFy is the validation to be in the linear scale of intensity. In the scope of this work we could not confirm that the relationship between intensity and protein abundance is really linear for our experimental setup.

9.3 Conclusion

In conclusion, we developed methods to normalize fluorescence images and presented a tool to reliably quantify cellular fluorescence in long-term time-lapse experiments. In this work, we have demonstrated that the application of our tool answered important biological questions of the molecular details during myeloid differentiation and clarified contradictory opinions of heterogeneity in ESCs. By publishing the methods and the software we contribute to a growing community and provide the instruments to answer long-standing biological questions, which can be addressed by continuous single-cell quantification.

Bibliography

- Abkowitz, J. L., Catlin, S. N., McCallie, M. T., and Gutter, P. (2002). Evidence that the number of hematopoietic stem cells per animal is conserved in mammals. *Blood*, 100(7):2665–7.
- Abranches, E., Bekman, E., and Henrique, D. (2013). Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression. *PloS one*, 8(3):e59928.
- Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L. A. M., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M., and Jacobsen, S. E. W. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell*, 121(2):295–306.
- Akashi, K., Traver, D., Miyamoto, T., and Weissman, I. L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature*, 404(6774):193–7.
- Al-Kofahi, Y., Lassoued, W., Lee, W., and Roysam, B. (2010). Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE transactions on bio-medical engineering*, 57(4):841–52.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). *Molecular Biology of the Cell, Fourth Edition*. Garland Science.
- Allan, C., Burel, J.-M., Moore, J., Blackburn, C., Linkert, M., Loynton, S., Macdonald, D., Moore, W. J., Neves, C., Patterson, A., Porter, M., Tarkowska, A., Loranger, B., Avondo, J., Lagerstedt, I., Lianas, L., Leo, S., Hands, K., Hay, R. T., Patwardhan, A., Best, C., Kleywegt, G. J., Zanetti, G., and Swedlow, J. R. (2012). OMERO: flexible, model-driven data management for experimental biology. *Nature methods*, 9(3):245–53.
- Arteta, C., Lempitsky, V., Noble, J. A., and Zisserman, A. (2012). Learning to detect cells using non-overlapping extremal regions. *Medical image computing and computer-assisted intervention : MICCAI ... International*

Conference on Medical Image Computing and Computer-Assisted Intervention, 15(Pt 1):348–56.

- Babaloukas, G., Tentolouris, N., Liatis, S., Sklavounou, A., and Perrea, D. (2011). Evaluation of three methods for retrospective correction of vignetting on medical microscopy images utilizing two open source software tools. *Journal of microscopy*, 244(3):320–4.
- Becker, A. J., McCulloch, E. A., and Till, J. E. (1963). Cytological Demonstration of the Clonal Nature of Spleen Colonies Derived from Transplanted Mouse Marrow Cells. *Nature*, 197(4866):452–454.
- Bendall, S. C., Simonds, E. F., Qiu, P., Amir, E.-a. D., Krutzik, P. O., Finck, R., Bruggner, R. V., Melamed, R., Trejo, A., Ornatsky, O. I., Balderas, R. S., Plevritis, S. K., Sachs, K., Pe'er, D., Tanner, S. D., and Nolan, G. P. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science (New York, N.Y.)*, 332(6030):687–96.
- Bergeest, J.-P. and Rohr, K. (2012). Efficient globally optimal segmentation of cells in fluorescence microscopy images using level sets and convex energy functionals. *Medical image analysis*, 16(7):1436–44.
- Bishop, C. M. (2007). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer.
- Bjornsson, C. S., Lin, G., Al-Kofahi, Y., Narayanaswamy, A., Smith, K. L., Shain, W., and Roysam, B. (2008). Associative image analysis: a method for automated quantification of 3D multi-parameter images of brain tissue. *Journal of neuroscience methods*, 170(1):165–78.
- Borisov, A., Tuv, E., and Runger, G. (2011). Active Batch Learning with Stochastic Query-by-Forest (SQBF). In *JMLR: Workshop and Conference Proceedings 16*, volume 16, pages 59–69.
- Botev, Z. I., Grotowski, J. F., and Kroese, D. P. (2010). Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5):2916–2957.
- Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., and Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, 122(6):947–56.
- Bray, M.-A., Fraser, A. N., Hasaka, T. P., and Carpenter, A. E. (2012). Workflow and metrics for image quality control in large-scale high-content screens. *Journal of biomolecular screening*, 17(2):266–74.

- Breiman, L. (1996). Bagging Predictors. *Machine Learning*, 24(2):123–140.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. (1984). *Classification and Regression Trees*. Chapman and Hall/CRC.
- Bryja, V., Bonilla, S., and Arenas, E. (2006). Derivation of mouse embryonic stem cells. *Nature protocols*, 1(4):2082–7.
- Buggenthin, F. (2011). *Computational prediction of hematopoietic cell fates using single cell time lapse imaging*. Diploma thesis, Technische Universität München, Ludwig-Maximilians-Universität München.
- Buggenthin, F., Marr, C., Schwarzfischer, M., Hoppe, P. S., Hilsenbeck, O., Schroeder, T., and Theis, F. J. (2013). An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC bioinformatics*, 14(1):297.
- Burtscher, I., Barkey, W., Schwarzfischer, M., Theis, F. J., and Lickert, H. (2012). The Sox17-mCherry fusion mouse line allows visualization of endoderm and vascular endothelial development. *Genesis (New York, N.Y. : 2000)*, 50(6):496–505.
- Burtscher, I. and Lickert, H. (2009). Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development (Cambridge, England)*, 136(6):1029–38.
- Campbell, N. A. and Reece, J. B. (2001). *Biology, 6th Edition*. Benjamin Cummings.
- Cantor, A. B. and Orkin, S. H. (2001). Hematopoietic development: a balancing act. *Current opinion in genetics & development*, 11(5):513–9.
- Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., Guertin, D. A., Chang, J. H., Lindquist, R. A., Moffat, J., Golland, P., and Sabatini, D. M. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome biology*, 7(10):R100.
- Carpenter, A. E., Kametsky, L., and Eliceiri, K. W. (2012). A call for bioimaging software usability. *Nature methods*, 9(7):666–70.
- Chalut, K. J., Höpfner, M., Lautenschläger, F., Boyde, L., Chan, C. J., Ekpenyong, A., Martinez-Arias, A., and Guck, J. (2012). Chromatin decondensation and nuclear softening accompany Nanog downregulation in embryonic stem cells. *Biophysical journal*, 103(10):2060–70.

- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*, 113(5):643–55.
- Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature*, 450(7173):1230–4.
- Chambers, I. and Tomlinson, S. R. (2009). The transcriptional foundation of pluripotency. *Development (Cambridge, England)*, 136(14):2311–22.
- Chan, K. K.-K., Zhang, J., Chia, N.-Y., Chan, Y.-S., Sim, H. S., Tan, K. S., Oh, S. K.-W., Ng, H.-H., and Choo, A. B.-H. (2009). KLF4 and PBX1 directly regulate NANOG expression in human embryonic stem cells. *Stem cells (Dayton, Ohio)*, 27(9):2114–25.
- Chen, H., Ray-Gallet, D., Zhang, P., Hetherington, C. J., Gonzalez, D. A., Zhang, D. E., Moreau-Gachelin, F., and Tenen, D. G. (1995). PU.1 (Spi-1) autoregulates its expression in myeloid cells. *Oncogene*, 11(8):1549–60.
- Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J., Loh, Y.-H., Yeo, H. C., Yeo, Z. X., Narang, V., Govindarajan, K. R., Leong, B., Shahab, A., Ruan, Y., Bourque, G., Sung, W.-K., Clarke, N. D., Wei, C.-L., and Ng, H.-H. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*, 133(6):1106–17.
- Chickarmane, V., Enver, T., and Peterson, C. (2009). Computational modeling of the hematopoietic erythroid-myeloid switch reveals insights into cooperativity, priming, and irreversibility. *PLoS computational biology*, 5(1):e1000268.
- Cohen, A. A., Geva-Zatorsky, N., Eden, E., Frenkel-Morgenstern, M., Issaeva, I., Sigal, A., Milo, R., Cohen-Saidon, C., Liron, Y., Kam, Z., Cohen, L., Danon, T., Perzov, N., and Alon, U. (2008). Dynamic proteomics of individual cancer cells in response to a drug. *Science (New York, N.Y.)*, 322(5907):1511–6.
- Cohen, A. R., Gomes, F. L. A. F., Roysam, B., and Cayouette, M. (2010). Computational prediction of neural progenitor cell fates. *Nature methods*, 7(3):213–8.
- Corish, P. and Tyler-Smith, C. (1999). Attenuation of green fluorescent protein half-life in mammalian cells. *Protein engineering*, 12(12):1035–40.

- Coutu, D. L. and Schroeder, T. (2013). Probing cellular processes by long-term live imaging - historic problems and current solutions. *Journal of cell science*, 126(August):3805–3815.
- Cross, M. A. and Enver, T. (1997). The lineage commitment of haemopoietic progenitor cells. *Current opinion in genetics & development*, 7(5):609–13.
- de Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., Meas-Yedid, V., Pankajakshan, P., Lecomte, T., Le Montagner, Y., Lagache, T., Dufour, A., and Olivo-Marin, J.-C. (2012). Icy: an open bioimage informatics platform for extended reproducible research. *Nature methods*, 9(7):690–6.
- Delassus, S., Titley, I., and Enver, T. (1999). Functional and molecular analysis of hematopoietic progenitors derived from the aorta-gonad-mesonephros region of the mouse embryo. *Blood*, 94(5):1495–503.
- Dietrich, J.-E. and Hiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. *Development (Cambridge, England)*, 134(23):4219–31.
- Dima, A. a., Elliott, J. T., Filliben, J. J., Halter, M., Peskin, A., Bernal, J., Kocielek, M., Brady, M. C., Tang, H. C., and Plant, A. L. (2011). Comparison of segmentation algorithms for fluorescence microscopy images of cells. *Cytometry. Part A : the journal of the International Society for Analytical Cytology*, 79(7):545–59.
- Dunn, G. A. and Jones, G. E. (2004). Cell motility under the microscope: Vorsprung durch Technik. *Nature reviews. Molecular cell biology*, 5(8):667–72.
- Eden, E., Geva-Zatorsky, N., Issaeva, I., Cohen, A., Dekel, E., Danon, T., Cohen, L., Mayo, A., and Alon, U. (2011). Proteome half-life dynamics in living human cells. *Science (New York, N.Y.)*, 331(6018):764–8.
- Eilken, H., Rieger, M., Hoppe, P., Hermann, A., Smejkal, B., Drew, E., Thum, M., Ninkovic, J., Beckervordersandforth, R., and Schroeder, T. (2011). Continuous long-term detection of live cell surface markers by in culture antibody staining. *Protocol Exchange*.
- Eilken, H. M., Nishikawa, S.-I., and Schroeder, T. (2009). Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature*, 457(7231):896–900.
- 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.,

- Stuurman, N., Stuurmann, N., Swedlow, J. R., Tomancak, P., and Carpenter, A. E. (2012). Biological imaging software tools. *Nature methods*, 9(7):697–710.
- Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science (New York, N.Y.)*, 297(5584):1183–6.
- Enver, T., Heyworth, C. M., and Dexter, T. M. (1998). Do stem cells play dice? *Blood*, 92(2):348–51; discussion 352.
- Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*, pages 226–231.
- Evans, T., Reitman, M., and Felsenfeld, G. (1988). An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proceedings of the National Academy of Sciences of the United States of America*, 85(16):5976–80.
- Faddah, D. a., Wang, H., Cheng, A. W., Katz, Y., Buganim, Y., and Jaenisch, R. (2013). Single-Cell Analysis Reveals that Expression of Nanog Is Biallelic and Equally Variable as that of Other Pluripotency Factors in Mouse ESCs. *Cell stem cell*, 13(1):23–9.
- Fidalgo, M., Faiola, F., Pereira, C.-F., Ding, J., Saunders, A., Gingold, J., Schaniel, C., Lemischka, I. R., Silva, J. C. R., and Wang, J. (2012). Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. *Proceedings of the National Academy of Sciences of the United States of America*, 109(40):16202–7.
- Filipczyk, A., Gkatzis, K., Fu, J., Hoppe, P. S., Lickert, H., Anastassiadis, K., and Schroeder, T. (2013). Biallelic expression of nanog protein in mouse embryonic stem cells. *Cell stem cell*, 13(1):12–3.
- Fischer, S. (2013). *Quantification of marker expression for endoderm and mesoderm segregation*. Bachelor thesis, Technische Universität München, Ludwig-Maximilians-Universität München.
- Foster, S. D., Oram, S. H., Wilson, N. K., and Göttgens, B. (2009). From genes to cells to tissues—modelling the haematopoietic system. *Molecular bioSystems*, 5(12):1413–20.
- Frischknecht, F., Gunzer, M., and Shorte, S. L. (2009). Retrospective: Birth of the cool - imaging and microbiology from Ibn al-Haytham to Jean Comandon. *Biotechnology journal*, 4(6):787–90.

- Gabor, D. (1946). Theory of communication. Part 1: The analysis of information. *Electrical Engineers - Part III: Radio and Communication Engineering, Journal of the Institution of*, 93(26):429–441.
- Garcia-Ojalvo, J. and Martinez Arias, A. (2012). Towards a statistical mechanics of cell fate decisions. *Current opinion in genetics & development*, 22(6):619–26.
- Gassmann, M., Grenacher, B., Rohde, B., and Vogel, J. (2009). Quantifying Western blots: pitfalls of densitometry. *Electrophoresis*, 30(11):1845–55.
- Gini, C. (1921). Measurement of Inequality of Incomes. *The Economic Journal*, 31:124 – 126.
- Glauche, I., Herberg, M., and Roeder, I. (2010). Nanog variability and pluripotency regulation of embryonic stem cells—insights from a mathematical model analysis. *PLoS one*, 5(6):e11238.
- Goldberg, A. D., Allis, C. D., and Bernstein, E. (2007). Epigenetics: a landscape takes shape. *Cell*, 128(4):635–8.
- Gordon, A., Colman-Lerner, A., Chin, T. E., Benjamin, K. R., Yu, R. C., and Brent, R. (2007). Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nature methods*, 4(2):175–81.
- Graf, T. (2002). Differentiation plasticity of hematopoietic cells. *Blood*, 99(9):3089–3101.
- Graf, T. and Enver, T. (2009). Forcing cells to change lineages. *Nature*, 462(7273):587–94.
- Grimmett, G. R. and Stirzaker, D. R. (2001). *Probability and Random Processes*. Oxford University Press, USA.
- Halter, M., Tona, A., Bhadriraju, K., Plant, A. L., and Elliott, J. T. (2007). Automated live cell imaging of green fluorescent protein degradation in individual fibroblasts. *Cytometry. Part A : the journal of the International Society for Analytical Cytology*, 71(10):827–34.
- Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6):610–621.
- Harper, C. V., Finkenstädt, B., Woodcock, D. J., Friedrichsen, S., Semprini, S., Ashall, L., Spiller, D. G., Mullins, J. J., Rand, D. A., Davis, J. R. E., and White, M. R. H. (2011). Dynamic analysis of stochastic transcription cycles. *PLoS Biology*, 9(4):e1000607.

- Hasemann, M. S., Schuster, M. B., Frank, A.-K., Theilgaard-Mönch, K., Pedersen, T. A., Nerlov, C., and Porse, B. T. (2012). Phosphorylation of serine 248 of C/EBP α is dispensable for myelopoiesis but its disruption leads to a low penetrant myeloid disorder with long latency. *PloS one*, 7(6):e38841.
- Heinäniemi, M., Nykter, M., Kramer, R., Wienecke-Baldacchino, A., Sinkkonen, L., Zhou, J. X., Kreisberg, R., Kauffman, S. A., Huang, S., and Shmulevich, I. (2013). Gene-pair expression signatures reveal lineage control. *Nature methods*, 10(6):577–83.
- Heinrich, S., Geissen, E.-M., Kamenz, J., Trautmann, S., Widmer, C., Drewe, P., Knop, M., Radde, N., Hasenauer, J., and Hauf, S. (2013). Determinants of robustness in spindle assembly checkpoint signalling. *Nature cell biology*, 15(11):1328–39.
- Henriksson, J., Hensch, J., Tong, Y. G., Johansson, A., Johansson, D., and Bürglin, T. R. (2013). Endrov: an integrated platform for image analysis. *Nature methods*, 10(6):454–6.
- Hernandez-Hernandez, A., Ray, P., Litos, G., Ciro, M., Ottolenghi, S., Beug, H., and Boyes, J. (2006). Acetylation and MAPK phosphorylation cooperate to regulate the degradation of active GATA-1. *The EMBO journal*, 25(14):3264–74.
- Herzenberg, L. A., Tung, J., Moore, W. A., Herzenberg, L. A., and Parks, D. R. (2006). Interpreting flow cytometry data: a guide for the perplexed. *Nature immunology*, 7(7):681–5.
- Hilsenbeck, O. (2011). *Automated construction of cell lineage trees from time-lapse microscopy data*. Bachelor thesis, Technische Universität München, Ludwig-Maximilians-Universität München.
- Hooke, R. (1667). *Micrographia: or, some physiological descriptions of minute bodies made by magnifying glasses. With observations and inquiries thereupon*. James Allestry.
- Hotelling, H. (1931). The Generalization of Student's Ratio. *The Annals of Mathematical Statistics*, 2(3):360–378.
- Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., and Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. *Genes & development*, 11(6):774–85.
- Huang, S., Guo, Y.-P., May, G., and Enver, T. (2007). Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental biology*, 305(2):695–713.

- Huth, J., Buchholz, M., Kraus, J. M., Molhave, K., Gradinaru, C., v Wichert, G., Gress, T. M., Neumann, H., and Kestler, H. A. (2011). TimeLapseAnalyzer: multi-target analysis for live-cell imaging and time-lapse microscopy. *Computer methods and programs in biomedicine*, 104(2):227–34.
- Huth, J., Buchholz, M., Kraus, J. M., Schmucker, M., von Wichert, G., Krndija, D., Seufferlein, T., Gress, T. M., and Kestler, H. A. (2010). Significantly improved precision of cell migration analysis in time-lapse video microscopy through use of a fully automated tracking system. *BMC cell biology*, 11:24.
- Iwasaki, H., Mizuno, S.-i., Mayfield, R., Shigematsu, H., Arinobu, Y., Seed, B., Gurish, M. F., Takatsu, K., and Akashi, K. (2005a). Identification of eosinophil lineage-committed progenitors in the murine bone marrow. *The Journal of experimental medicine*, 201(12):1891–7.
- Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E. a., Iwasaki-Arai, J., Mizuno, S.-I., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T., Fenyus, M. L., Elf, S., Chan, S., Kastner, P., Huettner, C. S., Murray, R., Tenen, D. G., and Akashi, K. (2005b). Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood*, 106(5):1590–600.
- Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S. L., and Danuser, G. (2008). Robust single-particle tracking in live-cell time-lapse sequences. *Nature methods*, 5(8):695–702.
- Jones, M. (2009). Stem cells diagram.
- Kalmar, T., Lim, C., Hayward, P., Muñoz Descalzo, S., Nichols, J., Garcia-Ojalvo, J., and Martinez Arias, A. (2009). Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS biology*, 7(7):e1000149.
- Kamentsky, L., Jones, T. R., Fraser, A., Bray, M.-A., Logan, D. J., Madden, K. L., Ljosa, V., Rueden, C., Eliceiri, K. W., and Carpenter, A. E. (2011). Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics (Oxford, England)*, 27(8):1179–80.
- Kanai-Azuma, M., Kanai, Y., Gad, J. M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P. P. L., and Hayashi, Y. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development (Cambridge, England)*, 129(10):2367–79.

- Kankaanpää, P., Paavolainen, L., Tiitta, S., Karjalainen, M., Päivärinne, J., Nieminen, J., Marjomäki, V., Heino, J., and White, D. J. (2012). BioImageXD: an open, general-purpose and high-throughput image-processing platform. *Nature methods*, 9(7):683–9.
- Karwacki-Neisius, V., Göke, J., Osorno, R., Halbritter, F., Ng, J. H., Weiß e, A. Y., Wong, F. C. K., Gagliardi, A., Mullin, N. P., Festuccia, N., Colby, D., Tomlinson, S. R., Ng, H.-H., and Chambers, I. (2013). Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. *Cell stem cell*, 12(5):531–45.
- Kaushansky, K., Lichtman, M., Beutler, E., Kipps, T., Prchal, J., and Seligsohn, U. (2010). *Williams Hematology, Eighth Edition*. McGraw-Hill Professional.
- Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S. H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell*, 132(6):1049–61.
- Kirstetter, P., Anderson, K., Porse, B. T., Jacobsen, S. E. W., and Nerlov, C. (2006). Activation of the canonical Wnt pathway leads to loss of hematopoietic stem cell repopulation and multilineage differentiation block. *Nature immunology*, 7(10):1048–56.
- Klein, J., Leupold, S., Biegler, I., Biedendieck, R., Münch, R., and Jahn, D. (2012). TLM-Tracker: software for cell segmentation, tracking and lineage analysis in time-lapse microscopy movies. *Bioinformatics (Oxford, England)*, 28(17):2276–7.
- Klein, S., Pluim, J. P. W., Staring, M., and Viergever, M. A. (2008). Adaptive Stochastic Gradient Descent Optimisation for Image Registration. *International Journal of Computer Vision*, 81(3):227–239.
- Komorowski, M., Finkenstädt, B., and Rand, D. (2010). Using a single fluorescent reporter gene to infer half-life of extrinsic noise and other parameters of gene expression. *Biophysical journal*, 98(12):2759–69.
- Kreutz, C., Bartolome Rodriguez, M. M., Maiwald, T., Seidl, M., Blum, H. E., Mohr, L., and Timmer, J. (2007). An error model for protein quantification. *Bioinformatics (Oxford, England)*, 23(20):2747–53.
- Krumsiek, J., Marr, C., Schroeder, T., and Theis, F. J. (2011). Hierarchical differentiation of myeloid progenitors is encoded in the transcription factor network. *PloS one*, 6(8):e22649.

- Kueh, H. Y., Champhekar, A., Nutt, S. L., Elowitz, M. B., and Rothenberg, E. V. (2013). Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science (New York, N.Y.)*, 341(6146):670–3.
- Kulesa, H., Frampton, J., and Graf, T. (1995). GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboplasts, and erythroblasts. *Genes & development*, 9(10):1250–62.
- Landgraf, D., Okumus, B., Chien, P., Baker, T. A., and Paulsson, J. (2012). Segregation of molecules at cell division reveals native protein localization. *Nature methods*, 9(5):480–2.
- Laslo, P., Spooner, C. J., Warmflash, A., Lancki, D. W., Lee, H.-J., Sciammas, R., Gantner, B. N., Dinner, A. R., and Singh, H. (2006). Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell*, 126(4):755–66.
- Lauritzen, S. L. (1996). *Graphical Models* (Google eBook). Oxford University Press.
- Lehmussola, A., Ruusuvoori, P., Selinummi, J., Huttunen, H., and Yli-Harja, O. (2007). Computational framework for simulating fluorescence microscope images with cell populations. *IEEE transactions on medical imaging*, 26(7):1010–6.
- Levy, K. J. and Narula, S. C. (1978). Testing Hypotheses concerning Partial Correlations: Some Methods and Discussion. *International Statistical Review / Revue Internationale de Statistique*, Vol. 46, No. 2, 46(2):215–218.
- Li, Y., Okuno, Y., Zhang, P., Radomska, H. S., Chen, H., Iwasaki, H., Akashi, K., Klemsz, M. J., McKercher, S. R., Maki, R. A., and Tenen, D. G. (2001). Regulation of the PU.1 gene by distal elements. *Blood*, 98(10):2958–65.
- Lim, J. S. (1990). *Two-dimensional signal and image processing*. Prentice Hall.
- Lindblad, J. and Bengtsson, E. (2001). A comparison of methods for estimation of intensity nonuniformities in 2D and 3D microscope images of fluorescence stained cells. In *12th Scandinavian Conference on Image Analysis (SCIA)*, number 1, pages 264–271.
- Ljosa, V., Sokolnicki, K. L., and Carpenter, A. E. (2012). Annotated high-throughput microscopy image sets for validation. *Nature methods*, 9(7):637.

- Loh, Y.-H., Ng, J.-H., and Ng, H.-H. (2008). Molecular framework underlying pluripotency. *Cell cycle (Georgetown, Tex.)*, 7(7):885–91.
- Lu, R., Markowitz, F., Unwin, R. D., Leek, J. T., Airoidi, E. M., MacArthur, B. D., Lachmann, A., Rozov, R., Ma'ayan, A., Boyer, L. A., Troyanskaya, O. G., Whetton, A. D., and Lemischka, I. R. (2009). Systems-level dynamic analyses of fate change in murine embryonic stem cells. *Nature*, 462(7271):358–62.
- Luo, Y., Lim, C. L., Nichols, J., Martinez-Arias, A., and Wernisch, L. (2012). Cell signalling regulates dynamics of Nanog distribution in embryonic stem cell populations. *Journal of the Royal Society, Interface / the Royal Society*.
- MacArthur, B. D. and Lemischka, I. R. (2013). Statistical mechanics of pluripotency. *Cell*, 154(3):484–9.
- MacArthur, B. D., Ma'ayan, A., and Lemischka, I. R. (2009). Systems biology of stem cell fate and cellular reprogramming. *Nature reviews. Molecular cell biology*, 10(10):672–81.
- MacArthur, B. D., Sevilla, A., Lenz, M., Müller, F.-J., Schuld, B. M., Schuppert, A. A., Ridden, S. J., Stumpf, P. S., Fidalgo, M., Ma'ayan, A., Wang, J., and Lemischka, I. R. (2012). Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity. *Nature cell biology*, 14(11):1139–47.
- Mancini, E., Sanjuan-Pla, A., Luciani, L., Moore, S., Grover, A., Zay, A., Rasmussen, K. D., Luc, S., Bilbao, D., O'Carroll, D., Jacobsen, S. E., and Nerlov, C. (2012). FOG-1 and GATA-1 act sequentially to specify definitive megakaryocytic and erythroid progenitors. *The EMBO journal*, 31(2):351–65.
- Marr, C., Strasser, M., Schwarzfischer, M., Schroeder, T., and Theis, F. J. (2012). Multi-scale modeling of GMP differentiation based on single-cell genealogies. *The FEBS journal*, 279(18):3488–500.
- Matas, J., Chum, O., Urban, M., and Pajdla, T. (2002). Robust Wide Baseline Stereo from Maximally Stable Extremal Regions. In *Proceedings of the British Machine Vision Conference 2002*, pages 36.1–36.10. British Machine Vision Association.
- Maurer, C. and Raghavan, V. (2003). A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):265–270.

- McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, 21(No. 2):239–245.
- McKercher, S. R., Torbett, B. E., Anderson, K. L., Henkel, G. W., Vestal, D. J., Baribault, H., Klemsz, M., Feeney, A. J., Wu, G. E., Paige, C. J., and Maki, R. A. (1996). Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *The EMBO journal*, 15(20):5647–58.
- Meijering, E., Dzyubachyk, O., Smal, I., and van Cappellen, W. A. (2009). Tracking in cell and developmental biology. *Seminars in cell & developmental biology*, 20(8):894–902.
- Meyer, F. (1994). Topographic distance and watershed lines. *Signal Processing*, 38(1):113–125.
- Mikkola, H. K. A. and Orkin, S. H. (2006). The journey of developing hematopoietic stem cells. *Development (Cambridge, England)*, 133(19):3733–44.
- Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, 113(5):631–42.
- Miyamoto, T., Iwasaki, H., Reizis, B., Ye, M., Graf, T., Weissman, I. L., and Akashi, K. (2002). Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Developmental cell*, 3(1):137–47.
- Miyanari, Y. and Torres-Padilla, M.-E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. *Nature*, 483(7390):470–3.
- Monaghan, A. P., Kaestner, K. H., Grau, E., and Schütz, G. (1993). Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. *Development (Cambridge, England)*, 119(3):567–78.
- Muñoz Descalzo, S., Rué, P., Faunes, F., Hayward, P., Jakt, L. M., Balayo, T., Garcia-Ojalvo, J., and Martinez Arias, A. (2013). A competitive protein interaction network buffers Oct4-mediated differentiation to promote pluripotency in embryonic stem cells. *Molecular systems biology*, 9(1):694.
- Muñoz Descalzo, S., Rué, P., Garcia-Ojalvo, J., and Martinez Arias, A. (2012). Correlations between the levels of Oct4 and Nanog as a signature for naïve pluripotency in mouse embryonic stem cells. *Stem cells (Dayton, Ohio)*, 30(12):2683–91.

- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series)*. The MIT Press.
- Muzzey, D. and van Oudenaarden, A. (2009). Quantitative time-lapse fluorescence microscopy in single cells. *Annual review of cell and developmental biology*, 25:301–27.
- Myers, G. (2012). Why bioimage informatics matters. *Nature methods*, 9(7):659–60.
- Navarro, P., Festuccia, N., Colby, D., Gagliardi, A., Mullin, N. P., Zhang, W., Karwacki-Neisius, V., Osorno, R., Kelly, D., Robertson, M., and Chambers, I. (2012). OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. *The EMBO journal*, 31(24):4547–62.
- Nerlov, C. and Graf, T. (1998). PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. *Genes & development*, 12(15):2403–12.
- Nerlov, C., Querfurth, E., Kulesa, H., and Graf, T. (2000). GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood*, 95(8):2543–51.
- Neumann, B., Walter, T., Hériché, J.-K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U., Cetin, C., Sieckmann, F., Pau, G., Kabbe, R., Wünsche, A., Satagopam, V., Schmitz, M. H. A., Chapuis, C., Gerlich, D. W., Schneider, R., Eils, R., Huber, W., Peters, J.-M., Hyman, A. A., Durbin, R., Pepperkok, R., and Ellenberg, J. (2010). Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature*, 464(7289):721–7.
- Nistér, D. and Stewénus, H. (2008). Linear time maximally stable extremal regions. *Computer Vision-ECCV 2008*, pages 183–196.
- Niwa, H., Miyazaki, J., and Smith, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature genetics*, 24(4):372–6.
- Niwa, H., Ogawa, K., Shimosato, D., and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature*, 460(7251):118–22.
- Nixon, M. and Aguado, A. S. (2012). *Feature Extraction & Image Processing for Computer Vision, Third Edition*. Academic Press.
- Nutt, S. L., Metcalf, D., D’Amico, A., Polli, M., and Wu, L. (2005). Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. *The Journal of experimental medicine*, 201(2):221–31.

- Okita, C., Sato, M., and Schroeder, T. (2004). Generation of optimized yellow and red fluorescent proteins with distinct subcellular localization. *BioTechniques*, 36(3):418–22, 424.
- Okuno, Y., Huang, G., Rosenbauer, F., Evans, E. K., Radomska, H. S., Iwasaki, H., Akashi, K., Moreau-Gachelin, F., Li, Y., Zhang, P., Göttgens, B., and Tenen, D. G. (2005). Potential autoregulation of transcription factor PU.1 by an upstream regulatory element. *Molecular and cellular biology*, 25(7):2832–45.
- Orkin, S. H. (2000). Diversification of haematopoietic stem cells to specific lineages. *Nature reviews. Genetics*, 1(1):57–64.
- Orkin, S. H. and Zon, L. I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. *Cell*, 132(4):631–44.
- Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.
- Pan, G., Li, J., Zhou, Y., Zheng, H., and Pei, D. (2006). A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 20(10):1730–2.
- Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M. M., and Choudhary, J. (2010). An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell stem cell*, 6(4):382–95.
- Peng, H., Bateman, A., Valencia, A., and Wren, J. D. (2012). Bioimage informatics: a new category in Bioinformatics. *Bioinformatics (Oxford, England)*, 28(8):1057.
- Petrou, M. and Petrou, C. (2010). *Image Processing: The Fundamentals*. Wiley.
- Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D’Agati, V., Orkin, S. H., and Costantini, F. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*, 349(6306):257–60.
- Piersma, S., Denham, E. L., Drulhe, S., Tonk, R. H. J., Schwikowski, B., and van Dijl, J. M. (2013). TLM-Quant: An Open-Source Pipeline for Visualization and Quantification of Gene Expression Heterogeneity in Growing Microbial Cells. *PloS one*, 8(7):e68696.

- Pronk, C. J. H., Rossi, D. J., Mansson, R., Attema, J. L., Norddahl, G. L., Chan, C. K. F., Sigvardsson, M., Weissman, I. L., and Bryder, D. (2007). Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell stem cell*, 1(4):428–42.
- Punga, S. (2013). *Supervised lipid raft cluster identification in hematopoietic stem cells*. Bachelor thesis, Technische Universität München, Ludwig-Maximilians-Universität München.
- Qiu, P., Simonds, E. F., Bendall, S. C., Gibbs, K. D., Bruggner, R. V., Linderman, M. D., Sachs, K., Nolan, G. P., and Plevritis, S. K. (2011). Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nature biotechnology*, 29(10):886–91.
- Rajaram, S., Pavie, B., Hac, N. E. F., Altschuler, S. J., and Wu, L. F. (2012). SimuCell: a flexible framework for creating synthetic microscopy images. *Nature methods*, 9(7):634–5.
- Ramakrishna, S., Suresh, B., Lim, K.-H., Cha, B.-H., Lee, S.-H., Kim, K.-S., and Baek, K.-H. (2011). PEST motif sequence regulating human NANOG for proteasomal degradation. *Stem cells and development*, 20(9):1511–9.
- Rieger, M. A., Hoppe, P. S., Smejkal, B. M., Eitelhuber, A. C., and Schroeder, T. (2009). Hematopoietic cytokines can instruct lineage choice. *Science (New York, N.Y.)*, 325(5937):217–8.
- Rieger, M. A. and Schroeder, T. (2007). Hämatopoetische Stammzellen. *BIOspektrum-Heidelberg*, 13(3):254.
- Rizzino, A. (2009). Sox2 and Oct-3/4: a versatile pair of master regulators that orchestrate the self-renewal and pluripotency of embryonic stem cells. *Wiley interdisciplinary reviews. Systems biology and medicine*, 1(2):228–36.
- Robertson, J. A. (2010). Embryo stem cell research: ten years of controversy. *The Journal of law, medicine & ethics : a journal of the American Society of Law, Medicine & Ethics*, 38(2):191–203.
- Roeder, I. and Glauche, I. (2006). Towards an understanding of lineage specification in hematopoietic stem cells: a mathematical model for the interaction of transcription factors GATA-1 and PU.1. *Journal of theoretical biology*, 241(4):852–65.
- Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S., and Elowitz, M. B. (2005). Gene regulation at the single-cell level. *Science (New York, N.Y.)*, 307(5717):1962–5.

- Rossi, L., Lin, K., Boles, N., Yang, L., King, K., Jeong, M., Mayle, A., and Goodell, M. (2012). Less Is More: Unveiling the Functional Core of Hematopoietic Stem Cells through Knockout Mice. *Cell Stem Cell*, 11(3):302–317.
- Scherf, N., Franke, K., Glauche, I., Kurth, I., Bornhäuser, M., Werner, C., Pompe, T., and Roeder, I. (2012a). On the symmetry of siblings: automated single-cell tracking to quantify the behavior of hematopoietic stem cells in a biomimetic setup. *Experimental hematology*, 40(2):119–30.e9.
- Scherf, N., Herberg, M., Thierbach, K., Zerjatke, T., Kalkan, T., Humphreys, P., Smith, A., Glauche, I., and Roeder, I. (2012b). Imaging, quantification and visualization of spatio-temporal patterning in mESC colonies under different culture conditions. *Bioinformatics (Oxford, England)*, 28(18):i556–i561.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7):676–82.
- Schneider, C. a., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7):671–675.
- Schroeder, T. (2008). Imaging stem-cell-driven regeneration in mammals. *Nature*, 453(7193):345–51.
- Schroeder, T. (2011). Long-term single-cell imaging of mammalian stem cells. *Nature methods*, 8(4 Suppl):S30–5.
- Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature*, 473(7347):337–42.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464.
- Schwarzfischer, M. (2009). *Single-cell analysis of multipotent hematopoietic progenitor cells*. Diploma thesis, Technische Universität München, Ludwig-Maximilians-Universität München.
- Schwarzfischer, M., Marr, C., Krumsiek, J., Hoppe, P. S., Schroeder, T., and Theis, F. J. (2011). Efficient fluorescence image normalization for time lapse movies. In *Microscopic Image Analysis with Applications in Biology 2011*, number x, pages 1–5.

- Scott, E. W., Simon, M. C., Anastasi, J., and Singh, H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science (New York, N.Y.)*, 265(5178):1573–7.
- Seita, J., Sahoo, D., Rossi, D. J., Bhattacharya, D., Serwold, T., Inlay, M. A., Ehrlich, L. I. R., Fathman, J. W., Dill, D. L., and Weissman, I. L. (2012). Gene Expression Commons: an open platform for absolute gene expression profiling. *PloS one*, 7(7):e40321.
- Seitz, A. (2013). *Development and Application of a Single Cell Quantification Tool for Microscopy Images*. Bachelor thesis, Technische Universität München, Ludwig-Maximilians-Universität München.
- Shi, W., Wang, H., Pan, G., Geng, Y., Guo, Y., and Pei, D. (2006). Regulation of the pluripotency marker Rex-1 by Nanog and Sox2. *The Journal of biological chemistry*, 281(33):23319–25.
- Shimomura, O., Johnson, F. H., and Saiga, Y. (1962). Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. *Journal of Cellular and Comparative Physiology*, 59(3):223–239.
- Shimoni, R., Pham, K., Yassin, M., Gu, M., and Russell, S. M. (2013). TACTICS, an interactive platform for customized high-content bioimaging analysis. *Bioinformatics (Oxford, England)*, 29(6):817–8.
- Shivanandan, A., Radenovic, A., and Sbalzarini, I. F. (2013). MosaicIA: an ImageJ/Fiji plugin for spatial pattern and interaction analysis. *BMC bioinformatics*, 14(1):349.
- Sieweke, M. H. and Allen, J. E. (2013). Beyond stem cells: self-renewal of differentiated macrophages. *Science (New York, N.Y.)*, 342(6161):1242974.
- Sigal, A., Milo, R., Cohen, A., Geva-Zatorsky, N., Klein, Y., Alaluf, I., Swerdlin, N., Perzov, N., Danon, T., Liron, Y., Raveh, T., Carpenter, A. E., Lahav, G., and Alon, U. (2006). Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins. *Nature methods*, 3(7):525–31.
- Sivia, D. S. (1996). *Data Analysis: A Bayesian Tutorial*. Oxford University Press.
- Smith, A. (2013). Nanog heterogeneity: tilting at windmills? *Cell stem cell*, 13(1):6–7.
- Smith, A. G. (1991). Culture and differentiation of embryonic stem cells. *Journal of tissue culture methods : Tissue Culture Association manual of cell, tissue, and organ culture procedures*, 13(2):89–94.

- Som, A., Harder, C., Greber, B., Siatkowski, M., Paudel, Y., Warsow, G., Cap, C., Schöler, H., and Fuellen, G. (2010). The PluriNetWork: an electronic representation of the network underlying pluripotency in mouse, and its applications. *PloS one*, 5(12):e15165.
- Sommer, C., Straehle, C., Kothe, U., and Hamprecht, F. A. (2011). Ilastik: Interactive learning and segmentation toolkit. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, number 1, pages 230–233. IEEE.
- Sternberg, S. (1983). Biomedical Image Processing. *Computer*, 16(1):22–34.
- Strasser, M., Theis, F. J., and Marr, C. (2012). Stability and multiattractor dynamics of a toggle switch based on a two-stage model of stochastic gene expression. *Biophysical journal*, 102(1):19–29.
- Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Developmental biology*, 56(1):110–56.
- Talbot, F. A. A. (1913). *Practical Cinematography and Its Applications*. W. Heinemann.
- Tamura, H., Mori, S., and Yamawaki, T. (1978). Textural Features Corresponding to Visual Perception. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(6):460–473.
- Tanaka, T. S. (2009). Transcriptional heterogeneity in mouse embryonic stem cells. *Reproduction, fertility, and development*, 21(1):67–75.
- Taniguchi, Y., Choi, P. J., Li, G.-W., Chen, H., Babu, M., Hearn, J., Emili, A., and Xie, X. S. (2010). Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533–8.
- Thomson, M., Liu, S. J., Zou, L.-N., Smith, Z., Meissner, A., and Ramanathan, S. (2011). Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell*, 145(6):875–89.
- Toutenburg, H. (2013). *Lineare Modelle: Theorie und Anwendungen (German Edition)*. Physica.
- Trott, J., Hayashi, K., Surani, A., Babu, M. M., and Martinez-Arias, A. (2012). Dissecting ensemble networks in ES cell populations reveals micro-heterogeneity underlying pluripotency. *Molecular bioSystems*, 8(3):744–52.
- Tsai, S. F., Strauss, E., and Orkin, S. H. (1991). Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as

- a positive regulator of its own promoter. *Genes & development*, 5(6):919–31.
- Tung, J. W., Heydari, K., Tirouvanziam, R., Sahaf, B., Parks, D. R., Herzenberg, L. A., and Herzenberg, L. A. (2007). Modern flow cytometry: a practical approach. *Clinics in laboratory medicine*, 27(3):453–68, v.
- Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*, 57(2):307–333.
- Waddington, C. H. (1957). *The strategy of the genes: a discussion of some aspects of theoretical biology*. Allen & Unwin.
- Wang, J. and Orkin, S. H. (2008). A protein roadmap to pluripotency and faithful reprogramming. *Cells, tissues, organs*, 188(1-2):23–30.
- Wang, J., Zhang, K., Xu, L., and Wang, E. (2011). Quantifying the Waddington landscape and biological paths for development and differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 108(20):8257–62.
- Warr, M. R., Pietras, E. M., and Passegué, E. (2011). Mechanisms controlling hematopoietic stem cell functions during normal hematopoiesis and hematological malignancies. *Wiley interdisciplinary reviews. Systems biology and medicine*, 3(6):681–701.
- Waters, J. C. (2009). Accuracy and precision in quantitative fluorescence microscopy. *The Journal of cell biology*, 185(7):1135–48.
- Weisberg, S. (2005). *Applied Linear Regression*. John Wiley & Sons.
- Wolf, D. E., Samarasekera, C., and Swedlow, J. R. (2007). Quantitative analysis of digital microscope images. *Methods in cell biology*, 81(06):365–96.
- Wolff, L. and Humeniuk, R. (2013). Concise review: erythroid versus myeloid lineage commitment: regulating the master regulators. *Stem cells (Dayton, Ohio)*, 31(7):1237–44.
- Xu, H., Baroukh, C., Dannenfelser, R., Chen, E. Y., Tan, C. M., Kou, Y., Kim, Y. E., Lemischka, I. R., and Ma’ayan, A. (2013). ESCAPE: database for integrating high-content published data collected from human and mouse embryonic stem cells. *Database : the journal of biological databases and curation*, 2013:bat045.
- Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K. L., Ema, H., and Nakauchi, H. (2013). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*, 154(5):1112–26.

- Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature*, 453(7194):519–23.
- Young, R. A. (2011). Control of the embryonic stem cell state. *Cell*, 144(6):940–54.
- Zhang, J., Campbell, R. E., Ting, A. Y., and Tsien, R. Y. (2002). Creating new fluorescent probes for cell biology. *Nature reviews. Molecular cell biology*, 3(12):906–18.
- Zhang, P., Andrianakos, R., Yang, Y., Liu, C., and Lu, W. (2010). Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. *The Journal of biological chemistry*, 285(12):9180–9.
- Zhang, P., Behre, G., Pan, J., Iwama, A., Wara-Aswapati, N., Radomska, H. S., Auron, P. E., Tenen, D. G., and Sun, Z. (1999). Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proceedings of the National Academy of Sciences of the United States of America*, 96(15):8705–10.
- Zhang, P., Zhang, X., Iwama, A., Yu, C., Smith, K. A., Mueller, B. U., Narravula, S., Torbett, B. E., Orkin, S. H., and Tenen, D. G. (2000). PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood*, 96(8):2641–8.
- Zhou, J. X. and Huang, S. (2011). Understanding gene circuits at cell-fate branch points for rational cell reprogramming. *Trends in genetics : TIG*, 27(2):55–62.
- Zimmermann, T. (2005). Spectral imaging and linear unmixing in light microscopy. *Advances in biochemical engineering/biotechnology*, 95:245–65.