Continuous Quantification of Transcription Factor Dynamics in Individual Hematopoietic Stem and Progenitor Cells

Philipp Hoppe

Dissertation der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

München, 2013

Diese Dissertation wurde angefertigt

unter der Leitung von Prof. Dr. Timm Schroeder

am Institut für Stammzellforschung

am Helmholtz Zentrum München

Erstgutachter:	Prof. Dr. Thomas Cremer
Zweitgutachter:	Prof. Dr. Angelika Böttger
Tag der Abgabe:	30. Juli 2013
Tag der mündlichen Prüfung:	11. September 2014

ERKLÄRUNG

Hiermit versichere ich ehrenwörtlich, dass meine Dissertation selbstständig und ohne unerlaubte Hilfsmittel angefertigt worden ist.

Die vorliegende Dissertation wurde weder ganz, noch teilweise bei einer anderen Prüfungskomission vorgelegt.

Ich habe noch zu keinem früheren Zeitpunkt versucht, eine Dissertation einzureichen oder an einer Doktorprüfung teilzunehmen.

München, den 2. Oktober 2014

Part of this work is published as follows:

Kimura A, Rieger MA, Simone JM, Chen W, Wickre MC, Zhu BM, **Hoppe PS**, O'Shea J, Schroeder T and Hennighausen L (2009) The transcription factors STAT5A/B regulate GM-CSF-mediated granulopoiesis *Blood*, 114: 4721 - 4728

Rieger MA, **Hoppe PS**, Smejkal BM, Eitelhuber AC and Schroeder T (2009) Hematopoietic cytokines can instruct lineage choice *Science*, 325: 217 – 218

Schwarzfischer M, Marr C, Krumsiek J, **Hoppe PS**, Schroeder T and Theis FJ (2011) Efficient fluorescent image normalization for time lapse movies *Proc. Microscopic Image Analysis with Applications in Biology*, Heidelberg, Germany, 2. Sept. 2011

Roeder I, Loeffler M, Glauche I, Bornhaeuser M, Buske P, Clarke G, Cross M, Efroni S, Ilan B, Essers M, Galle J, Giangreco A, Glauche I, Griessinger E, **Hoppe P**, Kobel S, Krinner A, Lauffenburger D, Lemischka I, LoCelso C, Loeffler M, Lutolf M, Moore K, Nazareth E, Oostendorp R, Palmer M, Pompe T, Rocheteau P, Roeder I, Scherf N, Schroeder T, Tajbakhsh S, Vilne B, van der Wath R, Waskow C, Werner C, Yaromina A, Zandstra P (2011) Towards a quantitative understanding of stem cell–niche interaction: Experiments, models, and technologies *Blood Cells, Molecules, and Diseases*, 46: 308-317

Eilken HM, Rieger MA, **Hoppe PS**, Hermann AC, Smejkal BM, Drew EC, Thum M, Ninkovic J, Beckervordersandforth RM and Schroeder T (2011) Continuous long-term detection of live cell surface markers by 'in culture' antibody staining *Protocols Exchange*, DOI: 10.1038/protex.2011.205

Nakles RE, Millman S, Cabrera MC, Johnson P, Mueller S, **Hoppe PS**, Schroeder T and Furth PA (2013) Time-lapse Imaging of Primary Preneoplastic Mammary Epithelial Cells Derived from Genetically Engineered Mouse Models of Breast Cancer *J Vis Exp*, 72: e50198

Buggenthin F, Marr C, Schwarzfischer M, **Hoppe PS**, Hilsenbeck O, Schroeder T and Theis FJ (2013) An automatic pipeline for robust cell detection in bright field images of high-throughput microscopy *BMC Bioinformatics*, 14: 297

Filipczyk A, Gkatzis K, Fu J, **Hoppe PS**, Lickert H, Anastassiadis K and Schroeder T (2013) Biallelic expression of nanog protein in mouse embryonic stem cells *Cell Stem Cell*, 13: 12-31

Hasemann MS, Lauridsen FKB, Waage J, Jakobsen JS, Frank AK, Schuster MB, Rapin N, Bagger FO, **Hoppe PS**, Schroeder T and Porse BT (2014) C/EBPα Is Required for Long-Term Self-Renewal and Lineage Priming of Hematopoietic Stem Cells and for the Maintenance of Epigenetic Configurations in Multipotent Progenitors *PLoS Genetics*, 10: e1004079

Tanaka Y, Sanchez V, Takata N, Yokomizo T, Yamanaka Y, Kataoka H, **Hoppe PS**, Schroeder T and Nishikawa SI (2014) Circulation-Independent Differentiation Pathway from Extraembryonic Mesoderm toward Hematopoietic Stem Cells via Hemogenic Angioblasts *Cell Reports*, 8: 31-39

Hoppe PS, Coutu DL, Schroeder T (2014) Single-cell technologies sharpen up mammalian stem cell research *Nature Cell Biology*, in press

1 Table of Contents

2 Abstract

The hematopoietic system is the best understood stem cell system in mammals. It is easy accessible and many decision steps need to take place during differentiation of multipotent hematopoietic stem cells into numerous mature blood cell types. Prospective isolation, *in vitro* differentiation experiments and *in vivo* transplantations have significantly contributed to the phenotypic definition of intermediate hematopoietic progenitor cells with restricted lineage potential. Lineage decisions at various branching points of hematopoiesis are commonly believed to be regulated by stochastically fluctuating transcription factor networks. The paradigmatic transcription factor pair PU.1 and Gata1 has specifically contributed to that hypothesis: PU.1 and Gata1 have been described as master regulators driving granulocytic-monocytic versus megakaryocytic-erythroid lineage choice, respectively. Lineage priming, reprogramming experiments, positive autoregulation and mutual cross-antagonism of PU.1 and Gata1 have led to the hypothesis that random 'noisy' fluctuations in PU.1 and Gata1 expression will allow one factor 'defeating' the other and thus to lineage choice. However, this hypothesis is based on limited expression data. So far, transcription factor expression mostly was measured at the mRNA level in discontinuous snapshot analysis. For a comprehensive understanding, quantification of protein levels before, during and after lineage choice would be required. Therefore, a knock-in of the fluorescent protein mCHERRY into the endogenous *Gata1* locus was performed and the mouse line was bred with an existing PU.1eYFP knock-in mouse line. Phenotypic analysis of the double knock-in line did not reveal any defects in the function of the transcription factors. The double knock-in line allowed the measurement of endogenous expression levels at any time in individual primary hematopoietic stem and progenitor cells. Flow cytometric analysis of double knock-in mice led to the identification and prospective isolation of several new myeloid hematopoietic subpopulations. Continuous long-term time-lapse imaging and tracking at the single-cell level allowed the generation of cell genealogy trees and continuous live quantification of PU.1 and Gata1 protein expression. Continuous single-cell data showed that future lineage choice does not correlate with an interdependent fluctuating network of PU.1 and Gata1. It is therefore not compulsory for, and likely not involved in, granulocytic-monocytic versus megakaryocytic-erythroid lineage choice. This corrects a central paradigm about the regulation of hematopoietic stem and progenitor cell lineage choice. In addition, the novel PU.1eYFP/Gata1mCHERRY double knock-in mouse line and the developed technology of

continuous quantification of transcription factor levels in living primary cells will be invaluable to further investigate what regulates hematopoiesis.

3 Introduction

3.1 Hematopoiesis

Hematopoiesis is the lifelong process of blood generation from hematopoietic stem cells (HSCs) which are both able to self-renew and differentiate into more than 10 different mature blood cell types. Among those are cells with basic function like oxygen supply, carbon dioxide removal or blood clotting and cells of the innate as well as the adaptive immune system. Every day of an adult human's life 10^{11} – 10^{12} cells are produced *de novo* (Kaushansky and Williams, 2010). After massive blood loss due to injury the blood system is able to rebuild itself within a relatively short period of time. As an easy accessible liquid organ, the blood system is the best investigated stem cell system in higher organisms and often serves as a model for other stem cell types. Therapeutic HSC transplantations are reliably performed in humans for more than 50 years. Despite of the clinical application there are still many open questions in HSC research. Due to experimental reasons like the easy and reproducible access to primary cells for basic research and the possibility of targeted genetic manipulations, the murine blood system is much better understood than the human. For therapeutic reasons, there is a huge interest of understanding how the highly complex and dynamic process of hematopoiesis is controlled. This includes both the developmental origin of hematopoietic cells and the regulation of adult hematopoiesis.

3.1.1 Embryonic Hematopoiesis

Embryonic hematopoiesis during development in mice occurs in different waves at different places (reviewed in Orkin and Zon, 2008). One distinguishes between primitive and definitive hematopoiesis. Primitive blood cells originate in the yolk sac and show unique morphological and molecular features compared to definitive blood cells. Their main function is the oxygen supply of the rapidly growing embryo. After the transient primitive phase definitive hematopoiesis takes place and generates blood types of adult phenotypes already. The sites of hematopoiesis then change during development from the AGM (aorta-gonads-mesonephros) region to the placenta, fetal liver, thymus, spleen and finally bone marrow, where hematopoiesis is maintained throughout the whole live of the organism. It is assumed that the

four latter sites only provide niches for maintenance and expansion for hematopoietic stem and progenitor cells (HSPCs) that have migrated there, instead of producing new cells *de novo*. To what extent HSCs arise at the other sites of embryonic hematopoiesis is not well understood. Endothelial cells are the definite immediate cellular precursor of blood cells which has been demonstrated recently (Eilken et al., 2009; Bertrand et al., 2010; Kissa and Herbomel, 2010).

3.1.2 Adult Hematopoiesis

Adult hematopoiesis is entirely maintained by HSCs that mainly reside in the bone marrow. After single cell transplants one individual HSC is defined as a cell that is able to reconstitute the whole blood system of a lethally irradiated recipient mouse over at least 16 weeks (reviewed in Ema et al., 2006). Only 1 out of about 100000 cells of unfractionated bone marrow is a HSC. Protocols for prospective isolation of HSCs have increased significantly during the past years (reviewed in Warr et al., 2011). Most current protocols include the marker combination 'Lineage', c-Kit and Sca-1 (Spangrude et al., 1988), where 'Lineage' resembles a cocktail of antibodies that are directed against antigens only expressed by mature cells ('lineage-positive'). HSCs are both positive for c-Kit, the receptor for the cytokine stem cell factor (SCF), and Sca-1 (stem cell antigen 1). Additional negative markers for better purification were described and included Thy-1.1 (Spangrude et al., 1988), CD34 (Osawa et al., 1996) and Flt3 (Adolfsson et al., 2001; Christensen and Weissman, 2001). A more recent protocol described the SLAM markers CD150 and CD48 that proved to be less strain and age dependent than Thy-1.1 and CD34 (Kiel et al., 2005).

Classically, the hematopoietic system has been divided into the lymphoid and myeloid system (Figure 3-1). This classical discrimination was confirmed by the discovery of the common-lymphoid progenitor (CLP) that is only able to give rise to B-, T- and natural killer (NK) cell lineages *in vivo* (Kondo et al., 1997), and the discovery of the presumptive common myeloid progenitor (CMP) that only exhibits lineage potential for the granulocytic-monocytic (GM) and megakaryocytic-erythroid (MegE) lineages, respectively (Akashi et al., 2000). From which cell type the CLP and CMP directly originate is an open question. Cells that do not have long-term reconstitution potential, but still can give to all mature blood lineages have been described as multipotent progenitors (MPPs). Due to the amount of available markers

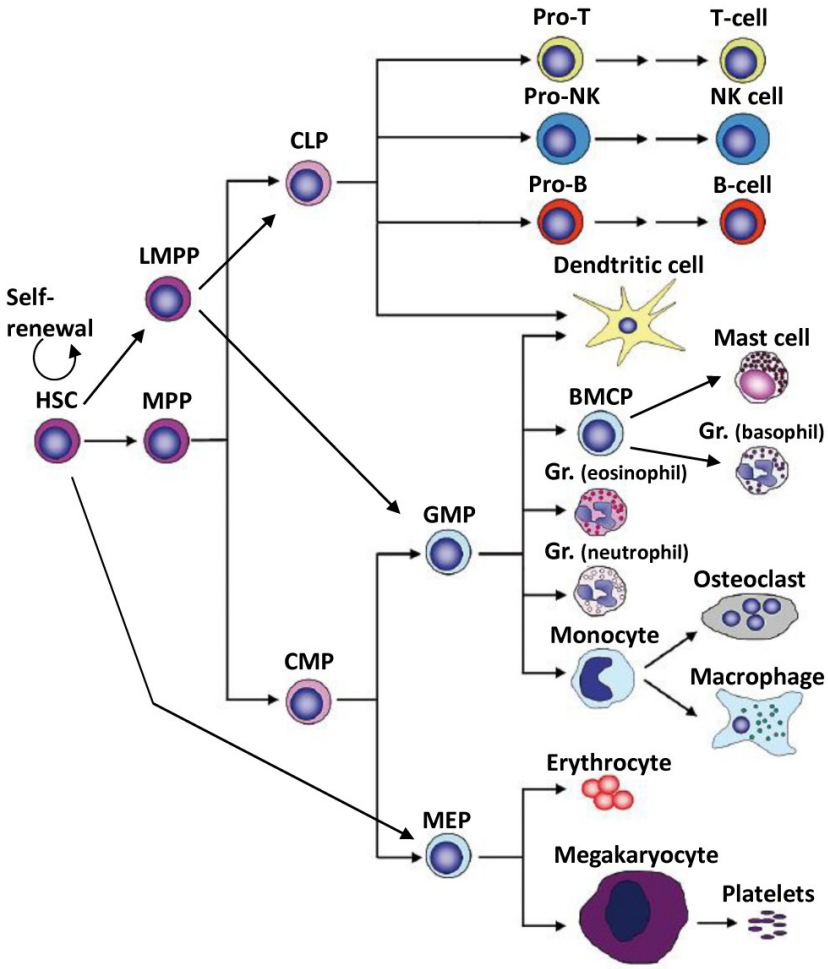


Figure 3-1: Hematopoietic lineage tree
Classically, the hematopoietic system is divided into the lymphoid and the myeloid system with the first branching points being the common lymphoid progenitor (CLP) and common myeloid progenitor (CMP). CLPs are able to give rise to T-cells, B-cells and natural killer (NK) cells. CMPs are further able to give rise to granulocyte-macrophage progenitors (GMPs) and megakaryocyte-erythrocyte progenitors (MEPs). GMPs then give rise to neutrophil, basophil and eosinophil granulocytes, osteoclasts, macrophages and mast cells. The progeny of MEPs are megakaryocytes and erythrocytes. Dendritic cells can both have lymphoid and myeloid origin. All hematopoietic cells develop from the HSCs which lose their self-renewal potential during differentiation but still remain multipotent (MPPs). MPPs are a mixture of numerous different phenotypes, whose hierarchical organization is not entirely clear yet. Besides the classical differentiation pathway an alternative route involving the LMPP (lymphoid-primed multipotent progenitor) without MegE-lineage potential has been described. The hematopoietic hierarchy is still incomplete and subjected to constant changes. Image adapted from Rieger and Schroeder, 2007.

and significant overlaps between their expressions, the hierarchical order of different MPP populations is not entirely clear. On a population level different MPP phenotypes show different proportions of cell-cycle stages *in vivo* that let them appear more immature (Wilson et al., 2008). Besides a shorter reconstitution potential of more mature MPPs the different lineage potential of those has to be taken into account: The upregulation of Flt3 among the $Lin^{-}c\text{-}Kit^{+}Sca\text{-}1^{+}$ cells has led to the description of the lymphoid-primed multipotent

progenitor (LMPP) that did not have significant erythroid and megakaryocytic potential anymore (Adolfsson et al., 2005). Conflicting data about $Flt3^+$ cells still retaining erythroid and megakaryocytic potential was published subsequently (Forsberg et al., 2006). In conclusion, there seems to be more than one pathway in adult hematopoiesis that allows the development of mature cell types. The question of this thesis deals with the lineage decision of a multipotent myeloid progenitor between the well-described and accepted granulocyte/macrophage progenitor (GMP) and megakaryocyte/erythrocyte progenitor (MEP). GMPs can further differentiate into dendritic cells, mast cells, basophilic/ eosinophilic/neutrophilic granulocytes and monocytes. MEPs only have lineage potential for erythrocytes and megakaryocytes that will further give rise to platelets.

3.2 Factors Influencing Hematopoiesis

3.2.1 Extrinsic Factors

Adult HSCs do not exist as shielded entities, but usually reside in a very specialized bone marrow environment called the niche. The concept of the niche was already postulated a long time ago by Ray Schofield (Schofield, 1978). Only many years later, the existence of a niche as specialized cells for germline stem cells in *Drosophila melanogaster* could be proven (reviewed in Jones and Wagers, 2008; Morrison and Spradling, 2008). Since then, many stem cell niches in various tissues in mammals, especially in the mouse, have been identified: Those for example include niches for muscle cells under the basal lamina of myofibers, for intestinal epithelium in the base of crypts, for interfollicular epidermis in the basal layer and for central nervous system stem cells in the subventricular zone. As opposed to those rather rigid tissues, HSCs reside in a liquid organ. HSCs do not constantly remain in the niche, but circulate through the peripheral blood (Wright et al., 2001). Within the bone marrow HSCs have been found to be located in close proximity to both endosteum and sinusoidal blood vessels in trabecular bone (Nilsson et al., 2001; Kiel et al., 2005; Adams and Scadden, 2006). At the cellular level osteoblasts and osteoclasts play an important role in the endosteal niche. Signals from osteoblasts positively influencing HSC expansion, maintenance and quiescence are mediated through Notch activation, BMP signaling, N-Cadherin and Tie2/angiopoietin-1 signaling (Calvi et al., 2003; Zhang et al., 2003; Arai et al., 2004). Ca^{2+}-detection of HSCs mediated through a calcium-sensing receptor (CaR) keeps them in close proximity to the endosteum and the chemokine CXCL12 (SDF-1) allows homing of HSCs via CXCR4

signaling (Wright et al., 2002; Adams et al., 2006; Sugiyama et al., 2006). Osteoclasts were shown to degrade important niche components accompanied by hematopoietic progenitor mobilization upon stress induction (Kollet et al., 2006). Subendothelial progenitor cells were found to be capable of re-establishing HSC niches at developing sinusoids after transplantation and releasing Angiopoietin-1 (Sacchetti et al., 2007). Remarkably, HSCs also can reside in the spleen or liver upon stress induction, supporting the theory of perivascular niches. Recently, it also has been shown that progeny of stem cells also participates in regulating tissue homeostasis directly in the niche (reviewed in Hsu and Fuchs, 2012). In conclusion, the adult HSC microenvironment is highly dynamic and complex and various models about 'the' niche exist which still need to be better refined.

Besides direct cell-to-cell contact and paracrine signaling in the niche, hematopoiesis is largely influenced by cytokines. In general, cytokines are important for steady-state hematopoiesis, stress response and immunomodulation. They are peptides or glycopeptides that act as signaling molecules both within short distances but also as systemic agents in the bloodstream. Cytokines control individual cellular processes like survival, proliferation, lineage commitment, activation and migration. Some cytokines are lineage restricted and act on more mature cells, others are able to influence multiple lineages in more immature cells. Stem cell factor (SCF) is a cytokine that is important both for hematopoietic stem and progenitor cells (Ashman, 1999). Its receptor c-Kit is an important marker for the discrimination and purification of those cells. Thrombopoietin (TPO) acts at various levels of hematopoiesis: Together with SCF, TPO was shown to make HSCs self-renew *in vitro*, but TPO is also a strong stimulus for megakaryocyte colony formation (Ema et al., 2000). A very lineage restricted cytokine is erythropoietin (EPO) which was initially purified from urine of anemic patients (reviewed in Krantz, 1991). It is responsible for the survival and maturation of committed erythroid progenitors (Wu et al., 1995). Other examples of rather lineage-restricted cytokines are the colony-stimulating factors (CSFs) G-CSF, M-CSF and GM-CSF that inherited their name from their ability to give rise to granulocyte and/or macrophage colonies *in vitro*, respectively (Metcalf, 2008). Interleukin-3 (IL-3) was originally also attributed to the group of CSFs, before it became clear that it allows the colony formation of not only granulocytes and macrophages, but also eosinophils, megakaryocytes and erythrocytes (reviewed in Zon, 2001). Interleukin-6 (IL-6) also has a broad spectrum of cellular targets and regulates both myelopoiesis and the innate as well as the adaptive immune response (Jones, 2005). Cytokines can be produces by a variety of tissue including myeloid cells, liver and kidney. Since the discovery of cytokines, the basic question if cytokines can

actively induce lineage choice (instructive) or only allow the survival of already progenitors (permissive) was unanswered (Enver et al., 1998). Recently, definite proof about an instructive role of the cytokines M-CSF and G-CSF could be provided (Rieger et al., 2009).

3.2.2 Intrinsic Factors

Like in many other cellular processes, epigenetics play an important role in hematopoiesis (reviewed in Butler and Dent, 2013). Epigenetics is defined as the stable and inheritable change of gene expression that is independent of alterations in the DNA sequence. A major molecular mechanism of epigenetics is the modulation of the chromatin structure. It involves posttranslational modifications of histones and chemical modifications of DNA bases. Dysregulation of epigenetic is often found in hematologic malignancies. In general, one distinguishes between two different states of chromatin: Euchromatin denotes an 'open' chromatin structure which makes DNA accessible for proteins like transcription factors and allows access of polymerase. It is accompanied by characteristic methylation and acetylation at specific lysine residues of histones and demethylated DNA. By contrast, heterochromatin denotes a 'closed' chromatin structure which is the result of a more dense packing of chromatin leading to restricted access of DNA-accessory molecules. Hallmarks are hypoacetylation, methylation at (other) amino acid residues and methylation of cytosin in CpG islands. Other modifications of DNA have recently been discovered, but their biological relevance still need to be elucidated. A myriad of chromatin-modifying enzymes is involved in establishing, maintaining and erasing chromatin changes. Conditional knock-out experiments have established an essential role of certain enzymes for HSC self-renewal and/or for specific hematopoietic lineages. By sorting HSCs, MPPs, CLPs, CMPs, GMPs and MEPs and subsequent epigenetic characterization it was shown that multipotent cells exhibited 'open' chromatin structures at different lineage-affiliated gene loci (Attema et al., 2007). During differentiation the 'open' chromatin was either maintained at loci important for certain lineages or chromatin changed into a 'closed' conformation in non-affiliated gene loci. In conclusion, different developmental stages are marked by different epigenetic modifications at various gene loci. Epigenetic is important for keeping cells in a multipotent and self-renewing status and for specification to certain lineages, as demonstrated by epigenetic dysregulation in cancers.

During recent years, non-coding RNAs have emerged as a new class of regulators in development (Mattick, 2007). Among those are microRNAs (miRNAs) which can regulate gene expression post-transcriptionally (reviewed in Filipowicz et al., 2008). MiRNAs are usually 21 nucleotides long and mediate translational repression or degradation of mRNAs through specific base-pair binding at their 3'-untranslated regions (3'-UTR). Predictions estimate that 30% of mammalian genes are regulated by miRNAs. Within the hematopoietic system, more than 100 different miRNAs are expressed (reviewed in O'Connell and Baltimore, 2012). In general, a single miRNA has many predicted target molecules. Additionally, some mRNAs have predicted binding sites for several miRNAs. In hematopoiesis miRNAs function at different stages, including HSCs and both during lymphoid and myeloid development. Global disruption of the miRNA machinery in the T- or B-cell lineage severely affects their differentiation and function. Within the myeloid system important miRNAs for both the GM and MegE lineages also have been identified. Knock-outs and overexpression of individual miRNAs showed a potential role in lineage-commitment by binding to lineage-specific transcription factors. Due to their redundancy, the specific influence of miRNAs on individual target mRNAs is difficult to elucidate and remains to be investigated.

Transcription factors are the ultimate determinants of hematopoiesis, because they have the capacity to directly regulate target gene expression. Transcription factors can either have a function in different HSPCs or they are very cell-type specific. They all are commonly believed to be part of a large transcription factor network (Laiosa et al., 2006). Most data about the importance of transcription factors for certain lineages was derived from knock-out experiments. Within the lymphoid part of the blood system, classic examples of lineage-specific essential transcription factors are Pax5 for B-cells and Gata3 for T-cells (Ting et al., 1996; Nutt et al., 1999; Mikkola et al., 2002). Another factor, E2A, has been shown to be important both for B- and T-cell lineages (Bain et al., 1994; Barndt et al., 2000). An example for a transcription factor being important both for HSPCs and certain myeloid lineages is Gata2 (Tsai and Orkin, 1997). C/EBPα was originally described to be essential for neutrophil development, but its disruption also leads to enhanced HSC self-renewal (Zhang et al., 1997; Zhang et al., 2004). Due to its central role in hematopoiesis, the transcription factor PU.1 is one of the best studied ones in hematopoiesis (reviewed in Gupta et al., 2009). It is important for the development of granulocytes, macrophages and B-cells, but also a critical role in HSCs has been revealed. Another very well studied factor is Gata1 which has been associated with the megakaryocytic-erythroid lineage (reviewed in Ferreira et al., 2005). Gata1 also

plays a role in the development of the eosinophilic lineage and mast cells. The existence of a transcription factor network and the diverse role of individual factors are nicely displayed by the network dynamics: Usually, certain lineages are associated by the (enforced) upregulation of one transcription factor, but different lineages can also develop depending on the sequence of transcription factor expression (Walsh et al., 2002; Chou et al., 2009). For lineage-determination between the GM and MegE lineage, PU.1 and Gata1 are commonly believed to be central transcription factors of a large network (Laiosa et al., 2006; Laslo et al., 2008; Krumsiek et al., 2011).

3.3 The Transcription Factors PU.1 and Gata1

3.3.1 Structure of PU.1 and Gata1

The transcription factor PU.1 (Purine-rich box 1) was discovered as a putative oncogene by insertional mutagenesis in murine erythroleukemia induced by the spleen focus forming virus (SFFV) (Moreau-Gachelin et al., 1988). Therefore, PU.1 is also known as Spi-1 or Sfpi-1 (SFFV proviral integration). The *PU.1* gene is located on mouse chromosome 2 and consists of 5 exons. So far, there has only been one transcript described. The protein belongs to the ETS (E-twenty six) family of transcription factors, consists out of 272 amino-acids and has a predicted molecular weight of 31 kDa. PU.1 contains 3 functional domains which include a transactivation domain, a PEST domain (rich in proline, glumatic acid, serine and threonine) and the helix-turn-helix DNA-binding ETS domain. Post-translational modifications, like the phosphorylation at serin 148 of PU.1 have been described (Pongubala et al., 1993), but its importance in PU.1 functionality in not clear.

The transcription factor Gata1 was first described as an erythroid-specific transcription factor binding to cis-elements of both human and chicken globin genes (Evans et al., 1988; Wall et al., 1988). The name “Gata” is derived from its binding specificity to the consensus sequence (A/T)GATA(A/G) by two conserved zinc finger domains. Gata1 is member of the Gata-family of transcription factors which contains five more structurally similar proteins (Gata2 – Gata6) out of which Gata1-Gata3 are highly specific to the hematopoietic system. However, Gata1 is also expressed in Sertoli cells (Ito et al., 1993). The murine *Gata1* gene is located on the X chromosome and consists out of 6 exons with the start codon being on exon 2. There is

some debate about additional 5'-UTR exons which leaves some doubts about the total length of Gata1 transcripts (Lowry and Mackay, 2006). Murine Gata1 (43 kDa) is a 413 amino acids long protein and consist of a N-terminal transactivation domain and 2 consecutive zinc-fingers, named N-finger and C-finger. Gata1 is subject to various post-translational modifications, including phosphorylation at 7 different serine residues (Crossley and Orkin, 1994), acetylation (Boyes et al., 1998) and SUMOylation (Collavin et al., 2004). The functional importance of these modifications is largely unknown. Possible roles in DNA binding specificity and transcriptional activation are being discussed.

3.3.2 PU.1 and Gata1 Knock-Outs

The functional importance for certain lineages can be demonstrated best by *in vivo* knock-out models. Mouse embryos that carry a homozygous knock-out of the PU.1 DNA-binding domain by insertional mutagenesis die during late gestation (Scott et al., 1994). The reason for death was a multilineage defect in monocytes, granulocytes, B-cells and T-cells, demonstrating the necessity of PU.1 in the proper development of those lineages. Megakaryocytic and erythroid progenitor cells were unaffected by the PU.1 knock-out, but erythroblasts maturation was impaired. Heterozygous PU.1 knock-out embryos did not show a phenotype. Another study using a similar approach reported similar effects (McKercher et al., 1996): Mice with a homozygous PU.1 knock-out died shortly after birth due to severe septicemia. Again, mature neutrophils, macrophages, B-cells and T-cells were missing. When animals were treated with antibiotics, survival was prolonged and cells with neutrophil features could be detected, showing that PU.1 might not be absolutely essential for myeloid lineage choice. Besides the importance for mature lineages, PU.1 was also shown to be important for the maintenance of the HSC pool in adult mice (Iwasaki et al., 2005). In that conditional knock-out model, CMPs and GMPs could still be identified and sorted, but cells failed to give rise to mature colonies *in vitro* and only consisted of myeloblasts.

The central importance of Gata1 for the erythroid lineage was first shown by targeted disruption of Gata1 in embryonic stem (ES) cells and their subsequent failure to give rise to mature red blood cells in chimeric animals (Pevny et al., 1991). *In vitro* differentiation experiments also showed that Gata1$^-$ ES cells were impaired in their erythroid development (Weiss et al., 1994; Pevny et al., 1995). Full embryonic Gata1 knock-outs died between

embryonic day 10.5 (E10.5) and E11.5 of gestation (Fujiwara et al., 1996). Those embryos completely lacked mature erythroid cells and instead showed an arrest at the proerythroblast stage, suggesting that Gata1 is not essential for erythroid lineage choice. Due to that severe phenotype the importance of Gata1 for other blood lineages had been elusive until the genetic manipulation of *Gata1* regulatory sequences led to a megakaryocyte-specific phenotype (Shivdasani et al., 1997). Affected mice showed massively reduced platelet numbers and megakaryocyte maturation defects. A conditional Gata1 knock-out model in adult mice confirmed the results and again led to a maturation arrest at the proerythroblast stage and thrombocytopenia (Gutierrez et al., 2008). Furthermore, mice failed to have an erythropoietic stress response upon treatment with a hemolytic agent.

In conclusion, both PU.1 and Gata1 are essential for embryonic and adult steady-state hematopoiesis. After knock-out of the transcription factors, mature cells of the respective lineages fail to develop. However, both transcription factors seem to be dispensable for the lineage choice in favor of the GM and MegE lineage, respectively.

3.3.3 Reprogramming Potential of PU.1 and Gata1

The power of individual transcription factors in reprogramming cells to certain lineages was first demonstrated by the generation of myoblasts from fibroblasts only by the overexpression of MyoD (Davis et al., 1987). Afterwards, the same potential was also shown for the transcription factor Gata1 whose enforced expression in a myeloid cell line induced megakaryocytic differentiation (Visvader et al., 1992). This observation was later confirmed and expanded by the overexpression of Gata1 in Myb-Ets-transformed chicken myeloblasts, a model for committed myelomonocytic cells (Kulessa et al., 1995). It could be shown that the enforced erythroid, megakaryocytic or eosinophilic lineage outcome was dependent on the dosage of Gata1. Additionally, Gata1 not only reprogrammed cells to other lineages, but actively repressed myelomonocytic markers.

Reprogramming potential could also be conferred to the transcription factor PU.1 in the same model system: Enforced expression of PU.1 led to an instructive differentiation along the GM lineage in still multipotent cells (Nerlov and Graf, 1998). PU.1 both upregulated GM lineage-specific genes and downregulated progenitor/thrombocyte-specific markers as well as Gata1. In the same publication, inducible forms of PU.1 were also used. PU.1 was fused to the

hormone-binding-domain of the human estrogen-receptor (ER), making its activation reversibly controllable by the application of estrogen or its derivates. Long-term activation of PU.1 led to GM cell fates again, whereas short-term activation led to the development of immature eosinophilic stages only. Gata1-ER fusions have also been described and proved their usefulness even in primary cells. Again, enforced expression of Gata1 led to the outcome of erythroid, eosinophilic and basophilic-like cell fates (Heyworth et al., 2002). Interestingly, reprogramming experiments with 2 transcription factors showed that the lineage outcome is dependent of the sequence of the expression of Gata2 and C/EBPα (Iwasaki et al., 2006). Therefore, inducible forms of transcription factors represent a powerful tool that makes their function easy to study in a timed, tunable and reversible way in the cell type of interest.

3.3.4 Interactions of PU.1 and Gata1

Along with the observation that PU.1 and Gata1 downregulate opposing lineage-specific genes, it could also be shown that both transcription factors directly bind to each other (Zhang et al., 1999; Nerlov et al., 2000). More in detail, *in vitro* protein interactions, electrophoresis mobility shift assays (EMSA), co-immunoprecipitations and luciferase reporter assays provided a detailed view about the biochemical interaction of PU.1 and Gata1. It could be demonstrated that the conserved C-finger of both Gata1 and Gata2 binds to the ETS domain of PU.1 which inhibits the transcriptional activity of PU.1. The inhibitory influence of PU.1 on Gata1 is mediated by the N-terminus of PU.1 (Zhang et al., 2000). By binding of PU.1 to the C-finger of Gata1, Gata1 binding to DNA is disabled and transcriptional activity is directly blocked. Additionally, PU.1 interacts with pRB and cooperatively represses Gata1 transcriptional activity and hence erythroid differentiation (Rekhtman et al., 2003). Direct negative influence of both transcription factors on the expression levels of the other factor could not be detected. To current knowledge the cross-antagonism is post-transcriptional only and it inhibits differentiation of the opposing lineage by inhibiting the central lineage-determining factors PU.1 and Gata1 in a dose-dependent manner. Moreover, PU.1 is also able to create a repressive chromatin structure by binding to Gata1 directly at its target genes (Stopka et al., 2005). The direct biochemical interaction was also studied detailed at the molecular level by extensive mutagenesis and NMR spectroscopy (Liew et al., 2006). The binding between PU.1 and Gata1 is of stable affinity which cannot be interrupted by individual amino acid changes. Interestingly, a part of the Gata1 C-finger shares sequence

homology with the PU.1 co-factor c-Jun, suggesting a competitive binding mechanism that allows Gata1 to inhibit PU.1 activity.

3.3.5 Autoregulation of PU.1 and Gata1

After the discovery of the importance of Gata1 for the erythroid lineage, attempts of delineating its transcriptional regulation were made. Based on the observation that Gata1 mRNA levels steadily increase during erythroid differentiation, promoter analysis of the *Gata1* gene pointed to a double GATA motif (Tsai et al., 1991). DNA footprinting experiments in an erythroleukemic cell line could indeed show Gata1 binding to its own promoter. Therefore, a positive feedback loop was suggested leading to the maintenance of differentiation by keeping the promoter state 'on'. Transgenic *in vivo* experiments in zebrafish embryos confirmed Gata1 binding to its own promoter and also showed that both the N- and C-finger are necessary for correct transcriptional function (Kobayashi et al., 2001). Furthermore, targeted deletion of the Gata1 binding site *in vivo* led to a selective loss of the eosinophilic lineage (Yu et al., 2002). Biochemically, a dimerization of Gata1 proteins seems to be important for full transcriptional activity at its own promoter (Crossley et al., 1995; Nishikawa et al., 2003).

Similarly to Gata1 and its autoregulatory potential, binding sites for PU.1 are found in its own murine and human promoters (Chen et al., 1995). Mutation of the -39 bp PU.1 binding site significantly abolished promoter activity. Correct myeloid gene expression is also importantly mediated by a -14kb upstream regulatory element (URE) (Li et al., 2001). Interestingly, that URE is sufficient for correct reporter gene expression *in vivo*. Further detailed analysis with DNA binding experiments showed that PU.1 binds to a conserved region within the URE (Okuno et al., 2005). Mutation of that site led to a significant decrease of PU.1 binding and reporter gene expression in cell lines. Like in the case of Gata1, another PU.1 mediated positive autoregulatory loop was suggested for correct expression of PU.1 in the respective cell types.

3.3.6 Mutations of PU.1 and Gata1

Targeted deletion of the aforementioned URE *in vivo* led to reduced PU.1 levels (20%) and predisposed mice to acute myeloid leukemia (Rosenbauer et al., 2004). Further analysis of the role of the URE showed its importance in regulating lymphoid development both in B- and T-cells. Depending on the lineage the URE had different functions either as an enhancer or a repressor (Rosenbauer et al., 2006). By this study it could be shown that a missing control element leading to dose-changes of one transcription factor can induce cancer in several murine blood lineages. Adult mice with induced deletion of PU.1 exon 5 all developed myeloid leukemia and started to die after 13 weeks (Metcalf et al., 2006). Several PU.1 mutations have been associated with AML in humans (Mueller et al., 2002) and PU.1 has been described as a tumor suppressor gene that is often mutated in radiation induced myeloid leukemia (Cook et al., 2004). Enforced expression of PU.1 in promyelocytic leukemia cells could rescue the phenotype. The potential of PU.1 as a therapeutic target in leukemic cells was supported by observations about terminal differentiation induced by lentiviral overexpression of PU.1 in human AML samples (Durual et al., 2007). More work of the group that discovered PU.1 as a putative oncogene could show that 50% of transgenic PU.1 overexpressing mice developed erythroleukemia, caused by a differentiation block at the proerythroblast stage (Moreau-Gachelin et al., 1996). Interestingly, excessive PU.1 expression leading to tumor formation could be overcome by enforced expression of Gata1 (Choe et al., 2003). Again, as a consequence the cells lost their tumorigenic potential and differentiated terminally. Gata1 mutations are also known to be involved in leukemia development (reviewed in Burda et al., 2010). Promoter manipulations leading to a $Gata1^{05/X}$ genotype generated female blood cells with either normal Gata1 levels or 5% Gata1 levels. Blood cells that only expressed 5% Gata1 levels developed a myelodysplastic syndrome, suggesting that low Gata1 levels are not sufficient for proper differentiation (Takahashi et al., 1997; Shimizu et al., 2004). Highly abundant Gata1 mutations are prevalent in children with Down syndrome and are associated with acute megakaryoblastic leukemia (Wechsler et al., 2002). A very common mutation is the production of a N-terminally truncated Gata1 protein lacking the transactivation domain. The phenotype of hyperproliferative megakaryoblasts could also be transferred to a transgenic mouse model harboring N-terminally truncated Gata1 (Li et al., 2005; Shimizu et al., 2009).

3.4 The 'PU.1/Gata1 Paradigm': A Stochastic Switch?

3.4.1 Lineage Priming

Northern blot expression analysis of several multipotent hematopoietic progenitor cell lines could show that PU.1, Gata1 and other lineage-affiliated genes like the erythropoietin receptor are already expressed before definite GM and erythroid lineage commitment of multipotent progenitor populations, respectively (Crotta et al., 1990; Heberlein et al., 1992; Cross et al., 1994). It was further observed that upon differentiation, lineage-affiliated genes exhibited a gradual upregulation in 'their' lineages and a gradual downregulation in lineages where they supposedly do not play a role. The co-expression of lineage-specific genes in presumptive uncommitted progenitor cells was called 'lineage priming'. In part, with highly sensitive reverse transcription polymerase chain reaction (RT-PCR) this lineage promiscuity could also be observed at the single-cell level in FDCPmix cells with the notion of heterogeneity between individual cells (Hu et al., 1997). Most abundant markers were the genes myeloperoxidase (MPO) and β-globin, specific for granulocytes and erythrocytes, respectively. One third of all cells investigated expressed both genes. However, the megakaryocyte-erythrocyte driving factor Gata1 was only expressed in 12% of all cells and PU.1 was not part of this study. Multilineage priming at the single cell level could also be shown for multipotent primary cells from the AGM region (Delassus et al., 1999). Again, the markers MPO and β-globin were used and 50% of all cells exhibited coexpression. Many of the multipotent cells also expressed the transcription factors PU.1 and Gata2. After the description of the CLP (Kondo et al., 1997) and the CMP (Akashi et al., 2000) the concept of lineage priming was also investigated in adult primary hematopoietic cells (Miyamoto et al., 2002). HSCs, CMPs, GMPs and MEPs were tested for gene expression by single-cell RT-PCR in two different "myeloid sets", each containing two GM (including PU.1) and two MegE affiliated genes. Individual GMPs and MEPs were mostly classified into strong GM and MegE-profiles, respectively. On the other hand, half of all CMPs tested showed a promiscuous profile, whereas the rest mostly exhibited a GM- or MegE-profile only. HSCs exhibited mostly a primitive profile (none of the 4 markers was expressed), but a significant amount of cells also showed a promiscuous profile. That led the authors of the study to the conclusion that a promiscuous gene profile starts already to emerge at HSC level. This RNA data from primary cells was confirmed on a larger scale with microarray analysis of HSCs, MPPs, CLPs and CMPs (Akashi et al., 2003). It was demonstrated that HSCs express genes of

all hematopoietic lineages and, surprisingly, genes that are not affiliated with the hematopoietic tissue. MPPs, as opposed to CMPs, still expressed myeloid and lymphoid genes. This increasingly restricted gene pattern during differentiation was proposed to support the hypothesis of a more open chromatin structure in more immature cells. Besides looking at the 'classical' differentiation pathway via CLPs and CMPs, lineage priming was also investigated in LMPPs (Mansson et al., 2007). Both large-scale and single-cell analysis confirmed multilineage priming in HSCs and showed downregulation of MegE genes in LMPPs, supporting the hypothesis of their existence. This phenomenon was also shown not to be adult-specific, because a similar LMPP could be identified in embryonic hematopoiesis, too.

3.4.2 The PU.1/Gata1 Network

In general, lineage commitment is considered as a stochastic cell-intrinsic process (Robb, 2007; Graf and Enver, 2009). After the discovery of primed multipotent progenitor cells the experimental observations were transformed into a model that assumed lineage priming in all multipotent cells characterized by low-level and/or sporadic transcription of lineage-affiliated genes associated with an open chromatin structure (Cross and Enver, 1997). Importantly, this model was marked by dynamics and fluctuations of different lineage-affiliated genes within thresholds before the actual commitment. Lineage priming was suggested to be "a ground state from which regulatory networks can develop through positive and negative feedback

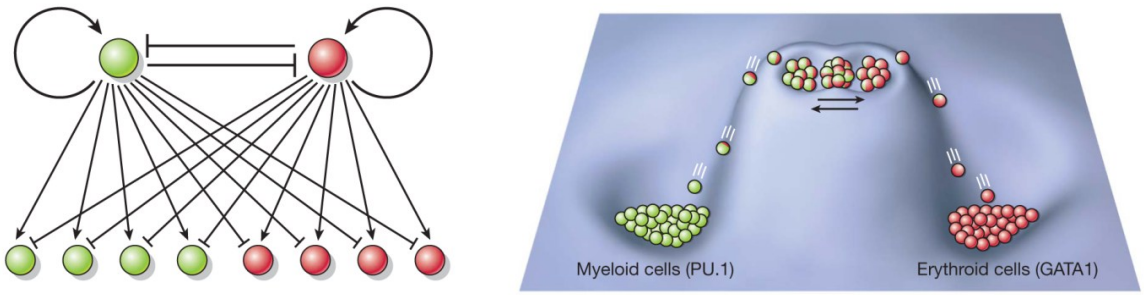


Figure 3-2: The 'PU.1/Gata1 Paradigm'
Green balls represent PU.1, red balls represent Gata1. Summary of activatory and inhibitory functions of PU.1 and Gata1 (left panel). Lineage commitment to one lineage is determined by transcription factor fluctuations and decided by one factor exceeding to other one (right panel). Image adapted from Graf and Enver, 2009.

loops". Along with the description of the mutual inhibitory influence of PU.1 and Gata1 the above mentioned model was specifically extended to the transcription factors themselves (Orkin, 2000; Cantor and Orkin, 2001): Due to the cross-antagonistic interplay of the transcription factors at low levels in uncommitted cells and, hence, the insignificant expression of lineage-specific target genes, cells were proposed to being kept multipotent. Lineage commitment then must be the consequence of stochastic upregulation of one factor above threshold levels or the induced upregulation because of external stimuli. As a result, one transcription factor stoichiometrically exceeds the other one and activates lineages-specific target genes (Figure 3-2). Since the existence of the concept of stochastic gene expression, the PU.1/Gata1 stochastic switch has become an important aspect of adult hematopoiesis suggested also in other reviews, where the ratio between PU.1 and Gata1 is depicted as a balance that governs lineage choice (Graf, 2002) and lineage promiscuity at sub-threshold levels precedes commitment (Miyamoto and Akashi, 2005). Whenever myeloid lineage choice from the HSCs is being discussed, the 'PU.1/Gata1 paradigm' is mentioned as the central element (Laiosa et al., 2006; Iwasaki and Akashi, 2007; Orkin and Zon, 2008).

3.4.3 Modeling Approaches

After the conceptual hypothesis of myeloid lineage determination involving fluctuating transcription factor networks, modeling approaches tried to support this from a mathematical perspective: Under the assumptions of both PU.1 and Gata1 being transcriptional activators for themselves and (probably indirectly) for each other, mutual inhibition mediated by two different complexes and several simplifications (e.g. no post-transcriptional regulation and time-delays), it was concluded that the suggested model of interacting transcription factors is able to explain a switching behavior during lineage commitment (Roeder and Glauche, 2006). Depending on the parameter choice of the model equations, lineage priming could be explained either as a one step process (from lineage priming to commitment) or a two-step-process in which lineage priming is preceded by an initialization. Mechanistically, in one instance increasing transcription rates were enough to induce differentiation, whereas in another instance previous perturbation of the system was needed for differentiation. In another even more simple mathematical model, taking into account mutual cross-inhibition and positive autoregulation only, the three cell states MPP, GM-committed cell and MegE-committed cell were pictured in a three-dimensional attractor landscape as stable valleys

(Huang et al., 2007). Differentiation was suggested to take place either by changes in the ordinary differential equations (which e.g. would correspond to a strong upstream or extracellular signal) or by an intermediate unstable state that is the result of a transformation of a stable valley (MPP) to a hill-top from which cells could tip into both again stable valleys (differentiation), potentially caused by stochastic fluctuations of transcription factors networks. In order to find out which model reflected more the process of myeloid differentiated, transcriptome analysis of the myeloid progenitor cell line FDCPmix was performed. The multipotent cells were differentiated into GM- or MegE-cells by different cytokine treatment and their transcriptome changes were compared during the differentiation process between 0 h and 168 h. Interestingly, during the first 48 h the transcriptomes changed very similarly, only later they developed differently. Thus, the authors' conclusion was that differentiation follows a two stage process including a destabilization event and the actual lineage commitment by either stochastic fluctuations or exposure to upstream signaling events. In another model, Chickarmane and colleagues suggested the existence of an additional factor X, because additional experimental data about PU.1-Gata1 interaction did not support the above mentioned theories of bistable switches (Chickarmane et al., 2009). Suppression of that factor X still could explain the primed PU.1/Gata1 state. They additionally suggested that another downstream antagonistic transcription factor pair (C/EBPα and FOG-1) inherits the information from PU.1 and Gata1 and thus reinforces lineage commitment.

Functional important transcription factor cross-antagonism has also been suggested at other branching points of hematopoiesis. In this context, the transcription factors Fli-1 and EKLF were reported to be involved in regulating lineage choice between the erythroid and megakaryocytic lineage (Starck et al., 2003). Likewise, the ratio of PU.1 and C/EBPα was reported to regulate monocytic versus granulocytic fate choice depending on the cytokine G-CSF (Dahl et al., 2003). Nonetheless, the molecular mechanism behind the suggested regulation remained covert, also because both PU.1 and C/EBPα are highly expressed in monocytic and granulocytic lineages. A deeper analysis of that cell fate decision led to the additional downstream transcription factors Egr-1,2/Nab-2 and Gfi-1 that again were suggested to form a gene regulatory network with bistable behavior (Laslo et al., 2006). More cross-antagonistic transcription factor pairs from many more different branching points of development are summarized in Graf and Enver, 2009.

In a general, metaphorical view, cell differentiation has long been regarded as a ball rolling down a three dimensional landscape with different, irreversible branching points corresponding to individual lineage choices during development (Waddington, 1957). Amazingly, this model is still valid today (Enver et al., 2009). Taking into account the knowledge about stem cells, progenitor cell types and mature cells, mathematical modeling can now shed some light about how the ball rolls down the hill and how its journey can be influenced to take different routes. However, mathematical modeling is always a simplified illustration of the truth and a can only integrate what is known to date. Still, it can be a powerful tool to setup hypothesis that later can be tested in practice in order to proof, contradict or refine the theoretical framework.

3.5 (Dis)proving the 'PU.1/Gata1 Paradigm'

3.5.1 Requirement for Continuous Single-Cell Analysis

The existence of a PU.1/Gata1 stochastic switch is an intuitive and attractive model about the control of lineage choice between the GM- and MegE-lineage, respectively. It is based on the experimental observations about 'lineage priming', the essentiality of both transcription factors for 'their' mature lineages, their reprogramming potential, their positive autoregulatory feedback loops and their potential of mutually inhibiting each other's function by molecular interaction. Mathematical modeling of all that plausible experimental data furthermore supported the hypothesis that transcription factor networks are involved in governing lineage choice. However, direct evidence about PU.1 and Gata1 indeed regulating lineage choice has never been provided.

The phenomenon of 'lineage priming' has only been shown at RNA-level, but never at protein level, first in a multipotent cell line, later in primary cells. So far, 'lineage priming' in terms of double expression of PU.1 and Gata1 at protein levels in multipotent primary cells has never been demonstrated. The convincing experimental data about PU.1/Gata1 molecular interaction was derived from *in vitro* pull-down assays, ectopic overexpressions in cell lines and leukemic cell lines which abnormally express both transcription factors at the same time. To date, no interaction of PU.1 and Gata1 has been shown in unmanipulated primary cells.

HSC differentiation has mostly been investigated at population level of defined subpopulations in order to get enough material for classical biochemical and molecular analysis like Western blotting or microarrays. A big issue about this approach is that presumptive homogeneous phenotypically identical populations can exhibit a large degree of heterogeneity (Huang, 2009). An easy example is shown in Figure 3-3: A classical population analysis (e.g. Western blot) time-course experiment can only display the population average of any marker of interest X which results in increasing strength of bands. Trying to explain that observation, it might be possible that a population changes homogeneously in the expression of the marker X which in a single-cell FACS analysis results in a gradual movement of the histogram along the axis. Or it could be the case that individual cells within the population start their expression of marker X earlier than others which in FACS analysis results in the appearance of two histograms. Thus, the biological conclusions from single-cell analysis can be completely different.

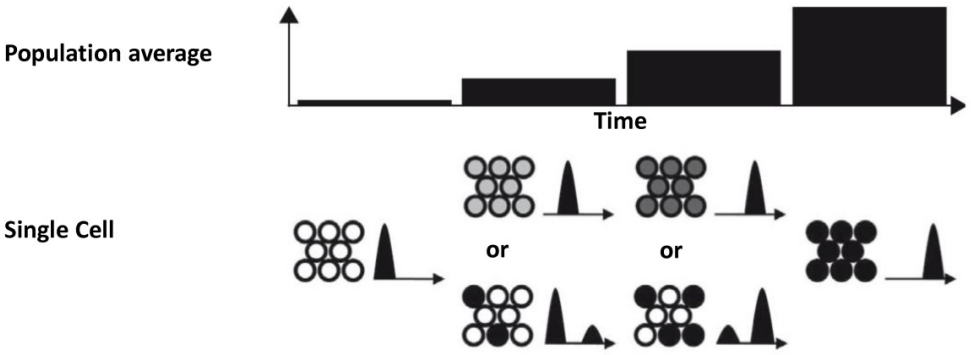


Figure 3-3: Single cell analysis can reveal heterogeneity of cell populations
As opposed to population analysis (e.g. Western blot) which only reports a population average of a cell population, single cell analysis (e.g. FACS analysis) can report heterogeneity among the cells leading ultimately to different biological conclusions. Image adapted from Schroeder, 2011.

The potential of FACS analysis as a very powerful technique can also be demonstrated by the following example: With the multipotent hematopoietic cell line EML it could be shown that cells exhibit large heterogeneity of the stem cell marker Sca-1 (Chang et al., 2008). By sorting and comparing Sca-1^{low} and Sca-1^{high} expressing cells of a clonal population, it was shown that both low and high Sca-1 expressing populations could reconstitute the original distribution within one week of culture. Interestingly, the two different populations had distinct transcriptomes and a different preference for the erythroid (Sca-1^{low}) and the GM (Sca-1^{high}) lineage accompanied by high Gata1 and high PU.1 expression, respectively.

Despite its potential in single-cell analysis, flow cytometry only allows the investigation of one or several markers at one single timepoint. Any potential fluctuations cannot be determined by that method, because the necessary continuous single-cell identity is being lost between timepoints. In two extreme scenarios, heterogeneity of a population can either be stable with very little or no fluctuation over time of a marker of interest or it can be very noisy with very high fluctuations within a short period of time. The only available method to observe any potential fluctuations at the single-cell level is continuous live-cell imaging.

In conclusion, in order to (dis)prove the 'PU.1/Gata1 paradigm' or investigate the role of transcription factor networks in general, one has to look continuously at protein levels in individual living unmanipulated primary cells before, during and after lineage commitment. Otherwise the foundation of the hypothesis, i.e. the behavior of PU.1 and Gata1 expression levels will be missed (Figure 3-4). During recent years, long-term time-lapse imaging has become a very powerful technique to observe HSPCs *in vitro* at the single-cell level (Eilken et al., 2009; Rieger et al., 2009; Schroeder, 2011). A novel approach including suitable hard- and software tools allows the observation of individual cells at high temporal resolution to keep their identity. With the help of the supervised tracking program 'Timm's Tracking Tool' (TTT), pedigree trees over many consecutive generations can be generated. Thus, this technique allows, amongst many other scientific questions, the observation of the

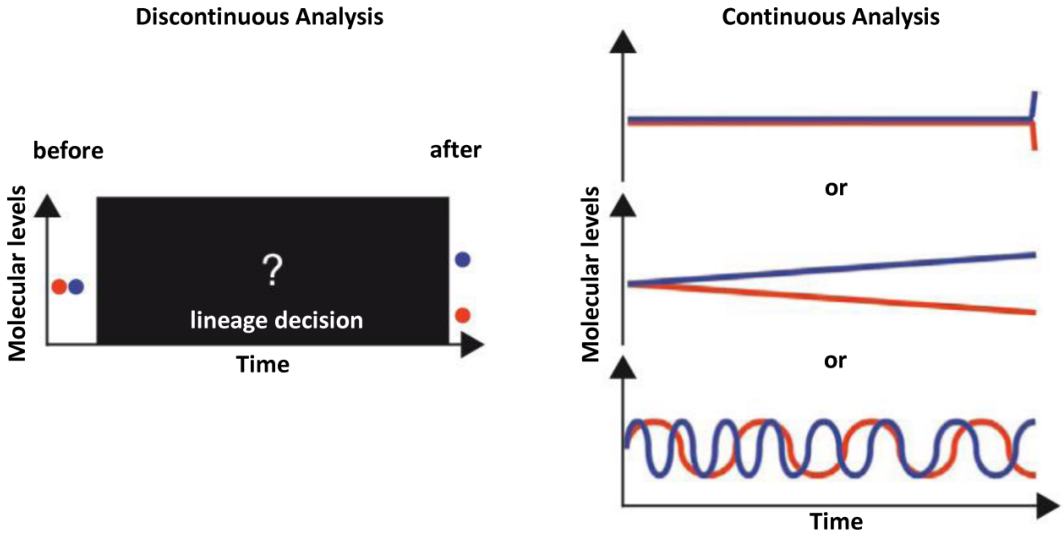


Figure 3-4: Only continuous single-cell analysis can reveal transcription factor behavior
Continuous single cell analysis at protein levels is the only possibility to reveal the behavior of the transcription factors and investigate their potential role in lineage decision. Image adapted from Schroeder, 2011.

differentiation process form individual sorted primary HSCs. In order to quantify PU.1 and Gata1 expression in a non-invasive way, protein levels can be quantified by fluorescence from transgenic reporter mice.

3.5.2 Existing PU.1 and Gata1 Fluorescent Reporter Mice

Fluorescent proteins have been proven to be a very powerful tool which culminated in the award of the Nobel Prize in chemistry to three scientists for the discovery and development of the green fluorescent protein (GFP) in 2008. GFP was first isolated from the jellyfish *Aequorea victoria* (Shimomura et al., 1962). After the cloning of GFP cDNA thirty years later, the application of GFP as a fluorescent reporter has become a standard method in cell biology (Prasher et al., 1992). It can be used as a transgenic marker for ectopic expression, live readouts of promoter activity and fusions to any protein of interest. Application of fluorescent proteins as biosensors by reporting cellular signaling have been developed recently (reviewed in Endele and Schroeder, 2012). Fluorescence can then be followed both in space and time by exposing cells to the correct excitation wavelengths and detecting a proper spectrum of emitting light. Variants of GFP with blue, cyan and yellow fluorescence (named accordingly BFP, CFP and YFP) have been cloned and allowed the simultaneous detection of more than one protein at the same time (summarized in Kremers et al., 2011). Additionally, improved versions of fluorescent proteins have been engineered (enhanced GFP, enhanced YFP) and the spectrum of fluorescent proteins was massively expanded with mutated variants discovered in other species.

In order to follow the expression of PU.1 and Gata1, several transgenic fluorescent reporter mice have been created: One PU.1 reporter line was made by a knock-in of eGFP into exon 1 of the *PU.1* gene which replaced the expression of PU.1 itself (Back et al., 2004). Mice were kept heterozygously, because homozygous PU.1$^{eGFP/eGFP}$ embryos resembled the severe phenotype from PU.1 knock-out mice (Scott et al., 1994; McKercher et al., 1996). Investigation of erythroid progenitors of PU.1$^{eGFP/eGFP}$ embryos suggested low level expression of PU.1 and an important functional role in those cells. Similar effects were observed in adult PU.1$^{wt/eGFP}$ heterozygous animals, too. In another report with the same mouse line PU.1^{GFP} was shown to be expressed in most GMPs, in 50% of CMPs, not in MEPs, in 50% of MPPs and in 25% of HSCs (Back et al., 2005). However, this knock-in approach is not suitable to follow PU.1 levels by fluorescence because of a potential different

half-life of eGFP and PU.1. Another PU.1 reporter mouse strain was created by knocking an IRES (internal ribosomal entry site)-GFP construct into the endogenous *PU.1* locus (Nutt et al., 2005). As a result two different proteins were produced from one mRNA. The expression of the upstream wildtype PU.1 was supposedly unaffected by the downstream insertion which was supported by the notion that homozygous PU.1$^{GFP/GFP}$ mice didn't have a phenotype. PU.1^{GFP} was expressed in all HSCs, CLPs, GMPs, but not in MEPs. Most CMPs also expressed PU.1 and Flt3^{-} CMPs exhibited a bimodal distribution with the GFP^{low} population being highly enriched for committed MegE progenitors. Notably, PU.1 and GFP exhibited a different half-life which made that mouse line not useful to quantify real endogenous PU.1 protein levels. Instead, the supposedly best solution available for following PU.1 was a direct fusion of PU.1 to eYFP by a knock-in of the eYFP reading frame into the endogenous *PU.1* locus, right after the last codon (Kirstetter et al., 2006). No data about correct functionality of PU.1eYFP has been published yet. The question if that mouse line does not have a phenotype and proves to be a reliable reporter of PU.1 expression is part of this thesis.

GFP Reporter mice for Gata1 have also been created (Iwasaki et al., 2005). In those animals the endogenous *Gata1* locus remained untouched and the reporter construct was a randomly integrated transgene containing a promoter construct with all three known DNase hypersensitivity regions, all 6 Gata1 exons and GFP replacing Gata1 Exon 2. The mouse strain was used to identify Gata1^{GFP} eosinophil-committed progenitor cells downstream of GMPs. Additionally, Gata1^{GFP} was reported to be expressed in freshly isolated MEPs, but not in GMPs. In order to quantify endogenous Gata1 levels, this mouse strain was not useful.

In a combined study using both the above mentioned PU.1$^{wt/eGFP}$ (eGFP knock-in into *PU.1* Exon 1) mice (Back et al., 2004) and the transgenic Gata1 GFP reporter mice (Iwasaki et al., 2005) it was suggested that the activation of either PU.1 and Gata1 marks the specification of multipotent progenitor cells into cell with granulocytic/monocytic/lymphoid and granulocytic/monocytic/megakaryocytic/erythroid potential, respectively (Arinobu et al., 2007). This could explain both the classical differentiation pathway through CLPs and CMPs and the existence of LMPPs. Due to the fact that both mouse lines expressed GFP as a reporter it was not possible to look at the simultaneous expression of PU.1 and Gata1 in individual cells. In order to do so, two different colors of fluorescent proteins have to be used.

4 Goals of the Thesis

Fluctuating transcription factor networks are commonly believed to regulate myeloid lineage choice. In this thesis, the widely accepted 'PU.1/Gata1 paradigm' should be investigated experimentally. Additionally to an existing PU.1eYFP knock-in mouse line, another knock-in line that simultaneously allows the readout of Gata1 in living HSPCs should be created. The influence of the transcription factor fusions on hematopoiesis in the newly created PU.1eYFP/Gata1mCHERRY knock-in line should then be investigated in order to evaluate the relevance of the fluorescent reporters. The relative expression levels of PU.1eYFP and Gata1mCHERRY in well-defined HSPCs should be determined and newly discovered myeloid subpopulations should be characterized. Establishing HSC culture conditions should further allow the development of all myeloid lineages within a reasonable amount of time and the choice of proper markers should permit the lineage readout *in vitro*. At last, transcription factor kinetics and dynamics in primary differentiating HSPCs should be determined by time-lapse imaging, single-cell tracking and continuous quantification of absolute molecule numbers in individual cell genealogies.

5 Results

5.1 Generation of Gata1mCHERRY Knock-In

5.1.1 Choice of Fluorescent Protein

There was already a mouse line existing that contained a knock-in of eYFP fused to the reading frame of the endogenous *PU.1* reading frame (Kirstetter et al., 2006). With the goal of generating another fluorescent protein knock-in into the *Gata1* locus that did not overlap in its spectral properties with eYFP, the red fluorescent proteins mCHERRY and tdTOMATO (Shaner et al., 2004) were chosen. Gata1 overexpression has been shown to reprogram primary hematopoietic cells with neutrophil and monocyte potential into cells with an erythroid, eosinophil or basophil phenotype (Heyworth et al., 2002). In order to choose a suitable fusion protein for the Gata1 knock-in similar *in vitro* reprogramming experiments were performed. Both fluorescent protein candidates were each cloned as N- and C-terminal fusion and tested for their reprogramming efficiency of primary LMPPs that have a strong bias for the GM lineage under permissive myeloid culture conditions (Adolfsson et al., 2005). Wildtype Gata1 overexpression in LMPPs from C57BL/6 mice led to an increase of multipotent and MegE colonies in subsequent colony-forming assays from 15.3% to 91.9%

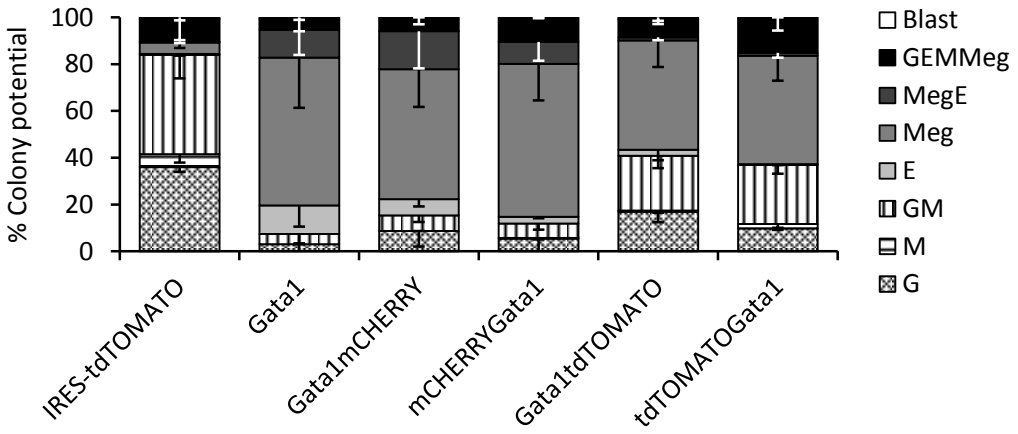


Figure 5-1: Gata1 in a fusion with mCHERRY reprograms LMPPs as efficiently as wildtype Gata1 LMPPs were sorted and infected with lentivirus containing the designated fusion proteins. After 24 h cells were subjected to a colony-forming assay that permits development of all myeloid lineages. Fluorescent colonies were scored after 1 week. Blast = blast colony; GEMMeg = granulocytic, erythroid, monocytic, megakaryocytic; MegE = megakaryocytic-erythroid; Meg = megakaryocytic; E = erythroid; GM = granulocytic-monocytic; M = monocytic; G = granulocytic. Data are mean ± standard deviation (n=3).

(Figure 5-1). Both the N- and C-terminal fusion of Gata1 with mCHERRY showed very similar reprogramming capacity leading to an increase of GEMMeg/MegE colonies to 88.1% and 84.7%, respectively. In constrast, both N- and C-terminal fusions of Gata1 with tdTOMATO were not as efficient in reprogramming and just lead to an increase of 63.0% and 58.4%, respectively. Although mCHERRY was directly fused to Gata1, it did not change the reprogramming capacity in primary wildtype LMPPs, whereas the fusion to the double-sized tandem protein tdTOMATO led to a lower reprogramming capacity. Since PU.1 and Gata1 directly interact with each other at the molecular level (Liew et al., 2006) and fusion proteins of both interactions partners could lead to a malfunction of the PU.1eYFP/Gata1-fluorescent protein complex, the Gata1 overexpression experiments were repeated in LMPPs from the PU.1eYFP mouse. Again, overexpression of Gata1 wildtype and Gata1 as a N- and C-terminal fusion protein with mCHERRY increased GEMMeg/MegE output from 3.7% to 80.8%, 81.7% and 89.1%, respectively, whereas Gata1 as a fusion with tdTOMATO just increased the output to 35.8% and 37.0% (Figure 5-2). Due to a lower reprogramming efficiency of tdTOMATO fusion proteins and in order not to change the expression levels of endogenous Gata1 expression levels by a different sequence at the start codon, it was decided to generate a C-terminal fusion knock-in with mCHERRY.

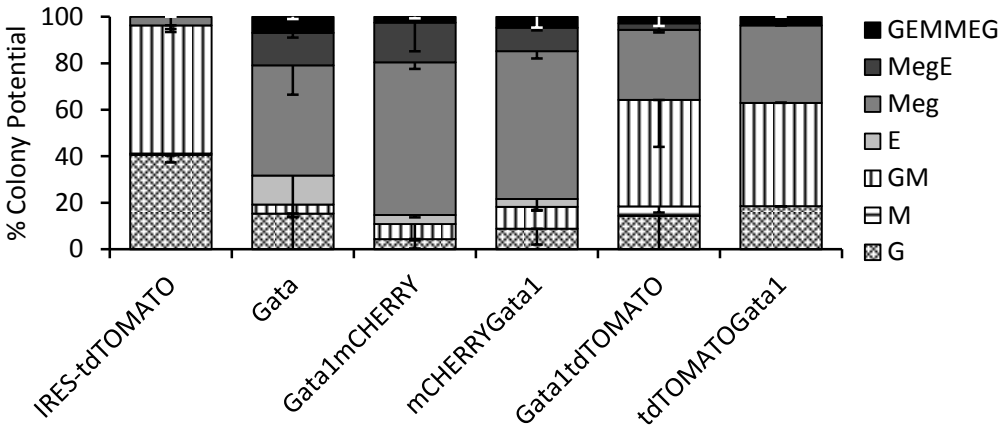


Figure 5-2: Gata1 in a fusion with mCHERRY reprograms PU.1eYFP$^+$ LMPPs as efficiently as wildtype Gata1

LMPPs of a PU.1eYFP mouse were sorted and infected with lentivirus containing the designated fusion proteins. After 24 h cell were subjected to a colony-forming assay that permits development of all myeloid lineages. Fluorescent colonies were scored after 1 week. GEMMeg = granulocytic, erythroid, monocytic, megakaryocytic; MegE = megakaryocytic-erythroid; Meg = megakaryocytic; E = erythroid; GM = granulocytic-monocytic; M = monocytic; G = granulocytic. Data are mean ± standard deviation (n = 2; except tdTOMATOGata1, n = 1).

5.1.2 Gata1mCHERRY Knock-In Strategy

The knock-in strategy was completely designed *in silico* (Figure 5-3): First suitable restriction enzyme sites were chosen that upon successful knock-in of mCHERRY would lead to a significant restriction fragment length polymorphism (RFLP). Next, potential unique genomic sequences were identified in order to detect RFLPs by Southern blotting and a specific radioactively labeled probe. After checking the specificity of putative Southern probes on

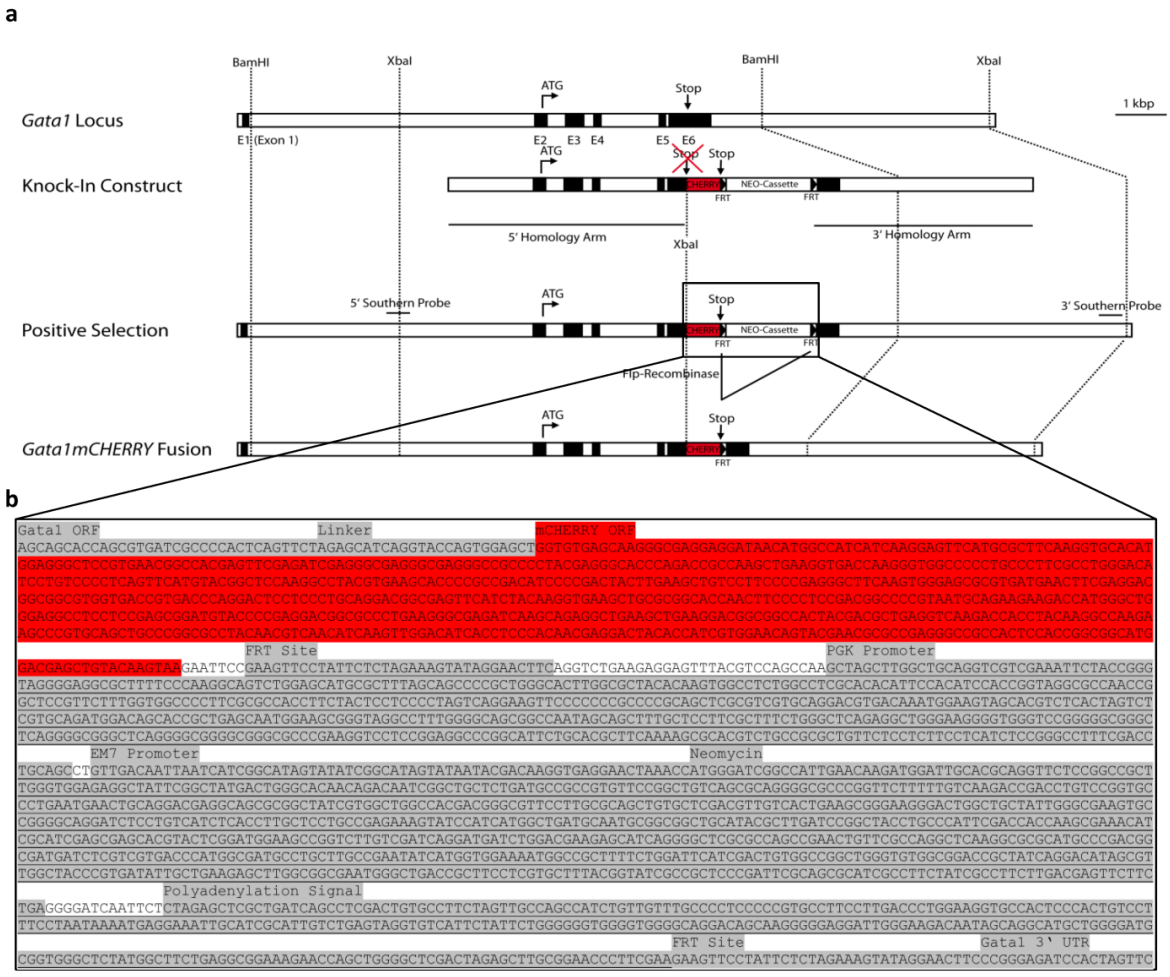


Figure 5-3: Gata1mCHERRY knock-in strategy
(**a**) Overview of Gata1mCHERRY knock-in strategy. BamHI and XbaI were chosen as suitable restriction enzymes in order to generate RFLPs for screening successful knock-ins. Genomic sequences for Southern probes were identified at designated locations. The final knock-in construct contained a 5.0 kilo base pairs (kbp) long 5' homology arm lasting until the last codon of *Gata1*, a short linker sequence (AGAGCATCAGGTACCAGTGGAGCT), the open reading frame (ORF) of mCHERRY, a FRT-flanked Neomycin-resistance cassette (including a eukaryotic and a prokaryoctic promoter and a polyadenylation signal) and a 4.6 kbp long 3' homology arm. (**b**) Complete sequence of endogenous *Gata1* locus after successful knock-in. Underlined bases were removed after Flp-e mediated recombination of FRT-sites.

digested genomic C57BL/6 DNA (data not shown), the knock-in construct was cloned using classical restriction enzyme mediated cloning and recombineering (Yu et al., 2000). After sequencing and confirming the accordance with the *in silico* prediction, 25 µg of the final plasmid were linearized, purified and frozen for subsequent electroporation of ES cells.

5.1.3 Screening of ES Cell Clones

Electroporation and selection of JM8.N4 ES cells (Pettitt et al., 2009) were performed by Dr. Antje Bürger (EUCOMM, Institute of Developmental Genetics, Helmholtz Center Munich). Briefly, ES cells were electroporated, selected in Neomycin containing medium and individual clones were picked and kept separated. ES cell clones were then multiplied and either frozen in liquid nitrogen or subjected to Southern screening by myself. First, ES cell clones were screened for correct integration at their 3'-end. *Gata1* is located on the X chromosome. Hence, positively targeted male ES cell clones resulted in a single band shift from 11.1 kbp to 5.7 kbp (Figure 5-4a). The targeting efficiency was 30%. Positive clones were further checked for correct integration in 5'-direction by another Southern blot. 21% of investigated clones showed the expected band shift from 9.9 kbp to 11.1 kbp (Figure 5-4b). Two double-positive ES cell clones were then selected for thawing and expansion (Dr. Antje Bürger) in order to aggregate them with E2.5 CD1 morulas and transfer into pseudopregnant

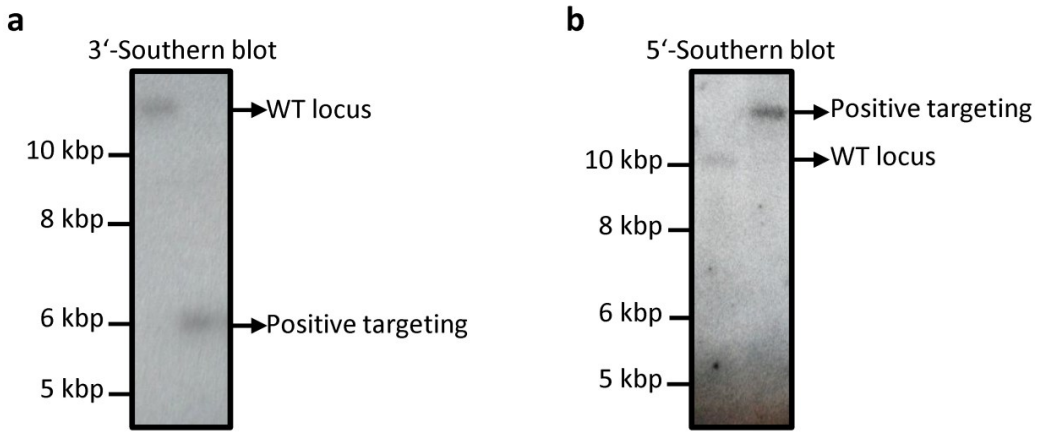


Figure 5-4: Positively targeted ES cell clones were identified by Southern blot
Genomic DNA from ES cell clones was purified, subjected to restriction digest, agarose gelelectrophoresis and Southern blotting. RFLPs were identified by radioactively labeled probes. One representative example is shown for 3'- (**a**) and 5'-Southern blots (**b**), respectively. The targeting efficiency in the 3'-Southern blot was 30% (29 out of 96 clones). Of those 24 were further checked for correct integration at the 5'-end. 21% (5) were positively targeted.

CD1 mice. ES cell aggregation was performed by Heide Oller (Institute of Diabetes and Regeneration Research, Helmholtz Center Munich).

5.1.4 Genotyping of Gata1mCHERRY Mice

The chimerism of the generated animals was estimated by coat color: Genetically pure CD1 mice have a white coat, whereas progeny of the targeted ES cell line (on C57BL/6 background) would have given rise to a dark coat. The chimerism of pups was up to 90% in favor of the targeted ES cell line, making it very likely that the JM8.N4 ES cells also gave rise to the germ line. In fact, progeny of one male animal mated with CD1 female mouse exclusively gave rise to dark colored pups, showing that Gata1mCHERRY^{+} animals are viable and fertile. The male chimera was then used for matings with a Flp-e deleter strain that constitutively expressed Flp-e recombinase (Turan and Bode, 2011) in order to excise the Neomycin cassette whose strong promoter activity could influence the expression levels of Gata1mCHERRY. Progeny was then genotyped by PCR for successful excision of the Neomycin resistance cassette, the presence/absence of the *Flp-e* transgene and the *Gata1* allele(s) (Figure 5-5).

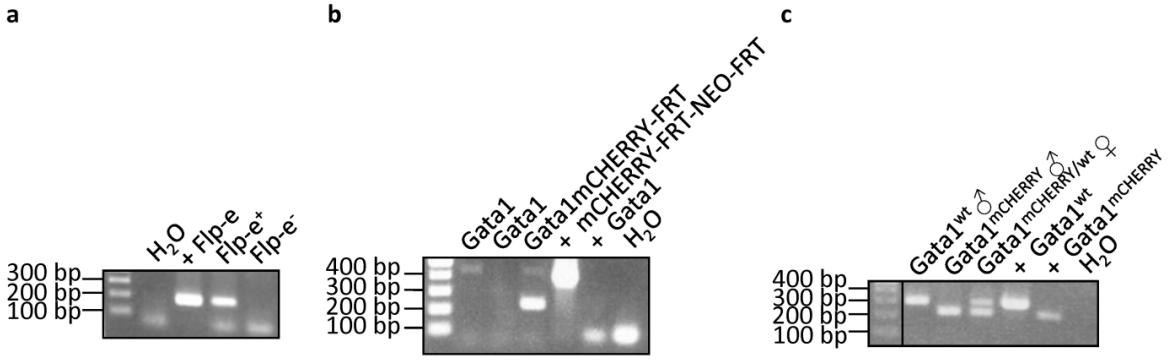


Figure 5-5: Genotyping of Gata1mCHERRY(-FRT-NEO-FRT) mice
Progeny of Gata1mCHERRY-NEO-FRT-NEO mouse mated with a Flp-e deleter strain was genotyped by PCR. (**a**) Presence of the *Flp-e* transgene led to a 176 bp band. (**b**) The successful excision of the Neomycin resistance cassette led to a 223 bp band instead of a 387 bp band in the case of its presence. (**c**) *Gata1* wildtype allele led to a 297 bp band, whereas *Gata1mCHERRY* allele led to 223 bp band. Note that the *Gata1* locus is on the X-Chromosome. '+' in normal writing resembles positive control. H_2O negative controls did not contain template DNA.

5.1.5 Comparison of Fetal Livers from Wildtype and Gata1mCHERRY Mice

Gata1 plays an important role during embryonic hematopoiesis (Pevny et al., 1991; Fujiwara et al., 1996). Since pups are born and survive into adulthood, Gata1 function in a fusion with mCHERRY is obviously not completely abolished. Still, Gata1 function could be impaired and alter normal (embryonic) hematopoiesis. In order to assess as early as possible if Gata1mCHERRY showed normal functionality and, hence, could be used as a reliable reporter, the composition of E14.5 fetal livers of Gata1mCHERRY^{+} pups was compared to C57BL/6 wildtype embryos of the same age. No major changes in the embryonic HSPC pool could be detected, demonstrating normal Gata1 function (Figure 5-6). HSC/MPP, CMP, GMP, MEP and most importantly erythroid progenitor numbers were not obviously altered. Therefore, Gata1mCHERRY knock-in mice were expected to also exhibit normal hematopoiesis in adult mice. The Gata1mCHERRY strain was then mated to the PU.1eYFP knock-in strain (Kirstetter et al., 2006).

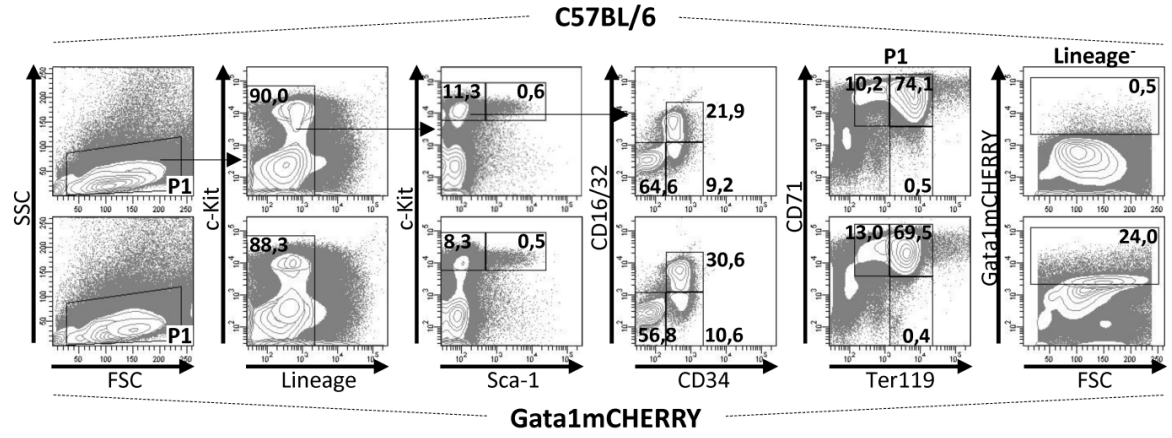


Figure 5-6: Gata1mCHERRY knock-in does not alter embryonic hematopoiesis
Fetal livers (FL) of E14.5 embryos were collected, subjected to Ficoll-density centrifugation, pooled (C57BL/6 7 FLs, Gata1mCHERRY 6 FLs) and analyzed by flow cytometry with typical markers to discriminate HSCs/MPPs (Lin^{-},Sca-1^{+},c-Kit^{+}), CMPs/GMPs/MEPs (Lin^{-},Sca-1^{-},c-Kit^{+}) and erythroid progenitors (CD71$^{+\ or\ -}$/Ter119$^{+\ or\ -}$). Shown are percentages of the parental gate. Upper panels: C57BL/6, lower panels: Gata1mCHERRY.

5.2 Functional Analysis of PU.1eYFP/Gata1mCHERRY Double Knock-In Mice

5.2.1 Progeny of PU.1$^{eYFP/wt}$Gata1$^{mCHERRY/wt}$ Mice

After successful removal of the Neomycin resistance cassette, Flp-e$^-$Gata1mCHERRY$^+$ mice were mated with PU.1eYFP$^+$ mice for the generation of a double knock-in mouse. In order to generate homo-/hemizygous offspring for PU.1eYFP and Gata1mCHERRY, PU.1eYFP heterozygous and Gata1mCHERRY heterozygous female mice were mated with PU.1eYFP heterozygous and Gata1mCHERRY hemizygous male mice. Genomic DNA from pups was genotyped and Mendelian ratios of the offspring were determined (Table 5-1). The analysis showed that homo-/hemizygous animals were born with the same frequency as their heterozygous or wildtype counterparts, pointing out that the functionality of the fusion proteins does not confer any advantage or disadvantage to embryos. Double homo-/hemizygous knock-in mice did not show any increased mortality in adulthood (spontaneous death rate < 2% out of 250 adult observed animals) and reached old ages without any obvious indications of hematopoietic diseases. All subsequent analyses were performed with adult PU.1eYFP homozygous, Gata1mCHERRY hemi-/homozygous animals at 12-18 weeks of age unless otherwise stated.

Table 5-1: Homo-/Hemizygous PU.1eYFP/Gata1mCHERRY mice are born at Mendelian ratios
Genotypes of in total 52 pups from the mating ♀ GATA-1$^{mCHERRY/WT}$ PU.1$^{eYFP/WT}$ x ♂ Gata1^{mCHERRY}Y PU.1$^{eYFP/WT}$. Off. = offspring; Freq. = actual frequency; Exp. = expected frequency.

Gata1 and PU.1	Off.	Freq.	Exp.
Gata1^{WT}Y PU.1$^{WT/WT}$	3	5,8%	6,3%
Gata1^{WT}Y PU.1$^{WT/eYFP}$	5	9,6%	12,5%
Gata1^{WT}Y PU.1$^{eYFP/eYFP}$	1	1,9%	6,3%
Gata1^{mCHERRY}Y PU.1$^{WT/WT}$	3	5,8%	6,3%
Gata1^{mCHERRY}Y PU.1$^{WT/eYFP}$	9	17,3%	12,5%
Gata1^{mCHERRY}Y PU.1$^{eYFP/eYFP}$	3	5,8%	6,3%
Gata1$^{mCHERRY/WT}$ PU.1$^{WT/WT}$	2	3,8%	6,3%
Gata1$^{mCHERRY/WT}$ PU.1$^{eYFP/WT}$	11	21,2%	12,5%
Gata1$^{mCHERRY/WT}$ PU.1$^{eYFP/eYFP}$	5	9,6%	6,3%
Gata1$^{mCHERRY/mCHERRY}$ PU.1$^{WT/WT}$	2	3,8%	6,3%
Gata1$^{mCHERRY/mCHERRY}$ PU.1$^{eYFP/WT}$	5	9,6%	12,5%
Gata1$^{mCHERRY/mCHERRY}$ PU.1$^{eYFP/eYFP}$	3	5,8%	6,3%

Gata1	Off.	Freq.	Exp.
Gata1^{WT}Y	9	17,3%	25,0%
Gata1^{mCHERRY}Y	15	28,8%	25,0%
Gata1$^{mCHERRY/WT}$	18	34,6%	25,0%
Gata1$^{mCHERRY/mCHERRY}$	10	19,2%	25,0%

PU.1	Off.	Freq.	Exp.
PU.1$^{WT/WT}$	10	19,2%	25,0%
PU.1$^{eYFP/WT}$	30	57,7%	50,0%
PU.1$^{eYFP/eYFP}$	12	23,1%	25,0%

5.2.2 Blood Counts of PU.1eYFP/Gata1mCHERRY Mice

PU.1 and Gata1 are central blood-specific transcription factors. Fusions with fluorescent proteins do not lead to increased death rates of animals but in their role as important regulators normal hematopoiesis could be altered. Therefore, peripheral blood counts of PU.1eYFP/Gata1mCHERRY were analyzed and compared to wildtype C57BL/6 mice. 2 individual mice from the PU.1eYPF/Gata1mCHERRY strain had increased white blood cell levels potentially due to an infection at the time of analysis (data not shown). Among the white blood cells the composition of lymphocytes, monocytes and granulocyctes was not altered. The amount of red blood cells was the same between two groups. The amount of hemoglobin in blood or red blood cells and the hematocrit were also comparable. Platelet counts were generally subjected to a high variance. Differences in the size of red blood cells and platelets could not be detected. In total, there was no significant difference between control and knock-in mice, leading to the conclusion that PU.1 and Gata1 as fusion proteins do not change steady-state hematopoiesis at the level of mature blood cells (Table 5-2). Neither did the results of the blood counts point to any anemia, specifically proofing the correct function of the erythroid master regulator Gata1.

Table 5-2: Homo-/Hemizygous PU.1eYFP/Gata1mCHERRY mice have normal bloodcounts
Peripheral blood counts from male C57BL/6 (n=6) and PU.1eYFP/Gata1mCHERRY (n=9) knock-in mice. A one-way multivariate analysis of variance (MANOVA) did not detect any significant difference between C57Bl/6 and PU.1eYFP/Gata1mCHERRY mice (p-value > 0.09).

Value	Unit	C57BL/6	PU.1eYFP/Gata1mCHERRY
White blood cells (wbc)	*/mm^3*	*$2.87 \cdot 10^3 \pm 1.72 \cdot 10^3$*	*$6.17 \cdot 10^3 \pm 3.85 \cdot 10^3$*
% Lymphocytes of wbc	*%*	*78.78 ± 4.27*	*80.79 ± 6.95*
% Monocytes of wbc	*%*	*4.57 ± 0.67*	*4.6 ± 1.39*
% Granulocytes of wbc	*%*	*16.65 ± 3.70*	*14.61 ± 6.08*
% Eosinophils of granulocytes	*%*	*5.8 ± 2.07*	*3.31 ± 2.05*
Red blood cells	*/mm^3*	*$10.11 \cdot 10^6 \pm 0.70 \cdot 10^6$*	*$9.13 \cdot 10^6 \pm 1.71 \cdot 10^6$*
Platelets	*/mm^3*	*$9.77 \cdot 10^3 \pm 6.81 \cdot 10^3$*	*$5.91 \cdot 10^3 \pm 3.74 \cdot 10^3$*
Hemoglobin	*g/dl*	*14.94 ± 1.05*	*13.01 ± 2.18*
Hematocrit	*%*	*52.65 ± 3.82*	*46.98 ± 9.10*
Mean corpuscular volume	*µm^3*	*52.17 ± 0.98*	*51.33 ± 1.00*
Mean corpuscular hemoglobin	*pg*	*14.78 ± 0.20*	*14.31 ± 0.37*
Mean corpuscular hemoglobin concentration	*g/dl*	*28.38 ± 0.73*	*27.84 ± 0.96*
Red cell distribution width	*%*	*12.95 ± 0.23*	*13.49 ± 0.24*
Mean platelet volume	*µm^3*	*6.31 ± 1.03*	*6.84 ± 0.37*

5.2.3 Bone Marrow Analysis of PU.1eYFP/Gata1mCHERRY Mice

In order to exclude that PU.1 and Gata1 as fusion proteins change hematopoiesis at the stem and progenitor cell level, bone marrow of PU.1eYFP/Gata1mCHERRY double knock-in mice was analyzed and compared to C57BL/6 wildtype bone marrow. First, the percentage of the

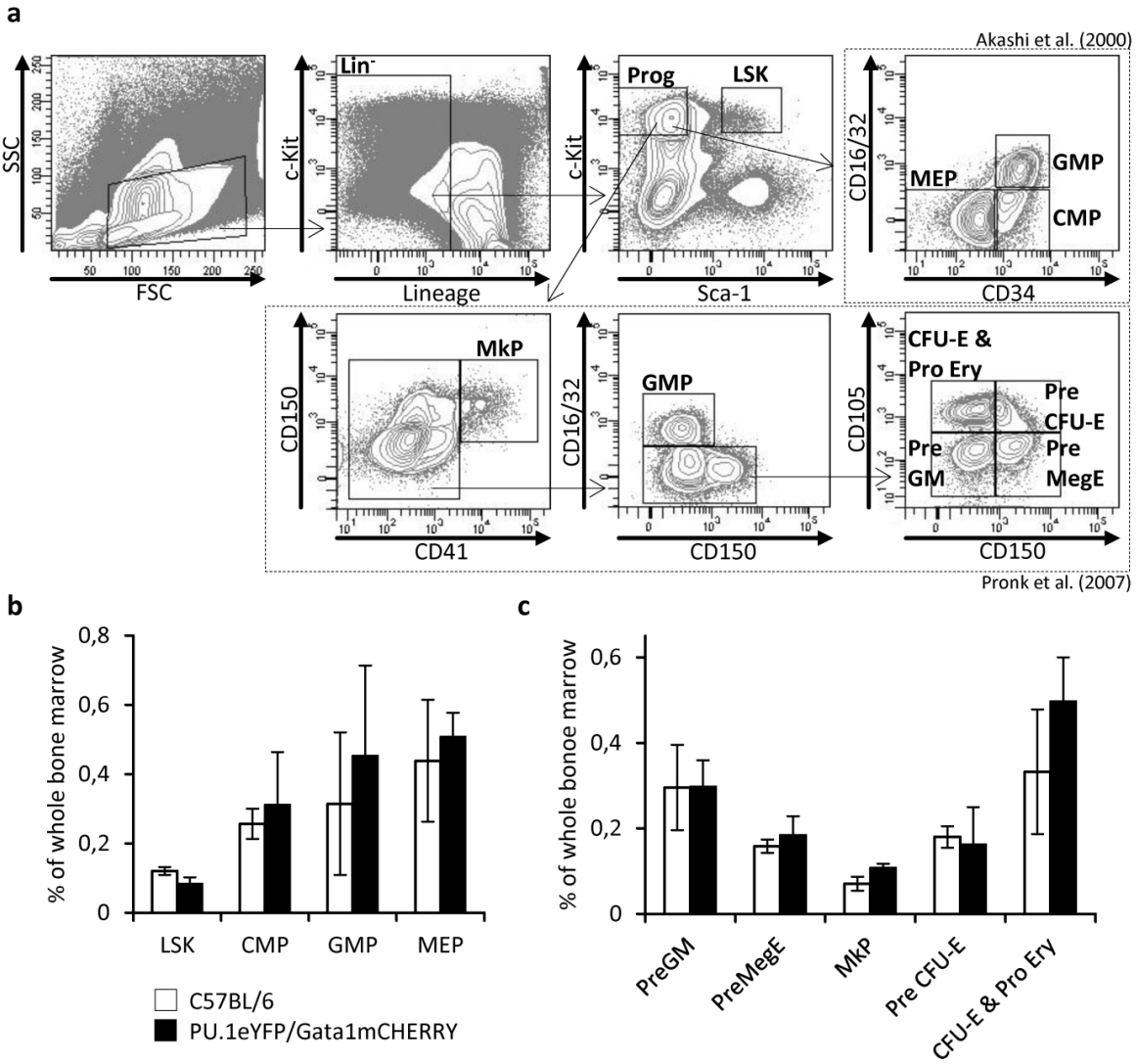


Figure 5-7: Composition of wildtype and PU.1eYFP/Gata1mCHERRY bone marrow is comparable (**a**) FACS gating strategy for the identification of HSPC subpopulations. (**b**) Comparison of C57BL/6 and PU.1eYFP/Gata1mCHERRY bone marrow according to the protocol of Akashi et al. (2000). (**c**) Comparison of C57BL/6 and PU.1eYFP/Gata1mCHERRY bone marrow according to the protocol of Pronk et. al (2007). A one-way multivariate analysis of variance (MANOVA) did not detect any significant difference between C57BL/6 and Pu.1eYFP/Gata1mCHERRY mice for (**b**) and a slight significant difference ($p < 0.03$) for (**c**) due to megakaryocyte progenitors. Shown are mean ± standard deviation ($n = 4$). Lin^- = lineage-negative cells; Prog = myeloid progenitors; MkP = megakaryocyte progenitors; CFU-E = colony-forming unit erythrocyte; Pro Ery = proerythroblasts.

'LSK' gate ($Lin^-Sca\text{-}1^+c\text{-}Kit^+$) (Osawa et al., 1996; Adolfsson et al., 2005; Kiel et al., 2005; Wilson et al., 2008) that contains the HSCs and MPPs, and the percentages of the myeloid progenitor subpopulations 'common myeloid progenitors' (CMPs), GMPs and MEPs were determined (Akashi et al., 2000). Additionally, the myeloid compartment was further analyzed following a more precise protocol to distinguish GM and MegE subpopulations (Pronk et al., 2007). Representative gating strategies for the identification of the respective subpopulations are shown in Figure 5-7a. The PU.1eYFP and Gata1mCHERRY double knock-in did not lead to an altered lineage output between the classical GM- and MegE-compartment (Figure 5-7b). Due to their reprogramming potential an increased activity of PU.1 or Gata1 by a fusion to a fluorescent protein would have been expected to lead to a more prominent lineage output of the respective lineage. A more refined investigation of the myeloid compartment could detect a slight significant difference ($p < 0.03$) between wildtype and PU.1eYFP/Gata1mCHERRY mice (Figure 5-7c), which was not obvious by just comparing individual populations from both mouse strains. By using another staining to analyze the erythroid compartment with the markers CD71 and Ter119 during different erythroblasts stages (Zhang et al., 2003) no significant difference could be detected (Figure 5-8a) again. The levels of T-cells, B-cells and mature GM cells in bone marrow were not altered either (Figure 5-8b). In total, the comparison of wildtype and PU.1eYFP/Gata1-

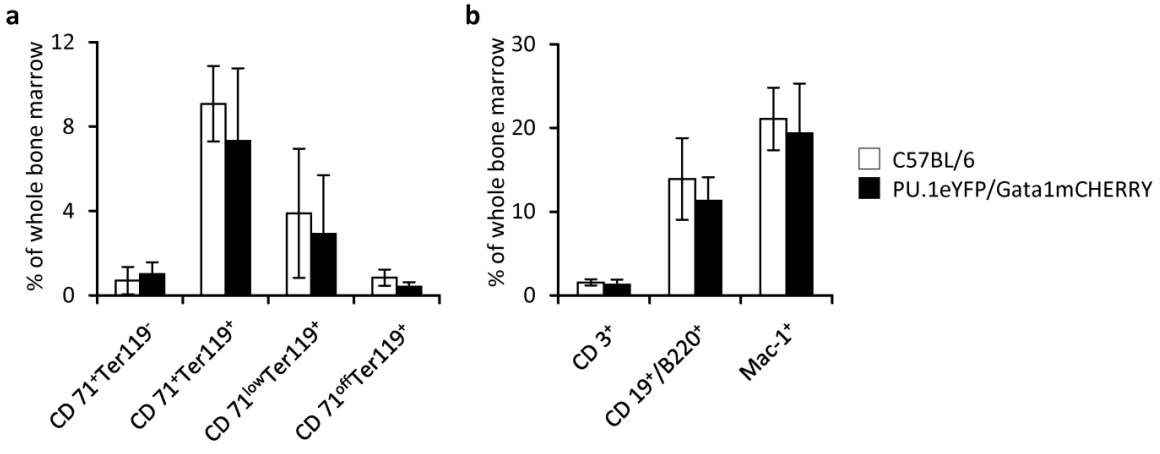


Figure 5-8: PU.1eYFP/Gata1mCHERRY mice have normal levels of erythroblasts, T-cells, B-Cells and mature GM-cells in bone marrow

(**a**) Comparison of bone marrow levels of different erythroblasts stages between wildtype and PU.1eYFP/Gata1mCHERRY mice. (**b**) Comparison of bone marrow levels of T-cells ($CD3^+$), B-cells ($CD19^+/B220^+$) and mature GM cells ($Mac\text{-}1^+$) between wildtype and PU.1eYFP/Gata1mCHERRY mice. Data are mean ± standard deviation (n = 4). A MANOVA for each panel did not detect any significant difference between wildtype and PU.1eYFP/Gata1mCHERRY mice.

mCHERRY bone marrow led to the conclusion that steady-state hematopoiesis at the progenitor cell level is not affected by fusions of the respective transcription factors to fluorescent proteins.

5.2.4 Biochemical Stability of PU.1 and Gata1 and Their Respective Fusions

Next, the biochemical stability of PU.1/PU.1eYFP and Gata1/Gata1mCHERRY, respectively, were compared in order to check if the fluorescent proteins change the half-life of the transcription factors. Freshly sorted PU.1eYFP$^+$ progenitor cells were distributed into equal aliquots and cultured with cycloheximide, a chemical inhibitor of eukaryotic proteinbiosynthesis. After 1, 2 and 3 hours whole cell lysates were prepared and subjected to polyacrylamidgelelectrophoresis (PAGE) and Western blotting. PU.1 and PU.1eYFP were detected using an anti-PU.1 antibody and a secondary antibody catalyzing a chemiluminescent reaction. PU.1eYFP levels of whole cell lysates were always higher than PU.1 levels (Figure 5-9a). If that was the case due to different expression levels or an artefact like different blotting efficiency for PU.1 and PU.1eYFP was indiscernible. Surprisingly, both

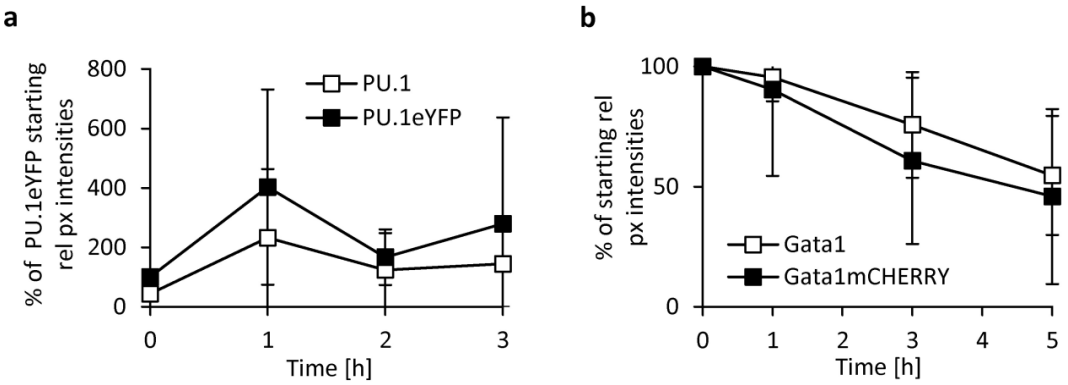


Figure 5-9: Transcription factors and their respective fusions have similar biochemical stabilities (**a**) PU.1$^+$ progenitor cells were sorted, resuspended in medium containing 50 µg/ml cycloheximide, distributed to equal aliquots and whole cell lysates were obtained after 0, 1, 2 and 3 h and stored at -20 °C until PAGE and Western blotting. (**b**) Equal numbers of cells from Gata1mCHERRY$^+$ and wildtype E14.5 fetal livers were mixed, cultured in medium containing 50 µg/ml cycloheximide, distributed to equal aliquots and whole cell lysates were obtained after 0, 1, 3 and 5 h and stored at -20 °C until PAGE and Western blotting. Protein levels were determined using chemiluminescence and quantification of intensity levels on a x-ray film. Shown are mean ± standard deviations (n = 3). A statistical test for differences between paired time-resolved observations did not detect any significant difference between WT and PU.1eYFP/GATA1mCHERRY mice (p>0.16) (Brand et al., 2013).

PU.1 and PU.1eYFP levels increased after 1 h. The general trend of both PU.1 and PU.1eYFP levels was comparable over 3 h of culture.

The very same approach for comparing the stability of Gata1 and Gata1mCHERRY was not possible to do: Since the *Gata1* locus is on the X chromosome, male mice are hemizygous for Gata1(mCHERRY) and cells from female mice exhibit random X chromosome inactivation at uncertain ratios. Compared to PU.1^{+} progenitors, Gata1^{+} progenitors are rarer and more fragile during the sorting procedure (data not shown). Therefore, it was decided to skip sorting and mix equal cell numbers of Gata1mCHERRY^{+} and wildtype E14.5 fetal liver cells which contain a high percentage of Gata1^{+} cells. Due to potential miscounts the expression levels were normalized each to initial Gata1 and Gata1mCHERRY levels at t = 0 h (Figure 5-9b). Again, the general trend of Gata1(mCHERRY) expression levels over 5 hours of culture was comparable and exhibited a constant decline in both cases.

5.2.5 Colony-Forming Potential of PU.1eYFP/Gata1mCHERRY Bone Marrow Cells

As a last functional analysis, the colony-forming potential of freshly isolated PU.1eYFP/ Gata1mCHERRY bone marrow cells was determined and compared to the potential of C57BL/6 wildtype cells. Possible malfunctions of the transcription factor fusions could still be masked *in vivo* through compensatory mechanisms of the organism. In fact, bulk bone marrow from the knock-in mice showed the same myeloid differentiation potential as bone marrow from wildtype mice (Figure 5-10). The assay used did not allow any lymphoid development. In conclusion, the comparison of double homo-/hemizygous PU.1eYFP/Gata1mCHERRY and C57Bl/6 mice in terms of their *in vivo* features like viable progeny, blood counts, bone marrow comparison at the progenitor cell level as well as *in vitro* features like biochemical stability and colony forming potential did not reveal any major differences. Therefore, the novel created PU.1eYFP/Gata1mCHERRY mouse strain was believed to be a truthful reporter for a live read-out of real endogenous transcription factor levels by fluorescence at any given timepoint.

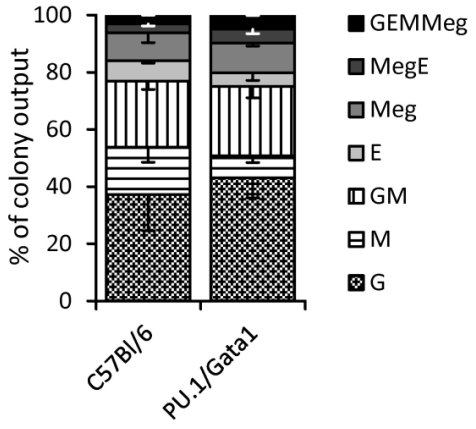


Figure 5-10: *In vitro* colony-forming potential of PU.1eYFP/Gata1mCHERRY cells is not altered $2 \cdot 10^4$ freshly isolated bone marrow cells from PU.1eYFP/Gata1mCHERRY and C57BL/6 mice were subjected to a colony-forming assay using Methocult M3434, containing serum and the cytokines SCF, IL-3, IL-6 and EPO. After 7 days of culture, colonies were scored according to their lineage potential. GEMMeg = granulocytic, erythroid, monocytic, megakaryocytic; MegE = megakaryocytic-erythroid; Meg = megakaryocytic; E = erythroid; GM = granulocytic-monocytic; M = monocytic; G = granulocytic. Data are mean ± standard deviation (n = 3). A Wilcoxon rank sum test did not detect any significant difference between C57Bl/6 and PU.1eYFP/Gata1mCHERRY mice for every population independently (p-value > 0.2).

5.3 Detection of Endogenous Transcription Factor Levels by Fluorescence

5.3.1 Relative Expression Levels of PU.1eYFP and Gata1mCHERRY in HSPCs

After generation of PU.1eYFP/Gata1mCHERRY double knock-ins and the analysis regarding the normal function of the fusion proteins *in vivo* and *in vitro*, relative protein levels of both transcription factors in HSPCs could be quantified by flow cytometry: Negative gates for the separation of signal from background noise were set in such a way that 'positive signal' from individual myeloid progenitor cell populations of wildtype C57BL/6 mice cells did not exceed 1.0% in one experiment (Figure 5-11). 84.7% of multipotent HSCs ($Lin^-Sca\text{-}1^+c\text{-}Kit^+CD150^+$ $CD48^-CD34^-$) express PU.1eYFP, but no distinct Gata1mCHERRY expression could be detected compared to wildtype cells. Of the remaining 15.3% most cells are very close to the positive gate, making it likely that they are in fact PU.1eYFP positive, too, but fluorescence levels were below the sensitivity threshold: PU.1eYFP in HSCs could always be detected when freshly sorted cell were imaged right away (see Figure 5-24 below). Of the 'LSK' gate which besides 10% HSCs contains all different MPPs nearly all cells expressed PU.1eYFP (96.1%), whereas Gata1mCHERRY could only be detected in a very small fraction of cells

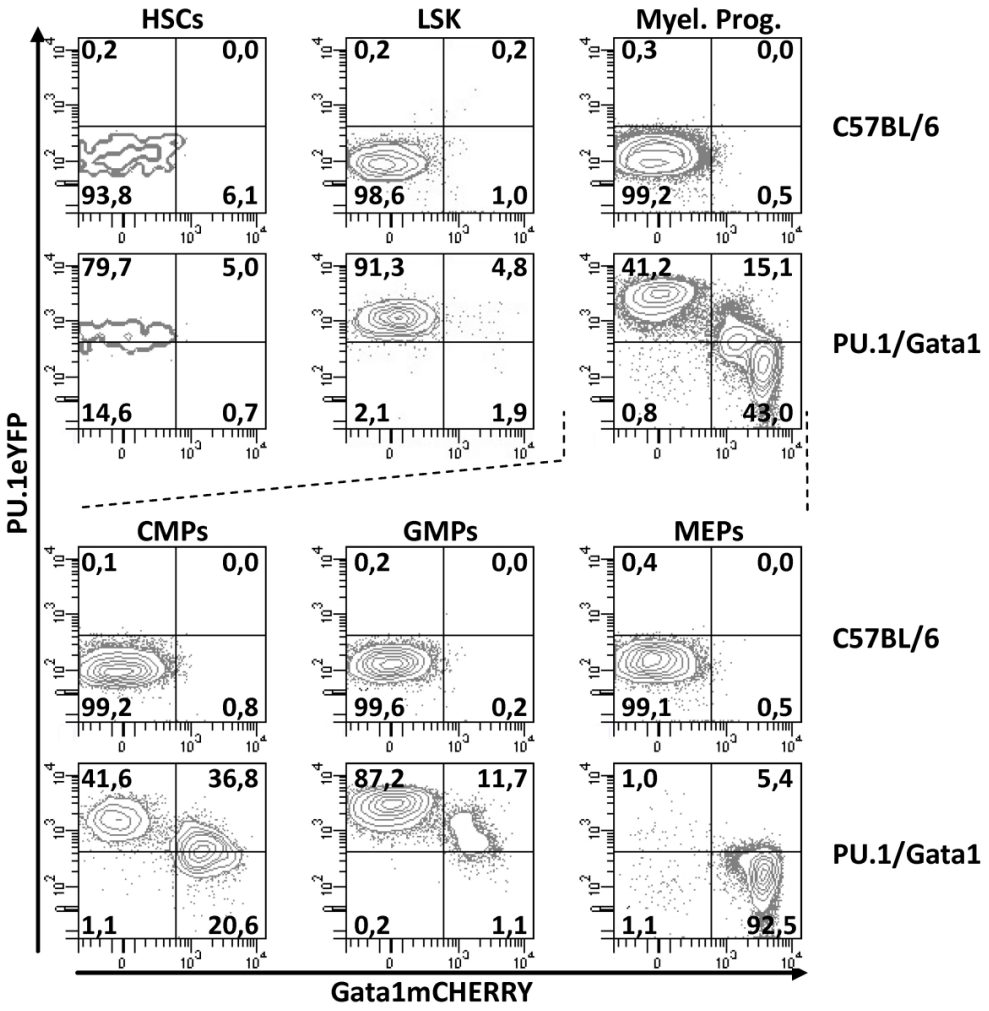


Figure 5-11: Simultaneous detection of PU.1eYFP and Gata1mCHERRY in primary cells reveals new hematopoietic subpopulations
Bone marrow from wildtype C57BL/6 mice and PU.1eYFP/Gata1mCHERRY double knock-in mice was isolated and stained with different antibodies for the identification of HSPC subpopulations. One representative FACS density plots shows PU.1eYFP and Gata1mCHERRY fluorescent intensity for each designated cell type. Numbers in plots show mean abundance (%) in the respective quadrants (n = 4).

(6.7% compared to 1.2% in wildtype cells). After gating for $Lin^{-}Sca\text{-}1^{-}c\text{-}Kit^{+}$ myeloid progenitor cells (Figure 5-7), 4 populations emerged: $PU.1eYFP^{+}Gata1mCHERRY^{-}$, $PU.1eYFP^{mid}Gata1mCHERRY^{mid}$, $PU.1eYFP^{low/-}Gata1mCHERRY^{mid}$ and $PU.1eYFP^{-}Gata1mCHERRY^{high}$ cells. When the antigens CD16/32 and CD34 where used to further subdivide the myeloid progenitors, it turned out that the previously reported CMP is not a homogeneous population, but a mixture of $PU.1eYFP^{+}Gata1mCHERRY^{-}$ and $PU.1eYFP^{low/-}Gata1mCHERRY^{mid}$ cells. GMPs are not homogeneous either, but they consist of mostly $PU.1eYFP^{+}Gata1mCHERRY^{-}$ cells (87.2%) and $PU.1eYFP^{mid}Gata1mCHERRY^{mid}$ cells

(11.7%). MEPs express nearly no PU.1eYFP (93.6%) but mostly only Gata1mCHERRY (92.5%). In conclusion, the simultaneous detection of PU.1eYFP and Gata1mCHERRY allowed the discovery of various hematopoietic subpopulations and the correction of current HSPC purification protocols.

5.3.2 Relative Expression Levels of PU.1eYFP and Gata1mCHERRY in the Myeloid Compartment

Additionally to the expression levels of PU.1eYFP and Gata1mCHERRY in the 'classic' myeloid subpopulations CMPs, GMPs and MEPs (Akashi et al., 2000), it was also interesting to further allocate the transcription factor levels in more refined myeloid subpopulations (Pronk et al., 2007) (for gating see Figure 5-7): Instead of using the marker CD34 for discriminating between the GM- and MegE-lineage, the markers CD150, CD105 and CD41 were used: Like in the case of 'classic' GMPs, it could be confirmed that the 'new' GMP population consists of PU.1eYFP^{high}Gata1mCHERRY^{-} cells and PU.1eYFP^{high}Gata1^{mid} cells

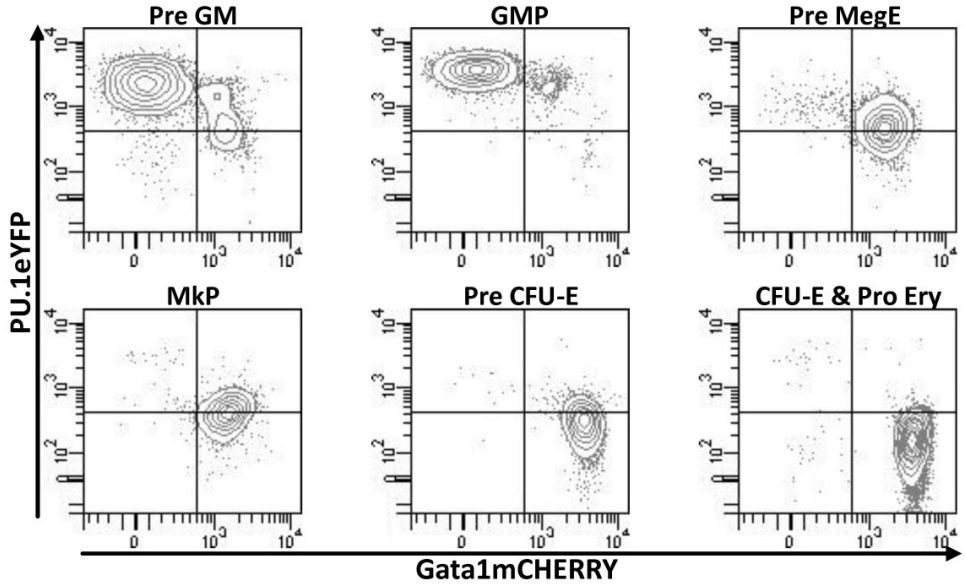


Figure 5-12: PU.1eYFP and Gata1mCHERRY expression levels can be allocated to certain myeloid subpopulations
PreGMs, GMPs, Pre MegEs, MkP (megakaryocytic progenitors), Pre CFU-Es (colony-forming unit erythrocyte), CFU-Es and Pro Erys (proerythroblasts) were identified as described (Pronk et al., 2007). Shown is one representative example of PU.1eYFP and Gata1mCHERRY expression among those subpopulations.

(Figure 5-12). In fact, CD16/32^{+}CD34^{+} GMPs and CD16/32^{+}CD150^{-} GMPs strongly overlap functionally (Pronk et al., 2007) and phenotypically (data now shown).

Both Pre MegEs and megakaryocytic progenitors (MkPs) clearly contribute to the newly discovered PU.1eYFP$^{low/-}$Gata1mCHERRY^{mid} population. The already erythroid committed progenitor populations Pre CFU-Es, CFU-Es and Pro Erys instead account for the PU.1eYFP^{-} Gata1mCHERRY^{high} population. Previously described Pre GMs are a mixture of mostly PU.1eYFP^{high}, PU.1eYFP^{mid}Gata1mCHERRY^{mid} and PU.1eYFP$^{low/-}$Gata1mCHERRY^{mid} cells. That nicely overlaps with the published data about Pre GMs having mostly GMP potential "although rare megakaryocytic and/or erythroid elements were also consistently observed from this cell population" (Pronk et al., 2007). In summary, the majority of myeloid subpopulations had a characteristic profile of the endogenous transcription factors PU.1eYFP and Gata1mCHERRY expression, which will be of use to further discriminate specific cell types more accurately on top of the usage of cell surface antigens.

5.3.3 Colony-Forming Potential of CMP Subpopulations

In order to determine if the newly discovered PU.1eYFP^{high} and Gata1mCHERRY^{mid} subpopulations of CMPs were functionally different, they were subjected to a colony-forming assay under permissive conditions. For that, PU.1eYFP^{high} or Gata1mCHERRY^{mid} CMPs were sorted from fresh bone marrow and seeded at low density in medium under conditions that allow the differentiation of all myeloid lineages in methylcellulose. After one week the lineage outcome of clonogenic colonies was scored by cell and colony morphology. Additionally, benzidine staining was applied to stain for erythroid cell specific hemoglobin. PU.1eYFP^{high} CMPs turned out to have granulocytic and/or monocytic lineage potential only (98.9%) and Gata1mCHERRY^{+} CMPs were highly restricted (96.0%) to the megakaryocytic and/or erythroid lineage (Figure 5-13). This finding has profound influence on the current model of hematopoietic hierarchy. Much published data relied on the existence of true CMPs (Akashi et al., 2000) which are now shown to rather be a mix of GMPs and MEPs.

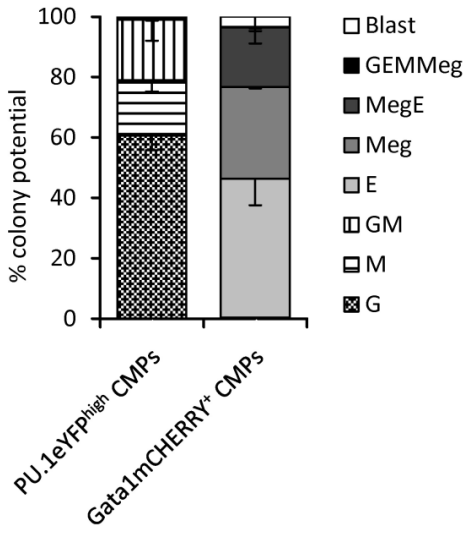


Figure 5-13: CMP subpopulations are restricted to one lineage already
PU.1eYFP^{high} and Gata1mCHERRY^{+} CMPs were sorted and seeded in colony-forming assay medium Methocult M3434 under permissive conditions (serum, SCF, IL-3, IL-6, EPO). Colony potential was scored after one week by cell morphologies, colony morphologies and benzidine staining. Data is normalized to the amount of counted colonies. The mean clonogenicity was 55.5% and 40.8% for PU.1eYPF^{high} and Gata1mCHERRY^{+} CMPs, respectively. Blast = myeloid blasts; GEMMeg = granulocytic, erythroid, monocytic, megakaryocytic; MegE = megakaryocytic-erythroid; Meg = megakaryocytic; E = erythroid; GM = granulocytic-monocytic; M = monocytic; G = granulocytic. Shown are mean ± standard deviation (n = 3).

5.3.4 Colony-Forming Potential of Myeloid Progenitors Discriminated by Only PU.1eYFP and Gata1mCHERRY

Next, it was investigated which lineage potential the newly discovered Lin^{-}Sca-1^{-}c-Kit^{+} PU.1eYFP^{mid} and Gata1mCHERRY^{mid} cell population has. FACS analysis showed that this population has a Pre GM or GMP phenotype. It could be possible that due to the so far missing simultaneous PU.1/Gata1 readout and the relative rareness among the myeloid progenitors this population might be a 'real' CMP with multilineage potential. Additionally, the *in vitro* colony potential of PU.1eYFP^{high}Gata1mCHERRY^{-}, PU.1eYFP$^{low/-}$ Gata1mCHERRY^{mid} and PU.1eYFP^{-}Gata1mCHERRY^{high} cells independently of any other markers than 'LSK' was determined: As expected PU.1eYFP^{high}Gata1mCHERRY^{-} cells exclusively gave rise to granulocytic and/or monocytic colonies (Figure 5-14). PU.1eYFP^{-} Gata1mCHERRY^{high} cells that have phenotypes of committed erythroid cells did not show any colony potential. PU.1eYFP$^{low/-}$Gata1mCHERRY^{mid} cells which strongly overlap with

Pre MegE and MkP cells mostly had megakaryocytic and/or erythroid lineage potential (82.8%). The remaining lineage potential was granulocytic/monocytic (8.4%) and GEMMeg (8.5%). Multilineage and megakaryocytic and/or erythroid potential could also be detected in PU.1eYFP^{mid}Gata1mCHERRY^{mid} cells (each 11.3%). Though, the majority of the cells gave rise to granulocytic and/or monocytic colonies only (76.4%) which is expected from cells that mostly express CD16/32, the marker for GMPs.

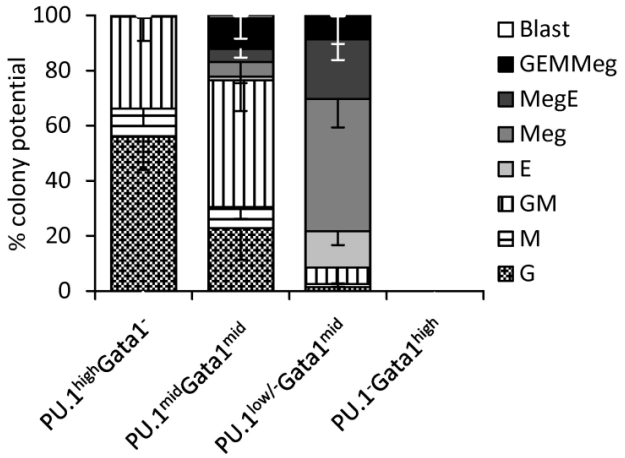


Figure 5-14: PU.1eYFP^{mid}Gata1mCHERRY^{mid} cells have mostly GM-lineage potential
Myeloid progenitor cells were sorted according to Lin$^-$Sca-1$^-$c-Kit$^+$ marker expression and designated fluorescent levels of PU.1eYFP and Gata1mCHERRY and subjected to a colony-forming assay in Methocult M3434 (serum, SCF, IL-3, IL-6, EPO). Colonies were scored after one week. Data are normalized to the amount of counted colonies. PU.1eYFP$^-$Gata1mCHERRY^{high} did not have colony potential anymore. Mean cloning efficiency of the other populations was between 32% and 40%. Blast = myeloid blasts; GEMMeg = granulocytic, erythroid, monocytic, megakaryocytic; MegE = megakaryocytic-erythroid; Meg = megakaryocytic; E = erythroid; GM = granulocytic-monocytic; M = monocytic; G = granulocytic. Shown are mean data ± standard deviations out of 3 independent experiments.

5.3.5 Fold-Changes of PU.1eYFP Levels Between Different HSPC Populations

Unlike Gata1, PU.1 is already expressed in HSCs. During the lineage decision process between the GM- and the MegE-lineage, PU.1 levels either increase or go down to levels below the detection threshold, respectively. In the case of differentiation into MegE-lineages it is not known when and how long before definitive lineage choice Gata1 is being expressed. At least along the 'classical' differentiation pathway via different multipotent progenitor levels and the 'CMP' it cannot be answered from snapshot analysis, if there is a cell population where PU.1 and Gata1 are equally expressed and stochastic fluctuation takes

place. It is also important to mention explicitly that analysis shown so far reflect endogenous transcription factor levels. Those levels do not differ by magnitudes, but only several-fold, which makes it difficult to separate different populations just by their expression levels. In the case of PU.1eYFP levels of yet uncommitted 'LSK' cells compared to committed GMP and MEP populations, the three different populations show overlapping histograms (Figure 5-15). In other words, single snapshot analysis of PU.1eYFP cannot lead to a definitive attribution to a cell population. One also has to bear in mind, that PU.1eYFP levels might change by 2-fold during one single cell-cycle simply due to the doubling of absolute PU.1eYFP numbers for expression maintenance during division. If there is any fluctuating transcriptional behavior it will be impossible to observe that in snapshot analysis. The only way to resolve those issues is to look continuously at the single cell level over time.

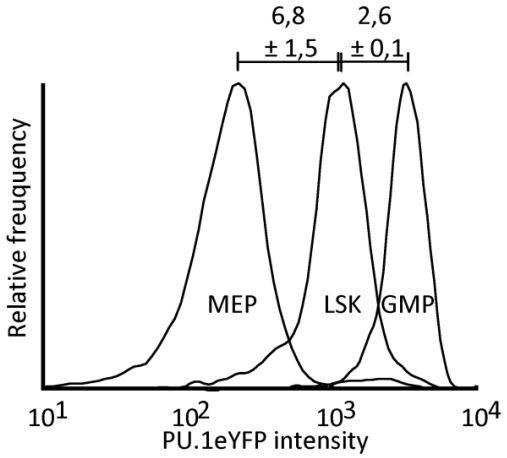


Figure 5-15: PU.1eYFP levels overlap between different cell populations in snapshot analysis
Designated cell populations were gated in FACS and PU.1eYFP distributions are being displayed as histograms. Data shows mean relative fluorescence differences ± standard deviations (n = 4).

5.4 Quantification of Transcription Factor Molecule Levels by Imaging

5.4.1 Imaging of Freshly Sorted Progenitor Cells

Different HSPC levels differ in their endogenous PU.1eYFP and Gata1mCHERRY levels by less than a 10-fold difference. In FACS-analysis different populations show partially overlapping fluorescence levels which leads to the question if those low differences can be resolved by epifluorescence microscopy in order to combine imaging quantification with

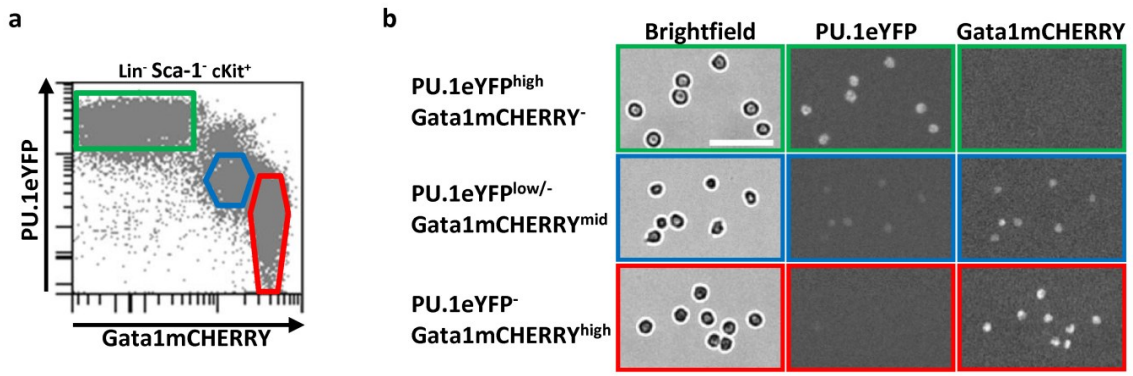


Figure 5-16: Quantification of transcription factor levels in imaging subjectively resembles intensity difference from flow cytometry
(**a**) Designated progenitor were sorted and subsequently imaged by epifluorescence microscopy (**b**). Scale bar is 50 µm.

well established long-term time-lapse microscopy (Eilken et al., 2009; Rieger et al., 2009). Therefore, PU.1eYFP^{high}Gata1mCHERRY^{-}, PU.1eYFP$^{low/-}$Gata1mCHERRY^{mid} and PU.1eYFP^{-}Gata1mCHERRY^{high} cells were sorted and subsequently imaged (Figure 5-16). The differences between PU.1eYFP^{high} and PU.1eYFP$^{low/-}$ cells as well as the difference between Gata1mCHERRY^{mid} and Gata1mCHERRY^{high} cells from flow cytometry could be seen by subjective imaging analysis. When fluorescence levels were tried to be quantified by pixel intensities, it turned out that this cannot happen without a suitable background subtraction: Fluorescence intensity quantifications were massively dependent on microscope setup, lifetime of the light source and exposure time. Additionally, the illumination within one field of view was highly uneven (data not shown). The center part of one picture was generally much better illuminated than the corners and edges. That means that quantification of the very same object either in the bright center or in a dim area would have led to completely different results. Taken into account that transcription factor and, hence, fluorescence levels just differ by on average 2-3 fold between different cycle stages or cell populations, a proper imaging quantification would not have been possible. That is why a background estimation algorithm was developed in cooperation with the group of Prof. Fabian Theis (Institute of Computational Biology, Helmholtz Center Munich) (Schwarzfischer et al., 2011). In brief, the algorithm works in such a way that the background is estimated in every single picture that was subjected to imaging analysis. This allows correct quantification of fluorescent levels independently of the position within the picture and independently of the acquisition time and technical variability. All further analysis described below was performed using that algorithm.

5.4.2 Fluorescent Fold-Changes in Flow Cytometry and Imaging

In order to determine if quantification of pixel intensities in imaging is an equally well and reliable method of fluorescence measurement as flow cytometry, the fold-changes of different hematopoietic progenitor cell populations were compared (Figure 5-17). The fold-changes in the PU.1eYFP fluorescent channel were well comparable between flow cytometry and imaging: The difference between PU.1eYFP^{high} and PU.1eYFP^{mid} expressing cells was 8.3-fold ± 2.4-fold and 8.8-fold ± 1.3-fold, respectively. The difference between PU.1eYFP^{mid} and PU.1eYFP^{-} cells was 2.6-fold ± 0.7-fold and 3.8-fold ± 1.3-fold in flow cytometry and imaging, respectively. Equally good comparison was the result in the Gata1mCHERRY fluorescent channel, too. The difference between Gata1mCHERRY^{high} and Gata1mCHERRY^{mid} cells in flow cytometry and imaging was 2.2-fold ± 0.3-fold and 2.3-fold ± 0.3-fold, respectively, and the difference between Gata1mCHERRY^{mid} and

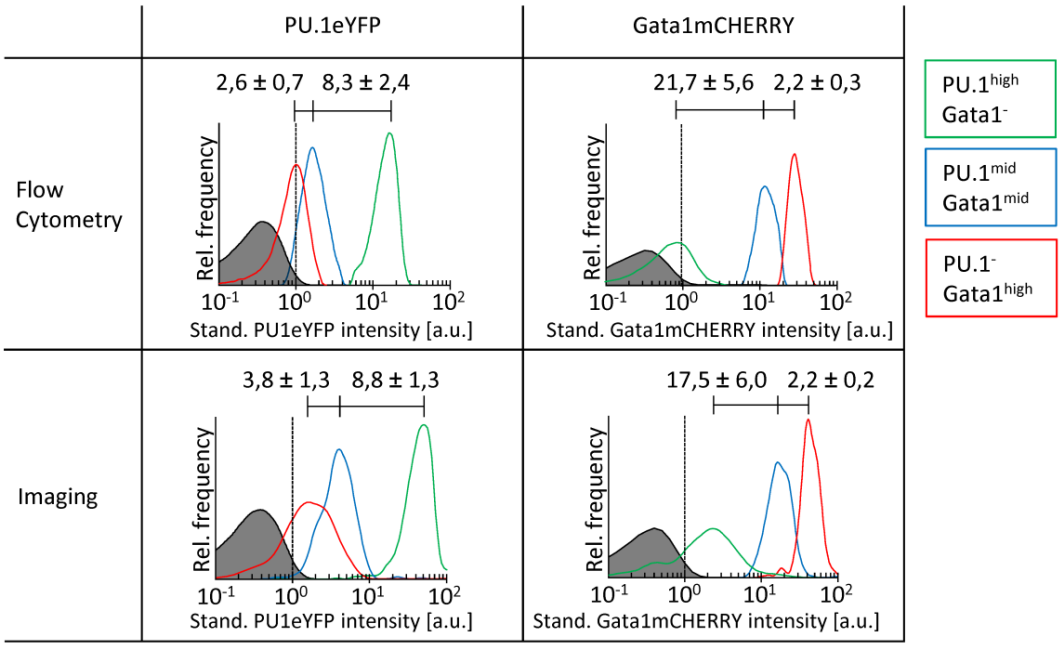


Figure 5-17: Fold-changes of fluorescent intensities are comparable between FACS and imaging Three different Lin^{-}Sca-1^{-}cKit^{+} progenitor cell populations (PU.1eYFP^{high}Gata1mCHERRY^{-}, PU.1eYFP^{mid} Gata1-mCHERRY^{mid}, PU.1eYFP^{-}Gata1mCHERRY^{high}) were freshly sorted and subsequently imaged. Fold-changes of fluorescent intensities of PU.1eYFP and Gata1mCHERRY both in flow cytometry and imaging were quantified and compared. Shaded histograms represent corresponding WT C57BL/6 cell populations (FACS) or typical cell areas on empty background (imaging). The dotted lines represent the 99%-quantile of negative cells (FACS) or background signal (imaging). Shown are means ± standard deviation from three independent experiments. Figure was provided and adapted from Michael Schwarzfischer (Institute of Computational Biology, Helmholtz Center Munich).

Gata1mCHERRY$^-$ cells was 21.7-fold ± 5.6-fold and 17.5-fold ± 6.0-fold, respectively. A difference between flow cytometry and imaging could be found in regard of fluorescent detection of supposedly PU.1eYFP$^-$ and Gata1mCHERRY$^-$ cells. Both 'negative' populations could be better discriminated from real negative signals (C57BL/6 cells or background pixels) in imaging than in flow cytometry. A more detailed analysis can be found in the next chapter.

5.4.3 Sensitivity and Dynamic Range in Flow Cytometry and Imaging

Here, sensitivity is defined as the percentage of cells of a certain population that is above the detection threshold (99% quantile) of 'negative' signal in a fluorescent channel. 'Negative' signal in flow cytometry means signal from wildtype C57BL/6 cells without fluorescent proteins and in imaging it means signal from background pixels from typical cell sizes. The dynamic range is defined as the fold-change (in decibel) between the 99% quantile of a fluorescent signal of a certain population and the detection threshold (99% quantile) of 'negative' signal. As expected, all PU.1eYFP^{high} cells, as determined by flow cytometry, were above the detection threshold in imaging (Figure 5-18). PU.1eYFP^{mid} cells whose distribution in flow cytometry is only partially above the detection threshold (67.8% ± 15.0%) are nearly entirely (99.5% ± 0.4%) above the detection threshold in imaging. Hence, imaging is even more sensitive than flow cytometry for PU.1eYFP signals. In the Gata1mCHERRY channel, all Gata1mCHERRY^{mid} and Gata1mCHERRY^{high} cells are positive both in flow cytometry and imaging. Among Gata1mCHERRY$^-$ cells in flow cytometry which are mostly below the detection threshold (0.9% ± 0.5%) a significant amount of signal could be detected in imaging (38.0% ± 31.6%). Although this seems to be a lot, the impact of this signal detection does not alter the fold-changes between different populations in imaging which are the same in flow cytometry (Figure 5-17). The dynamic range for negative/low, middle and high PU.1eYFP and Gata1mCHERRY signals is significantly higher in imaging compared to flow cytometry. In conclusion, fluorescence quantification by measuring pixel intensities in imaging proves to be an even more sensitive and better resolving method than fluorescence quantification by flow cytometry.

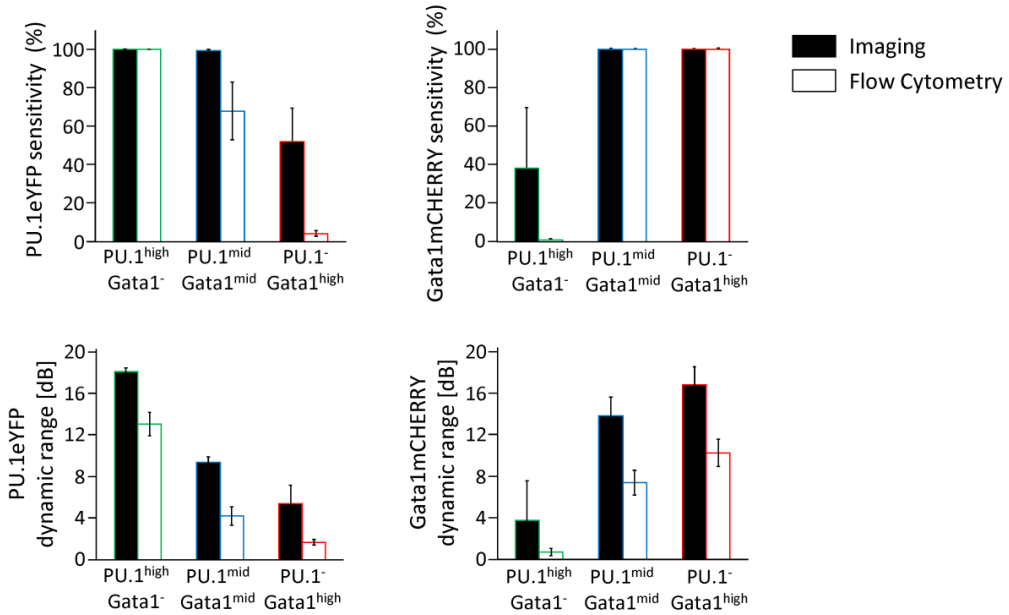


Figure 5-18: Imaging is more sensitive and has a larger dynamic range than flow cytometry The upper panels show the PU.1eYFP (upper left) and Gata1mCHERRY (upper right) sensitivity (% of signals above detection threshold) for different hematopoietic progenitor populations both in flow cytometry and imaging. The lower panels show the dynamic ranges (fold-change between signal and 'negative' signal) for PU.1eYFP (lower left) and Gata1mCHERRY (lower right) for flow cytometry and imaging, respectively. Data are mean ± standard deviation from three independent experiments. Figure was provided and adapted from Michael Schwarzfischer (Insitute of Computational Biology, Helmholtz Center Munich).

5.4.4 Estimation of Transcription Factor Molecule Numbers in Primary Cells

Modeling approaches can build an interesting framework to elucidate potential mechanisms that regulate transcription factor networks. They are partially built on experimental observations which are transformed into differential equations. A central aspect that has been elusive so far is how many transcription factor molecules exist in a single cell. In order to provide that central information (important transcription factor fluctuations would be more likely to happen at low molecule numbers), Western blot analysis with protein lysates from primary cells and a defined amount of recombinant protein were performed (Figure 5-19). Based on the known molecular weight of both the transcription factor fusions and the recombinant protein, molecule numbers per cell population (and per cell) could be calculated. For estimation of PU.1eYFP molecule numbers, PU.1eYFP^{high} expressing progenitor cells and commercially available GFP were chosen. Since eYFP and GFP are very closely related

fluorescent proteins, the respective antibody can detect its epitopes equally well on both molecules. The estimation of Gata1mCHERRY molecule numbers was supposed to happen likewise with recombinant mCHERRY protein, sorted primary cells and an anti-mCHERRY antibody. Gata1mCHERRY^{+} progenitor cells were sorted in the first place but they are more fragile than PU.1eYFP^{+} progenitor cells (data not shown). Additionally, the α-mCHERRY antibody turned out to be a magnitude worse than the α-GFP antibody. Therefore, fetal liver cells instead of bone marrow cells were chosen as source cells. Erythroid and, hence, Gata1mCHERRY^{+} cells are highly abundant in E14.5 fetal livers. By using whole litters of a pregnant mouse, enough cells could be provided and sorting times could be significantly decreased.

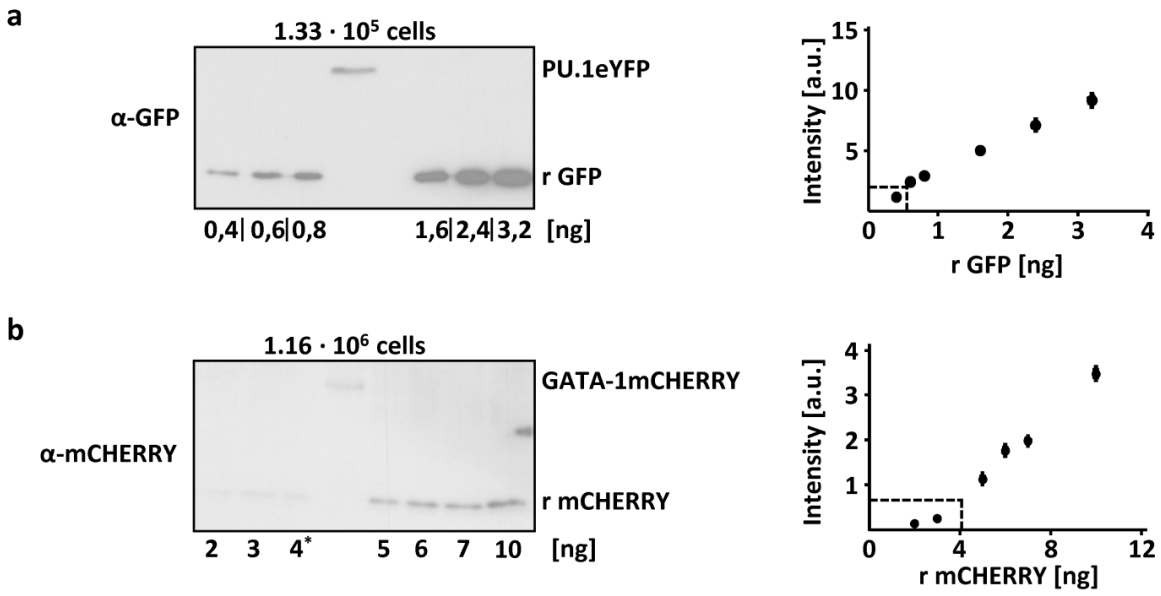


Figure 5-19: Western blot analysis allows estimation of molecule numbers per cell
(**a**) Whole cell lysates from 1.33 · 10^5 PU.1eYFP^{+} progenitor cells (Lin^{-} Sca-1^{-} c-Kit^{+}) were loaded onto a polyacrylamidgel together with defined amounts of recombinant GFP. PU.1eYFP and GFP-bands were detected using an anti-GFP antibody. (**b**) 1.16 · 10^6 Gata1mCHERRY^{+} FL cells were loaded onto a polyacrylamidgel together with defined amounts of recombinant mCHERRY. Signals were quantified using a secondary antibody labeled with horseradish peroxidase and exposure to x-ray films detecting chemiluminescence. Shown is one representative example each from three independent experiments. * Signal omitted due to uncomplete loading of sample.

5.4.5 Exact Calculation of Molecule Numbers in Flow Cytometry and Imaging

With the knowledge of mean transcription factor molecule numbers in freshly sorted primary cells and the relative fold-changes between different populations in flow cytometry, molecule numbers in any other cell population could be calculated (Table 5-3). Gata1mCHERRY^{+} sorted fetal liver cells for Western blot analysis expressed 23.0 ± 9.3 · 10^3 Gata1mCHERRY molecules per cell. PU.1eYFP^{+} progenitor cells expressed on average 43.1 ± 10.6 · 10^3 PU.1eYFP molecules. GMPs which highly overlap with that sorted population expressed 42.7 ± 11.7 · 10^3 molecules. The calculated molecule numbers for PU.1eYFP^{high} and PU.1eYFP^{mid} sorted cells (Figure 5-15) were 47.3 ± 12.9 · 10^3 and 5.9 ± 1.8 · 10^3, respectively. Gata1mCHERRY^{mid} and Gata1mCHERRY^{high} cells expressed 25.5 ± 12.3 · 10^3 and 54.6 ± 23.8 · 10^3 Gata1 molecules, respectively. HSCs expressed only 8.1 · 10^3 ± 2.1 · 10^3 molecules

Table 5-3: Average transcription factor protein numbers per cell
The first two cell populations were used to determine average PU.1eYFP and Gata1mCHERRY molecule numbers per cell, respectively (Figure 5-19). Based on that information and from relative relationships in flow cytometry and imaging, transcription factor molecule numbers were calculated in all other cell populations. Detection thresholds for PU.1eYFP and Gata1mCHERRY are also shown for flow cytometry and imaging. * Imaging data are from freshly sorted PU.1eYFP^{high} or Gata1mCHERRY^{high} expressing cells and were mapped to PU.1eYFP^{mid} and Gata1mCHERRY^{mid} signals, respectively; GMP molecule numbers in imaging were mapped to HSC numbers. Data shows mean ± total error (variability in Western blot and flow cytometry or imaging) from ≥ 3 independent experiments. n.d. = not determined; -- = cell population below detection threshold. Final protein numbers were provided and calculated by Michael Schwarzfischer (Institute of Computational Biology, Helmholtz Center Munich).

Cell Population	PU.1eYFP	Gata1mCHERRY
Flow Cytometry		
PU.1eYFP^{+} LK progenitor	*43.1 ± 10.6 · 10^3*	*n.d.*
E14.5 Gata1mCHERRY^{+} fetal liver	*n.d.*	*23.0 ± 9.8 · 10^3*
PU.1eYFP^{high}Gata1mCHERRY^{-}	*47.3 ± 12.9 · 10^3*	*--*
PU.1eYFP^{mid}Gata1mCHERRY^{mid}	*5.9 ± 1.8 · 10^3*	*25.5 ± 12.3 · 10^3*
PU.1eYFP^{-}Gata1mCHERRY^{high}	*--*	*54.6 ± 23.8 · 10^3*
HSC	*8.1 ± 2.1 · 10^3*	*--*
LSK	*16.4 ± 4.3 · 10^3*	*--*
GMP	*42.7 ± 11.7 · 10^3*	*--*
MEP	*--*	*49.4 ± 21.4 · 10^3*
Detection threshold (C57BL/6 GMPs)	***4.4 ± 1.2 · 10^3***	***6.5 ± 3.8 · 10^3***
Detection threshold (C57BL/6 MEPs)	***4.7 ± 1.5 · 10^3***	***8.4 ± 4.6 · 10^3***
Imaging		
PU.1eYFP^{high}Gata1mCHERRY^{-}	*47.3 ± 12.9 · 10^3**	*--*
PU.1eYFP^{mid}Gata1mCHERRY^{mid}	*5.5 ± 4.4 · 10^3*	*25.6 ± 16.3 · 10^3*
PU.1eYFP^{-}Gata1mCHERRY^{high}	*--*	*54.6 ± 23.8 · 10^3**
HSC	*8.1 ± 2.1 · 10^3**	*--*
GMP	*40.1 ± 4.7 · 10^3*	*--*
Detection threshold	***1.1 ± 2.0 · 10^3***	***1.9 ± 4.4 · 10^3***

per cell in flow cytometry and the whole LSK ($Lin^-Sca\text{-}1^+c\text{-}Kit^+$) population expressed 16.4 ± 4.3 · 10^3 on average. The concrete detection thresholds for PU.1eYFP in flow cytometry were 4.4 ± 1.2 · 10^3 and 4.7 ± 1.5 · 10^3 for C57BL/6 GMPs and MEPs, respectively. Different cell populations seem to have different autofluorescence levels in flow cytometry, especially in the Gata1mCHERRY channel: Detection thresholds were 6.5 ± 3.8 · 10^3 in GMPs and 8.4 ± 4.6 · 10^3 in MEPs, respectively. The previously described better sensitivity in imaging lowered the detection threshold down to 1.1 ± 2.0 · 10^3 for PU.1eYFP and 1.9 ± 4.4 · 10^3 for Gata1mCHERRY. Based on HSC molecule numbers at the time of sorting and the general proof about fluorescent fold-changes being reliable in imaging (Figure 5-17), estimation of molecule numbers could also be applied to below described long-term time-lapse imaging.

5.5 Time-lapse Imaging

5.5.1 Culture Conditions

The culture conditions to follow HSC differentiation in culture should allow 'permissive' development of all myeloid lineages. Cells were cultivated in defined serum-free expansion medium (SFEM) developed specifically for hematopoietic cells. Without addition of cytokines cells did not proliferate and had very low colony-forming potential (data not shown). Therefore, the cytokines SCF, IL-3, IL-6, EPO and TPO were chosen to be added to the medium. SCF is important for HSC survival, proliferation and differentiation. Moreover, it is expressed at least until GMP and MEP stages (Ashman, 1999). IL-3 and IL-6 are general differentiation and proliferation cytokines. EPO is essential for the development of mature erythroid lineages. TPO is similarly important for proper development of the megakaryocytic lineage. Additionally it is involved in HSCs maintenance (Metcalf, 2008). In order to test differentiation potential of freshly sorted HSCs *in vitro*, cells were cultured for 4 – 6 days and analyzed for expression of CD16/32 and Gata1mCHERRY (Figure 5-20a). CD16/32 is a well-established marker for defining fresh GMPs from bone marrow (Akashi et al., 2000). Since Gata1mCHERRY could only be detected in MegE cells (except the $PU.1^{high}Gata1mCHERRY^{mid}$ post-GMP population), it was hypothesized that Gata1mCHERRY expression itself comes along with lineage commitment. During a 7 day culture of HSCs, its expression does not overlap with the expression of Gata1mCHERRY (data not shown). $CD16/32^+$ cells (3.6%) and $Gata1mCHERRY^+$ (7.0%) cells could be

detected on day 4 of culture. The percentage of CD16/32$^+$ cells increased considerably on day 5 and 6 (28.9%; 42.4%), whereas the percentage of Gata1mCHERRY$^+$ cells remained relatively low (2.4%; 2.9%). The colony-forming potential of individual HSCs was analyzed after 5 – 7 days of liquid culture (Figure 5-20b). Most colonies still contained many cells without any obvious lineage commitment (CD16/32 expression or megakaryocyte morphology) at the time of analysis (data not shown). 75.1% of colonies contained GM-committed cells, but no megakaryocytes. On the contrary, 9.7% of all colonies contained megakaryocytes but no CD16/32$^+$ cells. 8.0% of all colonies could be attributed to both GM cells and megakaryocytes and 7.1% did not exhibit any lineage commitment yet. In conclusion, the chosen culture conditions allowed the development of both GM and megakaryocyte lineages. Variation of the culture conditions did not change lineage outcome significantly in colony assays (data now shown). Without further marker usage the potential detection of erythroid cells was not possible.

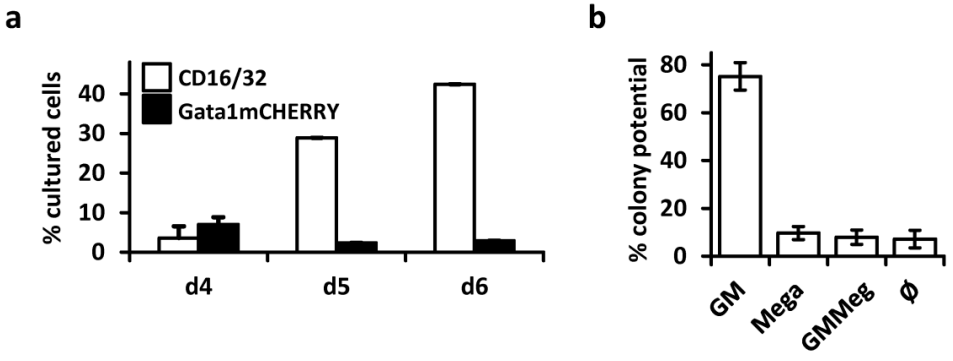


Figure 5-20: Culture condition allows development of both GM- and MegE-lineages
Analysis of HSCs cultured in SFEM and SCF, IL-3, IL-6, EPO, TPO and α-CD16/32-AlexaFluor 647 antibody. (**a**) Flow cytometric population analysis of HSCs cultured for 4 - 6 days (d4, d5, d6). CD16/32 and Gata1mCHERRY expression were interpreted as markers for GM- and MegE-lineages, respectively. Notably, more than 50% of all cultured cells did not express any of both markers on day 6 yet (day 4: mean from n = 3; day 5/6 n = 1). (**b**) Individual HSCs were cultured spatially separated and colony potential was examined after 5 – 7 days of culture. CD16/32 was used as a marker for GM-lineage and megakaryocyte morphology was used as marker for megkaryocytic colony potential. On day 5 - 7, the majority of all colonies still contained cells that could not attributed to any lineage yet. Figure denotes colonies that contain GM cells (GM), megakaryocytes (Mega), GM cells and megakaryocytes (GMMeg) or none of both lineages (Ø). Data are mean from 3 independent experiments with ≥ 79 colonies per experiment.

5.5.2 Markers for Lineage Commitment

In order to investigate if CD16/32 and Gata1mCHERRY expression could be used as truthful markers for the GM and the MegE lineage, respectively, positive cells were sorted on day 4 of culture and subjected to further analysis (Figure 5-21): As a gating control, cultures of C57BL/6 cells were kept in parallel (data not shown). The PU.1eYFP and Gata1mCHERRY negative gate was drawn in such a way that it contained all cells from the culture of wildtype cells without fluorescent proteins (Figure 5-21a). CD16/32 expression was positively correlated with levels of PU.1eYFP and mutually exclusive with Gata1mCHERRY expression. PU.1eYFP$^-$Gata1mCHERRY$^-$, PU.1eYFP$^-$Gata1mCHERRY$^+$ and PU.1eYFP$^+$ CD16/32$^+$ cells were subjected to a colony-forming assay that generally allows differentiation into all myeloid lineages (Figure 5-21b). PU.1eYFP$^-$Gata1mCHERRY$^-$ cells (compared to C57BL/6 culture) mostly (86.7%) gave rise to granulocytic and/or monocytic colonies. 8.4% of sorted cells gave rise to megakaryocytic colonies and 4.9% of cells gave rise to erythroid,

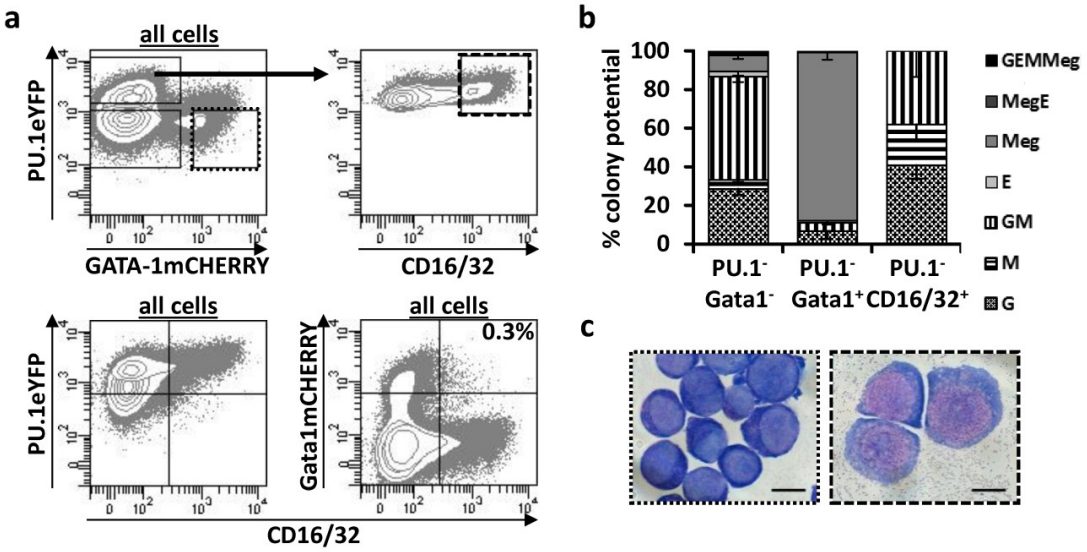


Figure 5-21: CD16/32 and Gata1mCHERRY expression can be used as markers for lineage commitment

(**a**) HSC cultures of PU.1eYFP/Gata1mCHERRY mice with CD16/32-AlexaFluor 647 live antibody staining were subjected to sorting on day 4 of culture. Gates were drawn according to C57BL/6 culture kept in parallel (data not shown). One representative example out of three independent experiments is shown. (**b**) PU.1eYFP$^-$Gata1mCHERRY$^-$ (clonogenicity 47.1%), PU.1eYFP$^-$ Gata1mCHERRY$^+$ (clono-genicity 17.2%) and PU.1eYFP$^+$CD16/32$^+$ cells (clonogenicity 47.0%) were subjected to a colony-forming assay under permissive conditions. GEMMeg = granulocytic, erythroid, monocytic, megakaryocytic; MegE = megakaryocytic-erythroid; Meg = megakaryocytic; E = erythroid; GM = granulocytic-monocytic; M = monocytic; G = granulocytic Data are mean ± standard deviation out of 3 independent experiments. (**c**) PU.1eYFP$^-$Gata1mCHERRY$^+$ (left panel) and PU.1eYFP$^+$CD16/32$^+$cells (right panel) were subjected to a cytospin analysis. Scalebar = 10 µm.

megakaryocytic/erythroid or granulocytic/erythroid/monocytic/megakaryocytic (GEMMeg) colonies. CD16/32 proved to be a very reliably marker for the GM-lineage, since all colonies (100%) derived from CD16/32$^+$ cells could be attributed to that lineage. 86.9% of PU.1eYFP$^-$ Gata1mCHERRY$^+$ cells gave rise to megakaryocytic colonies, the remaining ones were GM-colonies. Colonies with erythroid cells could not be found among those cells, although the medium allowed the development of erythroid colonies (compare to PU.1eYFP$^-$ Gata1mCHERRY$^-$ cells). Additionally, the clonogenicity of PU.1eYFP$^-$Gata1mCHERRY$^+$ cells (17.2%) was clearly lower than from PU.1eYFP$^-$Gata1mCHERRY$^-$ and PU.1eYFP$^+$CD16/32$^+$ cells (47.1% and 47.0%) which led to the assumption that Gata1mCHERRY$^+$ cells from HSC *in vitro* cultures lack erythroid colony-forming potential. Therefore, the sorted cells were subjected to a cytospin which allows classification of primary cells according to histological staining (Figure 5-21c). CD16/32$^+$ cells exhibited a myeloblast phenotype whereas Gata1mCHERRY$^+$ cells were significantly smaller (data not shown) and resembled an erythroblast phenotype. In conclusion, as soon as cells turn on CD16/32 *in vitro* they were committed to the GM-lineage. On the other hand, the vast majority of Gata1mCHERRY$^+$ cells was committed to the MegE-lineage. Hence, both events could be used as markers for live lineage readout in time-lapse imaging.

5.5.3 Generation of Cell Genealogy Trees

Time-lapse imaging and single-cell tracking of freshly sorted HSCs was performed for up to one week. Due to exponential proliferation and culture overgrowth, tracking was not possible longer because the time interval of brightfield pictures every 1.5 min was not sufficient to keep single-cell identity. The longer the movie ran, the higher the probability of losing individual cells was. Fluorescent pictures of PU.1eYFP and Gata1mCHERRY were acquired every 30 min and CD16/32 expression was checked every 4 h. Gata1mCHERRY expression, megakaryocyte morphology and CD16/32 expression were manually annotated (Figure 5-22). No obvious transient Gata1mCHERRY or CD16/32 positive stages could be observed. When either marker was started to be expressed, it remained detectable until the end of imaging. Like in flow cytometric analysis of HSC cultures Gata1mCHERRY and CD16/32 expression were mutually exclusive (compare to Figure 5-21a). Cell death was a very rare event. Only 3.7% of all tracked cells that were not lost in tracking during their life-time died. That included many cells at the end of movies when medium was used up. In general, sister cells

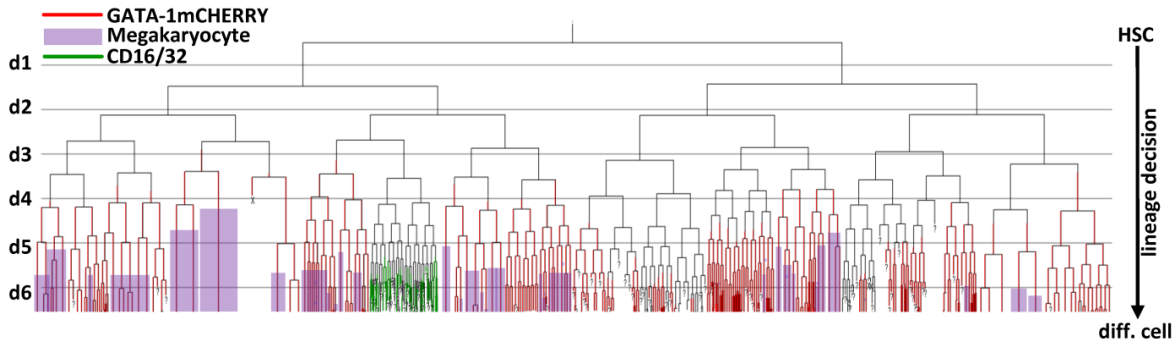


Figure 5-22: Time-lapse Imaging Allows Generation of Cell Genealogy Trees
HSCs were cultured for up to seven days, tracked at the single-cell level and cell genealogy trees were created. Up to 11 consecutive cell divisions could be followed. Manual annotations show Gata1mCHERRY expression (red), megakaryocyte morphology (purple) and CD16/32 expression (green). No transient Gata1mCHERRY or CD16/32 positive stages could be observed. 'X' = cell death; '?' = single cell identity was lost.

and even more distantly related cells exhibited a high degree of synchronism in their cell cycle lengths, but cells with asymmetric fate (like different cell-cycle length or different lineage outcome) could also be observed occasionally. In conclusion, this approach allowed the required quantification of PU.1 and Gata1 protein expression in living HSPCs throughout their differentiation into GM or MegE lineages.

5.5.4 Individual Time Courses of Transcription Factor Dynamics in Primary Cells

Based on the cell genealogies, PU.1eYFP and Gata1mCHERRY expression was quantified in differentiating multipotent HSCs before, during and after lineage choice. Endogenous protein levels with absolute molecule numbers were calculated and displayed (Figure 5-23). After cell divisions, PU.1eYFP and Gata1mCHERRY levels were symmetrically distributed between both daughter cells and signal intensity dropped by half. Asymmetric distribution of transcription factors could not be observed. In accordance with flow cytometric analysis, all HSCs and several daughter cell generations expressed PU.1eYFP. Cells were considered Gata1mCHERRY-positive when 5 consecutive timepoints were twice above the detection threshold of 1900 molecules. In the case of cells differentiating into the GM-lineage, Gata1mCHERRY was never expressed before cells were definitely committed (CD16/32 expression) (Figure 5-23, upper panel). In other words, multipotent HSCs and their progeny did not pass through a PU.1eYFP/Gata1mCHERRY double positive state during GM-lineage choice, which would be the prerequisite for PU.1/Gata1 interdependent fluctuations. As soon

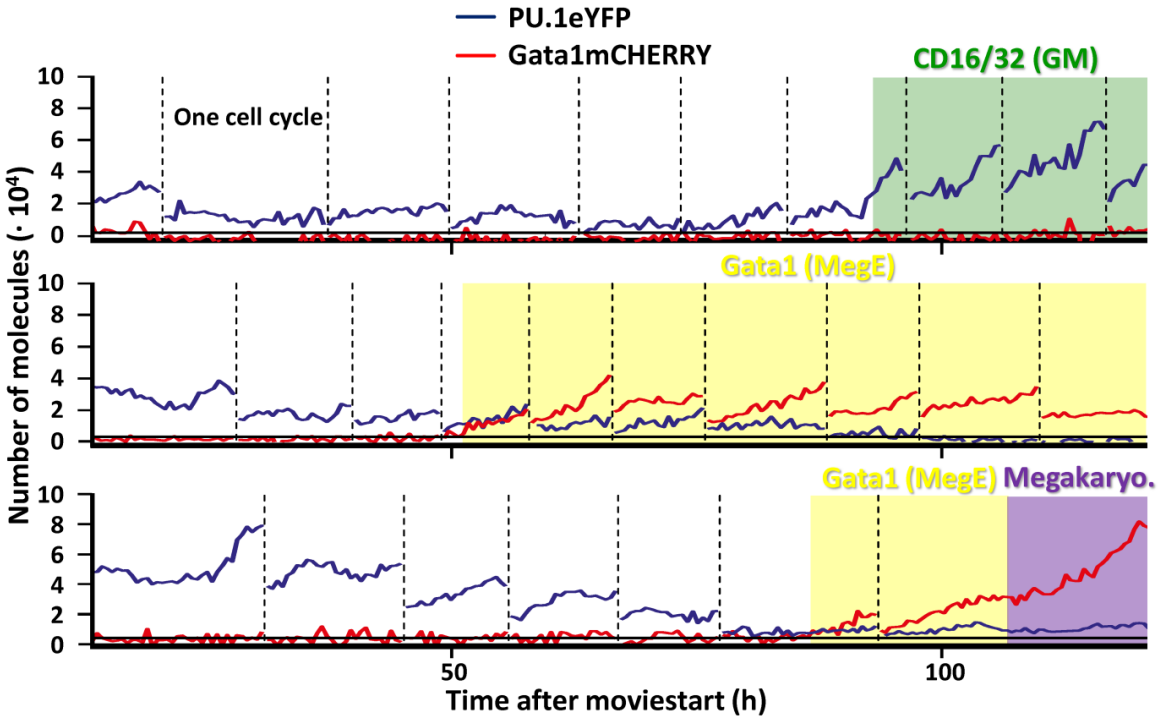


Figure 5-23: PU.1/Gata1 interdependent fluctuations do not precede GM/MegE lineage choice Panels show continuous PU.1eYFP (blue) and Gata1mCHERRY (red) protein number quantifications in consecutive daughter cells in a time interval of 30 min. Upper panel shows cell traces that differentiate into GM-cells (CD16/32^{+}, green). Middle panel show cells traces of a cell that will commit to the MegE-lineage (yellow) without further morphological features (Gata1mCHERRY wins). Lower panel shows a cell genealogy that differentiates into a megakaryocyte (Gata1mCHERRY wins, yellow, and megakaryocyte morphology, purple). All individual timepoints were manually inspected. If quantification was not possible at a certain timepoint, it was excluded from analysis. Horizontal black lines in graphs denote Gata1mCHERRY detection threshold of 2100 molecules. Vertical dashed lines denote cell divisions. For clarity, only one daughter cell is shown per generation, respectively.

as Gata1mCHERRY could be detected, cells and their progeny remained Gata1mCHERRY positive until the end of the movie and PU.1eYFP dropped below threshold levels or remained at low levels (Figure 5 23, middle and lower panel). Gata1mCHERRY expression onsets could be as early as 50 h after moviestart or anytime later until the end of the movie. A formal double-positive state with even stoichiometrically balanced PU.1eYFP/Gata1mCHERRY levels could be observed occasionally, but neither interdependent fluctuations nor another outcome than further Gata1mCHERRY upregulation and PU.1eYFP downregulation could be detected. This double-positive state was also observed in flow cytometric analysis of freshly isolated bone marrow: PU.1eYFPlow/-Gata1mCHERRYmid cells were already committed to the megakaryocytic-erythroid lineage (Figure 5-14). Hence, the double-positive state of differentiating HSCs and their progeny did not have any functional relevance. Furthermore, PU.1eYFP was already downregulated (see

also Figure 5-24 below) many generations before the actual Gata1mCHERRY onset which shows that Gata1mCHERRY was not the initial cause of PU.1eYFP downregulation (Figure 5-23, middle and lower panel).

5.5.5 Summary of Time Courses of Transcription Factor Dynamics in Primary Cells

The above mentioned examples only depict PU.1 and Gata1 expression in individual differentiating cells. To illustrate their expression behavior in a large number of differentiating cells, all single-cell time courses are summarized in Figure 5-24. GM-committed cells (normalized to onset CD16/32^{+} to t = 0 h) emerge as a population by constant upregulation of PU.1eYFP without an intermediate PU.1eYFP/Gata1mCHERRY double-positive state. The population of committed MegE-cells (Gata1mCHERRY^{+}) emerges out of the starting population first by

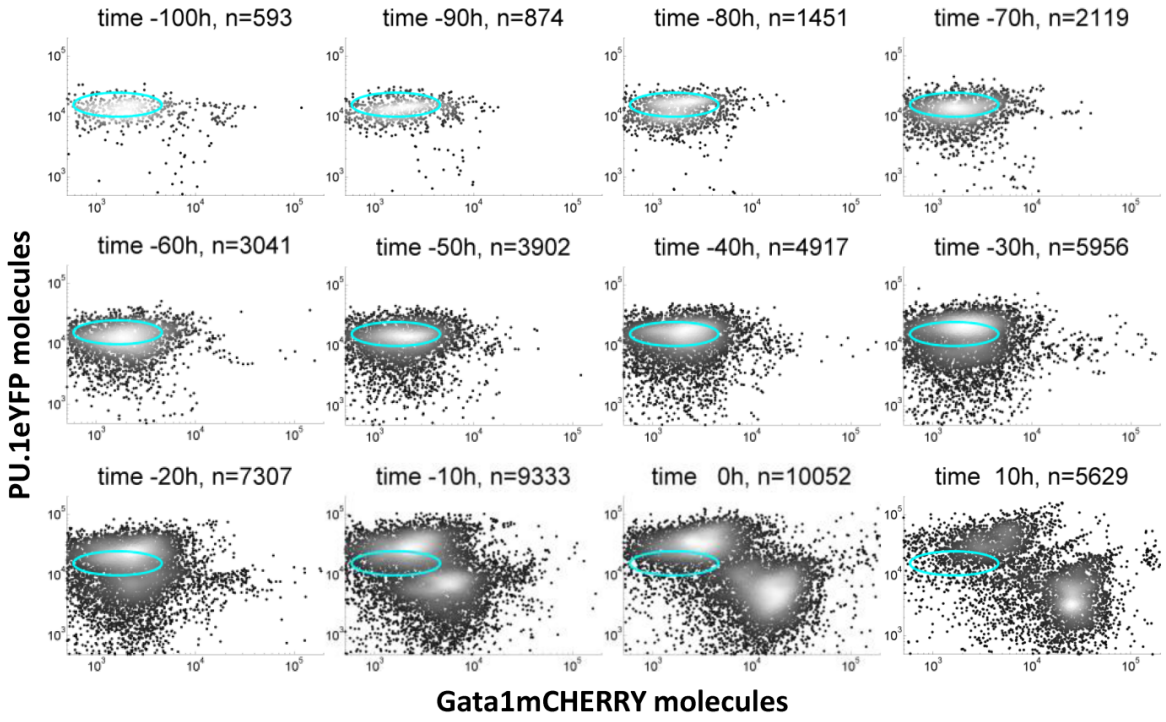


Figure 5-24: Lineage commitment in primary HSPCs takes place without passing through a PU.1eYFP/Gata1mCHERRY double positive state
Summary of all tracked and manually inspected single-cell time courses from continuous time-lapse imaging. Timepoints denote time relative to lineage commitment t = 0 h (Gata1mCHERRY or CD16/32 onset). The light blue oval indicates the starting population of PU.1eYFP^{low} HSCs/early MPPs. n-numbers show datapoints that contributed to the respective panel. Summary of 102 individual HSC colonies, with 167 cells in the first generation (after division of starting cell) with in total 349 Gata1mCHERRY onsets and 348 CD16/32 onsets from 3 independent experiments. Figure provided from Michael Schwarzfischer (Institute of Computational Biology, Helmholtz Center Munich).

downregulation of PU.1eYFP followed by upregulation of Gata1mCHERRY. The summary of all time-courses fit to the representative examples described above (Figure 5-23).

5.5.6 Stoichiometry between PU.1eYFP and Gata1mCHERRY

PU.1eYFP levels were always quantifiable until they reached the detection limit. In that case cells had already upregulated Gata1mCHERRY and were committed to the MegE-lineage, so detection issues for PU.1eYFP are not relevant for the question of myeloid choice. With the current experimental setup and sensitivity, no PU.1eYFP/Gata1mCHERRY double-positive stage could be observed in cells that differentiated into the GM-lineage. Therefore, the minimal molecule numbers of PU.1eYFP have to be taken into account and have to be compared to the detection limit of Gata1mCHERRY. 44.1% of all GM-traces increased their PU.1eYFP levels during differentiation compared to initial values in freshly sorted multipotent HSCs (Figure 5-25a). 16.5% of all cells GM-traces never dropped below 10^4 PU.1eYFP molecules. Surprisingly, 53.1% of all GM-traces 'dropped' their PU.1eYFP levels below initial values, before PU.1eYFP then was upregulated and CD16/32 onset could be detected later. Although lower than HSC starting values, 19.7% of all GM-traces never dropped below 10^4 molecules. The biological meaning of this phenomenon is not clear. Independently of 'dropping' behavior, only 1.8% of all GM-traces fell below 2000 molecules. All remaining GM-traces remain above 2000 molecules and, hence, the vast majority of GM-traces stoichiometrically outbalanced any potential Gata1mCHERRY molecules below the detection threshold of 1900 molecules all the time.

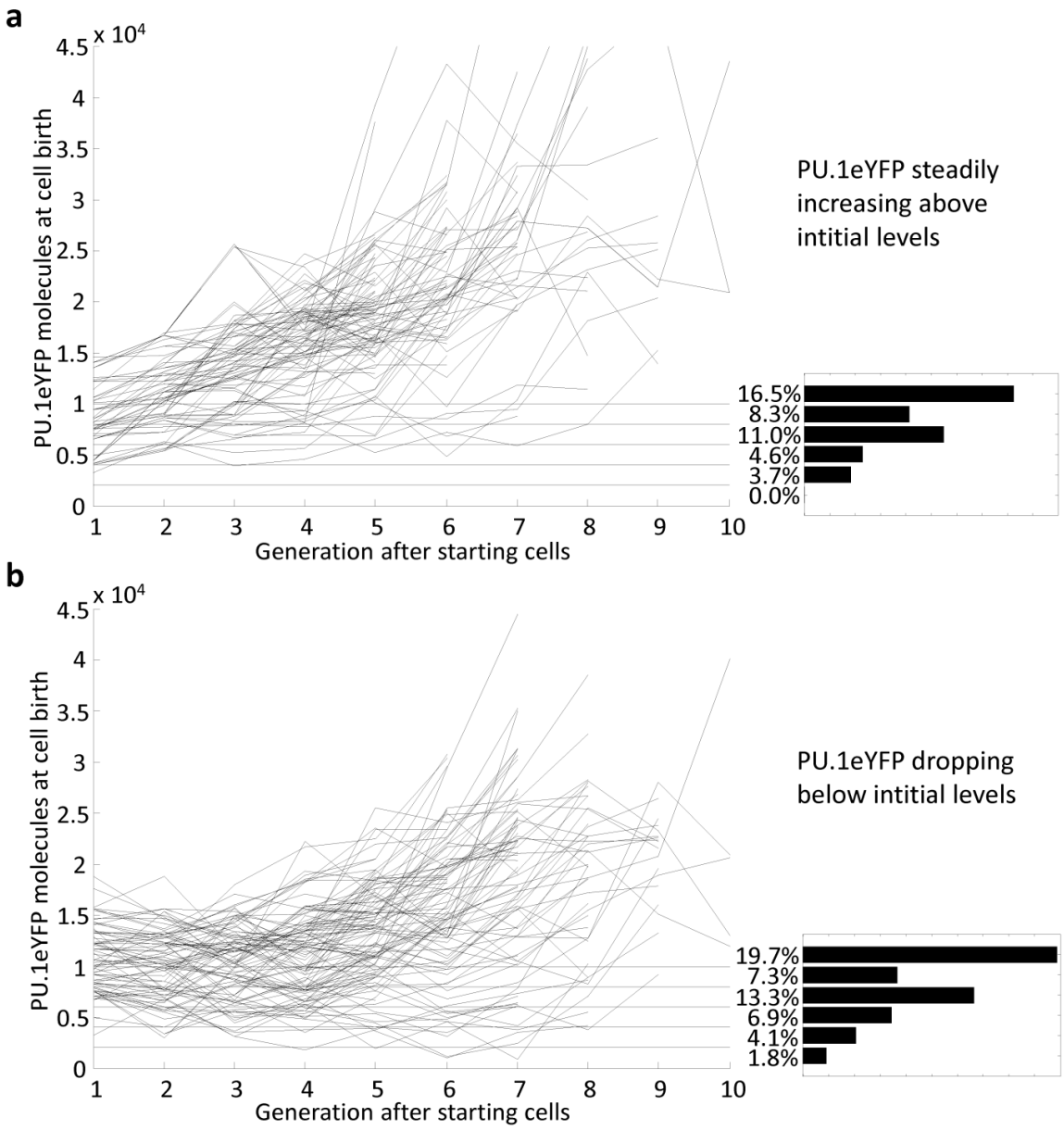
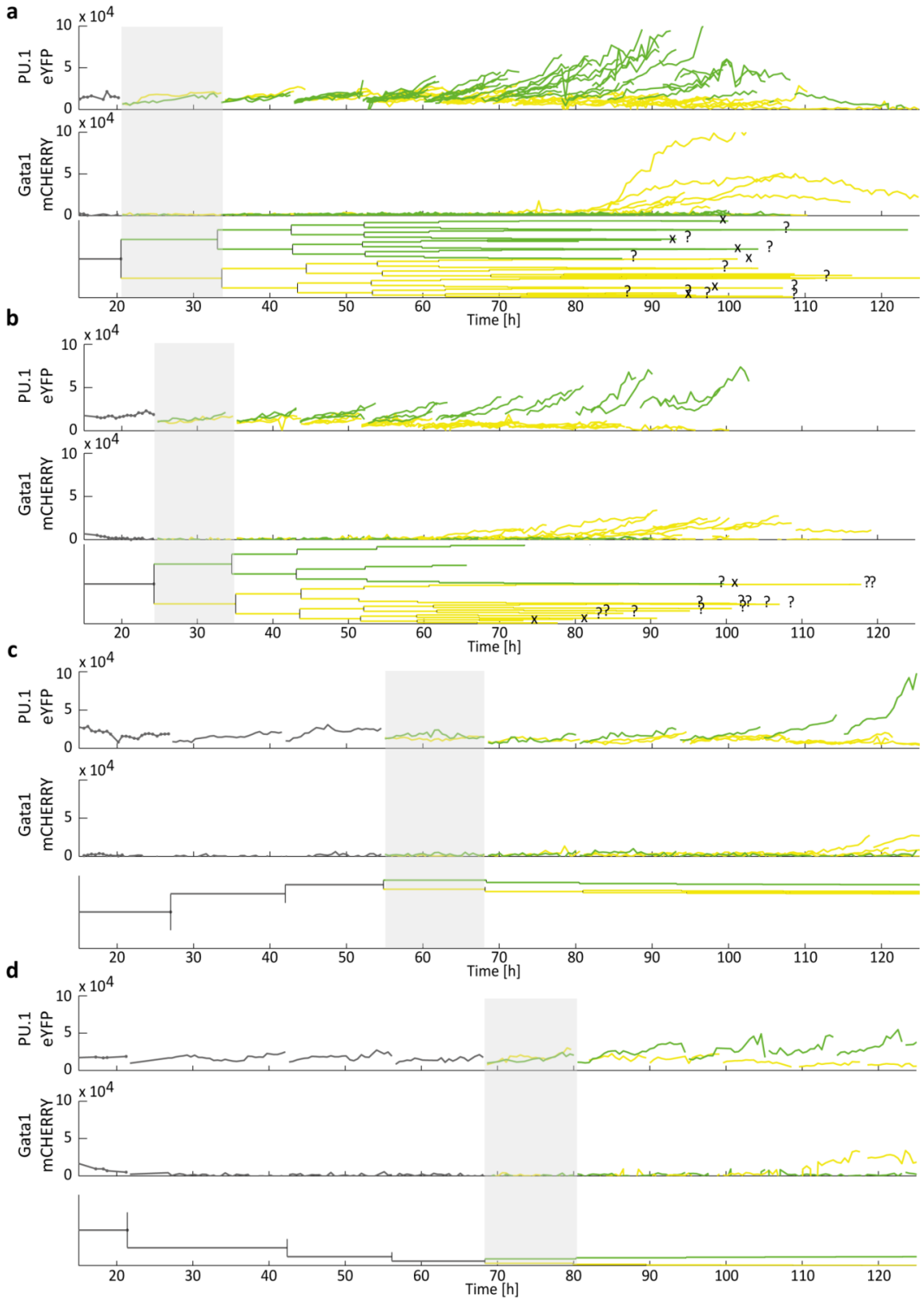


Figure 5-25: PU.1eYFP stoichiometrically exceeds potential present Gata1mCHERRY below detection levels in the vast majority of GM-traces

Panels include 218 traces of CD16/32-onsets from three independent experiments. The mean PU.1eYFP molecule numbers of the first three timepoints of every generation are plotted, in order to be independent from cell cycle stages. PU.1eYFP levels either never fell below the initial PU.1eYFP levels (**a**) or passed through an intermediate state with PU.1eYFP levels below the initial value (**b**). Initial values were calculated by setting the mean of the distribution of the initial PU.1eYFP value (mean of the first three timepoints of the start of the movie) to $8.1 \cdot 10^4$. Percentages denote traces of all GM-traces that are always above 10^4 or are below 10^4, $8 \cdot 10^3$, $6 \cdot 10^3$, $4 \cdot 10^3$ or $2 \cdot 10^3$, respectively, at least once. Individual CD16/32$^+$ can originate from the same starting cell. Cells with their lowest PU.1eYFP value in the last generation were excluded from the analysis (2.8%).

5.5.7 Transcription Factor Dynamics in Definite Multipotent Cells

In vitro differentiation experiments have been carried out with HSCs. Despite their multipotency, sister cells (even in early generations) and their progeny often but not always differentiated into the same lineage (see Figure 5-22). Manual tracking did not always lead to completely analyzed trees. Instead of generating complete trees, which is highly laborious, tracking was preferentially performed in many different colonies in order to investigate the transcription factor dynamics in many different clones. Due to time issues those trees were mostly fragmented. Still, some definite multipotent cells could be detected by time-lapse imaging whose progeny gave rise both to cells of the GM- and the MegE-lineage (Figure 5-26). The general observations mentioned above about no transient Gata1mCHERRY stages and upregulation of PU.1eYFP in the case of GM-lineage choice and downregulation of PU.1eYFP already before Gata1mCHERRY onset in the case of MegE-lineage choice were equally well observable here. Therefore, all data above (Figure 5-23 to Figure 5-25) can be used trustfully to interpret transcription factor dynamics in myeloid lineage choice, although the majority of genealogy trees were not analyzed completely (and therefore not definitely multipotent). Data from multilineage trees revealed that multipotent cells before lineage choice did not express any Gata1mCHERRY making fluctuating transcription factor networks again very unlikely to be the cause of myeloid lineage choice. Additionally, PU.1eYFP protein number dynamics behaved very similarly in sister cells whose progeny will either commit to the GM- or the MegE-lineage, respectively. Hence, PU.1eYFP upregulation does not seem to be the initial cause of GM-lineage commitment. Although the progeny of one sister cell will differentiate into the MegE-lineage, Gata1mCHERRY could not be detected during the life-time of that very same cell, but only up to 4 generations later only. This also strongly argues against Gata1 being the MegE lineage deciding. In summary, neither does a PU.1eYFP/Gata1mCHERRY double-positive multipotent cell state exist nor do PU.1eYFP and Gata1mCHERRY dynamics in committing cells point to any specific behavior although the progeny of sister cells will differentiate into opposing lineages. Therefore, it could be shown that, in contrast to the current opinion in the field, the PU.1-Gata1 stochastic switch is not responsible for GM versus MegE lineage choice.



(to be continued)

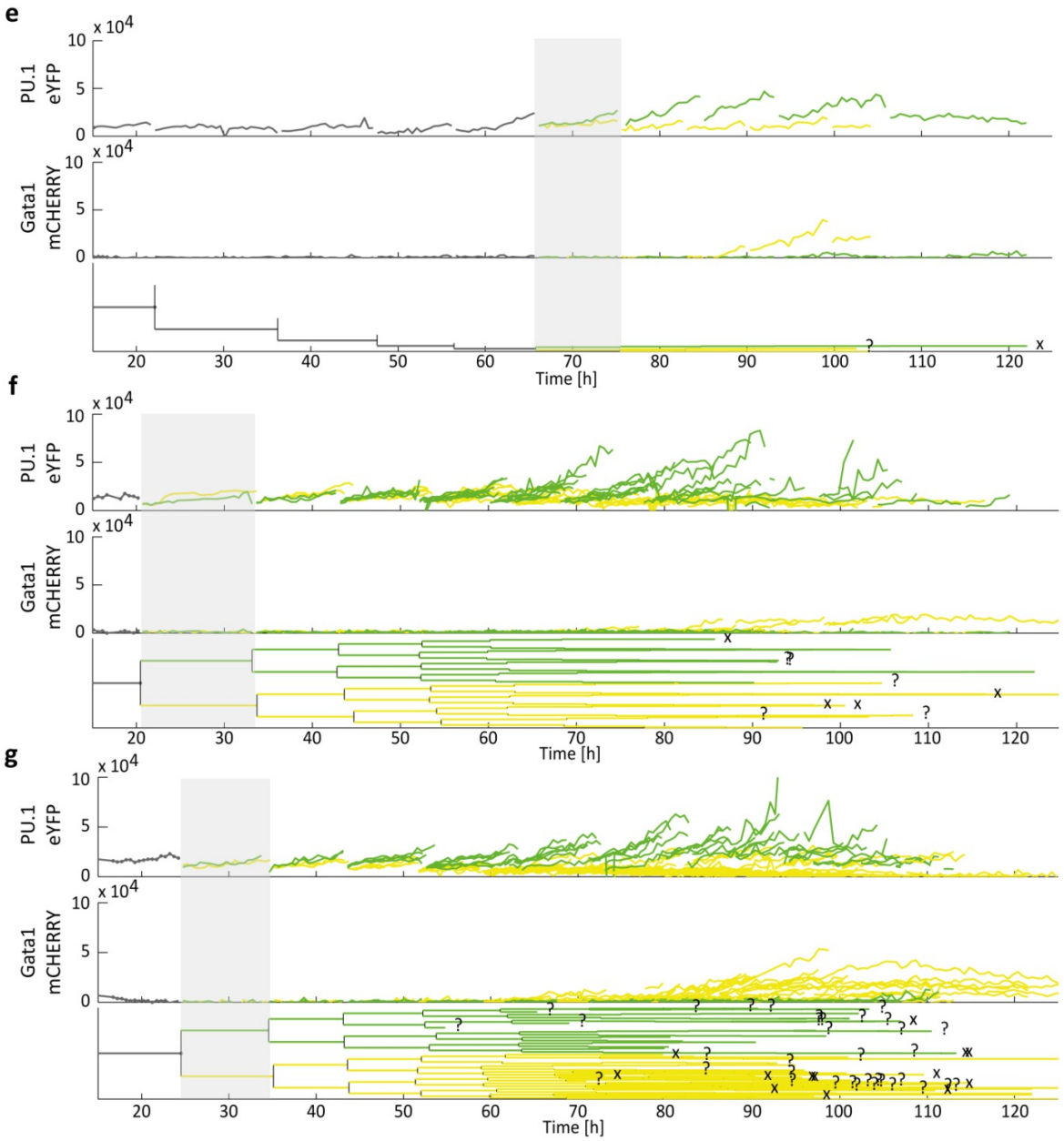


Figure 5-26: PU.1eYFP and Gata1mCHERRY do not exhibit different behavior in cells that will differentiate into GM- and MegE cells, respectively

Examples of multilineage trees that contain cells that will commit both to the GM- and the MegE-lineage. (**a**)-(**g**) Each panel contains (from bottom to top) the tracked and inspected cell genealogy tree, Gata1mCHERRY protein numbers and PU.1eYFP protein numbers. Grey cells and traces denote still multipotent cells. Green cells and traces denote cells that will commit to the GM-lineage. Yellow cells and traces denote cells that will commit to the MegE-lineage. Grey boxes denote cells that most likely will make the lineage decision. The generation before the grey boxes was definitely still multipotent. '?' = Cell lost during tracking; 'X' = Cell died during tracking.

6 Discussion

6.1 Experimental Investigation of the 'PU.1/Gata1 Paradigm'

The 'PU.1/Gata1 paradigm', suggesting PU.1 and Gata1 being a stochastic switch, has emerged as an interesting and intuitive model about myeloid lineage choice between the GM- and MegE-lineage, respectively. It is based on several experimental observations: (i) Both transcription factors have been shown to be the master-regulators of 'their' lineages. Knock-out experiments of PU.1 and Gata1 in both embryonic (complete knock-out) and adult (conditional knock-out) backgrounds have demonstrated their absolute necessity for proper development of 'their' mature lineages (Pevny et al., 1991; Scott et al., 1994; Fujiwara et al., 1996; McKercher et al., 1996; Iwasaki et al., 2005; Gutierrez et al., 2008). PU.1 and Gata1 are very powerful transcription factors, because upon ectopic overexpression they are able to reprogram other celltypes to 'their' lineages (Visvader et al., 1992; Kulessa et al., 1995; Nerlov and Graf, 1998; Heyworth et al., 2002). (ii) The phenomenon of 'lineage priming' has led to conclusion that opposing lineage-specific genes (like PU.1 and Gata1) are already expressed at low levels in presumptive multipotent cells (like HSCs, MPPs and CMPs). This assumption is mostly based on RNA-data from Northern blot analysis of cell populations and later from microarrays and from sensitive single-cell RT-PCR (Crotta et al., 1990; Heberlein et al., 1992; Cross et al., 1994; Hu et al., 1997; Delassus et al., 1999; Akashi et al., 2003). (iii) PU.1 and Gata1 can both positively autoregulate themselves via binding sites in their own promoters and UREs (Tsai et al., 1991; Chen et al., 1995; Crossley et al., 1995; Li et al., 2001; Nishikawa et al., 2003; Okuno et al., 2005) and also mutually inhibit each other's transcriptional activity by binding to each other (Zhang et al., 1999; Nerlov et al., 2000; Zhang et al., 2000; Rekhtman et al., 2003; Liew et al., 2006). Therefore, PU.1 and Gata1 have been suggested to be the central elements of a fluctuating transcription factor network that directly competes for myeloid lineage choice ultimately elicited by a stochastic event that leads to a strong imbalance and one factor dominating the other (Cross and Enver, 1997; Orkin, 2000; Cantor and Orkin, 2001; Graf, 2002; Laiosa et al., 2006; Iwasaki and Akashi, 2007; Orkin and Zon, 2008). As a result one factor is being highly upregulated and initiates lineage specific target gene expression along with strong repression of the other factor and opposing lineage specific target genes. Although very attractive, the PU.1/Gata1 stochastic switch model has never been proven experimentally and both conflicting and insufficient data leave the hypothesis at least questionable: (i) Knock-out experiments of PU.1 and Gata1

showed the essentiality of both factors for mature cells of 'their' lineages, but closer investigation allowed in both cases the discovery of lineage-committed progenitors cells implying that both factors might not be absolutely essential for myeloid lineage choice (Fujiwara et al., 1996; McKercher et al., 1996; Iwasaki et al., 2005; Iwasaki et al., 2005; Gutierrez et al., 2008). (ii) 'Lineage priming' in terms of simultaneous PU.1 and Gata1 expression has never been shown at the protein level. Some studies used GFP as a reporter for either PU.1 or Gata1 and could already show that not all HSCs, MPPs or CMPs expressed GFP reflecting PU.1 or Gata1 expression (Back et al., 2005; Nutt et al., 2005; Arinobu et al., 2007). (iii) All biochemical studies about direct interaction of PU.1 and Gata1 have been carried out in cell lines involving strong overexpression of both proteins or in leukemic cell lines which are (amongst other aspects) leukemic because of constant expression of one factor leading to a differentiation block and proliferation of committed blast cells (like PU.1 expression in erythroleukemia). (iv) To date, no experimental proof about fluctuating behavior of PU.1 and Gata1 has been adduced because cells (single or populations) have only been analyzed at single time points. Cell populations exhibit are large degree of heterogeneity and must not be expected to behave synchronously (Huang, 2009). In addition, any potential fluctuations have to be investigated at the single-cell level, otherwise the data will get lost in population noise.

In order to properly investigate the transcription factor dynamics, novel knock-in mouse lines were generated and used to simultaneously readout endogenous protein levels in individual living primary cells by fluorescence. By using advanced long-term time-lapse imaging at the single cell level which allows generation of cell genealogies over many generations and novel quantification methods, PU.1eYFP and Gata1mCHERRY expression levels were continuously quantified before, during and after the process of lineage choice. In cell genealogies that differentiated into GM-cells Gata1mCHERRY expression could never be detected. In cells with onset of Gata1mCHERRY expression, it was always followed by lineage commitment to the MegE-lineage. In conclusion, no data was found that supported the hypothesis that PU.1 and Gata1 as parts of a transcription factor network are directly involved in lineage choice. Importantly, the prerequisite of both factors being co-expressed (in stoichiometric numbers) before lineage choice was missing in the vast majority of cells. Additionally, PU.1eYFP downregulation before Gata1mCHERRY detection was often observed, demonstrating that other factor(s) than Gata1 are controlling that downregulation. Hence, the experimental analysis of PU.1 and Gata1 expression levels over time suggested that myeloid lineage choice is not essentially controlled by the PU.1/Gata1 stochastic switch.

The existence and relevance of the 'PU.1/Gata1 paradigm' remains highly questionable. Several aspects of the experimental approach and the biological conclusions will be discussed below.

6.2 Choice of Knock-In Approach and Effects on Functionality

For the relevance of this project, it was crucial to investigate real endogenous protein levels. Therefore, most of the published reporter lines were not suitable: Both existing PU.1 GFP-reporter lines were not made as knock-in fusion proteins, but either as GFP knock-in into exon 1 of PU.1, thereby abolishing and replacing PU.1 expression from one locus, or as a knock-in of an IRES-GFP construct into the endogenous *PU.1* locus (Back et al., 2005; Nutt et al., 2005). In the first case, mice were only heterozygously expressing PU.1, which might influence its proper function, and GFP has a different half-life and does not reflect endogenous PU.1 levels by altered decay rates. The latter approach allowed the generation of bicistronic mRNA and, hence, the possibility that both PU.1 and GFP are being transcribed from the same mRNA molecule. However, different decay rates of PU.1 and GFP again might not allow proper readout of PU.1 expression levels through GFP. Additionally, expression of the second gene (in this case GFP) was expected to be significantly lower compared the first gene (in this case PU.1) in the bicistronic mRNA (Mizuguchi et al., 2000), potentially leading to detection problems due to low GFP expression and to wrong conclusions.

Gata1 GFP reporter mice were generated using another approach (Iwasaki et al., 2005): The mice carried a randomly integrated transgenic construct harboring the Gata1 locus with a GFP knocked into exon 2 and all three known DNase hypersensitivity sites in the promoter region. Besides again potentially different half-lives of Gata1 and GFP another problem was that yet undiscovered regulatory elements might have been missed and not included into the transgenic construct. Additionally, the random integration into the genome might change expression of closely located genes. However, irrelevant GFP levels and again wrong conclusions about endogenous Gata1 expression would be the consequence.

Therefore, in order to read out endogenous transcription factor levels, the technical best solution is a fusion protein between the transcription factor and fluorescent protein homozygously expressed from the endogenous locus. In the case of PU.1, such a mouse line (PU.1eYFP) had been published (Kirstetter et al., 2006). Unfortunately, no data about the

functionality of the mouse line was available, but at least no phenotypes were known so far (personal communication with Prof. Dr. Claus Nerlov). In a complementary approach within this project a Gata1mCHERRY knock-in line was made. Although technically the best solution, the fusion of fluorescent proteins to the transcription factors must not change its functionality, otherwise it is worthless for proper biological conclusions. Interestingly, Gata1 in a fusion with the fluorescent protein tdTOMATO was not as efficient in reprogramming LMPPs to the megakaryocytic-erythroid lineage as wildtype Gata1 or Gata1 in a fusion with mCHERRY. Whether this effect was due to steric issues of altered protein stability cannot be answered here. This demonstrates that fluorescent proteins have to be chosen with care.

In general, eYFP and mCHERRY do not seem to be toxic to the cells, because heterozygous (and more importantly homo-/hemizygous) mice with fusion proteins are born. Fusion to fluorescent proteins does obviously not abolish the transcription factors function, otherwise no homo/hemizygous offspring would be born (Gata1) or survive long enough after birth (PU.1). Those phenotypes would be expected because of the embryonic knock-outs without any transcription factor being expressed at all. Blood counts of adult double homo-/hemizygous PU.1eYFP/Gata1mCHERRY mice were analyzed and no significant difference compared to wildtype C57BL/6 mice could be found. Since blood is a very complex organ with a lot of compensatory mechanisms, steady-state hematopoiesis at the stem and progenitor level was also investigated: Again, the abundance of many different well defined progenitor cell types was nicely comparable between knock-in and wildtype mice. Additionally, *in vitro* colony-forming potential of bulk bone marrow was not altered in knock-in mice. Dysregulated, hypormorphic mouse models of both PU.1 and Gata1 have been demonstrated to develop highly prevalent leukemia (Takahashi et al., 1997; Rosenbauer et al., 2004; Shimizu et al., 2004). On the other hand, deregulated ectopic expression of one transcription factor often leads to a differentiation block in the opposing lineage and to proliferation of blast cells, like PU.1 overexpression in erythroleukemia (Moreau-Gachelin et al., 1988; Moreau-Gachelin et al., 1996). Therefore, if the knock-in of open reading frames of fluorescent proteins into the endogenous loci had changed regulatory sequences in the 3'-UTRs or altered the overall stability of the fusion proteins compared to wildtype transcription factors, leukemias would have been expected to be observed. But neither blood counts nor bone marrow analysis pointed to any leukemia development. Additionally, double knock-in mice survived into old age and spontaneous death (e.g. due to leukemia) did not occur. Ultimately, the biochemical stability of wildtype transcription factors and their respective fusions was compared after cycloheximide treatment and no differences could be found.

Gata1 and Gata1mCHERRY exhibited a constant decay. Unexpectedly, PU.1 and PU.1eYFP showed even an increase in expression after the start of cycloheximide treatment. Importantly, the wildtype and fusion protein behaved the same way, showing that biochemical stability is not altered.

In conclusion, PU.1eYFP and Gata1mCHERRY fusion proteins proved themselves as fully functional transcription factors and reporter proteins at the same time. Since double knock-in mice expressed PU.1eYFP and Gata1mCHERRY homozygously (hemizygously for Gata1mCHERRY in male mice), no wildtype protein was produced anymore from the endogenous locus and all fluorescence intensity can directly be correlated with expression levels of the transcription factors. In imaging it could be observed that all fluorescence is exclusively located, as expected for active PU.1 and Gata1, in the nuclei of cells.

After verification of its normal functionality, the PU.1eYFP/Gata1mCHERRY double homo-/hemizygous knock-in mouse line was then used to elucidate the expression levels in defined HSPCs by flow cytometry. Unlike other PU.1-GFP reporter strains, PU.1eYFP could be detected in nearly all HSCs and MPPs. The other studies only observed subpopulations that expressed GFP, showing that the knock-in approach here is more suitable to readout complete PU.1 expression (Back et al., 2004; Nutt et al., 2005). Even GMPs which were exclusively PU.1eYFP$^+$ were not reported to be entirely GFP$^+$, leaving the value of previous reporter lines questionable at best. In terms of Gata1 expression the published data using a transgenic GFP reporter accorded with the Gata1mCHERRY expression in HSCs, MPPs and MEPs (Arinobu et al., 2007). In previous analyses at least 35% of CMPs were GFP$^+$, reading out either PU.1 or Gata1 expression. By now using two different fluorescent proteins, it was possible for the first time to readout PU.1 and Gata1 expression simultaneously. Interestingly, the CMP population was clearly divided into cells that exclusively expressed PU.1eYFP only or Gata1mCHERRY (and still low levels of PU.1eYFP). This clear separation became even more meaningful when those two distinct populations were sorted and used in a permissive colony-forming assay: The presumptive CMP was already pre-committed into the MegE- or the GM-lineage. This has far-reaching consequences for the interpretation of a lot of published data which assumed the CMP to be the true mother cell of GMPs and MEPs.

Evidently, the phenomenon of lineage priming could not be reproduced at the protein level by flow cytometric analysis. Simultaneous PU.1eYFP and Gata1mCHERRY expression could not be observed in any defined cell population between multipotent HSCs and lineage-restricted GMPs and MEPs. The only true PU.1eYFP/Gata1mCHERRY double-positive

population has a clear GMP phenotype and never significantly gave rise to MegE-cells in colony-forming assays. The low percentage of MegE-colonies was probably due its close proximity to the PU.1eYFP^{low}Gata1mCHERRY^{mid} cells in flow cytometry that overlap with the pre-MegE population. One observation of the knock-in approach was that endogenous transcription factor levels differ by only 2 – 3-fold between different defined HSPCs. As a consequence, the populations are not clearly discriminable by PU.1eYFP and Gata1mCHERRY levels without further markers. The PU.1eYFP/Gata1mCHERRY double positive GMP subpopulation will rather be important for the development of the post-GM lineages eosinophils, basophils and mast cells (Martin et al., 1990; Zon et al., 1993; Iwasaki et al., 2005). From the missing double expression of PU.1eYFP and Gata1mCHERRY in still multipotent cells it could be inferred that a simultaneous balanced expression of PU.1 and Gata1 cannot be the reason for cells remaining multipotent for all myeloid lineages. Alternatively, PU.1/Gata1 fluctuation could be a transient state that has to be 'turned on' in the first place. Therefore, the whole differentiation process from multipotent cells to committed GM- and MegE-lineages was followed continuously at the single-cell level *in vitro*.

6.3 Continuous Time-Lapse Imaging

After the discovery that CMPs are not true multipotent for all myeloid lineages but already pre-committed, they obviously were not a suitable population to follow *in vitro* differentiation and transcription factor dynamics. MPPs which constantly upregulate PU.1eYFP depending on their maturation stage have a strong bias for the GM-lineage. This phenomenon is also reflected by the existence of the LMPP (Adolfsson et al., 2005). Therefore, it was decided to use true multipotent long-term reconstituting HSCs for *in vitro* differentiation experiments. The protocol for the prospective isolation of HSCs combined two different approaches and purity of the stem cell population was expected to be at least 50%. This remarkably high percentage was deduced from single-cell transplantation experiments which demonstrate that one single HSC can reconstitute the whole blood system of lethally irradiated recipient mice (Osawa et al., 1996; Kiel et al., 2005). By this gold-standard proof of stemness it could be assumed that the majority of all HSCs used for time-lapse imaging experiments were truly multipotent.

One important issue of choosing HSCs as starting cells was the fact that HSCs would need more time to differentiate than for example MPPs because they are more immature. As a consequence, tracking of individual tree genealogies was much more laborious: If a single HSC and all its progeny divided symmetrically 10-times and no individual cells were lost during the tracking process, it would have led to 1024 (2^{10}) final cells and a tree consisting of in total 2047 cells. Just one more generation of cells would have led to 2048 (2^{11}) final cells and a tree with in total 4095 cells. In other words, every single generation more increases the tracking time by more than 2-fold. First, the bulk cell numbers increases by the factor 2 and second tracking of later generations is more difficult and time-consuming because of the exponential increase in cell density of the culture dish. Therefore, the culture conditions for time-lapse imaging experiments had to be chosen under certain aspects: (i) They had to allow the differentiation process in a reasonable amount time that allows keeping single-cell identity until the time point of lineage commitment. (ii) They had to allow development of all myeloid lineages. (iii) Ideally, culture condition should consist of defined components to later allow targeted manipulations of culture conditions in order investigate the individual effects of culture components. As a result, serum which is commonly used in *in vitro* differentiations was omitted. In serum-free cultures, cytokines are an essential component with allows differentiation of various lineages. Cytokines are highly pleiotropic and influence both many different hematopoietic cell types and many different cell processes like survival, proliferation, migration and maturation. To which extent cytokines control also differentiation in terms of lineage choice has been elusive for many years. Recently, the cytokines M-CSF and G-CSF have been demonstrated to be instructive, i.e. directly influencing lineage choice of GMPs (Rieger et al., 2009). Cytokine conditions were therefore chosen in a way that all lineages to be analyzed can be generated from HSCs. SCF was an obvious choice because it is essential for the culture of HSCs. Furthermore, its receptor c-Kit is expressed at least until GMP and MEP stages. It has general pro-survival and pro-proliferative effects. EPO was added to allow survival and proliferation of committed erythroid cells. Instructive influence on cells has never been demonstrated (Wu et al., 1995). The main reason for addition of TPO was its positive effect on development of the megakaryocytic lineage. TPO has also been shown to influence HSCs (Ema et al., 2000). In combination with SCF it was published as a condition that maintains HSCs. IL-3 and IL-6 were added as general pro-differentiation cytokines. They are mostly important for the development of GM-lineages. In combination with SCF and EPO, IL-3 and IL-6 (and serum) are used in commercially available standard assays to determine the myeloid potential of myeloid cells. In relation to the *in vitro*

differentiation of HSCs within one week of culture, the conditions led to a definite GM- or MegE-lineage choice in 50% of cells. Practically, under current physical culture conditions longer tracking is hardly possible. Therefore, the amount of true multilineage trees within 7 days of culture cannot be determined easily because individual colonies often did not contain cells with definite GM- or MegE-fate yet. Compared to Gatat1mCHERRY^{+} (marker for MegE-lineage) cells, CD16/32^{+} (marker for GM-lineages) cells were always more abundant. Here, it has to be taken into account that GM-committed cells never stopped proliferating after marker onset. In contrast, erythroid-committed cells did not survive long enough in the cultures and megakaryocytes at some point stopped proliferating but only 'divided' internally doing endomitosis. Therefore, the imbalance of Gata1mCHERRY^{+} and CD16/32^{+} cells on day 5 – 7 does not necessarily reflect an imbalance in lineage choice. Nonetheless, under the chosen culture conditions and many other similar conditions, there is a bias towards GM-cells at the expense of MegE-cells. The best readout for multipotent colonies under similar culture conditions is only after day 10 – 12 (personal communication with Dirk Löffler). Tracking that long and deep into individual tress is not feasible at the moment. Nevertheless, shorter tracking as performed here is good enough because the *in vitro* differentiation process from multipotent HSCs to committed myeloid cells could be followed.

Having investigated PU.1eYFP and Gata1mCHERRY transcription factor dynamics in individual HSCs under 'permissive' conditions, the influence of several culture components can now be tested experimentally and compared to the basic description of PU.1eYFP and Gata1mCHERRY behavior. Individual cytokines can either be omitted completely or added additionally or concentrations of cytokines can be altered and their influence on the transcription factor dynamics can be tested. Besides the manipulation of external signals, it is also possible to influence the transcription factor network directly: This includes both the direct manipulation of PU.1 and Gata1 by either (temporal) knock-down or (temporal) overexpression. Overexpression can nicely be achieved by lentiviral delivered tamoxifen sensitive ER fusion proteins: In theory, they consists out of the transcription factor, optionally a fluorescent protein (in eYFP/mCHERRY or any different color) and a tamoxifen-sensitive variant (ERT) of the estrogen receptor. Ideally, those proteins are expressed and stay in the cytoplasm. Upon tamoxifen-induction they move to the nucleus and can exert their function as transcription factors. The advantage of this system is its theoretical rapidness, titratability and reversibility. Alternatively, other transcription factors can be added or knocked-down and the influence on the core transcription factors PU.1 and Gata1 can be tested.

6.4 Endogenous Protein Numbers and Their Implications for Biology

In order to better understand the stoichiometric differences between PU.1eYFP and Gata1mCHERRY levels, Western blot analysis with a defined amount of PU.1eYFP$^+$ progenitor cells and Gata1mCHERRY$^+$ FL cells and defined amount of recombinant GFP or mCHERRY, respectively, were performed. The reason for choosing those populations was that they are quite abundant and sorting effort and time is reasonable for the amount of cells to receive. Additionally, both cell populations express high amounts of transcription factors. Still, the amount of protein was at the lower end of standard curves which were fairly at the detection limit of Western blotting. Despite of antibody optimization, signal intensities could not be increased. Protein lysate from even more cells could have been purified, but due to longer sorting times, cell survival would have decreased and cells would have been longer out of the bone marrow leaving the relevance of protein quantifications doubtful.

From the relative differences in flow cytometry molecule numbers from all other HSPCs of interest could be inferred. By using background calculation algorithms, quantification of endogenous transcription factor levels was also possible in imaging (Schwarzfischer et al., 2011). Although flow cytometry is very sensitive technology, image quantification even proved to have a higher dynamic range (fold-change to 'negative' signals) and more importantly sensitivity in quantifying low signals was even increased. However, every quantification technology has a certain detection limit. PU.1eYFP detection limit in imaging was $1.1 \pm 2.0 \cdot 10^3$. PU.1eYFP is expressed in HSCs and clearly above background levels. In image quantification, a PU.1eYFP$^-$ HSC has never been found. On the contrary, only 80% HSCs in flow cytometry were definitely PU.1eYFP$^+$. During differentiation into GM-cells PU.1eYFP intensity never dropped below the detection threshold. In cells differentiating into the MegE-lineage, PU.1eYFP was only not detectable anymore when cells already had committed and expressed Gata1mCHERRY. Therefore, the biological conclusion for PU.1eYFP levels dropping below detection levels is irrelevant for the questions addressed here because this happened only after lineage commitment to the MegE-lineage.

The detection threshold for Gata1mCHERRY in imaging was $1.9 \pm 4.4 \cdot 10^3$ molecules. In cells that differentiated into MegE cells, PU.1eYFP had already been downregulated before Gata1mCHERRY could be detected. Gata1 can prevent PU.1 from binding to its co-factor c-Jun which further could interfere which the positive autoregulatory loop of PU.1. Due to the stoichiometric imbalance in favor of PU.1eYFP, Gata1mCHERRY seems very unlikely to be

the initial cause of PU.1eYFP downregulation. During differentiation of cells into the GM-lineage Gata1mCHERRY was never detected and fluctuations of PU.1eYFP and Gata1mCHERRY as a mechanism seem to be very unlikely. Nonetheless, in order to critically interpret that result, one has to assume that – although unlikely – 'real' Gata1mCHERRY levels might always have been just below the detection threshold. Following this idea, half of all GM-traces never even dropped below the initial PU.1eYFP levels of $8.1 \cdot 10^3$ molecules. That means that more than half of all GM-traces always exhibited at least a 4-fold excess of PU.1eYFP over Gata1mCHERRY. Among the majority of all other GM-traces PU.1eYFP still exhibited a stoichiometric overbalance over Gata1mCHERRY. In conclusion, a theoretical stoichiometric balance of PU.1eYFP and Gata1mCHERRY in still multipotent cells was rarely the case (1.8%). Gata1mCHERRY was never transiently expressed but only always led to a sustained future expression and lineage choice into MegE cells, which was proven by sorting Gata1mCHERRY$^+$ cells and subjecting them to a cytospin or a colony-forming assay.

All above mentioned interpretations rely on the presence/absence of transcription factors in general and more specifically on the dynamics of their molecule numbers over time. Fluorescent signals were always detected in the nuclei of cells only and, therefore, also assumed to be active. In order to verify true transcriptional activity one could try to deliver specialized reporter constructs into primary cells. Such constructs could be derived from published constructs that harbor multiple binding sites for PU.1 or Gata1, a minimal promoter and the ORF for luciferase (Zhang et al., 1999; Zhang et al., 2000). Those constructs were used to determine transcriptional activity in cell lysates at the population level by a photometric assay. Instead of killing the cells and providing the substrate for the enzymatic reaction, the readout could be transferred to the single-cell level by cloning fluorescent reporters after the promoters. Then, by choosing another color than eYFP or mCHERRY transcriptional activity could be readout together with transcription factor levels and potential correlations with the protein numbers of PU.1 and Gata1 could be investigated.

6.5 Implications of Primary Data for Mathematical Modeling

Several studies tried to transform the experimental observations of lineage priming, positive autoregulation and mutual cross-antagonism of PU.1 and Gata1 into mathematical models in

order to test the hypothesis of a switch-like behavior by one factor being 'defeating' the other leading directly to lineage choice (Roeder and Glauche, 2006; Huang et al., 2007; Chickarmane et al., 2009; Duff et al., 2012; Strasser et al., 2012). All models built their assumptions, equations and parameters just on relative expression levels. After having quantified real endogenous protein dynamics and numbers for the first time here, those can be taken into account and confer much more relevance to modeling approaches. Besides now being able to continuously quantify protein number dynamics, recent technological advances have allowed to quantify specific mRNA numbers at the single-cell level by digital RT-PCR (Warren et al., 2006): Whole RNA samples from single primary cells are reversely transcribed into cDNA and diluted into 1200 portions on a microfluidic chip, so that each individual portion contains either none (mostly the case) or one cDNA template of the gene of interest. Then, highly sensitive PCR with specific primers amplifies the cDNA of interest. The amounts of successful PCR reactions from all individual microfluidic chambers corresponds to the numbers of original cDNAs and, hence, mRNAs: The mean amount of PU.1 mRNA in 21 individual tested HSCs was as low as 8.5 per cell. In that study GMPs were not investigated, but $Flt3^+$ CMPs which overlap well with GM pre-committed CMPs were reported to express 21.7 PU.1 mRNAs per cell (mean). That further means that one single mRNA serves as the template for 1000 PU.1 molecules in every HSC (mean PU.1eYFP expression level 8100). With that knowledge, the bottle neck of PU.1 production seems to be the amount of mRNAs. In order to get more insights into production rates, PU.1eYFP and Gata1mCHERRY primary could be exhibited to the established method of fluorescence recovery after photobleaching (FRAP) (reviewed in Ishikawa-Ankerhold et al., 2012): The existing fluorescent proteins linked to PU.1 or Gata1 can be bleached in live-cell imaging and fluorescence will only be detected again when new molecules will be produced. From the dynamics of fluorescent intensities, production rates in any cell-type of choice can be inferred. Complementary to that, protein decay rates can also be quantified by fluorescent signal decline in time-lapse imaging at the single-cell level (data now shown, manuscript in preparation). Instead of using magnitudes more of cells for classical determination of decay rates in Western blot, around 100 cells are already enough to quantify protein decay in imaging. Besides the advantage of saving a lot of time and effort for population analysis, single-cell analysis has the additional advantage of being able to detect heterogeneities in protein decay rates from a presumptive homogeneous cell population. In conclusion, primary data from single-cell analysis will contribute to better be able to model biological observations in a systems biology approach. As a result modeling approaches are well able to

propose further experiments in order to answer specific experimental question which then will again help to improve the models.

Although the PU.1/Gata1 switch does not seem the essential mechanism of myeloid lineage choice, both factors have central importance for the development of their 'lineages'. Instead of necessarily being the decision makers, the PU.1/Gata1 interaction network seems more importantly to be a lock-down mechanism that allows terminal differentiation into GM- or MegE-cells in order to prevent blast stages from uncontrolled proliferation. Indeed, PU.1 and Gata1 expression levels are often deregulated in myeloid leukemias (Burda et al., 2010). In addition, the theoretical frame-work of a stochastic switch behavior can still be valid for other lineage choices and in other cellular systems. Furthermore, the dynamics of PU.1 and Gata1 and the importance of a switch-like behavior in regard of differentiation into eosinophils, basophils and mast cells from GMPs still need to be elucidated using the same technical approach as applied to the investigation of GM- versus MegE-lineage choice. Additionally, other knock-in lines with interesting transcription factor candidates (like Gata2, C/EBPα, Runx1, SCL) with different colors can still be made and added to the already existing PU.1eYFP/Gata1mCHERRY knock-in line. Since emission spectra of fluorescent molecules can share a significant of overlap, sophisticated imaging approaches like linear unmixing have to be used (Zimmermann, 2005). That approach allows the correct calculation of signal contribution of individual fluorophores to emission spectra. Moreover, new technologies like strong LED excitation sources of defined wavelengths will contribute to make excitation of fluorophores more specific and, hence, live-cell imaging less toxic.

6.6 Relevance for Myeloid Lineage Choice *in vivo*

The investigation of the 'PU.1/Gata1 paradigm' here has been carried out by isolating primary HSCs and differentiating them *in vitro* under constant cytokine conditions that are very unlikely to overlap with the real *in vivo* conditions. In order to confer more realistic conditions for culturing primary cells the time-lapse imaging of suspension cells at the single-cell level can also be done on bone marrow stromal cell lines (Song et al., 2010). In another recent study, blood cells have been shown to emerge from endothelial cells between stroma cells using time-lapse imaging (Eilken et al., 2009). The stromal cell line OP9 (Kodama et al., 1994) used in that study is a widely accepted model for *in vitro* differentiation experiments

from ES cells into blood (Nakano et al., 1994; Nakano et al., 1996) or for adult bone marrow progenitors into lymphoid cells (Wang et al., 2005; Wang et al., 2006). Unlike other stromal cells lines, OP9 cells have not been shown to support maintenance of adult HSCs. In contrast, the stromal cell lines PA6 (Kodama et al., 1982) and AFT024 (Moore et al., 1997) have been shown to both maintain adult HSCs and allow proliferation and development of mature cell lineages (Kodama et al., 1984; Szilvassy et al., 1996; Shimizu et al., 2008). Of course, *in vitro* differentiation on stromal cell lines can additionally be influenced by the addition of cytokines or other compounds of interest. Additionally, large scale transcriptome analysis and genetic manipulation of the stromal cells allows testing the influence of individual factors expressed by the niche cells on *in vitro* hematopoiesis.

Ultimately, the observation of transcription factor dynamics at the single-cell *in vivo* would be highly desirable. Due to missing labeling technology and many other obvious difficulties, *in vivo* imaging of transcription factor dynamics in humans is not possible. In contrast, *in vivo* imaging of HSCs in their long bone marrow niche has been performed using highly invasive fiber-optics insertion through the femoral head (Lewandowski et al., 2010) or by two-photon imaging of mechanically thinned tibiae of anesthetized mice (Kohler et al., 2009). In another less invasive study, labeled HSCs were imaged in the calvarium bone marrow of transplanted mice (Lo Celso et al., 2009). The temporal resolution at imaging timepoints in any of those studies was within the minute range, but the duration between the timepoints was at least several days witch would not allow to follow individual cells long enough. The biggest obstacle for continuous long-term imaging is the frequency and duration of individual anesthesias which need to be high and long, respectively, in order to keep single-cell identity. Additionally, since long-term continuous imaging has never been performed, cells might be highly mobile and get lost by changing their location to other areas than the imaging site or even to other bones through the circulation.

Another model organism that has itself proven to be highly suitable for imaging is the zebrafish (*Danio rerio*). The fish are relatively small, transparent and allow a deep tissue penetration of light. Moreover, they develop and proliferate fast and are susceptible to general genetic manipulation as well as screening approaches. During the last decade, zebrafish has also become a valuable model for normal and malignant hematopoiesis (reviewed in Jing and Zon, 2011). Valuable assays like clonal analysis of hematopoietic progenitors that are commonly used in mouse hematopoiesis research have been established for the zebrafish model system, too (Stachura et al., 2011). Definitive hematopoiesis in zebrafish has further

been shown to be initiated through a committed erythromyeloid progenitor (EMP) (Bertrand et al., 2007) and other myeloid progenitors can be identified by forward/side scatter characteristics and transgenic reporter lines (Traver et al., 2003). PU.1 and Gata1 are highly conserved transcription factors and their interplay has also been suggested to determine myelo-erythroid progenitor cell fate in zebrafish (Rhodes et al., 2005). In order to elucidate the 'PU.1/Gata1 paradigm' *in vivo* endogenous transcription factor levels would need to be quantified in the same approach like it has been done in this thesis. In zebrafish, endogenous PU.1 and Gata1 transcription factor levels have not been investigated due to missing protocols. Although targeted knock-in approaches by homologous recombination are commonly used in mouse hematopoietic research, this method is not well established in zebrafish. Other approaches like zinc-finger nuclease mediated mutation protocols have been provided recently and might allow more targeted genetic manipulations possible in future (Jing and Zon, 2011). In conclusion, to date it is not possible yet to investigate endogenous transcription factor levels continuously *in vivo*. Further technological developments including mostly reliable continuous long-term deep tissue imaging in mice or the generation of knock-in lines in zebrafish are necessary to be able to elucidate the 'PU.1/Gata1 paradigm' in the model organism of choice.

6.7 Upstream Regulators of PU.1 and Gata1

In vitro differentiation experiments showed that no clear PU.1/Gata1 double positive state precedes myeloid lineage choice. Rather than being decision makers between the MegE- and the GM-lineage, their role in myelopoiesis seems to be restricted to executing lineage choice by upregulating lineage-specific genes (and downregulating opposing lineage specific genes). Investigation of transcription factor dynamics in multipotent HSCs and their progeny has revealed that lineage choice in favor of the GM-lineage is preceded by upregulation of PU.1 without any transient Gata1 expression. Cells were committed to the MegE lineage as soon as Gata1 could be detected whose expression mostly happened after PU.1 had already been downregulated (independently of Gata1). Those observations immediately lead to the question which factors upstream of PU.1 and Gata1 regulate their expression and, ultimately, which factors are the decision-makers (reviewed in Wolff and Humeniuk, 2013).

Cytokines play an important role in hematopoiesis. This becomes even more apparent under stress-induced hematopoiesis, e.g. after massive blood loss or during infections. Steady-state hematopoiesis obviously must react to external cues. After decades of intensive debate, only recently it has been shown that the cytokines M-CSF and G-CSF can instruct cell fate in GMPs (Rieger et al., 2009). It was also suggested that subjection of M-CSF to HSCs both *in vitro* and *in vivo* directly leads to an increase in PU.1 levels and further to a lineage-bias in differentiation (Mossadegh-Keller et al., 2013). The influence of more cytokines on endogenous transcription factor levels and, hence, lineage choice has not been revealed, yet. Cytokine signaling cannot be regarded as a strict unidirectional process, but its significance can be strongly influence by internal factors: The transcription factor MafB was demonstrated to regulate the sensitivity of M-CSF signaling on multipotent progenitors (Sarrazin et al., 2009). $MafB^{-/-}$ HSCs exhibited a competitive repopulation advantage of the myelomonocytic compartment that still was dependent of M-CSF and PU.1. Therefore, cytokine signaling and transcription factors are no independent entities, but exhibit large interdependency as part of huge networks. Candidates that might be involved in regulation myeloid lineage choice can be inferred from myeloid leukemias. Both PU.1 and Gata1 themselves are often dysregulated in leukemias (Wolff and Humeniuk, 2013), but also many other genes are dysregulated in hematologic malignancies which also can lead to imbalances in the MegE/GM lineage switch. An example of another type of internal factor that might be upstream of PU.1 and Gata1 and might regulate myeloid lineage choice is the tumor-suppressor gene p15Ink4b. It was originally described as a cyclin-dependent kinase inhibitor and is very often deregulated in patients with AML (Drexler, 1998). Knockout experiments of p15Ink4b *in vivo* and *in vitro* showed a lineage-bias in favor of the granulocyte-monocyte lineage at the expense of the erythroid compartment (Rosu-Myles et al., 2007). P15Ink4b is higher expressed in MEPs than in GMPs and ectopic expression in primary multipotent cells leads to an increase in mitogen-activated protein kinase/extracellular signal-regulated kinase (MEK/ERK) signaling and to an increase in Gata1 expression as well as a decrease in PU.1 expression. Observations from zebrafish have suggested the factor transcription intermediate factor-1γ (tif1γ) to be another modulator of the PU.1/Gata1 switch by positively regulating Gata1 and negatively regulating PU.1 (Monteiro et al., 2011). A signaling molecule having opposing function by negatively regulates Gata1 expression and positively regulating PU.1 is Desert Hedgehog (Dhh) (Lau et al., 2012). Dhh-deficient CMPs preferentially differentiated into MegE cells than into GM cells. Additionally, Dhh negatively regulated erythroblast differentiation and erythroid progenitor populations were increased in Dhh deficient mice.

In summary, unlike hypothesized for many years, PU.1 and Gata1 do not seem to be the central transcription factors deciding myeloid lineage choice. Their role rather seems to be restricted to carrying out lineage choice and locking it down. Along with the discovery of many factors that can influence myeloid choice, there does not seem to be a master regulator of GM- or MegE-lineage choice. Individual cytokines, signaling components, transcription factors, epigenetic regulators and other molecules all function as a small part in a huge network. Technically, the PU.1eYFP/Gata1mCHERRY double knock-in mouse is a perfect tool in order to directly image and quantify endogenous transcription factor levels upon exposure to any external cue of interest or coordinated genetic manipulation controlled both in time and space in freshly isolated primary cells. Further technological development will allow the quantification of endogenous transcription levels and their consequences in lineage outcome also *in vivo*. Therefore, the knowledge about the influence of individual molecules on PU.1 and Gata1 and myeloid lineage choice will increase drastically in near future. The integration of new experimental data into mathematical models and the interpretation of the data from the point of systems biology will help to (i) suggest more relevant experiments, (ii) further understand what exactly orchestrates myeloid lineage choice and (iii) elucidate how it might be manipulated in order to influence hematopoiesis and attenuate hematopoietic diseases.

7 Experimental Procedures

7.1 Molecular Biology

7.1.1 Cloning Strategy

All cloning strategies were designed *in silico* using Clone Manager Professional 9 software (Scientific & Educational Software, Cary, USA) based on complete sequences of available plasmids and bacterial artificial chromosomes (BACs).

7.1.2 Restriction Digests and Ligations

Restriction digests and ligations were carried out using enzymes and suitable buffers from New England Biolabs (Ipswich, USA) or Thermo Fisher Scientific (Waltham, USA) according to manufacturer's instructions.

7.1.3 Polymerase Chain Reaction (PCR)

PCR primers were designed using Clone Manager Professional 9 software. Annealing temperature was chosen according to the software. PCR was carried out according to manufacturer's instruction using Taq polymerase from Thermo Fisher Scientific (Cat. Nr. EP0072, Waltham, USA) for genotypings or Advantage Polymerase 2 from Clontech (Cat. Nr. 639206, Mountain View, USA) for cloning.

7.1.4 Agarosegel

DNA fragments from genomic digests, restriction digests and PCR products were separated on 0.7% to 1.5% agarose (Cat. Nr. 840004, Biozym, Oldendorf, Germany) gels prepared in TAE-buffer composed of 40 mM tris-(hydroxymetyhl)-aminomethane (TRIS) (Cat. Nr.

5429.3, Roth, Karlsruhe, Germany), 20 mM acetic acid (Cat. Nr. 1000632511, Merck, Darmstadt, Germany) and 1 mM ethylenediaminetetraacetate (EDTA) (Cat. Nr. 8043.2, Roth, Karlsruhe, Germany). Agarosegels were stained with 1% ethidium bromide solution (Cat. Nr. 2218.2, Roth, Karlsruhe, Germany) at a concentration of 6 µl per 100 ml agarose solution. DNA fragments were separated with voltages between 30 V and 150 V in TAE-buffer depending on the size of gels.

7.1.5 Purification of DNA Fragments

DNA fragments were cut out from agarose gels using a scalpel and purified using the QIAquick Gel Extraction Kit (Cat. Nr. 28704, Qiagen, Hilden, Germany) and QIAEX II Gel Extraction Kit (Cat. Nr. 20021, Qiagen, Hilden, Germany). Purification of individual PCR fragments was performed using the QIAquick (Cat. Nr. 28104, Qiagen, Hilden, Germany). DNA fragments were resuspended in H_2O bidest and DNA concentration was measured on a NanoDrop spectrophotometer nd-1000 (Thermo Scientific, Waltham, USA).

7.1.6 Isolation of Plasmid DNA

High copy plasmids from DH5α *Escherichia coli* bacteria were purified from 5 ml overnight cultures using the QIAprep Spin Miniprep Kit (Cat. Nr. 27104, Qiagen, Hilden, Germany) or a modified protocol replacing columns with an isopropanol (Cat. Nr. 6752.2, Roth, Karlsruhe, Germany) precipitation step and self-made buffers replacing manufacturer's buffers. Purification from 150 ml overnight cultures was performed using Qiagen Plasmid Maxi Kit (Cat. Nr. 12165, Qiagen, Hilden, Germany).

7.1.7 Isolation of BAC DNA

Bacteria containing BACs were inoculated in 5 ml of LB medium (Cat. Nr. 12780-029, Invitrogen, Karlsruhe, Germany) with 25 µg/ml chloramphenicol (Cat. Nr. 3886.2, Roth,

Karlsruhe, Germany). Overnight cultures were centrifuged at 4680 rcf for 15 min at 4 °C. Supernatant was removed and pellet was resuspended in 250 µl buffer P1 (QIAprep Spin Miniprep Kit). 250 µl buffer P2 were added and bacteria were lysed for 5 min at RT. 250 µl of buffer P3 were added and the solution was incubated for 5 min on ice. Solution was centrifuged for 5 min at maximum speed at 4 °C in a table-top centrifuge. Supernatant was transferred to a new tube and the centrifugation step was repeated. 750 µl isopropanol were added, mixed and incubated for 10 min on ice. The DNA precipitate was centrifuged for 10 min in a table-top centrifuge at 4 °C. Supernatant was removed and the pellet was washed with 1 ml of 70% ethanol (Cat. Nr. 1.00983.2500, Merck, Darmstadt, Germany). After another centrifugation step, the pellet was air-dried and resuspended in 30 µl H_2O bidest.

7.1.8 Transformation of Bacteria

Transformation was performed by thawing chemical competent DH5α bacteria frozen at – 80 °C for 15 minutes on ice. Up to 100 ng of plasmid or ligation cocktail were added to bacteria suspension and incubated 30 minutes on ice. After heatshock treatment of bacteria for 90 s at 42 °C, the suspension was cooled 2 min on ice. Afterwards, 1 ml LB-medium was added to the cells and cells were incubated shaking 1 h at 37 °C. Then, different dilutions of bacteria suspension were plated on LB agar (Cat. Nr. 244520, Becton Dickinson, Heidelberg, Germany) with adequate antibiotic for selection of successfully transformed bacteria. LB agar plates were incubated overnight at 37 °C.

7.1.9 Phenol-Chloroform Extraction

The volume of the DNA solution was increased to 900 µl with H_2O bidest. An equal volume of phenol-chloroform-isoamylalcohol (25/24/1) (Cat. Nr. A156.3, Carl Roth, Karlsruhe, Germany) was added to the DNA and vortexed for 10 s. The solution was centrifuged in a table-top centrifuge for 5 min at maximum speed. The aqueous layer was carefully transferred to a new tube. In order to remove residual phenol, an equal volume of chloroform was added, vortexed and centrifuged again. The aqueous layer was transferred to a new tube again and the steps were repeated 2 more times.

7.1.10 Ethanol Precipitation

3 M sodium acetate (10% of sample volume) (Cat. Nr. 6773.1, Roth, Karlsruhe, Germany) was added to DNA solution. Then, 100% ethanol (2 sample volumes) was added to the sample. Solution was incubated at -20 °C overnight. DNA sample was centrifuged in a table-top centrifuge at maximum speed for 15 minutes at 4 °C. Supernatant was removed and the pellet was washed with ice-cold 70% ethanol. DNA solution was centrifuged again and supernatant was removed. The pellet was air-dried and resolved in H_2O bidest.

7.1.11 Generation of Gata1mCHERRY Knock-In Construct

The Gata1mCHERRY knock-in construct was generated using classical cloning and recombineering (Yu et al., 2000). Recombineering allows site directed exchange of large DNA fragments depending on homology sites at the end of linearized DNA. Recombineering was used for the retrieval of the *Gata1* locus from a BAC using homology sequences defined by the primer pairs 5'-GTTTGTCCACTGACCTCCAGA TAG-3'/5'-CCAAAGCTACTGGC TTCCTCTG-3' and 5'-TCCCTCCCTCTTTCCCATTCTTC-3'/5'-TAAGTATGCTCCCGC AAGATGACCTG-3'. The sequence of the *Gata1* locus was changed using the primer pairs 5'-GTAGCAGCAGTAGTGGGAATTGTG-3'/5'-TGCCACAAGGTCAAGGCTATTCTG-3' and 5'-AGGTACACAGAATAGCCTTGACCTTGTG-3'/5'-AAAGCCAGCCTAGGCTG CATAG-3'. The change included the removal of the endogenous stop codon of *Gata1* and the addition of a short linker sequence (AGAGCATCAGGTACCAGTGGAGCT) and the open reading frame of mCHERRY (Shaner et al., 2004) and a FRT-site flanked neomycin resistance cassette with a eukaryotic promoter (PGK), a eukaryotic promoter (EM7) and a polyadenylation signal.

7.1.12 BAC Electroporation

BACs containing the *Gata1* locus (Cat. Nr. RPCIB731C02198Q, Source BioScience, Berlin, Germany) were electroporated into EL350 bacteria. This strain was derived from E. coli DH10B and contains a defective λ prophage with the recombination proteins *exo*, *bet* and *gam*

which are controlled by a temperature-sensitive repressor (Lee et al., 2001). For preparation of electrocompetent bacteria an overnight culture in 5 ml LB medium was inoculated at 32 °C (to prevent expression of heat-shock inducible genes). The next day, the culture was diluted 1:50 in 50 ml LB medium and grown until an optical density (OD_{600}) of 0.6. The culture was incubated 5 min on ice and centrifuged at 4680 rcf for 10 min at 4 °C. All subsequent steps for removing residual salt were carried out on ice. The supernatant was exhausted and the pellet was resuspended in 1.8 ml ice-cold 10% glycerol and the solution was transferred to a 2 ml Eppendorf tube. The sample was centrifuged at 4 °C in a table-top centrifuge at maximum speed for 20 s. Supernatant was exhausted and the pellet was resuspended in 1.8 ml 10% ice-cold glycerol. The washing step was repeated 3 more times. After the last wash, the pellet was resuspended in a total volume of 100 µl 10% glycerol. 3 µg of BAC DNA were mixed with 50 µl of competent bacteria and left 5 minutes on ice. The bacteria were transferred to a pre-cooled 1 mm cuvette (Cat. Nr. 165-2089, Bio-Rad, Munich, Germany) and electroporated at 1700 V, 25 µF and 200 Ω. For successful electroporation, the resulting time constant should be close to 5 ms. 1 ml of LB medium was immediately added and the bacteria were shaken at 32 °C for 1 h. Bacteria were then plated on LB agar plates containing 30 µg/ml chloramphenicol and incubated at 32 °C.

7.1.13 Recombineering

For site-directed recombineering, electrocompetent bacteria were prepared as described above with an additional heat-shock step (7.1.12): After an OD_{600} = 0.6 the 50 ml culture was divided in 2 aliquots. One was left at RT, the other one was placed in a 42 °C water bath and shaken for 15 min. Both cultures were then kept one ice and electrocompetent bacteria were prepared as described. One 25 ml culture was sufficient for 4 electroporations. 2 electroporations (heat-shocked and not heat-shocked) with 15 ng of linearized plasmid were carried out. Cultures were plated at different dilutions on LB agar containing suitable antibiotics and incubated at 32 °C overnight.

7.1.14 Isolation of Genomic DNA from ES Cell Clones

Individual ES cell clones were grown to confluence in gelatin-coated (Cat. Nr. G1890, Sigma-Aldrich, Taufkirchen, Germany) 96-well plates (Cat. Nr. 167008, Thermo Scientific, Waltham, USA). Proteinase K (Cat. Nr., 3115801001, Roche, Penzberg, Germany) was freshly added to the lysis buffer to a final concentration of 1 mg/ml. The lysis buffer consisted of 10 mM TRIS (pH = 7.5), 10 mM EDTA, 10 mM sodium chloride (NaCl) (Cat. Nr. 3957.1, Roth, Karlsruhe, Germany) and 0.5% N-lauroylsarcosine sodium salt (Cat. Nr. L5777, Sigma-Aldrich, Taufkirchen, Germany). The medium was exhausted and ES cells were washed twice with DPBS (Cat. Nr. 14040-091, Invitrogen, Karslruhe, Germany). 50 µl of lysis buffer were added to each well and incubated overnight at 55 °C in a humidified chamber. On the next day, 150 µl 5 M NaCl were added to 10 ml icecold 100% ethanol. 100 µl of NaCl/ethanol were added to each well and plates were incubated at room temperature for 2 h. DNA emerged as a filamentous network at the bottom of each well. The plates were carefully inverted and dried on a paper towel. DNA was washed 3 times with 150 µl 70% ethanol. After drying DNA was resuspended in 25 µl H_2O bidest.

7.1.15 Restriction Digest of Genomic DNA

Restriction digest of genomic DNA from ES cell clones was performed using 10 U EcoRI (or XbaI), the appropriate buffer, 1 mM spermidine (Cat. Nr. 7161.3, Roth, Karlsruhe, Germany), 100 µg/ml RNase A (Cat. Nr. 19101, Qiagen, Hilden, Germany) and H_2O bidest in a total volume of 50 µl per well.

7.1.16 Southern Blot

Digested genomic DNA was separated on a 0.7% agarose gel at 30 V overnight. A picture with a fluorescent ruler was taken the next day. Gel was shaken 15 minutes in depurination solution (0.4 M hydrochloric acid) (Cat. Nr. 1.00319.2511, Merck, Darmstadt, Germany) and washed with H_2O bidest. Then, gel was incubated 45 min in denaturation solution (0.2 M sodium hydroxide, 0.6 M NaCl) (sodium hydoxide, Cat. Nr. 9356.1, Roth, Karlsruhe,

Germany) and washed with H_2O bidest. At last, gel was incubated 45 min in neutralization solution (1 M TRIS, 0.6 M NaCl) and washed again with H_2O bidest. Gel was blotted onto a Hybond-N+ nylon membrane (Cat. Nr. RPN203B, GE Healthcare, Freiburg, Germany) with 20x saline-sodium citrate (SSC) buffer (3 M NaCl, 342 mM trisodium citrate (Cat. Nr. 3580.4, Roth, Karlsruhe, Germany), pH = 7) overnight. Membranes were air-dried and baked for 30 min at 80 °C to fix the DNA.

7.1.17 Generation and Hybridization of Southern Probes

Southern probes were generated by PCR using the Gata1 BAC as a template and the specific primer pairs 5'-CAGCCACTGCCCAAATAGGTGGAG-3'/5'-ATTTCCCTAAAACTGTGC TGATGTG-3' (upstream Southern probe) and 5'-CTGAAGTGGTGCTCTGGACTTTAC-3'/5'-TGAGGAAGAGGGAAGGATGTGAAG-3' (downstream Southern probe). PCR product was run on an agarose gel, purified and labeled with radioactive [α-32P] deoxycytidine triphosphate (3000 Ci/mmol, 10 mCi/ml) (Cat. Nr. NEG513H400UC, Perkin Elmer LAS, Rodgau-Jügesheim, Germany) using the Prime-It II Random Primer Labeling Kit (Cat. Nr. 300385, Agilent Technologies, Santa Clara, USA) according to manufacturer's instructions. 1 µl of the radioactive labeled probe was measured and cleared for usage, if liquid scintillation counts exceeded 300000 counts per minute on a Triathler Multilabel Tester (Cat.Nr. 425-004, Hidex, Turku, Finland). The probe was then prepared for hybridization: 500 µl of 10 mg/ml herring sperm single-stranded DNA (Cat. Nr. ab46666, Abcam, Cambridge, United Kingdom) was incubated 10 min in a 100 °C water bath and mixed with the desired amount of probe. 50 µl of 10 M NaOH was slowly added during constant shaking followed by 300 µl 2M TRIS, pH = 8. At last 475 µl of 1 M hydrochloric acid (HCl) (Cat. Nr. P074.3, Roth, Karlsruhe, Germany) were added very slowly. Denatured probe was added to membranes at least 3 h prehybridized with hybridization juice at 65 °C (1000000 counts of the probe per 1 ml hybridization juice). Hybridization juice consisted of 1 M NaCl, 5% 1 M TRIS (pH = 7.5), 0.1 g/ml dextran sulfate (Cat. Nr. 5956.3, Roth, Karlsruhe, Germany), 1% sodium dodecyl sulfate (SDS) (Cat. Nr. 2326.1, Roth, Karlsruhe, Germany) and 250 µg/ml herring sperm single-stranded DNA. Probe was hybridized rotating overnight at 65 °C. Probe was washed off with 2x SSC / 0.5% SDS (preincubated at 65 °C) at 65 °C. Membranes were constantly measured and cleared for exposure to Kodak BioMax MS films (Cat. Nr. z363057, Sigma-Aldrich, Taufkirchen, Germany) if beta radiation was between 30 – 50 counts per

second measured on contamination monitor (Cat. Nr. LB 122, Berthold Technologies, Bad Wildbad, Germany). Membranes were wrapped in plastic foil and exposed in a film cassette at – 80 °C for at least 3 days. X-ray films were processed using an AGFA Curix 60 developer machine (Cat. Nr. 09712043, Siemens, Munich, Germany).

7.1.18 Genotyping

PU.1eYFP mice were genotyped using the primers 5'-TGCGCAACTACGGCAAGAC-3' (forward), 5'-GGGCGACGGGTTAATGCTATG-3' (wildtype allele reverse) and 5'-TCAGCTCGATGCGGTTCAC-3' (knock-in allele reverse) leading to a 176 bp band (wildtype) or a 546 bp and 947 bp bands (knock-in), respectively.

Gata1mCHERRY mice were genotyped using the primers 5'-AGGTACTGCCCACCTCTAT C-3' (wildtype forward), 5'-GCATGGACGAGCTGTACAAG-3' (knock-in forward) and 5'-GCAGGAGAATGGGAAATGTG-3' (reverse) leading to a 297 bp (wildtype) or a 223 bp (and a 1101 bp band) (knock-in), respectively.

Successful removal of the neomycin resistance cassette was verified by the primers 5'-GCATGGACGAGCTGTACAAG-3' (forward), 5'-CTGCACGAGACTAGTGAGAC-3' (neomycin reverse) and 5'-GCAGGAGAATGGGAAATGTG-3' (Gata1 3'-UTR reverse), leading to 387 bp band (and a 2020 bp band) in the case of non-removal and a band of 223 bp band in the case of Flp-e mediated recombination.

7.1.19 Sequencing

Capillary sequencing was performed in the in-house sequencing facility using the BigDye Terminator v3.1 Cycle Sequencing Kit (Cat. Nr. 4337455) on a 3730 DNA Analyzer (both Applied Biosystem, Foster City, USA).

7.2 Generation of Lentivirus

7.2.1 Virus Production

Vesicular stomatitis virus glycoprotein (VSV-g) pseudotyped lentivirus was produced in human embryonic kidney (HEK) 293T cells. In comparison to vectors pseudotyped with murine ecotropic envelope, VSV-g lentivirus always resulted in higher titers (Schambach et al., 2006) and better transduction rates of primary cells (data not shown). HEK cells were kept in Dulbecco's modified Eagle medium (DMEM) (Cat. Nr. 41966-029, Invitrogen, Karlsruhe, Germany) with 10% fetal calf serum (FCS) (Cat. Nr. S1900-500, Lot Nr. S05130S1900, Biowest SAS, Nuaillé, France). For each virus cells were seeded at a density of $5 \cdot 10^6$ cells / 60 cm^2 in 4 dishes (Cat. Nr. 150350, Thermo Scientific, Waltham, USA). The next day each plate was cotransfected with 4 different plasmids in 0.5 ml 250 mM calcium chloride ($CaCl_2$) HEPES buffered saline (HBS) (Cat. Nr. CAPHOS, Sigma-Aldrich, Taufkirchen, Germany). The plasmids added were 2,5 µg pRSV_rev (containing rev, regulator of expression of virion proteins), 5 µg pMDLg_pRRE (containing gag = structural protein and pol = lentivirus-specific enzymes and a rev responsive element), 1 µg pMD2.VSV-g (containing envelope protein) and 10 µg of the plasmid with the gene of interest derived from pRRL.PPT.SFFV.GFP.PRE (Schambach et al., 2006). Prior to transfection, the medium was replaced by 10 ml of transfection medium which contained DMEM, 10% FCS, 0.1 mM non-essential amino-acids (NEAA) (Cat. Nr. 11140-035, Invitrogen, Karlsruhe, Germany), 1 mM sodium pyruvate (Cat. Nr. S8636, Sigma-Aldrich, Taufkirchen, Germany), 20 mM HEPES (Cat. Nr. 15630-056, Invitrogen, Karlsruhe, Germany) and 100 U/ml and 100 µg/ml penicillin/streptomycin (Cat. Nr. 15140122, Invitrogen, Karlsruhe, Germany). Cells were incubated at least 6 hs before medium was changed. 2 days after transfection, the virus containing medium from 4 dishes was collected, centrifuged at 240 rcf at 4 °C and the supernatant was filtered through a 0.2 µm filter (Cat. Nr. 17805, Sartorius, Göttingen, Germany). Supernatant was then centrifuged at 50000 rcf at 4°C for at least 1 h. The medium was exhausted and pellets were resuspended in 200 µl StemSpan serum free expansion medium (SFEM) (Cat. Nr. 09650, Stem Cell Technologies, Vancouver Canada), aliquoted in 10 µl vials and frozen at – 80 °C.

7.2.2 Virus Titration

Virus titer was determined by transducing NIH-3T3 (mouse embryonic fibroblasts) and quantifying infected cells with a fluorescent marker. NIH-3T3 cells were kept in DMEM/10% FCS and seeded at a density $1.5 \cdot 10^4$ cells per well of a 24-well plate (Cat. Nr. 131068, Thermo Scientific, Waltham, USA). The next day a 1:10 dilution series of lentivirus was prepared between 10^{-1} and 10^{-4} in DMEM/10% FCS. Cells were infected with each dilution in triplicates. 2 days later, medium was removed, cells were trypsinized (Cat. Nr. 25300-054, Invitrogen, Karlsruhe, Germany), counted and analyzed on a FACSCalibur or FACSAriaI (Beckton Dickinson, San Jose, USA). The amount of positive cells from dilutions which resulted in 3% - 30% positive cells was used for titer calculation.

7.2.3 Virus Infection

Cells were sorted, counted and cultured in 100 µl SFEM containing 100 ng/ml SCF (Cat. Nr. 250-03), 10 ng/ml IL-3 (Cat. Nr. 213-13) and 10 ng/ml IL-6 (Cat. Nr. 200-06, all PeproTech, Hamburg, Germany) in a well of a 96-well plate. Lentivirus was added at a multiplicity of infection (MOI) of 60 – 100. Cells were incubated at 37 °C for 24 h before further usage.

7.3 Isolation of Primary Cells

7.3.1 Mouse Lines

Mouse lines used for experiments included C57BL/6J wildtype mice from an in-house breeding facility, PU.1eYFP knock-in mice (Kirstetter et al., 2006) and self-generated Gata1mCHERRY knock-in mice made from JM8.N4 (C57BL/6N) ES cells (Pettitt et al., 2009). PU.1eYFP and Gata1mCHERRY mice were crossed in order to generate double knock-in mice.

7.3.2 Fetal Liver Preparation

Fetal livers from C57BL/6 and Gata1mCHERRY E14.5 day embryos were isolated, pooled and singularized by vigorous pipetting of the cell suspension. Cells were washed with DPBS, centrifuged at 300 rcf for 5 min at 4 °C and resuspended in 50 µl (per fetal liver) FACS buffer containing DPBS, 10% FCS, 0.1% sodium azide (NaN_3) (Cat. Nr. S2002, Sigma-Aldrich, Taufkirchen, Germany) and 10 mM EDTA.

7.3.3 Bone Marrow Preparation

For analysis and sorting of mononuclear cells from murine bone marrow, femurs and tibiae of adult mice were isolated, flushed with 6 ml DPBS and isolated from the interphase of a Histopaque-1083 Ficoll gradient (Cat. Nr., 10831, Sigma-Aldrich, Taufkirchen, Germany) after centrifugation for 15 min at 490 rcf at room temperature with minimal acceleration and no brake.

Alternatively, femurs, tibiae, humeri, hip bones and vertebrae from adult mice were isolated and muscle tissue was removed with a scalpel. Bones were crushed in cold 2% FCS in DPBS and filtered through a 40 µm filter (Cat. Nr. 352340, Becton Dickinson, Heidelberg, Germany) in a total volume of 50 ml. The cell suspension was centrifuged at 300 rcf for 5 min at 4 °C. If cells were prepared for sorting, the pellet was resuspended in twice the volume of ACK lysis buffer (Cat. Nr. 10-548E, Lonza, Basel, Switzerland) and incubated for 2 min at room temperature. The volume was increased to 50 ml with DPBS and cells were counted in a hemocytometer.

7.3.4 Staining of Primary Cells

Isolated cells were centrifuged at 300 rcf for 5 min at 4 °C and resuspended in 200 µl (per mouse) FACS buffer. For analysis and sorting of HSPCs, 1 µl biotinylated lineage-antibody was added per 10^7 cells and incubated for 20 min on ice. Lineage-antibodies included CD3 (Cat. Nr. 13-0031-85), CD19 (Cat. Nr. 13-0193-85), B220 (Cat. Nr. 13-0452-86), Gr1 (Cat. Nr. 13-5931-85), Mac-1 (Cat. Nr. 13-0112-85), Ter119 (Cat. Nr. 13-5921-85) and CD41 (Cat.

Nr. 13-0411-85, all eBioscience, San Diego, USA). If lineage-specific antigens were stained with a color, they were omitted in the lineage-mix. CD41 was omitted when MegE progenitor cells were sorted. If cells were prepared for sorting, a magnetic depletion of lineage positive cells was performed: Cells were washed, centrifuged and resuspended in 200 µl (per mouse) FACS buffer. Streptavidin-coated magnetic beads (Cat. Nr. HP57.1, Roth, Karlsruhe, Germany) were added to the cells (1 µl per 10^7 cells) and incubated for 7 min on ice. Cells were then incubated for 7 min on an EasySep magnet (Cat. Nr. 18001, Stem Cell Technologies, Vancouver, Canada) and lineage-negative cells were decanted. Cells were washed, counted, centrifuged and resuspended in 100 µl (per mouse) FACS buffer. Fluorescence labeled antibodies and fluorescence labeled streptavidin (2-3 µl per 10^7 cells) were added to the cells and incubated for at least 30 min on ice. Antibodies used for staining included CD16/32-AlexaFluor 700 (Cat. Nr. 56-0161-82), CD19-APC (Cat. Nr. 17-0191-81), CD34-eFluor660 (Cat. Nr. 50-0341-82), CD41-PerCP-eFluor710 (Cat. Nr. 46-0411-82), CD48-PerCP-eFluor710 (Cat. Nr. 46-0481-82), CD71-PE (Cat. Nr. 12-0711-81), CD71-PerCP-eFluor710 (Cat. Nr. 46-0711-82), B220-PE (Cat. Nr. 12-0452-83), Flt3-PE (Cat. Nr. 12-1351-83), c-Kit-PE-Cy7 (Cat. Nr. 25-1171-82), c-Kit-APC-eFluor780 (Cat. Nr. 47-1172-82), Gr-1-PE-Cy7 (Cat. Nr. 25-5931-81), Mac1-PE-Cy7 (Cat. Nr. 25-0112-81), Sca-1-PerCP-Cy5.5 (Cat. Nr. 45-5981-82), Ter119-PE (Cat. Nr. 12-5921-81), Ter119-APC-eFluor780 (Cat. Nr. 47-5921-82, all eBiocience, San Diego, USA), CD3-Biotin, CD16/32-APC (Cat. Nr. 101326), CD105-APC (Cat. Nr. 120414,), CD150-PE (Cat. Nr. 115904) and Sca-1-PB (Cat. Nr. 108120, all Biolegend, San Diego, USA). Streptavidin for staining biotin-labeled lineage mix was either APC-eFluor780 (Cat. Nr. 47-4317-82, eBioscience, San Diego, USA), PB (Cat. Nr. S-11222) or AlexaFluor 430 (Cat. Nr. S-11237, both Invitrogen, Karlsruhe, Germany). After staining, cells were washed, centrifuged, resuspended in FACS Buffer (200 µl per mouse) and filtered through a 35 µm filter (Cat. Nr. 352235, Becton Dickinson, Heidelberg, Germany) into polypropylene round-bottom tubes (Cat. Nr. 352063, Becton Dickinson, Heidelberg, Germany).

7.3.5 Flow Cytometric Analysis and Sorting of Primary Cells

All flow cytometric analysis and sorting was performed on a FACSAriaI or FACSAriaIII (Becton Dickinson, Heidelberg, Germany) equipped with a 405 nm, 488 nm and 633 nm Laser (both machines) and a 561 nm laser (only FACSAriaIII) using a 70 µm nozzle and

manually adjusted compensations necessary for multi-color staining. Filters for detection of fluorescent colors with the 405 nm laser were 450/40 (PB) and 530/30 (AlexaFluor 430). Filters for the 488 nm laser were 530/30 (eYFP), 576/15 (PE), 695/40 (PE-Cy5.5) and 780/60 (PE-Cy7). Filters for the 561 nm laser were 582/15 (PE), 610/20 (mCHERRY), 670/14 (PE-Cy5) and 780/60 (PE-Cy7). Filters for the 633 nm laser were 660/20 (APC, AlexaFluor 647, eFluor660), 730/45 (AlexaFluor 700) and 780/60 (APC-Cy7, APC-eFluor780). Cells were sorted in 'purity' or '4-way-sort' mode into 4 °C cold SFEM or Iscove's Modified Dulbecco's Medium (IMDM) (Cat. Nr. P04-20451S1, PAN Biotech, Aidenbach, Germany) containing 100 ng/ml SCF and additionally 5 U/ml EPO (Cat. Nr. C-60023, Promokine, Heidelberg, Germany) for sorting of erythroid cells. A sorting purity of at least 95% was always confirmed by reanalysis of sorted populations. Flow cytometric was analyzed using FACSDiva software version 6.1.3 (Becton Dickinson, Heidelberg, Germany).

7.3.6 Blood Counts

Peripheral blood was collected by cardiac puncture immediately after sacrifice of mice and transferred to EDTA-coated tubes (Cat. Nr. 078035, Kabe Labortechnik, Nümbrecht-Elsenroth, Germany). Blood counts were quantified by Sebastian Kaidel (Institute of Experimental Genetics, Helmholtz Center Munich) on an Abc Animal Blood Counter (scil animal care company, Viernheim, Germany).

7.4 Cell Culture

7.4.1 Liquid Culture

Freshly sorted cells were cultured in SFEM containing 100 ng/ml SCF, 10 ng/ml IL-3, 10 ng/ml IL-6, 5 U/ml EPO, 100 ng/ml TPO (Cat. Nr. 315-14, PeproTech, Hamburg, Germany) and 100 U/ml and 100 µg/ml penicillin/streptomycin, respectively. Cell culture dishes for population culture were 24-well plates and 1536-well glass bottom dishes (Cat. Nr. 783892, Greiner Bio-One, Frickenhausen, Germany) for single-cell cultures. For time-lapse movies, cells were cultured in fibronectin-coated (Cat. Nr IHMFBN, Innovative Research, Novi,

USA) ibidi µ-slides VI$^{0.4}$ (Cat. Nr. 80606, Ibidi, Martinsried, Germany). Cells were incubated at 37 °C and 5% CO_2.

7.4.2 Live Continuous In Culture Antibody Staining

CD16/32 antibody (Cat. Nr. 553142, BD Pharmingen, Heidelberg, Germany) was self-labeled with AlexaFluor 647 Labeling Kit (Cat. Nr. A-20186, Invitrogen, Karlsruhe, Germany) according to manufacturer's protocol. CD16/32-AlexaFluor 647 antibody was added to cell cultures at a final concentration of 10 – 50 ng/ml.

7.4.3 Colony Assay

Freshly isolated bone marrow or freshly sorted cells were sorted, counted and seeded at suitable densities in 35 mm dishes (Cat. Nr. 430165, Corning, Amsterdam, Netherlands) with MethoCult GF M3434 colony assay medium (Stem Cell Technologies, Vancouver, Canada) containing 100 U/ml and 100 µg/m penicillin/streptomycin. Colonies with at least 30 cells (or at least 1 megakaryocyte) were scored after 4 – 14 days according to cell morphology and colony morphology. For staining of erythroid colonies, a benzidine staining was performed: Stock solution was prepared by dissolving 0.1 g dibenzidine hydrochloride (Cat. Nr. B3383, Sigma-Aldrich, Taufkirchen, Germany) in 1.46 ml glacial acetic acid (Cat. Nr. 1.00063.1000, Merck, Darmstadt, Germany) and 48.54 ml H_2O bidest. In order to prepare the working solution 1 ml stock solution were mixed with 20 µl 30% hydrogen peroxide (H_2O_2) (Cat. Nr. 216763, Sigma-Aldrich, Taufkirchen, Germany). A few drops of the working solution were added on top of the colony assay medium. The hemoglobin of erythroid colonies stained blue after 2 – 3 minutes.

7.5 Cytospin

7.5.1 Spinning

Cells were spun down on object slides using Hettich cytospin equipment (Cat. Nr 1662, 1668, 5280, Hettich, Tuttlingen, Germany) and a Rotanta 460 R centrifuge (Cat. Nr. 5660, Hettich, Tuttlingen, Germany). Up to 1 ml of cell suspension were centrifuged at 270 rcf for 3 min at room temperature. Medium was exhausted until 100 µl were left and the cyto-chamber was removed. Remaining medium was removed by centrifugation at 1100 rcf for 1 min. Cells were air-dried before staining.

7.5.2 Staining and Analysis

Cytoplasm and granulae of spun cells were stained for 4 min with May-Gruenwald-Solution (Cat. Nr. T863.1, Roth, Karlsruhe, Germany). Cells were washed twice with H_2O bidest and nuclei were stained for 16 min with a 5% Giemsa-Solution (Cat. Nr. 1.09204.0500, Merck, Darmstadt, Germany). Cells were washed three times with H_2O bidest, air-dried and covered with Pertex mounting medium (Cat. Nr. PER20000, Medite, Burgdorf, Germany). Cytospins were analyzed using a 63x oil-immersion objective and an Axiovert 200M inverted microscope. Pictures were taken using the RGB camera AxioCam MRc5 and a 1x TV-adapter (all Zeiss, Hallbergmoos, Germany).

7.6 Western Blot

7.6.1 Cycloheximide Treatment

For comparison of biochemical stability of wildtype transcription factors and transcription factor fusions with fluorescent proteins cells were subjected to cycloheximide treatment which blocks proteinbiosynthesis. For comparison of PU.1/PU.1eYFP stability, PU.1eYFP$^+$ progenitor cells from PU.1/PU.1eYFP heterozygous mice were sorted, cultured in IMDM containing 100 ng/ml SCF, 10 ng/ml IL-3, 10 ng/ml IL-6 and 50 µM cycloheximide (Cat. Nr.

C4859-1ML, Sigma-Aldrich, Taufkirchen, Germany) and distributed into equal aliquots. For comparison of Gata1/Gata1mCHERRY stability, equal numbers of E14.5 fetal liver cells from wildtype C57BL/6 and Gata1mCHERRY embryos were mixed and cultured in IMDM containing 100 ng/ml SCF, 10 ng/ml IL-3, 10 ng/ml IL-6, 5 U/ml EPO and 50 µM cycloheximide. At designated time points, cells were washed, spun down at 300 rcf for 5 min at 4 °C and resuspended in 30 µl Laemmli-buffer (Laemmli, 1970). Cell lysates were boiled at 100 °C for 5 minutes and stored at – 20 °C until further usage.

7.6.2 SDS-Polyacrylamidgelectrophoresis

Proteins from cell lysates and recombinant GFP (Cat. Nr. 632373, Clontech, Mountain View, USA) or mCHERRY (Cat. Nr. ABIN412973, antibodies-online.com, Aachen, Germany) were separated by SDS-polyacrylamidgelectrophoresis. Gels were prepared and mounted according to manufacturer's instructions (Cat. Nr. 165-8001, Bio-Rad, Munich, Germany). Collecting gel buffer was 0.5 M TRIS (pH = 6.8) with 0.4% SDS. Resolving gel buffer was 1.3 M TRIS (pH = 8.8) with 1.0% SDS. The resolving gel was prepared by mixing 3 ml of resolving gel buffer with 5 ml H_2O and 4 ml 30% bisacrylamid (Cat. Nr. 3029.3, Roth, Karlsruhe, Germany). Polymerisation was started by adding 16 µl tetramethylethylendiamin (TEMED) (Cat. Nr. 2367.3, Roth, Karlsruhe, Germany) and 52 µl 25% ammonium persulfate (APS) (Cat. Nr. GENOP011-B, VWR, Darmstadt, Germany). Gel was poured and layered with 100% isopropanol. After polymerisation of the resolving gel, the 5% collecting gel was prepared: 1.25 ml collecting gel buffer were mixed with 2.92 ml H_2O bidest and 0.83 ml 30% bisacrylamid. At last, 10 µl TEMED and 25 µl APS were added. After removal of isopropanol, the collecting gel was poured and supplied with a comb generating 10 wells. Gels were run in buffer consisting of 25 mM TRIS, 19.2 mM glycine (Cat. Nr. T873.2, Roth, Karlsruhe, Germany) and 0.5% SDS. Gels were run at 80 V until proteins reached the resolving gel and voltage was increased to 100 V. PAGERuler (Cat. Nr. 10398469, Thermo Scientific, Waltham, USA) was used as a protein standard.

7.6.3 Western Blotting

After separation of proteins, gels were blotted by wet blotting onto a PVDF membrane according to manufacturer's instructions (Cat. Nr. 162-0177, Bio-Rad, Munich, Germany). Transfer buffer was 0.16 M glycine and 25 mM TRIS. PVDF membrane was wetted in methanol (Cat. Nr. 1.06009.2500, Merck, Darmstadt, Germany) prior to blotting. Gel, membrane, sponges and filter papers were preincubated in transfer buffer for 10 minutes. Blotting was performed at 100V for 2 hours or at 30V overnight with a – 20 °C thermal pack or at 4 °C.

7.6.4 Protein Detection and Processing

All protein detection was performed in TRIS-buffered saline (TBS) consisting of 20 mM TRIS and 0.14 M NaCl (pH = 7.5) under light shaking. PVDF membranes were blocked with 5% milk powder (Cat. Nr. T145.1, Roth, Karlsruhe, Germany) in TBS for 1 h at room temperature. The following primary antibodies were used: α-PU.1 (Cat. Nr. 2258, Cell Signaling, Frankfurt am Main, Germany), α-Gata1 (Cat. Nr. sc-265, Santa Cruz, Heidelberg, Germany), α-GFP (Cat. Nr. 11814460001, Roche, Penzberg, Germany) and α-mCHERRY (Cat. Nr. ab125096, Abcam, Cambridge, United Kingdom). Primary antibodies were incubated overnight at 4 °C in a dilution of 1:1000 in 5% milk (Cat. Nr. T145.2, Roth, Karlsruhe, Germany) in TBS. Membranes were washed 3 times for 5 min in 0.1% Tween-20 (Cat. Nr. 9127.1, Roth, Karlsruhe, Germany) in TBS (TBS-T). The following horseradish peroxidase linked secondary antibodies were used: α-mouse IgG (Cat. Nr. NA931-1ML), α-rat IgG (Cat. Nr. NA935) and α-rabbit IgG (Cat. Nr. NA934-1ML, all GE Healthcare Life Sciences, Freiburg, Germany). Secondary antibodies were incubated for 1 h at room temperature at a dilution of 1:10000 – 1:15000 in 5% milk in TBS and excess antibodies were removed by washing 6 times for 5 min in TBS-T. Specific proteins were detected by chemiluminescent signals provided by an enzymatic reaction. Substrate was provided by the ECL+ Kit (Cat. Nr. RPN2132, GE Healthcare, Freiburg, Germany) and used according to manufacturer's instructions. Signals were detected using Fuji medical X-ray films (Cat. Nr. RF11, A. Hartenstein, Würzburg, Germany). Films were processed on an AGFA Curix 60 Developer Machine. Films were scanned and signals were quantified using ImageJ (http://rsbweb.nih.gov/ij/) software drawing manual gates around bands and subtracting the

average of the same background area above and below the bands. The mean out of three manual drawings was used for further calculations.

7.6.5 Calculation of PU.1eYFP and Gata1mCHERRY Protein Numbers

Calculation of PU.1eYFP and Gata1mCHERRY protein numbers will be described elsewhere (Schwarzfischer et al., manuscript in preparation) and was not part of this thesis. Briefly, calculated protein data from Western blot analysis of defined amount of PU.1eYFP^{+} or Gata1mCHERRY^{+} cells were assigned to other defined HSPC populations by their average fold-changes in flow cytometry. Error estimates were derived using propagation of uncertainty. Average HSC protein numbers for PU.1eYFP in movies were mapped to the signal intensities of HSCs in the first three timepoints of the movies. Protein numbers for Gata1mCHERRY were estimated by mapping the average protein number of Gata1mCHERRY^{mid} cells to the average intensity of cells which have been marked as Gata1mCHERRY-positive in the time-lapse movies. A cell was marked as Gata1mCHERRY positive if its expression exceeded twice the negative gate for more than 5 consecutive time points.

7.7 Time-Lapse Imaging

7.7.1 Movie Acquisition

Time-lapse movies were acquired on the inverted epifluorescence microscope Zeiss Axio Observer (Zeiss, Hallbergmoos, Germany) equipped with AxioVision Software 4.9. A self-written macro was used to control hardware equipment. Cells were constantly kept at 37 °C in a custom-made Plexiglas housing by the Heating Unit XL S (Cat. Nr. 411857-9030-000, Zeiss, Hallbergmoos, Germany) and directly supplied with pre-mixed gas consisting of 5% CO_2, 5% O_2 and 90% N_2 (Praxair, Düsseldorf, Germany). Images were acquired with a 10x Fluar objective (Cat. Nr. 440135-0000-000, Zeiss, Hallbergmoos, Germany) and a 0.63x TV-Adapter (Cat. Nr., 426113-0000-000 Zeiss, Hallbergmoos, Germany). The brightfield lightsource was a halogen lamp and images were acquired every 1.5 min with an Axiocam HRm (Cat. Nr. 426511-9901-000, Zeiss, Hallbergmoos, Germany) camera at a resolution of

1388 x 1040 pixels in lossless TIF or PNG-format. Fluorescent light was provided by a HXP 120 light source (Cat. Nr. 423013-9010-000, Zeiss, Hallbergmoos, Germany) with a Osram HXP-R 120W/45C VIS bulb (Cat. Nr. 882772, Osram, Munich, Germany). Filters for excitation/emission of different fluorophores were 46 HE (Cat. Nr. 489046-9901-000, Zeiss, Hallbergmoos, Germany) for eYFP, 43 HE (Cat. Nr. 489043-9901-000, Zeiss, Hallbergmoos, Germany) for mCHERRY and AHF Cy5 (Cat. Nr. F46-006, AHF Analysentechnik, Tübingen, Germany) for AlexaFluor 647. Excitation times were 1.5 s for eYFP, 0.8 s for mCHERRY and 0.6 s for AlexaFluor 647. Hardware autofocus (Cat. Nr. 410133-0506-000, Zeiss, Hallbergmoos, Germany) was applied just prior to eYFP image acquisition. EYFP and mCHERRY images were acquired every 30 min and AlexaFluor 647 images were acquired every 4 hours.

7.7.2 Tracking Software

Time-lapse data was analysed using self-written tracking software (TTT, Timm's Tracking Tool) (Eilken et al., 2009; Rieger et al., 2009) on Fujitsu Siemens (Munich, Germany) workstations with up to 96 GB RAM. Briefly, the software allows manual tracking of individual colonies at the single-cell level which results in cell genealogies with exact information about many parameters, e.g. division, cell cycle length, apoptosis or cell movement. Onsets of Gata1mCHERRY and CD16/32 expression as well as appearance of morphological distinct megakaryocytes were manually annotated. Any cell with insecure identity (loss of single cell data) was excluded from analysis.

7.7.3 Image Quantification Software

Background and position-dependent gain from every acquired fluorescent picture was quantified using a recently developed algorithm (Schwarzfischer et al., 2011). Based on corrected backgrounds, fluorescent signals were quantified using new interactive software (QTF) which allows inspection and manual correction of every single fluorescent data point (developed in collaboration with Prof. Fabian Theis Group, Institute of Computational Biology Helmholtz Center Munich, manuscript in preparation). Only data points that were manually inspected were used for further analysis.

7.8 Statistical Analysis

A one-way multivariate analysis of variance (MANOVA) was used for comparing C57BL/6 and PU.1eYFP/Gata1mCHERRY bone marrow composition and blood counts. A Wilcoxon rank sum test was used for comparing *in vitro* colony forming potential of C57BL/6 and PU.1eYFP/Gata1mCHERRY bulk bone marrow. A statistical test for differences between paired time-resolved observations was used for comparing biochemical stability of wildtype PU.1/Gata1 and PU.1eYFP/Gata1mCHERRY, respectively (Brand et al., 2013). A significance level of 5% was used for every test.

8 References

Adams, G. B., K. T. Chabner, I. R. Alley, D. P. Olson, Z. M. Szczepiorkowski, M. C. Poznansky, C. H. Kos, M. R. Pollak, E. M. Brown and D. T. Scadden (2006). "Stem cell engraftment at the endosteal niche is specified by the calcium-sensing receptor." Nature **439**(7076): 599-603.

Adams, G. B. and D. T. Scadden (2006). "The hematopoietic stem cell in its place." Nat Immunol **7**(4): 333-337.

Adolfsson, J., O. J. Borge, D. Bryder, K. Theilgaard-Monch, I. Astrand-Grundstrom, E. Sitnicka, Y. Sasaki and S. E. Jacobsen (2001). "Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity." Immunity **15**(4): 659-669.

Adolfsson, J., R. Mansson, N. Buza-Vidas, A. Hultquist, K. Liuba, C. T. Jensen, D. Bryder, L. Yang, O. J. Borge, L. A. Thoren, K. Anderson, E. Sitnicka, Y. Sasaki, M. Sigvardsson and S. E. Jacobsen (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., X. He, J. Chen, H. Iwasaki, C. Niu, B. Steenhard, J. Zhang, J. Haug and L. Li (2003). "Transcriptional accessibility for genes of multiple tissues and hematopoietic lineages is hierarchically controlled during early hematopoiesis." Blood **101**(2): 383-389.

Akashi, K., D. Traver, T. Miyamoto and I. L. Weissman (2000). "A clonogenic common myeloid progenitor that gives rise to all myeloid lineages." Nature **404**(6774): 193-197.

Arai, F., A. Hirao, M. Ohmura, H. Sato, S. Matsuoka, K. Takubo, K. Ito, G. Y. Koh and T. Suda (2004). "Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche." Cell **118**(2): 149-161.

Arinobu, Y., S. Mizuno, Y. Chong, H. Shigematsu, T. Iino, H. Iwasaki, T. Graf, R. Mayfield, S. Chan, P. Kastner and K. Akashi (2007). "Reciprocal activation of GATA-1 and PU.1 marks initial specification of hematopoietic stem cells into myeloerythroid and myelolymphoid lineages." Cell Stem Cell **1**(4): 416-427.

Ashman, L. K. (1999). "The biology of stem cell factor and its receptor C-kit." Int J Biochem Cell Biol **31**(10): 1037-1051.

Attema, J. L., P. Papathanasiou, E. C. Forsberg, J. Xu, S. T. Smale and I. L. Weissman (2007). "Epigenetic characterization of hematopoietic stem cell differentiation using miniChIP and bisulfite sequencing analysis." Proc Natl Acad Sci U S A **104**(30): 12371-12376.

Back, J., D. Allman, S. Chan and P. Kastner (2005). "Visualizing PU.1 activity during hematopoiesis." Exp Hematol **33**(4): 395-402.

Back, J., A. Dierich, C. Bronn, P. Kastner and S. Chan (2004). "PU.1 determines the self-renewal capacity of erythroid progenitor cells." Blood **103**(10): 3615-3623.

Bain, G., E. C. Maandag, D. J. Izon, D. Amsen, A. M. Kruisbeek, B. C. Weintraub, I. Krop, M. S. Schlissel, A. J. Feeney, M. van Roon and et al. (1994). "E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements." Cell **79**(5): 885-892.

Barndt, R. J., M. Dai and Y. Zhuang (2000). "Functions of E2A-HEB heterodimers in T-cell development revealed by a dominant negative mutation of HEB." Mol Cell Biol **20**(18): 6677-6685.

Bertrand, J. Y., N. C. Chi, B. Santoso, S. Teng, D. Y. Stainier and D. Traver (2010). "Haematopoietic stem cells derive directly from aortic endothelium during development." Nature **464**(7285): 108-111.

Bertrand, J. Y., A. D. Kim, E. P. Violette, D. L. Stachura, J. L. Cisson and D. Traver (2007). "Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo." Development **134**(23): 4147-4156.

Boyes, J., P. Byfield, Y. Nakatani and V. Ogryzko (1998). "Regulation of activity of the transcription factor GATA-1 by acetylation." Nature **396**(6711): 594-598.

Brand, T., I. Kondofersky, K. Ehlers, W. Römisch-Margl, A. Bangert, F. Stückler, J. Krumsiek, A. Artati, C. Prehn, J. Adamski, G. Kastenmüller, C. Fuchs, F. J. Theis, H. Hauner and H. Laumen (2013). "Effect of dietary standardization on postprandial metabolomics profiles in healthy individuals." (Manuscript in preparation).

Burda, P., P. Laslo and T. Stopka (2010). "The role of PU.1 and GATA-1 transcription factors during normal and leukemogenic hematopoiesis." Leukemia **24**(7): 1249-1257.

Butler, J. S. and S. Y. Dent (2013). "The role of chromatin modifiers in normal and malignant hematopoiesis." Blood **121**(16): 3076-3084.

Calvi, L. M., G. B. Adams, K. W. Weibrecht, J. M. Weber, D. P. Olson, M. C. Knight, R. P. Martin, E. Schipani, P. Divieti, F. R. Bringhurst, L. A. Milner, H. M. Kronenberg and D. T. Scadden (2003). "Osteoblastic cells regulate the haematopoietic stem cell niche." Nature **425**(6960): 841-846.

Cantor, A. B. and S. H. Orkin (2001). "Hematopoietic development: a balancing act." Curr Opin Genet Dev **11**(5): 513-519.

Chang, H. H., M. Hemberg, M. Barahona, D. E. Ingber and S. Huang (2008). "Transcriptome-wide noise controls lineage choice in mammalian progenitor cells." Nature **453**(7194): 544-547.

Chen, H., D. Ray-Gallet, P. Zhang, C. J. Hetherington, D. A. Gonzalez, D. E. Zhang, F. Moreau-Gachelin and D. G. Tenen (1995). "PU.1 (Spi-1) autoregulates its expression in myeloid cells." Oncogene **11**(8): 1549-1560.

Chickarmane, V., T. Enver and C. Peterson (2009). "Computational modeling of the hematopoietic erythroid-myeloid switch reveals insights into cooperativity, priming, and irreversibility." PLoS Comput Biol **5**(1): e1000268.

Choe, K. S., F. Radparvar, I. Matushansky, N. Rekhtman, X. Han and A. I. Skoultchi (2003). "Reversal of tumorigenicity and the block to differentiation in erythroleukemia cells by GATA-1." Cancer Res **63**(19): 6363-6369.

Chou, S. T., E. Khandros, L. C. Bailey, K. E. Nichols, C. R. Vakoc, Y. Yao, Z. Huang, J. D. Crispino, R. C. Hardison, G. A. Blobel and M. J. Weiss (2009). "Graded repression of PU.1/Sfpi1 gene transcription by GATA factors regulates hematopoietic cell fate." Blood **114**(5): 983-994.

Christensen, J. L. and I. L. Weissman (2001). "Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells." Proc Natl Acad Sci U S A **98**(25): 14541-14546.

Collavin, L., M. Gostissa, F. Avolio, P. Secco, A. Ronchi, C. Santoro and G. Del Sal (2004). "Modification of the erythroid transcription factor GATA-1 by SUMO-1." Proc Natl Acad Sci U S A **101**(24): 8870-8875.

Cook, W. D., B. J. McCaw, C. Herring, D. L. John, S. J. Foote, S. L. Nutt and J. M. Adams (2004). "PU.1 is a suppressor of myeloid leukemia, inactivated in mice by gene deletion and mutation of its DNA binding domain." Blood **104**(12): 3437-3444.

Cross, M. A. and T. Enver (1997). "The lineage commitment of haemopoietic progenitor cells." Curr Opin Genet Dev **7**(5): 609-613.

Cross, M. A., C. M. Heyworth, A. M. Murrell, E. O. Bockamp, T. M. Dexter and A. R. Green (1994). "Expression of lineage restricted transcription factors precedes lineage specific differentiation in a multipotent haemopoietic progenitor cell line." Oncogene **9**(10): 3013-3016.

Crossley, M., M. Merika and S. H. Orkin (1995). "Self-association of the erythroid transcription factor GATA-1 mediated by its zinc finger domains." Mol Cell Biol **15**(5): 2448-2456.

Crossley, M. and S. H. Orkin (1994). "Phosphorylation of the erythroid transcription factor GATA-1." J Biol Chem **269**(24): 16589-16596.

Crotta, S., S. Nicolis, A. Ronchi, S. Ottolenghi, L. Ruzzi, Y. Shimada, A. R. Migliaccio and G. Migliaccio (1990). "Progressive inactivation of the expression of an erythroid transcriptional factor in GM- and G-CSF-dependent myeloid cell lines." Nucleic Acids Res **18**(23): 6863-6869.

Dahl, R., J. C. Walsh, D. Lancki, P. Laslo, S. R. Iyer, H. Singh and M. C. Simon (2003). "Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBPalpha ratio and granulocyte colony-stimulating factor." Nat Immunol **4**(10): 1029-1036.

Davis, R. L., H. Weintraub and A. B. Lassar (1987). "Expression of a single transfected cDNA converts fibroblasts to myoblasts." Cell **51**(6): 987-1000.

Delassus, S., I. Titley and T. Enver (1999). "Functional and molecular analysis of hematopoietic progenitors derived from the aorta-gonad-mesonephros region of the mouse embryo." Blood **94**(5): 1495-1503.

Drexler, H. G. (1998). "Review of alterations of the cyclin-dependent kinase inhibitor INK4 family genes p15, p16, p18 and p19 in human leukemia-lymphoma cells." Leukemia **12**(6): 845-859.

Duff, C., K. Smith-Miles, L. Lopes and T. Tian (2012). "Mathematical modelling of stem cell differentiation: the PU.1-GATA-1 interaction." J Math Biol **64**(3): 449-468.

Durual, S., A. Rideau, S. Ruault-Jungblut, D. Cossali, P. Beris, V. Piguet and T. Matthes (2007). "Lentiviral PU.1 overexpression restores differentiation in myeloid leukemic blasts." Leukemia **21**(5): 1050-1059.

Eilken, H. M., S. Nishikawa and T. Schroeder (2009). "Continuous single-cell imaging of blood generation from haemogenic endothelium." Nature **457**(7231): 896-900.

Ema, H., Y. Morita, S. Yamazaki, A. Matsubara, J. Seita, Y. Tadokoro, H. Kondo, H. Takano and H. Nakauchi (2006). "Adult mouse hematopoietic stem cells: purification and single-cell assays." Nat Protoc **1**(6): 2979-2987.

Ema, H., H. Takano, K. Sudo and H. Nakauchi (2000). "In vitro self-renewal division of hematopoietic stem cells." J Exp Med **192**(9): 1281-1288.

Endele, M. and T. Schroeder (2012). "Molecular live cell bioimaging in stem cell research." Ann N Y Acad Sci **1266**: 18-27.

Enver, T., C. M. Heyworth and T. M. Dexter (1998). "Do stem cells play dice?" Blood **92**(2): 348-351; discussion 352.

Enver, T., M. Pera, C. Peterson and P. W. Andrews (2009). "Stem cell states, fates, and the rules of attraction." Cell Stem Cell **4**(5): 387-397.

Evans, T., M. Reitman and G. Felsenfeld (1988). "An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes." Proc Natl Acad Sci U S A **85**(16): 5976-5980.

Ferreira, R., K. Ohneda, M. Yamamoto and S. Philipsen (2005). "GATA1 function, a paradigm for transcription factors in hematopoiesis." Mol Cell Biol **25**(4): 1215-1227.

Filipowicz, W., S. N. Bhattacharyya and N. Sonenberg (2008). "Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?" Nat Rev Genet **9**(2): 102-114.

Forsberg, E. C., T. Serwold, S. Kogan, I. L. Weissman and E. Passegue (2006). "New evidence supporting megakaryocyte-erythrocyte potential of flk2/flt3+ multipotent hematopoietic progenitors." Cell **126**(2): 415-426.

Fujiwara, Y., C. P. Browne, K. Cunniff, S. C. Goff and S. H. Orkin (1996). "Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1." Proc Natl Acad Sci U S A **93**(22): 12355-12358.

Graf, T. (2002). "Differentiation plasticity of hematopoietic cells." Blood **99**(9): 3089-3101.

Graf, T. and T. Enver (2009). "Forcing cells to change lineages." Nature **462**(7273): 587-594.

Gupta, P., G. U. Gurudutta, D. Saluja and R. P. Tripathi (2009). "PU.1 and partners: regulation of haematopoietic stem cell fate in normal and malignant haematopoiesis." J Cell Mol Med **13**(11-12): 4349-4363.

Gutierrez, L., S. Tsukamoto, M. Suzuki, H. Yamamoto-Mukai, M. Yamamoto, S. Philipsen and K. Ohneda (2008). "Ablation of Gata1 in adult mice results in aplastic crisis, revealing its essential role in steady-state and stress erythropoiesis." Blood **111**(8): 4375-4385.

Heberlein, C., K. D. Fischer, M. Stoffel, J. Nowock, A. Ford, U. Tessmer and C. Stocking (1992). "The gene for erythropoietin receptor is expressed in multipotential hematopoietic and embryonal stem cells: evidence for differentiation stage-specific regulation." Mol Cell Biol **12**(4): 1815-1826.

Heyworth, C., S. Pearson, G. May and T. Enver (2002). "Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells." EMBO J **21**(14): 3770-3781.

Hsu, Y. C. and E. Fuchs (2012). "A family business: stem cell progeny join the niche to regulate homeostasis." Nat Rev Mol Cell Biol **13**(2): 103-114.

Hu, M., D. Krause, M. Greaves, S. Sharkis, M. Dexter, C. Heyworth and T. Enver (1997). "Multilineage gene expression precedes commitment in the hemopoietic system." Genes Dev **11**(6): 774-785.

Huang, S. (2009). "Non-genetic heterogeneity of cells in development: more than just noise." Development **136**(23): 3853-3862.

Huang, S., Y. P. Guo, G. May and T. Enver (2007). "Bifurcation dynamics in lineage-commitment in bipotent progenitor cells." Dev Biol **305**(2): 695-713.

Ishikawa-Ankerhold, H. C., R. Ankerhold and G. P. Drummen (2012). "Advanced fluorescence microscopy techniques--FRAP, FLIP, FLAP, FRET and FLIM." Molecules **17**(4): 4047-4132.

Ito, E., T. Toki, H. Ishihara, H. Ohtani, L. Gu, M. Yokoyama, J. D. Engel and M. Yamamoto (1993). "Erythroid transcription factor GATA-1 is abundantly transcribed in mouse testis." Nature **362**(6419): 466-468.

Iwasaki, H. and K. Akashi (2007). "Myeloid lineage commitment from the hematopoietic stem cell." Immunity **26**(6): 726-740.

Iwasaki, H., S. Mizuno, Y. Arinobu, H. Ozawa, Y. Mori, H. Shigematsu, K. Takatsu, D. G. Tenen and K. Akashi (2006). "The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages." Genes Dev **20**(21): 3010-3021.

Iwasaki, H., S. Mizuno, R. Mayfield, H. Shigematsu, Y. Arinobu, B. Seed, M. F. Gurish, K. Takatsu and K. Akashi (2005). "Identification of eosinophil lineage-committed progenitors in the murine bone marrow." J Exp Med **201**(12): 1891-1897.

Iwasaki, H., C. Somoza, H. Shigematsu, E. A. Duprez, J. Iwasaki-Arai, S. Mizuno, Y. Arinobu, K. Geary, P. Zhang, T. Dayaram, M. L. Fenyus, S. Elf, S. Chan, P. Kastner, C. S. Huettner, R. Murray, D. G. Tenen and K. Akashi (2005). "Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation." Blood **106**(5): 1590-1600.

Jing, L. and L. I. Zon (2011). "Zebrafish as a model for normal and malignant hematopoiesis." Dis Model Mech **4**(4): 433-438.

Jones, D. L. and A. J. Wagers (2008). "No place like home: anatomy and function of the stem cell niche." Nat Rev Mol Cell Biol **9**(1): 11-21.

Jones, S. A. (2005). "Directing transition from innate to acquired immunity: defining a role for IL-6." J Immunol **175**(6): 3463-3468.

Kaushansky, K. and W. J. Williams (2010). Williams hematology. New York, McGraw-Hill Medical.

Kiel, M. J., O. H. Yilmaz, T. Iwashita, O. H. Yilmaz, C. Terhorst and S. J. Morrison (2005). "SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells." Cell **121**(7): 1109-1121.

Kirstetter, P., K. Anderson, B. T. Porse, S. E. Jacobsen and C. Nerlov (2006). "Activation of the canonical Wnt pathway leads to loss of hematopoietic stem cell repopulation and multilineage differentiation block." Nat Immunol **7**(10): 1048-1056.

Kissa, K. and P. Herbomel (2010). "Blood stem cells emerge from aortic endothelium by a novel type of cell transition." Nature **464**(7285): 112-115.

Kobayashi, M., K. Nishikawa and M. Yamamoto (2001). "Hematopoietic regulatory domain of gata1 gene is positively regulated by GATA1 protein in zebrafish embryos." Development **128**(12): 2341-2350.

Kodama, H., M. Nose, S. Niida, S. Nishikawa and S. Nishikawa (1994). "Involvement of the c-kit receptor in the adhesion of hematopoietic stem cells to stromal cells." Exp Hematol **22**(10): 979-984.

Kodama, H., H. Sudo, H. Koyama, S. Kasai and S. Yamamoto (1984). "In vitro hemopoiesis within a microenvironment created by MC3T3-G2/PA6 preadipocytes." J Cell Physiol **118**(3): 233-240.

Kodama, H. A., Y. Amagai, H. Koyama and S. Kasai (1982). "A new preadipose cell line derived from newborn mouse calvaria can promote the proliferation of pluripotent hemopoietic stem cells in vitro." J Cell Physiol **112**(1): 89-95.

Kohler, A., V. Schmithorst, M. D. Filippi, M. A. Ryan, D. Daria, M. Gunzer and H. Geiger (2009). "Altered cellular dynamics and endosteal location of aged early hematopoietic progenitor cells revealed by time-lapse intravital imaging in long bones." Blood **114**(2): 290-298.

Kollet, O., A. Dar, S. Shivtiel, A. Kalinkovich, K. Lapid, Y. Sztainberg, M. Tesio, R. M. Samstein, P. Goichberg, A. Spiegel, A. Elson and T. Lapidot (2006). "Osteoclasts degrade endosteal components and promote mobilization of hematopoietic progenitor cells." Nat Med **12**(6): 657-664.

Kondo, M., I. L. Weissman and K. Akashi (1997). "Identification of clonogenic common lymphoid progenitors in mouse bone marrow." Cell **91**(5): 661-672.

Krantz, S. B. (1991). "Erythropoietin." Blood **77**(3): 419-434.

Kremers, G. J., S. G. Gilbert, P. J. Cranfill, M. W. Davidson and D. W. Piston (2011). "Fluorescent proteins at a glance." J Cell Sci **124**(Pt 2): 157-160.

Krumsiek, J., C. Marr, T. Schroeder and F. J. Theis (2011). "Hierarchical differentiation of myeloid progenitors is encoded in the transcription factor network." PLoS One **6**(8): e22649.

Kulessa, H., J. Frampton and T. Graf (1995). "GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblasts, and erythroblasts." Genes Dev **9**(10): 1250-1262.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-685.

Laiosa, C. V., M. Stadtfeld and T. Graf (2006). "Determinants of lymphoid-myeloid lineage diversification." Annu Rev Immunol **24**: 705-738.

Laslo, P., J. M. Pongubala, D. W. Lancki and H. Singh (2008). "Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system." Semin Immunol **20**(4): 228-235.

Laslo, P., C. J. Spooner, A. Warmflash, D. W. Lancki, H. J. Lee, R. Sciammas, B. N. Gantner, A. R. Dinner and H. Singh (2006). "Multilineage transcriptional priming and determination of alternate hematopoietic cell fates." Cell **126**(4): 755-766.

Lau, C. I., S. V. Outram, J. I. Saldana, A. L. Furmanski, J. T. Dessens and T. Crompton (2012). "Regulation of murine normal and stress-induced erythropoiesis by Desert Hedgehog." Blood **119**(20): 4741-4751.

Lee, E. C., D. Yu, J. Martinez de Velasco, L. Tessarollo, D. A. Swing, D. L. Court, N. A. Jenkins and N. G. Copeland (2001). "A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA." Genomics **73**(1): 56-65.

Lewandowski, D., V. Barroca, F. Duconge, J. Bayer, J. T. Van Nhieu, C. Pestourie, P. Fouchet, B. Tavitian and P. H. Romeo (2010). "In vivo cellular imaging pinpoints the role of reactive oxygen species in the early steps of adult hematopoietic reconstitution." Blood **115**(3): 443-452.

Li, Y., Y. Okuno, P. Zhang, H. S. Radomska, H. Chen, H. Iwasaki, K. Akashi, M. J. Klemsz, S. R. McKercher, R. A. Maki and D. G. Tenen (2001). "Regulation of the PU.1 gene by distal elements." Blood **98**(10): 2958-2965.

Li, Z., F. J. Godinho, J. H. Klusmann, M. Garriga-Canut, C. Yu and S. H. Orkin (2005). "Developmental stage-selective effect of somatically mutated leukemogenic transcription factor GATA1." Nat Genet **37**(6): 613-619.

Liew, C. W., K. D. Rand, R. J. Simpson, W. W. Yung, R. E. Mansfield, M. Crossley, M. Proetorius-Ibba, C. Nerlov, F. M. Poulsen and J. P. Mackay (2006). "Molecular analysis of the interaction between the hematopoietic master transcription factors GATA-1 and PU.1." J Biol Chem **281**(38): 28296-28306.

Lo Celso, C., H. E. Fleming, J. W. Wu, C. X. Zhao, S. Miake-Lye, J. Fujisaki, D. Cote, D. W. Rowe, C. P. Lin and D. T. Scadden (2009). "Live-animal tracking of individual haematopoietic stem/progenitor cells in their niche." Nature **457**(7225): 92-96.

Lowry, J. A. and J. P. Mackay (2006). "GATA-1: one protein, many partners." Int J Biochem Cell Biol **38**(1): 6-11.

Mansson, R., A. Hultquist, S. Luc, L. Yang, K. Anderson, S. Kharazi, S. Al-Hashmi, K. Liuba, L. Thoren, J. Adolfsson, N. Buza-Vidas, H. Qian, S. Soneji, T. Enver, M. Sigvardsson and S. E. Jacobsen (2007). "Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors." Immunity **26**(4): 407-419.

Martin, D. I., L. I. Zon, G. Mutter and S. H. Orkin (1990). "Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages." Nature **344**(6265): 444-447.

Mattick, J. S. (2007). "A new paradigm for developmental biology." J Exp Biol **210**(Pt 9): 1526-1547.

McKercher, S. R., B. E. Torbett, K. L. Anderson, G. W. Henkel, D. J. Vestal, H. Baribault, M. Klemsz, A. J. Feeney, G. E. Wu, C. J. Paige and R. A. Maki (1996). "Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities." EMBO J **15**(20): 5647-5658.

Metcalf, D. (2008). "Hematopoietic cytokines." Blood **111**(2): 485-491.

Metcalf, D., A. Dakic, S. Mifsud, L. Di Rago, L. Wu and S. Nutt (2006). "Inactivation of PU.1 in adult mice leads to the development of myeloid leukemia." Proc Natl Acad Sci U S A **103**(5): 1486-1491.

Mikkola, I., B. Heavey, M. Horcher and M. Busslinger (2002). "Reversion of B cell commitment upon loss of Pax5 expression." Science **297**(5578): 110-113.

Miyamoto, T. and K. Akashi (2005). "Lineage promiscuous expression of transcription factors in normal hematopoiesis." Int J Hematol **81**(5): 361-367.

Miyamoto, T., H. Iwasaki, B. Reizis, M. Ye, T. Graf, I. L. Weissman and K. Akashi (2002). "Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment." Dev Cell **3**(1): 137-147.

Mizuguchi, H., Z. Xu, A. Ishii-Watabe, E. Uchida and T. Hayakawa (2000). "IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector." Mol Ther **1**(4): 376-382.

Monteiro, R., C. Pouget and R. Patient (2011). "The gata1/pu.1 lineage fate paradigm varies between blood populations and is modulated by tif1gamma." EMBO J **30**(6): 1093-1103.

Moore, K. A., H. Ema and I. R. Lemischka (1997). "In vitro maintenance of highly purified, transplantable hematopoietic stem cells." Blood **89**(12): 4337-4347.

Moreau-Gachelin, F., A. Tavitian and P. Tambourin (1988). "Spi-1 is a putative oncogene in virally induced murine erythroleukaemias." Nature **331**(6153): 277-280.

Moreau-Gachelin, F., F. Wendling, T. Molina, N. Denis, M. Titeux, G. Grimber, P. Briand, W. Vainchenker and A. Tavitian (1996). "Spi-1/PU.1 transgenic mice develop multistep erythroleukemias." Mol Cell Biol **16**(5): 2453-2463.

Morrison, S. J. and A. C. Spradling (2008). "Stem cells and niches: mechanisms that promote stem cell maintenance throughout life." Cell **132**(4): 598-611.

Mossadegh-Keller, N., S. Sarrazin, P. K. Kandalla, L. Espinosa, E. R. Stanley, S. L. Nutt, J. Moore and M. H. Sieweke (2013). "M-CSF instructs myeloid lineage fate in single haematopoietic stem cells." Nature **497**(7448): 239-243.

Mueller, B. U., T. Pabst, M. Osato, N. Asou, L. M. Johansen, M. D. Minden, G. Behre, W. Hiddemann, Y. Ito and D. G. Tenen (2002). "Heterozygous PU.1 mutations are associated with acute myeloid leukemia." Blood **100**(3): 998-1007.

Nakano, T., H. Kodama and T. Honjo (1994). "Generation of lymphohematopoietic cells from embryonic stem cells in culture." Science **265**(5175): 1098-1101.

Nakano, T., H. Kodama and T. Honjo (1996). "In vitro development of primitive and definitive erythrocytes from different precursors." Science **272**(5262): 722-724.

Nerlov, C. and T. Graf (1998). "PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors." Genes Dev **12**(15): 2403-2412.

Nerlov, C., E. Querfurth, H. Kulessa and T. Graf (2000). "GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription." Blood **95**(8): 2543-2551.

Nilsson, S. K., H. M. Johnston and J. A. Coverdale (2001). "Spatial localization of transplanted hemopoietic stem cells: inferences for the localization of stem cell niches." Blood **97**(8): 2293-2299.

Nishikawa, K., M. Kobayashi, A. Masumi, S. E. Lyons, B. M. Weinstein, P. P. Liu and M. Yamamoto (2003). "Self-association of Gata1 enhances transcriptional activity in vivo in zebra fish embryos." Mol Cell Biol **23**(22): 8295-8305.

Nutt, S. L., B. Heavey, A. G. Rolink and M. Busslinger (1999). "Commitment to the B-lymphoid lineage depends on the transcription factor Pax5." Nature **401**(6753): 556-562.

Nutt, S. L., D. Metcalf, A. D'Amico, M. Polli and L. Wu (2005). "Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors." J Exp Med **201**(2): 221-231.

O'Connell, R. M. and D. Baltimore (2012). "MicroRNAs and hematopoietic cell development." Curr Top Dev Biol **99**: 145-174.

Okuno, Y., G. Huang, F. Rosenbauer, E. K. Evans, H. S. Radomska, H. Iwasaki, K. Akashi, F. Moreau-Gachelin, Y. Li, P. Zhang, B. Gottgens and D. G. Tenen (2005). "Potential autoregulation of transcription factor PU.1 by an upstream regulatory element." Mol Cell Biol **25**(7): 2832-2845.

Orkin, S. H. (2000). "Diversification of haematopoietic stem cells to specific lineages." Nat Rev Genet **1**(1): 57-64.

Orkin, S. H. and L. I. Zon (2008). "Hematopoiesis: an evolving paradigm for stem cell biology." Cell **132**(4): 631-644.

Osawa, M., K. Hanada, H. Hamada and H. Nakauchi (1996). "Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell." Science **273**(5272): 242-245.

Pettitt, S. J., Q. Liang, X. Y. Rairdan, J. L. Moran, H. M. Prosser, D. R. Beier, K. C. Lloyd, A. Bradley and W. C. Skarnes (2009). "Agouti C57BL/6N embryonic stem cells for mouse genetic resources." Nat Methods **6**(7): 493-495.

Pevny, L., C. S. Lin, V. D'Agati, M. C. Simon, S. H. Orkin and F. Costantini (1995). "Development of hematopoietic cells lacking transcription factor GATA-1." Development **121**(1): 163-172.

Pevny, L., M. C. Simon, E. Robertson, W. H. Klein, S. F. Tsai, V. D'Agati, S. H. Orkin and F. Costantini (1991). "Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1." Nature **349**(6306): 257-260.

Pongubala, J. M., C. Van Beveren, S. Nagulapalli, M. J. Klemsz, S. R. McKercher, R. A. Maki and M. L. Atchison (1993). "Effect of PU.1 phosphorylation on interaction with NF-EM5 and transcriptional activation." Science **259**(5101): 1622-1625.

Prasher, D. C., V. K. Eckenrode, W. W. Ward, F. G. Prendergast and M. J. Cormier (1992). "Primary structure of the Aequorea victoria green-fluorescent protein." Gene **111**(2): 229-233.

Pronk, C. J., D. J. Rossi, R. Mansson, J. L. Attema, G. L. Norddahl, C. K. Chan, M. Sigvardsson, I. L. Weissman and D. Bryder (2007). "Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy." Cell Stem Cell **1**(4): 428-442.

Rekhtman, N., K. S. Choe, I. Matushansky, S. Murray, T. Stopka and A. I. Skoultchi (2003). "PU.1 and pRB interact and cooperate to repress GATA-1 and block erythroid differentiation." Mol Cell Biol **23**(21): 7460-7474.

Rhodes, J., A. Hagen, K. Hsu, M. Deng, T. X. Liu, A. T. Look and J. P. Kanki (2005). "Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish." Dev Cell **8**(1): 97-108.

Rieger, M. A., P. S. Hoppe, B. M. Smejkal, A. C. Eitelhuber and T. Schroeder (2009). "Hematopoietic cytokines can instruct lineage choice." Science **325**(5937): 217-218.

Rieger, M. A. and T. Schroeder (2007). "Hämatopoeische Stammzellen" (german) Biospektrum **3.2007**: 254-256.

Robb, L. (2007). "Cytokine receptors and hematopoietic differentiation." Oncogene **26**(47): 6715-6723.

Roeder, I. and I. Glauche (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." J Theor Biol **241**(4): 852-865.

Rosenbauer, F., B. M. Owens, L. Yu, J. R. Tumang, U. Steidl, J. L. Kutok, L. K. Clayton, K. Wagner, M. Scheller, H. Iwasaki, C. Liu, B. Hackanson, K. Akashi, A. Leutz, T. L. Rothstein, C. Plass and D. G. Tenen (2006). "Lymphoid cell growth and transformation are suppressed by a key regulatory element of the gene encoding PU.1." Nat Genet **38**(1): 27-37.

Rosenbauer, F., K. Wagner, J. L. Kutok, H. Iwasaki, M. M. Le Beau, Y. Okuno, K. Akashi, S. Fiering and D. G. Tenen (2004). "Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1." Nat Genet **36**(6): 624-630.

Rosu-Myles, M., B. J. Taylor and L. Wolff (2007). "Loss of the tumor suppressor p15Ink4b enhances myeloid progenitor formation from common myeloid progenitors." Exp Hematol **35**(3): 394-406.

Sacchetti, B., A. Funari, S. Michienzi, S. Di Cesare, S. Piersanti, I. Saggio, E. Tagliafico, S. Ferrari, P. G. Robey, M. Riminucci and P. Bianco (2007). "Self-renewing osteoprogenitors in bone marrow sinusoids can organize a hematopoietic microenvironment." Cell **131**(2): 324-336.

Sarrazin, S., N. Mossadegh-Keller, T. Fukao, A. Aziz, F. Mourcin, L. Vanhille, L. Kelly Modis, P. Kastner, S. Chan, E. Duprez, C. Otto and M. H. Sieweke (2009). "MafB restricts M-CSF-dependent myeloid commitment divisions of hematopoietic stem cells." Cell **138**(2): 300-313.

Schambach, A., M. Galla, U. Modlich, E. Will, S. Chandra, L. Reeves, M. Colbert, D. A. Williams, C. von Kalle and C. Baum (2006). "Lentiviral vectors pseudotyped with murine ecotropic envelope: increased biosafety and convenience in preclinical research." Exp Hematol **34**(5): 588-592.

Schofield, R. (1978). "The relationship between the spleen colony-forming cell and the haemopoietic stem cell." Blood Cells **4**(1-2): 7-25.

Schroeder, T. (2011). "Long-term single-cell imaging of mammalian stem cells." Nat Methods **8**(4 Suppl): S30-35.

Schwarzfischer, M., C. Marr, J. Krumsiek, P. S. Hoppe, T. Schroeder and F. J. Theis. "Efficient fluorescence image normalization for time lapse movies." In Proc. Microscopic Image Analysis with Applications in Biology: September 2011, Heidelberg, Germany.

Scott, E. W., M. C. Simon, J. Anastasi and H. Singh (1994). "Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages." Science **265**(5178): 1573-1577.

Shaner, N. C., R. E. Campbell, P. A. Steinbach, B. N. Giepmans, A. E. Palmer and R. Y. Tsien (2004). "Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein." Nat Biotechnol **22**(12): 1567-1572.

Shimizu, N., S. Noda, K. Katayama, H. Ichikawa, H. Kodama and H. Miyoshi (2008). "Identification of genes potentially involved in supporting hematopoietic stem cell activity of stromal cell line MC3T3-G2/PA6." Int J Hematol **87**(3): 239-245.

Shimizu, R., E. Kobayashi, J. D. Engel and M. Yamamoto (2009). "Induction of hyperproliferative fetal megakaryopoiesis by an N-terminally truncated GATA1 mutant." Genes Cells **14**(9): 1119-1131.

Shimizu, R., T. Kuroha, O. Ohneda, X. Pan, K. Ohneda, S. Takahashi, S. Philipsen and M. Yamamoto (2004). "Leukemogenesis caused by incapacitated GATA-1 function." Mol Cell Biol **24**(24): 10814-10825.

Shimomura, O., F. H. Johnson and Y. Saiga (1962). "Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea." J Cell Comp Physiol **59**: 223-239.

Shivdasani, R. A., Y. Fujiwara, M. A. McDevitt and S. H. Orkin (1997). "A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development." EMBO J **16**(13): 3965-3973.

Song, Y., A. Bahnson, N. Hall, H. Yu, H. Shen, D. Koebler, R. Houck, Y. Xie and T. Cheng (2010). "Stem cell traits in long-term co-culture revealed by time-lapse imaging." Leukemia **24**(1): 153-161.

Spangrude, G. J., S. Heimfeld and I. L. Weissman (1988). "Purification and characterization of mouse hematopoietic stem cells." Science **241**(4861): 58-62.

Stachura, D. L., O. Svoboda, R. P. Lau, K. M. Balla, L. I. Zon, P. Bartunek and D. Traver (2011). "Clonal analysis of hematopoietic progenitor cells in the zebrafish." Blood **118**(5): 1274-1282.

Starck, J., N. Cohet, C. Gonnet, S. Sarrazin, Z. Doubeikovskaia, A. Doubeikovski, A. Verger, M. Duterque-Coquillaud and F. Morle (2003). "Functional cross-antagonism between transcription factors FLI-1 and EKLF." Mol Cell Biol **23**(4): 1390-1402.

Stopka, T., D. F. Amanatullah, M. Papetti and A. I. Skoultchi (2005). "PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure." EMBO J **24**(21): 3712-3723.

Strasser, M., F. J. Theis and C. Marr (2012). "Stability and multiattractor dynamics of a toggle switch based on a two-stage model of stochastic gene expression." Biophys J **102**(1): 19-29.

Sugiyama, T., H. Kohara, M. Noda and T. Nagasawa (2006). "Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches." Immunity **25**(6): 977-988.

Szilvassy, S. J., K. P. Weller, W. Lin, A. K. Sharma, A. S. Ho, A. Tsukamoto, R. Hoffman, K. R. Leiby and D. P. Gearing (1996). "Leukemia inhibitory factor upregulates cytokine expression by a murine stromal cell line enabling the maintenance of highly enriched competitive repopulating stem cells." Blood **87**(11): 4618-4628.

Takahashi, S., K. Onodera, H. Motohashi, N. Suwabe, N. Hayashi, N. Yanai, Y. Nabesima and M. Yamamoto (1997). "Arrest in primitive erythroid cell development caused by promoter-specific disruption of the GATA-1 gene." J Biol Chem **272**(19): 12611-12615.

Ting, C. N., M. C. Olson, K. P. Barton and J. M. Leiden (1996). "Transcription factor GATA-3 is required for development of the T-cell lineage." Nature **384**(6608): 474-478.

Traver, D., B. H. Paw, K. D. Poss, W. T. Penberthy, S. Lin and L. I. Zon (2003). "Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants." Nat Immunol **4**(12): 1238-1246.

Tsai, F. Y. and S. H. Orkin (1997). "Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation." Blood **89**(10): 3636-3643.

Tsai, S. F., E. Strauss and S. H. Orkin (1991). "Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter." Genes Dev **5**(6): 919-931.

Turan, S. and J. Bode (2011). "Site-specific recombinases: from tag-and-target- to tag-and-exchange-based genomic modifications." FASEB J **25**(12): 4088-4107.

Visvader, J. E., A. G. Elefanty, A. Strasser and J. M. Adams (1992). "GATA-1 but not SCL induces megakaryocytic differentiation in an early myeloid line." EMBO J **11**(12): 4557-4564.

Waddington, C. H. (1957). "The Strategy of Genes." London: Allen & Unwin.

Wall, L., E. deBoer and F. Grosveld (1988). "The human beta-globin gene 3' enhancer contains multiple binding sites for an erythroid-specific protein." Genes Dev **2**(9): 1089-1100.

Walsh, J. C., R. P. DeKoter, H. J. Lee, E. D. Smith, D. W. Lancki, M. F. Gurish, D. S. Friend, R. L. Stevens, J. Anastasi and H. Singh (2002). "Cooperative and antagonistic interplay between PU.1 and GATA-2 in the specification of myeloid cell fates." Immunity **17**(5): 665-676.

Wang, H., L. J. Pierce and G. J. Spangrude (2005). "Lymphoid potential of primitive bone marrow progenitors evaluated in vitro." Ann N Y Acad Sci **1044**: 210-219.

Wang, H., L. J. Pierce and G. J. Spangrude (2006). "Distinct roles of IL-7 and stem cell factor in the OP9-DL1 T-cell differentiation culture system." Exp Hematol **34**(12): 1730-1740.

Warr, M. R., E. M. Pietras and E. Passegue (2011). "Mechanisms controlling hematopoietic stem cell functions during normal hematopoiesis and hematological malignancies." Wiley Interdiscip Rev Syst Biol Med **3**(6): 681-701.

Warren, L., D. Bryder, I. L. Weissman and S. R. Quake (2006). "Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR." Proc Natl Acad Sci U S A **103**(47): 17807-17812.

Wechsler, J., M. Greene, M. A. McDevitt, J. Anastasi, J. E. Karp, M. M. Le Beau and J. D. Crispino (2002). "Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome." Nat Genet **32**(1): 148-152.

Weiss, M. J., G. Keller and S. H. Orkin (1994). "Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells." Genes Dev **8**(10): 1184-1197.

Wilson, A., E. Laurenti, G. Oser, R. C. van der Wath, W. Blanco-Bose, M. Jaworski, S. Offner, C. F. Dunant, L. Eshkind, E. Bockamp, P. Lio, H. R. Macdonald and A. Trumpp (2008). "Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair." Cell **135**(6): 1118-1129.

Wolff, L. and R. Humeniuk (2013). "Erythroid versus Myeloid Lineage Commitment: Regulating the Master Regulators." Stem Cells.

Wright, D. E., E. P. Bowman, A. J. Wagers, E. C. Butcher and I. L. Weissman (2002). "Hematopoietic stem cells are uniquely selective in their migratory response to chemokines." J Exp Med **195**(9): 1145-1154.

Wright, D. E., A. J. Wagers, A. P. Gulati, F. L. Johnson and I. L. Weissman (2001). "Physiological migration of hematopoietic stem and progenitor cells." Science **294**(5548): 1933-1936.

Wu, H., X. Liu, R. Jaenisch and H. F. Lodish (1995). "Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor." Cell **83**(1): 59-67.

Yu, C., A. B. Cantor, H. Yang, C. Browne, R. A. Wells, Y. Fujiwara and S. H. Orkin (2002). "Targeted deletion of a high-affinity GATA-binding site in the GATA-1 promoter leads to selective loss of the eosinophil lineage in vivo." J Exp Med **195**(11): 1387-1395.

Yu, D., H. M. Ellis, E. C. Lee, N. A. Jenkins, N. G. Copeland and D. L. Court (2000). "An efficient recombination system for chromosome engineering in Escherichia coli." Proc Natl Acad Sci U S A **97**(11): 5978-5983.

Zhang, D. E., P. Zhang, N. D. Wang, C. J. Hetherington, G. J. Darlington and D. G. Tenen (1997). "Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice." Proc Natl Acad Sci U S A **94**(2): 569-574.

Zhang, J., C. Niu, L. Ye, H. Huang, X. He, W. G. Tong, J. Ross, J. Haug, T. Johnson, J. Q. Feng, S. Harris, L. M. Wiedemann, Y. Mishina and L. Li (2003). "Identification of the haematopoietic stem cell niche and control of the niche size." Nature **425**(6960): 836-841.

Zhang, J., M. Socolovsky, A. W. Gross and H. F. Lodish (2003). "Role of Ras signaling in erythroid differentiation of mouse fetal liver cells: functional analysis by a flow cytometry-based novel culture system." Blood **102**(12): 3938-3946.

Zhang, P., G. Behre, J. Pan, A. Iwama, N. Wara-Aswapati, H. S. Radomska, P. E. Auron, D. G. Tenen and Z. Sun (1999). "Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1." Proc Natl Acad Sci U S A **96**(15): 8705-8710.

Zhang, P., J. Iwasaki-Arai, H. Iwasaki, M. L. Fenyus, T. Dayaram, B. M. Owens, H. Shigematsu, E. Levantini, C. S. Huettner, J. A. Lekstrom-Himes, K. Akashi and D. G. Tenen (2004). "Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha." Immunity **21**(6): 853-863.

Zhang, P., X. Zhang, A. Iwama, C. Yu, K. A. Smith, B. U. Mueller, S. Narravula, B. E. Torbett, S. H. Orkin and D. G. Tenen (2000). "PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding." Blood **96**(8): 2641-2648.

Zimmermann, T. (2005). "Spectral imaging and linear unmixing in light microscopy." Adv Biochem Eng Biotechnol **95**: 245-265.

Zon, L. I. (2001). Hematopoiesis : a developmental approach. New York, Oxford University Press.

Zon, L. I., Y. Yamaguchi, K. Yee, E. A. Albee, A. Kimura, J. C. Bennett, S. H. Orkin and S. J. Ackerman (1993). "Expression of mRNA for the GATA-binding proteins in human eosinophils and basophils: potential role in gene transcription." Blood **81**(12): 3234-3241.

9 Abbreviations

°C	degrees Celsius
µ	micro
Ω	Ohm
AGM	aorta-gonads-mesonephros
AML	acute myeloid leukemia
APC	allophycocyanin
BAC	bacterial artificial chromosome
BFP	blue fluorescent protein
bp	base pair
Cat Nr	catalogue number
CD	cluster of differentiation
CFP	cyan fluorescent protein
CFU	colony-forming unit
CLP	common lymphoid progenitor
CMP	common myeloid progenitor
CSF	colony-stimulating factor
d	day
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DPBS	Dulbecco's phosphate-buffered saline
E	embryonic day
E	erythrocyte
e.g.	exempli gratia
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoresis mobility shift assay
EPO	erythropoietin
ER	estrogen receptor

ES	embryonic stem
F	Farad
FACS	fluorescence-activated cell sorting
FCS	fetal calf serum
FL	fetal liver
FLP-e	enhanced FLP recombinase
Flt3	fms-related tyrosine kinase 3
FSC	forward scatter
g	gram
G	granulocyte
GB	gigabyte
G-CSF	granulocyte colony-stimulating factor
GEMMeg	granulocytic, erythroid, monocytic, megakaryocytic
GFP	green fluorescent protein
GM	granulocytic-monocytic
GM-CSF	granulocyte macrophage colony-stimulating factor
GMP	granulocyte-macrophage progenitor
h	hour
HEK	human embryonic kidney
HSC	hematopoietic stem cell
HSPC	hematopoietic stem and progenitor cell
IL	interleukin
IRES	internal ribosomal entry site
l	liter
LB	Luria broth
Lin	Lineage
LMPP	lymphoid-primed multipotent progenitor
LSK	$Lineage^{-}Sca\text{-}1^{+}c\text{-}Kit^{+}$
M	macrophage
m	meter

M	molar
M-CSF	macrophage colony-stimulating factor
Meg	megakaryocytic
MegE	megakaryocytic-erythroid
MEP	megakaryocyte-erythrocyte progenitor
min	minute
miRNA	microRNA
MkP	megakaryocytic progenitors
ml	milliliter
mM	millimolar
MOI	multiplicity of infection
MPO	myeloperoxidase
MPP	multipotent progenitor
mRNA	messenger RNA
ms	millisecond
NaCl	sodium chloride
NaOH	sodium hydroxide
ng	nanogram
NK	natural killer cell
nm	nanometer
OD	optical density
PAGE	polyacrylamide gelelectrophoresis
PB	Pacific Blue
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PE	phycoerythrin
rcf	radial centrifugal force
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RT	room temperature

RT-PCR	reverse transcriptase polymerase chain reaction
s	second
Sca-1	stem cell antigen 1
SCF	stem cell factor
SDS	sodium dodecyl sulfate
SFEM	serum-free expansion medium
SFFV	spleen focus forming virus
SSC	saline sodium citrate
SSC	side scatter
TAE	TRIS acetate EDTA
TBS	TRIS-buffered saline
TBS-T	TRIS-buffered saline, Tween
TPO	thrombopoietin
TRIS	tris-(hydroxymethyl)-aminomethane
URE	upstream regulatory element
UTR	untranslated region
V	Volt
VSV-g	vesicular stomatitis virus G protein
WT	wildtype
YFP	yellow fluorescent protein

Acknowledgements

First, I would like to thank Timm Schroeder for always believing in me and giving me the opportunity of working on this interesting project. Thank you for all the support that significantly contributed to this PhD-project, especially for teaching me how to do science and how to communicate it.

Thanks to Prof. Thomas Cremer for officially supervising me.

Thanks to Ingo Burtscher for explaining how to design a knock-in mouse. I would also like to thank Antje Bürger for the targeting of the embryonic stem cells and Heide Oller for doing the embryonic stem cell aggregation that led to the generation of the Gata1mCHERRY mouse.

Thanks to Angelika Ziegler, Sandra Ammersdörfer, Bianca Vogel and Christian Raithel for their excellent technical support.

I would like to specially thank Michael Schwarzfischer for the nearly daily close interactions that successfully led the development of background correction algorithms and to the program QTF which allows the supervised quantification of fluorescent signal intensities in imaging. I would also like to thank Carsten Marr, Michael Strasser, Felix Buggenthin and Fabian Theis for their contributions to the informatics and mathematical framework of this project.

I would also like to thank my colleagues Max Endele, Daniel Coutu, Joost van den Berg, Martin Etzrodt, Dirk Löffler, Konstantinos Kokkalaris Adam Filipczyk, Simon Hastreiter and Laura Skylaki for the numerous discussions that helped to improve the project and for the great atmosphere in the lab. Thanks to all members of the Institute of Stem Cell Research for their helpful comments in seminars. I would also like to thank my former colleague Nadine Moritz for her direct support and Michael Rieger for introducing me as a little student into science.

At last, I would like to thank my friends and family, especially my parents for their constant support.

Thank you, Joana, for always being there for me! Future work won't be as nice knowing that I can't just go to the next room to find you. Luckily we will spend the rest of our lives together. I love you!