CROSS-JOURNAL FOCUS

Stem Cells
CROSS-JOURNAL FOCUS

Stem Cells

EDITORS

**Thomas Schwarz-Romond**  
Senior Editor  
t.schwarz-romond@embojournal.org | T +49 6221 8891 407

Thomas studied Wnt-signal transduction in the lab of Walter Birchmeier at the Max-Delbrueck Center in Berlin and received his PhD from the FU-Berlin in 2003. He continued to investigate this exciting topic as a postdoctoral fellow with Mariann Bienz at the MRC-LMB in Cambridge and joined *The EMBO Journal* in 2006.

**Barbara Pauly**  
Senior Editor  
pauly@embo.org | T +49 6221 8891 109

Barbara joined *EMBO Reports* in September 2008. She completed her PhD at the University of Munich, focusing on signal transduction in the fresh water polyp Hydra. She worked at the University of California at Berkeley as a post-doctoral researcher, studying the role of the actin cytoskeleton in endocytosis in mammalian cells.

**Céline Carret**  
Editor  
c.carret@embomolmed.org | T +49 6221 8891 310

Céline Carret completed her PhD at the University of Montpellier, France, characterising host immunodominant antigens to fight babesiosis, a parasitic disease caused by a unicellular Apicomplexan parasite closely related to the malaria agent Plasmodium. She further developed her post-doctoral career on malaria working at the Wellcome Trust Sanger Institute in Cambridge, UK and Instituto de Medicina Molecular in Lisbon, Portugal. Céline joined *EMBO Molecular Medicine* as a Scientific Editor in March 2011.

**Maria Polychronidou**  
Editor  
maria.polychronidou@embo.org | T +49 6221 8891 410

Maria received her PhD from the University of Heidelberg, where she studied the role of nuclear membrane proteins in development and aging. During her post-doctoral work, she focused on the analysis of tissue-specific regulatory functions of Hox transcription factors using a combination of computational and genome-wide methods.
Neutral competition of stem cells is skewed by proliferative changes downstream of Hh and Hpo.
Amoyel M, Simons BD, Bach EA.
DOI: 10.15252/embj.201387500 | Published 04.08.2014

The Hippo signaling pathway in stem cell biology and cancer.
Mo JS, Park HW, Guan KL.
DOI: 10.15252/embr.201438638 | Published 13.05.2014

Heterotrimeric G proteins control stem cell proliferation through CLAVATA signaling in Arabidopsis.
DOI: 10.15252/embr.201438660 | Published 26.09.2014

Targeted gene therapy and cell reprogramming in Fanconi anemia.
DOI: 10.15252/emmm.201303374 | Published 23.05.2014

Intercellular network structure and regulatory motifs in the human hematopoietic system.
DOI: 10.15252/msb.20145141 | Published 15.07.2014

For further reading please see inside back cover
Neutral competition of stem cells is skewed by proliferative changes downstream of Hh and Hpo

Marc Amoyel1,–, Benjamin D Simons2,3,4 & Erika A Bach1,5,∗∗

Abstract

Neutral competition, an emerging feature of stem cell homeostasis, posits that individual stem cells can be lost and replaced by their neighbors stochastically, resulting in chance dominance of a clone at the niche. A single stem cell with an oncogenic mutation could bias this process and clonally spread the mutation throughout the stem cell pool. The Drosophila testis provides an ideal system for testing this model. The niche supports two stem cell populations that compete for niche occupancy. Here, we show that cyst stem cells (CySCs) conform to the paradigm of neutral competition and that clonal deregulation of either the Hedgehog (Hh) or Hippo (Hpo) pathway allows a single CySC to colonize the niche. We find that the driving force behind such behavior is accelerated proliferation. Our results demonstrate that a single stem cell colonizes its niche through oncogenic mutation by co-opting an underlying homeostatic process.

Keywords competition, Hedgehog, Hippo, stem cell, testis

Subject Categories Cell Cycle; Development & Differentiation; Stem Cells

DOI 10.15252/embj.201387500 | Received 25 November 2013 | Revised 2 July 2014 | Accepted 7 July 2014 | Published online 4 August 2014

The EMBO Journal (2014) 33: 2295–2313

See also: ER Morrissey & L Vermeulen (October 2014)

Introduction

The ability of a stem cell to continually generate offspring for tissue maintenance depends on its ability to remain and renew at the niche. A critical consideration is whether stem cells are eternal and always divide invariantly or whether they function as members of an equipotent population, within which a single stem cell could be lost and replaced stochastically by a neighbor. Recent work has revealed that the latter, termed neutral competition, is an emerging feature of stem cell homeostasis. This model states that individual stem cells can be stochastically lost and replaced by their neighbors, resulting in chance dominance of a clone at the niche. Neutral competition has been established for both vertebrates and invertebrates and in several different tissues (Clayton et al., 2007; Klein et al., 2010; Lopez-Garcia et al., 2010; Snippert et al., 2010; Doupe et al., 2012; de Navascues et al., 2012). However, the fact that loss and gain of stem cells occurs opens the possibility of a transformed stem cell exploiting this process in its favor and achieving clonal dominance. Such behavior theoretically could underlie the observation of tumor-initiating cells in certain types of cancer (Reya et al., 2001) and has recently been reported for mouse intestinal stem cells (Vermeulen et al., 2013; Snippert et al., 2014).

The Drosophila testis provides an ideal system for analyzing single stem cell behavior. The niche (called the hub) supports two stem cell populations, germ line stem cells (GSCs) and somatic cyst stem cells (CySCs) (Fig 1A and de Cuevas & Matunis, 2011; Hardy et al., 1979). GSCs give rise to sperm, while CySCs produce somatic cyst cells, which ensheath developing germ cells and are required for germ cell differentiation. Each testis niche harbors approximately 9–14 GSCs, which divide with oriented mitosis perpendicular to the niche, such that one offspring, likely to remain in contact with the niche, self-renews while the other, physically displaced from niche signals, begins differentiation (Yamashita et al., 2003; Sheng & Matunis, 2011). Serially reconstructed electron micrographs of wild-type testes revealed ~13 somatic cells, presumed to be the CySCs, in contact with the hub in young adults (Hardy et al., 1979). Most current studies rely on immunofluorescence of nuclear factors in presumptive CySCs and their daughters. The best molecular marker of CySCs is Zfh1, which labels the nucleus of ~44 cells in wild-type testes (Leatherman & Dinardo, 2008; Inaba et al., 2011; Amoyel et al., 2013). This value substantially overestimates the true number of CySCs and includes post-mitotic daughter cells that no longer contact the niche. Finally, there is no evidence for oriented division among CySCs (Cheng et al., 2011), raising the possibility that this population may be subject to different regulation than GSCs. Stem cell loss and replacement has been observed in Drosophila gonads, in both somatic and germ lineages, but its significance remains under debate (Margolis & Spradling, 1995; Xie & Spradling, 1998, 2000; Zhang & Kalderon, 2001; Wallengraf et al., 2006; Nystul & Spradling, 2007). It remains to...
Gain of Hh or Yki cause niche competition

Marc Amoyel et al.

The Authors

The EMBO Journal

© 2014 The Authors
be resolved whether loss of stem cells reflects their loss of fitness or represents a normal homeostatic process of neutral competition.

The molecular signals governing self-renewal at the testis niche have been well characterized (de Cuevas & Matunis, 2011). GSCs are maintained by Bone Morphogenetic Protein (BMP) signals originating from both the hub and CySCs (Shivdasani & Ingham, 2003; Kawase et al., 2004; Leatherman & Dinardo, 2010). CySCs require at least two signaling inputs, the Janus Kinase/Signal Transducer and Activator of Transcription (JAK/STAT) and Hedgehog (Hh) pathways, in order to self-renew (Kiger et al., 2001; Leatherman & Dinardo, 2008; Michel et al., 2012; Amoyel et al., 2013). Ligands for both pathways, Unpaired (Upd) and Hh, respectively, are produced by the hub cells (Forbes et al., 1996; Kiger et al., 2001; Tulina & Matunis, 2001; Dinardo et al., 2011). Two known targets are expressed in CySCs in response to JAK/STAT pathway activation, Zfh1 and Chimo1. Overexpression in CySCs of the JAK Hopscotch (Hop) or of either pathway target results in autonomous hyper-proliferation of CySCs and non-autonomous hyper-proliferation of GSCs, due to BMP production by the CySCs (Leatherman & Dinardo, 2008, 2010; Wang et al., 2008; Flaherty et al., 2010). Conversely, Hh activation only regulates the self-renewal and numbers of the CySCs, without affecting the GSC niche (Amoyel et al., 2013).

Although both stem cells co-exist at the same niche, and although the CySCs are a necessary component of the niche for GSCs, these two populations compete for access to the niche, as revealed by analysis of the mutant phenotype of the JAK/STAT negative feedback regulator Socs36E (Jissignon et al., 2009; Singh et al., 2010). This reduction of GSCs in Socs36E mutants was attributed to increased JAK/STAT signaling in Socs36E mutant CySCs, leading to upregulation of integrin-based adhesion and enabling the mutant cells to displace wild-type GSCs and CySCs from the niche.

Here, we characterize CySC behavior by clonal analysis. We found that the behavior of CySCs was consistent with them being lost and replaced stochastically, as predicted by the neutral competition model. For this study, we made clones homozygous mutant for patched (ptc), which encodes the Hh receptor (Chen & Struhl, 1996); loss of ptc causes constitutive activation of the pathway. We found that ptc mutant CySCs outcompeted both wild-type CySCs and GSCs for niche access. We determined that this phenotype was due to biased competition, skewing normal behavioral dynamics in favor of the mutant cell. We showed that adhesion and JAK/STAT signaling could not cause stem cells to acquire colonizing capabilities. Rather, we showed that simply accelerating proliferation was sufficient to cause a single CySC and its descendants to outcompete wild-type CySCs and GSCs. Furthermore, we established a critical role for the conserved growth regulatory Hippo pathway in regulating competition and self-renewal in CySCs independent of Hh signaling. Thus, we demonstrate that proliferation is the key driver of somatic stem cell behavior and provide a model for how oncogenic mutations can spread throughout a stem cell pool by exploiting a fundamental homeostatic process of stochastic stem cell replacement.

**Results**

**Characterizing the CySC pool**

We first attempted to use molecular markers to sub-divide the somatic population near the niche. We reasoned that only a subset of the ~44 Zfh1-positive cells could constitute the true stem cell pool. We therefore examined whether markers of self-renewal pathways in CySCs—Ptc for Hh and Stat92E for JAK/STAT—were co-expressed. We only found expression of these markers in Zfh1-positive somatic cells located one cell diameter from the hub. Within this group, only one subset co-expressed Ptc and Stat92E (Fig 1B′−′′′, red arrowhead), while others expressed only one or neither (Fig 1B′−′′′, yellow arrowhead and arrow, respectively). This analysis suggests that using the best available molecular markers may not be the most robust method to identify CySCs. Since membrane contact with the niche appears to be the defining feature of stemness in the *Drosophila* testis (Hardy et al., 1979; de Cuevas & Matunis, 2011), we estimated the actual number of CySCs by generating single-cell control
MARCM clones (Lee & Luo, 1999), expressing a membrane-targeted GFP (Fig 1C). We used this clonal method because labeling all somatic cell membranes did not allow us to determine whether an individual cell contacts the hub or not. Only 30.5% of Zhhl-positive clones (29/95 single cell clones) had membrane extensions contacting the hub (Fig 1C’, arrow). Extrapolating this proportion to an average of 43 ± 7 Zhhl-positive cells per testis that we counted in these samples (n = 59), we estimated 13 CySCs per testis, consistent with the 12.6 value that has been previously reported (Hardy et al., 1979). In the genotype we examined, there were 13.2 GSCs (n = 34). In the Drosophila testis, stem cells are actively dividing, and within the somatic lineage, only CySCs divide (Hardy et al., 1979; Inaba et al., 2011). As further confirmation of the number of CySCs, we examined markers of cycling cells, PCNA-GFP to mark cells in S-phase (Thacker et al., 2003) and Cyclin B (CycB) for G2/M. We found 11.2 somatic cells one cell diameter away from the niche undergoing replication that were positive for PCNA (Supplementary Fig S1A–′A″, arrow). In the same testes, 9.2 out of 12.2 total GSCs on average expressed PCNA-GFP, suggesting a 1.3:1 ratio of CySCs to GSCs and by extrapolation a total of ~15 CySCs. Similarly, in an unrelated genetic background that contained on average 7.9 GSCs, we observed 5.6 GSCs and 5.6 CySCs expressing CycB (Supplementary Fig S1B–′B″, arrows). Taken together, these data suggest that GSCs and CySCs exist in a ratio close to 1:1.

Two different models have been proposed to explain stem cell behavior in actively cycling homeostatic tissues; in the first, stem cells are invariant and divide asymmetrically to self-renew and are only rarely lost in cases of damage or loss of fitness. In the other, asymmetry is achieved only at the level of the stem cell population. In the latter case, stem cell populations are dynamic, and their clonal make-up changes according to stochastic variations such that some clones are lost entirely while others expand to occupy empty stem cell berths, through a process termed ‘neutral competition’ (Simons & Clevers, 2011).

We tested these models by generating control FRT<sup>73</sup> MARCM clones that mis-expressed only membrane CD8-GFP and scored the number of labeled somatic cells contacting the hub. While the membrane labeling of clones allows for direct identification of CySCs, this methodology has two drawbacks. First, CySCs outside the clones (which are unmarked) have to be scored more subjectively by their position relative to the hub. Second, once many cells around the niche are labeled, it becomes difficult to distinguish the membranes of individual cells, resulting in a slight overestimation of the total number of CySCs (~16–21 obtained by this method versus ~13 obtained above). Therefore, to circumvent this uncertainty, we monitored both the total number of GFP-labeled and unlabeled cells considered to be contacting the hub and used these values to determine the fraction of labeled CySCs as a percentage in each testis.

At 2 days post-clone induction (dpci), we found few GFP-labeled CySCs, consistent with a low clone induction rate (Fig 1D and H, Supplementary Materials and Methods, see below). To characterize CySC dynamics, we separated testes according to whether they maintained at least one GFP-expressing cell in contact with the hub (termed ‘persisting’) and those in which all GFP-expressing cells had detached from the hub (termed ‘differentiating’). We observed empirically that the mean fraction of labeled CySCs in persisting clones increased steadily as a function of time (Fig 1G), while the number of labeled CySCs in individual clones varied considerably between samples at the same time point, as exemplified by the 14 dpci samples shown in Fig 1E and F. The increased number of labeled CySCs in persisting clones is inconsistent with the model of invariant asymmetric stem cell division as in this scenario this parameter should not change over time. However, the observed change is consistent with CySCs undergoing loss and replacement (Fig 1G).

We next subjected these data to a quantitative analysis, using a parallel approach to that developed to study stem cell dynamics in the murine intestinal crypt (Lopez-Garcia et al., 2010). The assumptions contained in the model are the following: (1) CySCs form a single equipotent population in which any cell has an equal chance of being lost and replaced; (2) in line with the geometry of the testis, the ensemble of CySCs can be approximated as a one-dimensional ‘necklace’ of cells around the niche; (3) as CySCs proliferate, some lose contact with the niche and differentiate, and this is perfectly compensated by the duplication of a neighboring CySC to maintain a constant total number of CySCs. By contrast, an asymmetrical CySC division leaves the number of labeled CySCs unchanged (Fig 1A). As a simplification, we do not take into account GSCs, which are simply regarded as a separate lineage with their own fate behavior. In this modeling scheme, clonal dynamics of the CySC compartment is dependent upon only two parameters—the CySC loss/replacement rate, λ, and the total number of CySCs contacting the hub, N (see Supplementary Materials and Methods). Since CySC division leaves that are asymmetric fate outcome do not change the number of marked CySCs in a clone, the clonal fate data are insensitive to the CySC division rate.

To implement the modeling scheme, it is important to define the labeling efficiency of the FRT<sup>73</sup> MARCM system. From the
frequency of unlabeled testes at 2 dpci, we estimated a labeling efficiency of around 10% for each of the 13 CySCs following the heat shock. At this level of induction, we therefore expect that testes will experience multiple induction events, leading to isolated “clusters” of GFP-expressing cells. By comparing the predicted frequency of clusters with direct measurements at 2 dpci, a least-squares fit suggested a labeling efficiency of \( q = 11\% \) (Fig 1H), consistent with the observed frequency of unlabeled testes. With the labeling
efficiency and CySC number defined, only the CySC loss/replacement rate, \( \lambda \), remained to be determined.

To fix the loss/replacement rate, we compared the predictions of the model with measurements of the average fraction of labeled CySCs in persisting clones using an induction frequency of 10%. Adjusting the loss/replacement rate, we found that the mean fraction of labeled CySCs could be well reproduced by a loss/replacement rate of \( \lambda = 0.84 \pm 0.05 \) per day (Fig 1G, compare lines and boxes). However, alongside the mean fraction, the model also predicts the variation in the size distribution of individual clones as a function of time post-induction. Taking the inferred loss/replacement rate and induction frequency, we found that the model provides an excellent prediction of the measured cumulative clone size distribution (Fig 1I, compare lines to boxes), defined as the fraction of testes that have a fraction of labeled CySCs larger than the given value. (We note that N and B exist in a fixed ratio \( N^2/B \), (see Supplementary Materials and Methods), meaning that the fit to a model with a different N would generate quite reasonable agreement with the data. However, for a larger N, we would require a proportionately larger \( \lambda \), potentially in excess of the CySC division rate.) Taken together, these empirical and modeling data strongly suggest that the ~13 CySCs in a wild-type testis produce equivalent offspring which have stochastic fates.

Given the potential uncertainty of the membrane labeling method as a means to identify CySCs, we chose to challenge our findings by following a second unbiased (albeit less direct) approach to measuring CySC number and scoring clones. We generated control MARCM clones on two separate chromosome arms (FRT40A and FRT42D) that mis-expressed only nuclear GFP and scored the size of CySC clones relative to the total Zfh1-positive population and the clone recovery rate (percentage of testes with any marked clone positive for Zfh1) at various times after clone induction. We assumed that each CySC contributes equally to the total Zfh1-positive pool and counted the number of Zfh1-positive cells (N) that were labeled with the clone marker and expressed this as a fraction N/43 where 43 was the average number of Zfh1-positive cells found per testis (see above). At 2 dpci, we found few GFP-labeled Zfh1-positive cells, consistent with a low clone induction rate (\( q = 0.18 \) for FRT40A and \( q = 0.3 \) for FRT42D) (Supplementary Fig S2A and B and Supplementary Materials and Methods)). At 14 dpci, individual clones varied considerably between samples at the same time point (Supplementary Fig S2C and D). Using this method, we obtained similar results to those observed for membrane labeling; the mean fraction of labeled CySCs increased as a function of time (Supplementary Fig S2E and F) and the number of testes harboring labeled clones decreased over time (Supplementary Fig S2G and H). Quantitative analysis of nuclear GFP MARCM control clones (Supplementary Fig S2E–J) revealed that they obeyed similar neutral drift dynamics to the membrane CD8-GFP MARCM control clones (Fig 1D–I). Importantly, we were able to infer a CySC loss/replacement rate of around once per day, which is comparable to the loss/replacement rate of 0.84 per day inferred from the earlier modeling scheme (see above). Using two different labeling methods, generating clones on two chromosome arms, scoring CySCs by two independent methods, we reach a similar conclusion: CySCs are lost and replaced stochastically and obey neutral drift dynamics.

**ptc mutant CySCs skew neutral drift dynamics and outcompete wild-type CySCs**

The dynamics of neutral stem cell competition have been reported in mammalian and Drosophila stem cells (Wallenfang et al, 2006; Clayton et al, 2007; Klein et al, 2010; Lopez-Garcia et al, 2010; Snippert et al, 2010; de Navascues et al, 2012), but mutations that co-opt the homeostatic mechanisms underlying this process for the benefit of the mutant cell have only recently been described (Vermeulen et al, 2013; Snippert et al, 2014). We and others previously showed that Hh signal reception is required for the maintenance of CySC fate. CySCs that are unable to transduce the Hh signal are lost from the niche and differentiate (Michel et al, 2012; Amoyel et al, 2013). Here, we studied the effect of clonal gain of Hh signaling by making clones homozygous mutant for patched (ptc). Cells lacking ptc function can no longer inhibit Smoothened activity and experience sustained ligand-independent Hh signal transduction (Ingham et al, 1991; Chen & Struhl, 1996). We examined FRT42D ptc mutant CD8-GFP MARCM clones as compared to the appropriate control, that is, FRT42D CD8-GFP MARCM control clones. Similar to control, we found few GFP-labeled ptc mutant CySCs at 2 dpci (Fig 2A). In contrast to control clones, ptc mutant clones contained more labeled CySCs on average by 14 dpci and were often seen to take over the entire somatic lineage, presumably by causing the displacement of wild-type CySCs (Fig 2B, compare boxes in Fig 2C to those in Fig 1G). We counted unlabeled CySCs in control and ptc samples and found that there were significantly fewer when ptc mutant clones were present (Fig 2E, \( P < 0.004 \)). These results indicated that ptc mutant CySCs expanded at the expense of their wild-type neighbors. However, like control clones, the frequency of persistent ptc mutant CySC clones decreased over time (Fig 8E, red line). These results

Figure 3. Clonal overactivation of the Hh pathway, but not JAK/STAT, causes niche competition phenotypes.

A No increase in Stat92E staining (red, single channel in A′) was seen in a ptc mutant CySC (green, arrow, single channel in A′) compared to neighboring wild-type CySCs (arrowhead) at 2 dpci. Somatic cells were labeled with Tj (blue, single channel in A′). See also Supplementary Fig S5 for Stat92E staining in clones at 7 and 14 dpci.

B–G MARCM clones at 14 dpci, with single channels showing the clone marker GFP in B′–F′. Vasa in red and Zfh1 (B, C or Tj) (D–F) in blue. Control clones (B, C) showed variation in the number of cells labeled. Overexpression of Ci/ATC (E) recapitulated the ptc mutant phenotype (compare with Fig 2B), but overexpression of Hop did not (F). Hop overexpression activated JAK/STAT signaling (G), as seen by stabilization of Stat3 protein (blue, single channel in C′) in the clones (green, arrows).

H Number of GSCs at 14 dpci when CySC clones of the indicated genotype were induced. Hop overexpression did not affect GSC number, while C/ATC and ptc RNAi-expressing clones caused loss of GSCs, similar to ptc mutant clones (see Fig 2F). Asterisks denote statistically significant change from control: n = 15, 24, 27, 8 for control, UAS-Hop, UAS-C/ATC, UAS- ptc RNAi, respectively. Error bars denote SEM.

I Number of GSCs at 14 dpci when control or ptc mutant CySCs were present, showing an enhancement of GSC loss when ptc mutant clones were induced in a background lacking one copy of the Stat92E gene. Asterisks denote statistically significant change from the ptc mutant clones alone. n = 48 (control), 36 (control; Stat92E+), 49 (ptc), 20 (ptc; Stat92E+/−). Error bars denote SEM.

Gain of Hh or Yki cause niche competition

Marc Amoyel et al
and D). Using this method, we obtained similar results to those adjusting the loss/replacement rate, we found that the mean fraction of labeled CySC number and scoring clones. We generated control MARCM clones on two separate chromosome arms (Fig 1J, compare lines to boxes), defined as the size distribution (Fig 1I, compare lines to boxes). However, alongside the mean fraction, the model also exist in a fixed ratio $N_2/C_6 = 0.84$ per day (Fig 1G, compare lines and $E$). We counted unlabeled CySCs in control and mutants, showing an enhancement of GSC loss when mutant CySCs were present, showing an enhancement of GSC loss when mutant clones were induced in a ptc mutant background lacking one copy of the maternal $ptc$ allele (Fig 1E). Importantly, we were able to determine whether CySCs contributed equally to the total Zfh1-positive pool and counted the number of Zfh1-positive cells found per testis (see above).

**Figure 1:**
- **A**, **A’**, **A”**, **A”’**: Ptc clones at 2 dpci.
- **B**, **B’**, **B”**, **B”’**: Control clone Vasa (2h1) at 14 dpci.
- **C**, **C’**, **C”**, **C”’**: Control clone Vasa (2h1) at 14 dpci.
- **D**, **D’**, **D”**, **D”’**: UAS-CiAct clone Vasa at 14 dpci.
- **E**, **E’**, **E”**, **E”’**: UAS-Stat92E clone Vasa at 14 dpci.
- **F**, **F’**, **F”**, **F”’**: UAS-Hop clone Vasa at 14 dpci.
- **G**, **G’**, **G”**, **G”’**: UAS-Hop clone Vasa at 14 dpci.
- **H**: Number of GSCs.
- **I**: Number of GSCs.
show that 

\textit{ptc} clones could differentiate, indicating that they are not locked in a perpetual state of stem cell self-renewal, which is consistent with a prior report (Michel et al., 2012). Indeed, we observed differentiating \textit{ptc} mutant cyst cells ensheathing spermatogonial cysts similar to control clones (Supplementary Fig S3A and B). This represents a different situation from a previously observed instance of stem cell competition, where stem cells in the ovary that cannot differentiate eventually replaced their wild-type neighbors (Jin et al., 2008). Thus, these data suggest that \textit{ptc} mutant CySCs have a competitive advantage over wild-type CySCs in effecting stem cell replacement.

To assess whether the dynamics of \textit{ptc} mutant clones represent a biasing of the neutral competition process toward persistence, we sought for the simplest revision of the neutral drift model which could accommodate the observed behavior. In particular, we assumed that, following the loss of a CySC (control or \textit{ptc} mutant) through commitment to differentiation, a neighboring \textit{ptc} mutant CySC will have a higher chance of replacing it through symmetric cell division than a wild-type neighboring CySC. We also assumed that the competitive advantage of the \textit{ptc} mutant CySC is sustained since the loss rate of CySCs is not differentially affected by \textit{ptc} mutation. Once again, using the frequency of unlabeled tests at 2 dpci, we inferred a CySC labeling efficiency of around 10%, similar to the control. Then, taking the loss/replacement rate of wild-type CySCs to be unperturbed from its control value, by adjusting the bias of \textit{ptc} mutant CySCs away from loss and toward replacement (by around 35%), we found a good agreement between the model dynamics and the experimental data (Fig 2C, compare lines and boxes, and Supplementary Materials and Methods). Significantly, taking these model parameters, comparison of the cumulative clone size distribution revealed an excellent agreement of the model prediction with the data over the range of time points (Fig 2D, compare lines and boxes).

Once again, we repeated this experiment using the alternative labeling and scoring method in which CySC number is estimated by the labeled fraction of Zfh1-expressing cells. We analyzed \textit{FRT42D} \textit{ptc} mutant nuclear GFP MARCM clones as compared to \textit{FRT42D} nuclear GFP MARCM control clones (compare Supplementary Fig S4A and B to Supplementary Fig S2B and D). We found that the number of these \textit{ptc} mutant CySCs increased faster than control CySCs and that they were lost less frequently (compare boxes in Supplementary Fig S4C and D to those in Supplementary Fig S2F and H), similar to the results obtained for the membrane labeling experiment. Finally, we carried out modeling using the same bias as before (Supplementary Fig S4C–E, lines) and found that the computational output for number and distribution of CySCs as well as clone recovery rate was well matched to the experimental data for \textit{FRT42D} \textit{ptc} mutant nuclear GFP MARCM clones. Taken together, the data indicate that the behavior of \textit{ptc} mutant clones is consistent with a biasing of competition between stem cells.

**\textit{ptc} mutant CySCs outcompete wild-type GSCs**

As a readout for the competitive activity of \textit{ptc} mutant CySCs, we also quantified the number of GSCs (defined as Vasa-positive cells in contact with the niche) in testes with control or \textit{ptc} mutant CySCs at 14 dpci. Indeed, we found that GSC number was significantly reduced ($P < 0.0001$) non-autonomously when \textit{ptc} mutant CySC clones were present (Fig 2F, red bar). At the same time, colonizing CySCs contacted the hub in place of the outcompeted GSCs (Fig 2B, arrow, Supplementary Fig S4B), similar to the phenotype described for \textit{Socs36E} (Issigonis et al., 2009). GSC loss was only observed once the majority of CySCs were replaced by \textit{ptc} mutant CySCs (Supplementary Fig S4F), suggesting a selective outcome of first wild-type CySCs and then wild-type GSCs by \textit{ptc} mutant CySCs. The fact that \textit{ptc} mutant CySCs had normal levels of factors that mediate GSC extended niche function, that is, Stat92E and Zfh1 (Leatherman & Dinardo, 2010) (Fig 3A, Supplementary Fig S5, arrows, for Stat92E; Fig 2B, Supplementary Fig S4B, Fig 3B, arrows for Zfh1), strongly suggests that GSC loss is not due to lack of appropriate support from \textit{ptc} mutant CySCs. Thus, gain of Hh signaling results in niche colonization by the mutant cell, as a consequence of the displacement of resident wild-type CySCs and GSCs at the niche.

**JAK/STAT signaling, adhesion and cell competition factors are not causal to niche competition**

Several possibilities could explain niche colonization by \textit{ptc} mutant CySCs. We ruled out the trivial explanation that the niche size was altered in tests with \textit{ptc} mutant clones (Supplementary Table S1). We next tested whether an increase in integrin-based adhesion downstream of Stat92E, as proposed for \textit{Socs36E} mutants (Issigonis et al., 2009), caused \textit{ptc} mutant CySCs to anchor more securely to the niche. We previously found no epistatic relationship between constitutively active form of the Hh signal transducer Cubitus inter-ruptus (CiAct) (Price & Kalderon, 1999) or of an RNAi hairpin against G3PDH, a nuclear Junctin (Jin et al., 2008), caused \textit{ptc} mutant CySC clones to be less able to displace resident wild-type CySCs at the niche. We therefore asked whether the levels of Stat92E were down-regulated by \textit{ptc} mutation.

\textbf{Figure 4. Increased adhesion is not causal to niche competition.}

\textbf{A, B} \textit{ptc} mutant clones did not upregulate adhesion molecules. No change in JPS-integrin (A, red, single channel in A) or in DE-cadherin expression (B, blue, single channel in B) was seen at the hub in testes with \textit{ptc} mutant clones (green). Vasa labels germ cells in red (B), Tj labels somatic cells in blue (A). The hub is indicated with a dotted line.

\textbf{C, D} GFP-positive MARCM clones (green, single channels in C, D) overexpressing JPS-integrin (C) or DE-cadherin (D) did not outcompete neighboring wild-type CySCs or GSCs. Vasa labels germ cells in red, Tj labels somatic cells in blue. The hub is indicated with a dotted line.

\textbf{E, F} Control (E) and rhea mutant (F) MARCM clones showing marked CySCs which contacted the hub at 7 dpci (arrows). Vasa labels germ cells in red, Tj labels somatic cells in blue. The hub is indicated with a dotted line.

\textbf{G} CySC clone recovery rates at 2 (blue bars) and 7 (red bars) dpci for control (left) and \textit{rhea} mutant (right). The presence of \textit{rhea} mutant clones at the niche at the 7-day time point indicates that \textit{rhea} was not required in CySCs for self-renewal; $n = 38$ and 24 for control at 2 and 7 dpci, respectively, and $n = 9$ and 49 for \textit{rhea} at 2 and 7 dpci, respectively.

\textbf{H} Number of GSCs present when CySC clones of the indicated genotype were generated at 14 dpci. Overexpression of JPS-integrin, TalinH or DE-cadherin (Tj) did not affect GSC numbers. $n = 15, 19, 25, 17$ for control, UAS-JPS-integrin, UAS-TalinH or UAS-DE-cadherin, respectively. Error bars denote SEM.

\textbf{I} Number of GSCS when \textit{ptc} mutant CySC clones were present along with a single mutant copy of the indicated genes at 14 dpci. Reduction of \textit{α-cat} had no effect on the \textit{ptc} phenotype, while one \textit{rhea} allele partly suppressed GSC loss. \textit{GSC loss} $n = 48$ (control), 21 (control; \textit{rhea}0/1); 49 (ptc), 26 (ptc; \textit{α-cat}+/); 19 (ptc; \textit{rhea}AA/AA); 35 (ptc; \textit{rhea}AA/AA). Error bars denote SEM.
the niche. We previously found no epistatic relationship between Hh and STAT signaling in the testis (Amoyel et al., 2013). Consistent with this, we found that Stat92E levels were unchanged in ptc mutant CySCs (Fig 3A–A’, arrow, Supplementary Fig S5). We used the MARCM technique to assess whether the clonal overexpression of various factors could induce niche competition. As expected, control clones that only overexpressed GFP had variable clone sizes (Fig 3B and C). By contrast, clonal overexpression of either a constitutively active form of the Hh signal transducer Cubitus interruptus (Ci'Act) (Price & Kalderon, 1999) or of an RNAi hairpin against ptc recapitulated the ptc mutant phenotype (Fig 3D and E) and caused a statistically significant reduction of GSCs (Fig 3H, P < 0.0001 for both Ci'Act and ptc RNAi), thus validating our technique. Surprisingly, clonal hyper-activation of Stat92E by misexpression of Hop did not cause CySC clones to compete with either wild-type CySCs or with GSCs (Fig 3F and H), despite clearly
Figure 5. ptc mutant CySCs proliferate faster than controls.

A, B There was an increase in the S-phase index in CySCs mutant for ptc. Quantification of S-phase in control (A) or ptc mutant (B) clones. Clones expressing GFP (green, single channel A), B’ were labeled with Tj (red, single channel A’, B”) and EdU (blue, single channel A”, B”). Triply labeled cells (yellow arrowheads) were counted as a ratio of total cells double positive for GFP and Tj, with quantification shown in (E).

C stg-GFP (green, single channel C) was upregulated in ptc mutant CySCs (yellow arrowheads). Zh/t1 (red, single channel C’) labels CySCs, and their offspring and clones are identified by loss of the Igal marker (blue, single channel C”).

D PCNA-GFP (green, single channel D”) was upregulated in ptc mutant clones. Clones are labeled by loss of Igal (blue, single channel D”, Zht1 (red, single channel D”) marks CySCs and their offspring. Arrow shows control CySC, and arrowhead shows a ptc mutant CySC.

E S-phase index. See legend of (A) above. Asterisks denote statistically significant change from control. Error bars denote SEM.

F Quantification of PCNA-GFP fluorescence intensity in control or ptc mutant CySCs. n = 11 for both genotypes. An asterisk denotes statistically significant change from control. Error bars denote SEM.

Gain of Hh or Yki cause niche competition

Marc Amoyel et al.

The EMBO Journal

© 2014 The Authors

Elevated levels of Stat92E in the clone (Fig 3G, arrows), a well-established readout of Stat92E activity (Chen et al., 2002). We reasoned that if the ptc mutant phenotype was due to elevated Stat92E levels in CySCs, we could suppress ptc-dependent GSC loss by removing a copy of Stat92E. The number of GSCs in Stat92E+/− heterozygotes was indistinguishable from wild type (Fig 3I, dark green bar). However, GSC loss was actually enhanced when ptc mutant clones were induced in a Stat92E+/− heterozygous background (Fig 3I, red bar, P < 0.008), presumably due to the role of Stat92E in GSC-hub adhesion (Leatherman & Dinardo, 2010).

We also examined whether the ptc mutant phenotype could be ascribed to changes in cell-matrix (integrin) or cell-cell (cadherin) adhesion. We did not detect changes in JPS-integrin in ptc clones (Fig 4A, arrow), in contrast to the observations reported for Socs36E mutants (Iggois et al., 2009). Furthermore, clonal overexpression of JPS-integrin, or a dominant-active form of Talin (TalinH), which strengthens integrin adhesion (Tanentzapf & Brown, 2006), neither recapitulated the ptc phenotype nor induced competition with CySCs and GSCs (Fig 4C and H). Importantly, we found that rhea, which encodes the Drosophila Talin, was dispensable for CySC self-renewal (Fig 4E-G). DE-cadherin levels did not change in ptc mutant clones (Fig 4B, arrow). Moreover, clonal mis-expression of DE-cadherin also did not cause niche colonization (Fig 4D and H). Furthermore, competition caused by ptc mutant clones was not altered by reducing the genetic dose of a-Catenin, which connects Dc-cadherin to the cytoskeleton (Sarpal et al., 2012) (Fig 4I).

Although one mutant allele of rhea partially suppressed the ptc phenotype (Fig 4I, P < 0.65 for rhea+/ and P < 0.051 for rhea−/−), this is likely to be an indirect effect of loosening the tethering of the hub to the muscle sheath and allowing more stem cells to surround the hub (Tanentzapf et al., 2007). Consistent with this, there were more GSCs in testes from rhea+/− heterozygous animals (Fig 4I). These data strongly suggest that increased adhesion does not skew neutral drift dynamics in CySCs.

An alternative explanation for the ptc phenotype is that ptc mutant CySCs induce death in neighboring wild-type cells, akin to classical cell competition in which more robust cells kill and take the place of weaker cells (Amoyel & Bach, 2014). A key process in cell competition is ribosomal function, which in turn is dependent on optimal levels of the cellular growth regulator dMyc and of ribosomal subunits, encoded by Minute genes (M) (Morata & Ripoll, 1975; de la Cova et al., 2004; Moreno & Basler, 2004). Clonal overexpression of dMyc, which causes cell competition in imaginal discs (de la Cova et al., 2004), did not cause niche colonization or loss of GSCs (Supplementary Fig S6A and B). Similarly, testes from a M/+ animal harboring wild-type clones (labeled M+), which normally predominate in a M/+ heterozygous background, contained a normal complement of GSCs (Supplementary Fig S6B). Finally, we found no evidence of cell death in testes with ptc mutant clones (Supplementary Fig S6C and D), and removing a copy of the pro-apoptotic gene hid (which suppresses dMyc-dependent cell competition) did not suppress ptc-dependent competition (Supplementary Fig S6E, red bar).

ptc mutant CySCs proliferate faster than controls

Having ruled out increased JAK/STAT signaling or adhesion as causal factors in niche competition, we reasoned that proliferation might be a driving force of clone dominance within the stem cell pool. We therefore tested whether ptc mutant CySCs had an altered rate of proliferation relative to control clones and, in doing so, might outcompete wild-type CySCs in the race to replace neighbors. Labeling with 5-ethylcytosine (2′-deoxyuridine (EdU) revealed that ptc mutant CySCs had a higher S-phase index than control clones (Fig 5A, B), suggesting that ptc mutant CySCs proliferate faster than controls.

Increased proliferation downstream of ptc is necessary and sufficient for colonizing behavior

We next addressed if the competitive behavior of ptc mutant CySCs depended on their ability to increase their proliferation rate. To accomplish this, we removed one copy of stg and counted the number of labeled CySCs and of GSCs at the niche. In a stg+/− heterozygous background, the number of ptc mutant CySCs was significantly reduced (Fig 6A, P < 0.034), suggesting that ptc mutant CySCs have a reduced competitive advantage when stg is limiting. In addition, in a stg+/− background, the outcompetition of GSCs by ptc mutant CySCs was significantly suppressed (Fig 6B, red bar, P < 0.008). We note that the number of GSCs was not changed in stg+/− heterozygotes when control clones were present (Fig 6B, dark green bar). These data indicate that increased proliferation downstream of Ptc is necessary for niche competition in the
DE-cadherin to the cytoskeleton (Sarpal et al. 2007). Importantly, we found that DE-cadherin also did not cause niche colonization (Fig 4D and H). Nonetheless, we observed a stronger integrin adhesion (Tanentzapf & Brown, 2006), neither the hub nor the muscle sheath were altered by reducing the genetic dose of ptc (Issigonis et al. 2009). Furthermore, clonal overexpression of PS-integrin, or a dominant-active form of Talin (TalinH), which encodes the PS-integrin, or a dominant-active form of Talin (TalinH), which was dispensable for CySC self-renewal (Fig 4E).

We also examined whether the hub (Tanentzapf et al., 2010) and CySCs had a higher S-phase index than control clones (Fig 5A, B). There was an increase in the S-phase index in CySCs mutant for ptc mutant clones. Clones are labeled by loss of the b-gal marker (blue, single channel C). However, GSC loss was actually enhanced when ptc mutant clones were induced in a Minute animal harboring wild-type clones (labeled with a green bar). However, GSC loss was actually enhanced when ptc mutant clones were induced in a Minute animal harboring wild-type clones (labeled with a green bar). Nonetheless, GSC loss was actually enhanced when ptc mutant clones were induced in a Minute animal harboring wild-type clones (labeled with a green bar).

Gain of Hh or Yki cause niche competition (Leatherman & Dinardo, 2008). The EMBO Journal Vol 33 | No 20 | 2014

© 2014 The Authors
Drosophila testis and that CySC–CySC and CySC–GSC competitive interactions are related, making GSC number a good readout for CySC competitiveness.

To corroborate the hypothesis that increased proliferation is necessary for niche competition by CySCs and to determine which pathways are normally active in CySCs, we examined other cellular growth and proliferation factors for their ability to rescue the ptc mutant phenotype when reduced. Removing one copy of the gene E2f, which encodes an S-phase regulator (Duronio et al., 1995), partially suppressed the loss of GSCs; similar genetic interactions
were found with cdk2, which encodes a cyclin-dependent kinase (Lehner & O’Farrell, 1990) (Fig 6C, P < 0.04, and P < 0.02, respectively). We note that GSC number was not changed in E2f/+ or in any of the heterogeneous backgrounds tested below (Fig 6C). In Drosophila, cellular growth and proliferation are genetically separable (Neufeld et al, 1998), so we also tested whether increased cellular growth was required for CySC colonization. Removal of one copy of the gene encoding the Drosophila Insulin receptor, InR, or genes encoding its effectors Akt1 and S6k, did not suppress niche colonization by ptc mutant clones (Fig 6C) (Chen et al, 1996; Montagne et al, 1999; Verdu et al, 1999). In fact, clonal mis-expression of Drosophila Phosphoinositide 3-kinase (PI3K) Dp110, or clonal loss of PI3K pathway inhibitors Tsc1 or Pten, was incompatible with CySC fate (Supplementary Fig S7 and Supplementary Table S2) (Leevers et al, 1996; Goberdhan et al, 1999; Potter et al, 2001; Tapon et al, 2001). Consistently, we also recovered fewer dMyc-expressing CySC clones at 14 dpc compared to control (Supplementary Table S2). Thus, cell cycle progression is essential for CySCs to gain an advantage over their neighbors at the niche, while excessive activation of cellular growth pathways like PI3K and dMyc is detrimental to CySC function.

We next tested whether increasing proliferation was sufficient to cause niche competition by expressing the G1/S-phase promoting factor CyclinE (CycE) and G2/M-phase promoting factor Stg together in a clonal fashion. In imaginal discs, clonal overexpression of these factors together led to marked acceleration of the cell cycle and increased cell number (Neufeld et al, 1998). We found that CycE+Stg overexpressing clones with at least one labeled CySC grew at a faster rate than control clones, indicating that CycE+Stg overexpression also led to cell cycle acceleration in the testis (Fig 6D). Strikingly, CycE+Stg overexpressing clones outcompeted both wild-type CySCs and GSCs at the niche (Fig 6E and F, P < 0.004), in a manner reminiscent of ptc mutant clones. Combined, these data indicate that proliferation downstream of ptc is necessary and sufficient to induce competition at the niche. Thus, altering the rate of cell division skews the stochastic process of stem cell loss and replacement at the niche in favor of the faster proliferating CySCs, and disrupting the normal homeostatic balance between GSCs and CySCs, in favor of the latter.

**The Hippo pathway regulates proliferation, self-renewal, and niche competition independently of Hh**

As a proof of concept for the central role of proliferation in niche competition, we examined a universal regulator of proliferation, the Hippo (Hpo) pathway (Pan, 2010) using clonal assays. Hpo represses the activity of the transcriptional co-activator Yorkie (Yki), the Drosophila homolog of Yes-associated protein (YAP), which is oncogenic in flies and mice (Dong et al, 2007). We noted that this pathway was active in the soma, as seen by expression of the pathway target expanded (ex)-lacZ (Hamaratoglu et al, 2006) (Fig 7A). Next, we generated hpo mutant clones and measured the number of mutant CySCs at several time points. We note that FRT42D CD8-GFP hpo mutant MARCM clones were induced at rates comparable to FRT42D CD8-GFP control MARCM clones (compare Fig 7B to Fig 1D). Strikingly, hpo mutant clones displayed overproliferation and colonized the niche at the expense of wild-type stem cells (Fig 7C and F). Importantly, hpo mutant cyst cells differentiated normally and were readily observed ensheathing spermatogonial cysts (Supplementary Fig S3C).

We applied the same quantitative analysis to hpo mutant clones as described above for ptc mutant clones. Noting that the labeling efficiency of the CySC was comparable to that of the control and ptc mutant (at around 10%), we used the same strategy to analyze the clonal fate data. In doing so, we found that the behavior of hpo mutant clones was consistent with a bias in neutral drift in favor of the mutant cell (Fig 7D and E, compare lines and boxes), quantitatively similar to the trend we found for ptc mutant clones (compare Fig 7D and E with Fig 2C and D). Indeed, within error bars, we could discern no distinction between the bias for ptc and hpo mutants. Furthermore, hpo mutant CySCs displaced GSCs from the niche (Fig 7F, P < 0.0001 hpo versus control), similar to the competitive behavior of ptc mutant clones (Fig 7F, P < 0.0001 ptc versus control). Like ptc-dependent niche colonization, the loss of GSCs caused by hpo mutant CySCs could be suppressed by removing one copy of stg (Fig 7F, dark red bar, P < 0.0001 hpo versus hpo; stg/+).

We next tested the role of the Hpo pathway effector Yki in niche competition. Clonal mis-expression of an activated form of Yki (YkiAct) (Oh & Irvine, 2008) resulted in CySC clones that outcompeted wild-type CySCs and GSCs at the niche (Fig 7G and H, P < 0.0014). Consistent with an essential role of yki in CySCs, we found that yki was required autonomously for self-renewal in CySCs but not in GSCs (Fig 8A, E and F), the latter consistent with a prior report (Sun et al, 2008). Finally, we addressed whether Hh and Hpo, two proliferative pathways in CySCs, were epistatic. To test this, we generated clones that were mutant for both ptc and yki, with the expectation that loss of yki would suppress the competitiveness of ptc mutant CySCs. Indeed, CySCs lacking ptc and yki did not overproliferate and colonize the niche (Fig 8B, C and E, compare red to purple line), indicating that Hpo is epistatic to Hh signaling in the tests. However, we observed no change in
Gain of Hh or Yki cause niche competition

The EMBO Journal

Marc Amoyel et al

Exploratory cell dynamics and niche competition

Characterizing the testis stem cell niche

Discussion

Figure 8D, arrow), suggesting no direct link between these pathways (Fig 8D, arrow), suggesting no direct link between these pathways.
expression of the Yki target gene ex-lacZ in ptc mutant clones (Fig 8D, arrow), suggesting no direct link between these pathways in this tissue. These data establish Yki as a central regulator of somatic stem cell fate in the testis and suggest a parallel requirement for the Hh and Hpo pathways in CySC proliferation, through independent or convergent control of cell cycle progression genes.

Discussion

In this study we characterized the behavior of somatic CySCs in the Drosophila testis and explored the molecular mechanisms that regulate their ability to compete with their neighbors for limited space at the niche. We found that single stem cell clones bias stem cell replacement dynamics in their favor, leading to non-neutral competition, when they had increases in Hh signaling, Yki activity or in the rate of proliferation, but not when JAK/STAT signaling or adhesion were dys-regulated. Furthermore, we found that the dynamics of CySCs were well-described by a model in which they were continually and stochastically lost and replaced, leading to neutral drift dynamics and a consolidation of clonal diversity.

This observation contrasts with the dynamics of GSC offspring fate choices, where oriented divisions and mother centromere retention determine which cells remain as stem cells and which are thrust out of the niche to differentiate (Yamashita et al., 2003, 2007; Sheng & Matunis, 2011). However, careful analysis of GSC dynamics has suggested that they also undergo neutral competition, albeit at a slower loss/replacement rate than CySCs (Wallenfang et al., 2006; Sheng & Matunis, 2011; Salzmann et al., 2013). Thus, within the same stem cell niche, the markedly different strategies for self-renewal are in use, exemplified by the requirement for yki in CySC self-renewal, but not in GSC self-renewal (this study and Sun et al., 2008). This is particularly surprising as the two stem cell populations are by necessity linked, in that they need to produce offspring in the correct ratio, as well as the fact that CySCs support GSC self-renewal through BMP production (Leatherman & Dinardo, 2010). It has been hypothesized that the careful choice of stem cell retention in the GSC pool is a requirement of their role in preserving the genetic integrity of the species (Yuan & Yamashita, 2010). CySCs are under no such constraint, and moreover, need to proliferate twice as fast in order to produce two cyst cells for every germ cyst (Inaba et al., 2011). Thus it may be that the functional imperatives of the tissue (e.g., careful replication of DNA versus rapid production of offspring) determine which type of self-renewal strategy a stem cell adopts.

Characterizing the testis stem cell niche

Our study revealed an unexpected ratio of CySCs to GSCs, close to 1:1 and different from the 2:1 ratio described by Hardy et al. However we note that both studies find the same number of CySCs (approximately 13), and that the difference resides in the number of GSCs. Indeed, Hardy et al. find a ratio of 1.3 CySCs:1 GSC in larval testes which increases to 1.8:1 in young adults, due entirely to a drop in the number of GSCs (Hardy et al., 1979). This may be a function of the genetic background used by these authors, as we established our 1:1 ratio through three different experiments in distinct genetic stocks. Although the analysis of the data is consistent with neutral competition between 13 equipotent CySCs, by the nature of the neutral competition model, we cannot rule out the possibility that the stem cell compartment is heterogeneous with cells moving reversibly between states in which they become primed for duplication or loss, as recently defined in the mouse intestinal crypt (Ritsma et al., 2014). In this case, the effective number of CySCs may be smaller than the observed figure of N = 13, while the true loss/replacement rate, λ, might be proportionately adjusted to a lower value such that the ratio N²/λ remains constant.

Mechanisms of niche competition by CySCs

Our results also show that the predominant force driving niche colonization by CySCs is proliferation. How proliferation causes stem cells to replace neighbors more efficiently is not established by this study. However, we hypothesize that in such a competitive situation, the rate of stem cell loss is not altered but the overproliferating mutants simply produce more offspring, which are in the right place to fill a vacant seat at the niche. It remains possible that a mechanism of active displacement is involved in CySC dominance (i.e., the colonizing stem cells crowd out the wild-type ones), and live-imaging of competing clones might distinguish between passive replacement and active displacement.

A related issue is how CySCs outcompete GSCs. We found that GSC loss is only observed after most of the CySC pool is comprised of colonizing mutant CySCs (Supplementary Fig S4F). We therefore favor the model that competition among CySCs for niche space...
Gain of Hh or Yki cause niche competition

Marc Amoyel et al
precedes that between CySCs and GSCs. It is unclear whether the numerous offspring of the competitive CySCs are passively replacing GSCs that have spontaneously left a vacancy at the niche, or whether colonizing CySCs actively push the GSCs out of the niche. The latter scenario is reminiscent of competition among GSCs in the Droso-
phila ovary, where the contact area between the GSC and niche depended on DE-cadherin. GSCs that elevated cadherins adhered better to the niche and caused the physical displacement of neighbors (Jin et al., 2008; Tian et al., 2012). We explored the contribution of integrin- and cadherin-based adhesion and found that neither affected the competitiveness of CySCs. Moreover, we found that integrin binding was entirely dispensable for CySC self-renewal, unlike cadherin (Voog et al., 2008). Importantly, clonal gain in integrin or cadherin did not lead to niche colonization, indicating that they are not instructive for CySC maintenance. Moreover, we found no role for JAK/STAT signaling in inducing competition at the niche. The fact that neither Stat92E nor integrin was causal to colonization in clonal assays is surprising because both were ascribed critical roles in CySC-dependent niche competition (Issigonis et al., 2009). The reasons for the difference in results by our group and the previous study are not entirely clear. However, we note that gain of Stat92E activity in CySCs in an otherwise wild-type background leads to expansion (not loss) of GSCs because JAK/STAT signaling in CySCs enables their extended niche function to support GSC self-renewal (Leatherman & Dinardo, 2008, 2010). The latter niche role is specific to CySCs and cannot be fulfilled by Hh signal- ing (Leatherman & Dinardo, 2008, 2010). The latter niche role is specific to CySCs and cannot be fulfilled by Hh signaling, another CySC self-renewal pathway (Amoyel et al., 2013). Moreover, our clonal assays (as opposed to lineage mis-expression) are able to recapitulate the constant justling for space at the niche that normally occurs. Regardless, our findings establish that competition and self-renewal are two facets of the same homeostatic process (i.e., proliferation) and that colonizing stem cells have not acquired a new cellular property, but are simply better at self-renewing.

Our study exemplifies how corrupting the naturally occurring process of neutral competition endows a stem cell with greater competitiveness, enabling it to gain dominance within a tissue. Such behavior may be relevant to the early steps of oncogenesis driven by tumor-initiating cells, which have stem cell-like properties (Reya et al., 2001), as in the case of carcinoma, glioma and leukemia caused by sustained Hh signaling (Clement et al., 2007; Zhao et al., 2009; Youssel et al., 2012). The process described here of biasing neutral drift by stem cells harboring oncogenic mutations and the mechanism underlying it appear to be conserved (Vermeulen et al., 2013; Snippert et al., 2014). Taken together, these findings may explain observations such as field cancerization, in which a molecular lesion spreads through a tissue, causing multiple foci of the primary tumor (Vanharanta & Massague, 2012).

**Materials and Methods**

Fly stocks and genotypes are described in Supplementary Materials and Methods. For ptc mutants, ptc<sup>yki</sup> was used in all experiments shown, but similar results were obtained with ptc<sup>hpo<sub>212</sub></sup>. hpo<sup>212</sup> phenotypes were confirmed using hpo<sup>52-47</sup>.

Freshly eclosed adult males were aged for 1 day and then heat shocked for 1 h at 37°C to induce clones and raised at 25°C until the appropriate time for dissection. For self-renewal assays, CySCs were scored as Zfh1-positive or Tj-positive cells one cell diameter away from the hub, and GSCs as Vasa-positive cells in contact with the hub. For control, ptc or hpo CySCs, the method of counting is detailed in the text.

Dissections and immunohistochemistry were performed as previously described (Flaherty et al., 2010). Primary antibodies used were rabbit anti-GFP (1:500, Invitrogen), mouse anti-GFP (1:500, Invitrogen), chicken anti-fl-galactosidase (1:250, Immunology Consultants Lab), goat anti-Vasa (1:400, Santa Cruz), rabbit anti-Zfh1 (1:5,000, gift of Ruth Lehmann), guinea pig anti-Zfh1 (1:1,000, gift of James Skeath), guinea pig anti-Tj (1:3,000, gift of Dorothea Goedt), rabbit anti-Stat92E (1:1,000), mouse anti-Ptc (1:100, DSHB), rat anti-DE cadherin (1:50, DSHB), mouse anti-βPS-integrin (1:20, DSHB), rabbit anti-cleaved caspase 3 (1:50, Cell Signaling).

For 5-ethynyl-2′-deoxyuridine (EdU, Invitrogen) labeling, samples were incubated for 30 min before fixation in Ringer’s medium containing 10 μM EdU. Testes were fixed and processed for 5-ethynyl-2′-deoxyuridine (EdU, Invitrogen) labeling, samples were incubated for 30 min before fixation in Ringer’s medium containing 10 μM EdU. Testes were fixed and processed normally for antibody labeling and then treated per manufacturer’s instructions.

For statistical tests, we used the GraphPad Prism software. To compare two samples, we used the Mann–Whitney U-test to determine significance; for multiple conditions, we used the Kruskal–Wallis test and the Sidak’s multiple comparisons test for post hoc analysis.

The mathematical model is described in Supplementary Materials and Methods.

**Supplementary information** for this article is available online: http://embj.embopress.org

**Acknowledgements**

We thank R. Lehmann, F. Schöck, L. Johnston, D. Kalderon, U. Tepass, J. Treisman, S. Grewal, Y. Yamashita, T. Harris, B. Ohlstein, L. Butitta, D. Godt, Bloomington, and DSHB for antibodies and reagents. We thank members of the Bach laboratory for fruitful discussions. We are extremely grateful to Esteban Mazzoni, Andrew Tomlinson, and Gary Struhl for their generosity in the aftermath of Supertornado Sandy. BDS acknowledges the support of the Wellcome Trust (Grant Number 098357/Z/12/Z). Work in the Bach laboratory is...
supported by grants from the NIH (R01-GM085075-05 and R01-GM085075-05S1) and NYSTEM (C028132 and C024284).

Author contributions
MA designed, carried out, and analyzed the experiments. EAB contributed to experimental design and analysis. BDS contributed to experimental design and provided the mathematical analysis. MA, BDS, and EAB wrote the manuscript.

Conflict of interest
The authors declare that they have no conflict of interest.

References
Gain of Hh or Yki cause niche competition

The EMBO Journal

Marc Amoyel et al


License: This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
The Hippo signaling pathway in stem cell biology and cancer

Jung-Soon Mo, Hyun Woo Park & Kun-Liang Guan*

Abstract

The Hippo signaling pathway, consisting of a highly conserved kinase cascade (MST and Lats) and downstream transcription co-activators (YAP and TAZ), plays a key role in tissue homeostasis and organ size control by regulating tissue-specific stem cells. Moreover, this pathway plays a prominent role in tissue repair and regeneration. Dysregulation of the Hippo pathway is associated with cancer development. Recent studies have revealed a complex network of upstream inputs, including cell density, mechanical sensation, and G-protein-coupled receptor (GPCR) signaling, that modulate Hippo pathway activity. This review focuses on the role of the Hippo pathway in stem cell biology and its potential implications in tissue homeostasis and cancer.

Keywords: cancer; Hippo pathway; regeneration; stem cell; YAP

See the Glossary for abbreviations used in this article.

Introduction: The Hippo signaling pathway

The Hippo pathway is evolutionally conserved and regulates diverse cellular processes, including cell survival, proliferation, differentiation, and organ size. This pathway was initially characterized through clonal genetic screens identifying genes involved in tissue growth control in Drosophila melanogaster. In Drosophila, the core components of the Hippo pathway include the kinase cascade of Ste20-like kinase Hpo (Hippo) and NDR family kinase Wts (Warts) [1–7]. Hpo complexes with the scaffolding protein Sav (Salvador) to phosphorylate and activate Wts, which then forms a complex with its regulatory protein Mats (Mob as tumor suppressor) [8–10]. When in complex with Mats, Wts directly phosphorylates the transcriptional coactivator Yki (Yorkie), sequestering it in the cytoplasm by promoting its interaction with 14-3-3 [11–15]. Conversely, when the Hippo pathway is inactivated, unphosphorylated Yki translocates into the nucleus where it associates with the TEAD/TEF family transcription factor Sd (Scalloped) to initiate gene expression, promoting cell survival and proliferation [16,17]. Yorkie can also bind to other DNA binding proteins including Mad, Homothorax (Hth), and teashirt to promote gene expression [18,19].

The Hippo pathway is a tumor suppressor pathway because mutations in these regulatory pathway components result in an overgrowth phenotype.

In mammals, the Hippo pathway consists of the serine/threonine kinases MST1/2 (mammalian Ste2-like kinases, Hpo orthologs) and LATS1/2 (large tumor suppressor kinase 1/2, Wts orthologs) [7,20–22]. Activation of the Hippo pathway results in the inactivation of YAP (Yes-associated protein, Yki ortholog) by LATS1/2-mediated direct phosphorylation on YAP Ser127 (in humans). Phosphorylated YAP is sequestered in the cytoplasm via binding to 14-3-3 and is degraded in a ubiquitin-proteasome-dependent manner, which depends on phosphorylation of YAP Ser381 and Ser384 [23]. Conversely, dephosphorylated YAP acts mainly through TEAD family transcription factors to promote cell proliferation and organ growth [24]. TAZ (transcriptional coactivator with PDZ binding motif), a paralog of YAP in mammals, is regulated by the LATS1/2 in a similar manner. YAP/TAZ are the major downstream mediators of the Hippo pathway. Besides the TEAD family transcription factors, YAP/TAZ also interacts with other transcription factors including Smad, Runx1/2, p73, ErbB4, Pax3, and T-box transcription factor 5 (TBX5) to mediate the transcription of a diverse array of genes, although the biological functions of these other transcription factors in mediating Hippo signaling are less clear [25].

Although the core signaling cascade from Hpo (MST1/2) to Yki (YAP) is well understood, the upstream regulators of the Hippo pathway are just beginning to be delineated. Interestingly, accumulating evidence from both Drosophila and mammals has shown that apical–basal polarity proteins may regulate the Hippo pathway by controlling YAP/TAZ localization. For instance, earlier studies in Drosophila implicated the apical membrane-associated FERM-domain proteins Mer (Merlin) and Ex (Expanded), which are apical tumor suppressors, and the WW and C2 domain-containing protein Kibra (kidney and brain protein) as components upstream of Hpo. The Mer/Ex/Kibra complex recruits Hpo to the plasma membrane to enhance its kinase activity [26–30]. The apical transmembrane protein Crb (Crumbs) also interacts with Ex and modulates its (Ex) localization and stability [31–35]. Similar to Crb, the Scribble (Scrible) complex (Scrib/Dlg/Lgl) and Par3 polarity complex (Par3/Par6/αPKC) have been implicated in the regulation of the Hippo pathway activity [34,36,37]. In addition, a multitude of other cellular junction
The Hippo pathway in stem cell biology

Jung-Soon Mo, Hyun Woo Park & Kun-Liang Guan*

The Hippo signaling pathway in stem cell biology and cancer. 642

Keywords cancer; Hippo pathway; regeneration; stem cell; YAP

cations in tissue homeostasis and cancer. of the Hippo pathway in stem cell biology and its potential impli-
sensation, and G-protein-coupled receptor (GPCR) signaling, that regeneration. Dysregulation of the Hippo pathway is associated

Abstract

The Hippo pathway in stem cells and cancer

The first cell differentiation event in mammalian development occurs during preimplantation, when the outer blastomeres of the embryo form an outer epithelial trophoderm (TE) that envelops the remaining blastomeres, the inner cell mass (ICM). The TE is necessary for implantation and later contributes to the placenta. Embryonic stem cells (ESCs) are pluripotent cells, derived from the ICM of an early blastocyst, that have the potential to self-renew and differentiate into different cell types and tissues. This pluripotent capacity raises hope for their potential application in regenerative medicine [69].

The association between Hippo signaling and stem cell-like properties has been previously shown. For example, Yap−/− embryos arrest during development around E8.5 and display a yolk sac vascular defect [70]. The Yap target transcription factors, TEADs, are the earliest genes expressed at high levels during embryo development, and TEAD4 is required for specification of the TE lineage during preimplantation of the mouse embryo [71–73]. At the blastocyst stage, TEAD4 promotes expression of multiple genes associated with trophoblast specification, including Cdx2 and Gata3, which are selectively expressed only in blastomeres destined to become TE [71,74,75] (Fig 2A). Moreover, it has been shown that this activity of TEAD4 is dependent on Yap localization in the nucleus, which is modulated by cell–cell contact and LATS1/2-mediated phosphorylation. This finding suggests that Yap localization is essential for TEAD4 activity and cell fate specification [76]. Additionally, NF2 (Neurofibromin 2) and AMOT, two upstream components of the Hippo pathway, facilitate Yap phosphorylation via LATS1/2 during cell fate specification of mouse preimplantation development [77,78] (Fig 2B). Although TAZ is also highly enriched in the developing mouse embryo, inactivation of the gene encoding TAZ, Wt1r, results in only minor skeletal defects and the development of renal cysts, and these mice still grow to adulthood [79]. Altogether, these results demonstrate a critical role for Yap and TEADs in the process of cell fate determination in early mouse embryos [67,73].

Recently, the Hippo pathway has also emerged as a crucial regulator of pluripotency in vitro [80,81]. Initially, BMP and LIF
signals were shown to maintain mouse ESCs in an undifferentiated, pluripotent state, whereas human ESCs require FGF, BMP, and TGF-β/activin [82–84]. Fine-tuning these multiple signaling pathways is crucial in maintaining the balance between differentiation and self-renewal in ESCs. Supporting the role for YAP and TEADs in maintaining pluripotency, the high expression of YAP and TEAD2 in ESCs, neural stem cells, and hematopoietic stem cells initially placed these genes into a general ‘stemness’ transcriptional signature based on transcriptional profiling [85]. Tamm et al found that YAP and TEAD2 could activate the expression of ESC master transcriptional regulators Oct4 and Nanog in mammalian ESCs. Furthermore, restricting YAP and TEAD2 expression or inhibiting TEAD function resulted in differentiation toward the endoderm lineage [86]. Conversely, YAP protein and mRNA levels are significantly decreased with the loss of pluripotent markers during ESC differentiation [87]. In addition, YAP is sequestered and thereby inactivated in the cytoplasm, and consequently, a large number of genes important for stem cell maintenance and function, including PcG, Nanog, Oct3/4, and Sox2, are repressed.

Additional evidence for the role of YAP in pluripotency is seen in induced pluripotent stem cells (iPSCs). The seminal findings by Yamanaka’s group demonstrated that mouse somatic cells can be reprogrammed into iPSCs by inducing the activity of four
transcription factors: Sox, Oct3/4, c-Myc, and KLF4 [88]. YAP is activated during the reprogramming of human embryonic fibroblasts into iPSCs, and the addition of YAP to Sox2, Oct4, and KLF4 increases iPSC’s reprogramming efficiency in mouse embryonic fibroblasts, further confirming a positive role of YAP in stemness [87].

Moreover, it has been reported that the Hippo pathway can interact with other pathways to promote and maintain pluripotency. For example, TAZ associates with Smad2/3 to maintain the nuclear accumulation of Smad complexes, thereby promoting expression of pluripotency markers (Oct4, Nanog) in response to TGF-β stimulation [80]. Another piece of evidence linking the Hippo and TGF-β/BMP pathways is the finding that YAP binds Smad1 to regulate the induction of Id family members for mESC maintenance upon stimulation with BMP [81]. Finally, TAZ has been identified as a coactivator of Pax3-dependent transcription, which influences the expression of various genes during embryogenesis [89].

Thus, in ESCs, YAP/TAZ promotes stemness directly, as well as indirectly by mediating TGF-β/BMP or LIF signaling, through regulating the expression of genes responsible for maintaining pluripotency both in vivo and in vitro [86,87]. These studies implicate the Hippo pathway with those involved in maintaining ESC pluripotency and controlling cell fate specification in development. In conclusion, YAP, TAZ, and TEAD proteins seem to be key regulators for maintaining the pluripotent properties of both ESCs and iPSCs in mammals. The transcription coactivator activity of YAP/TAZ is similarly required for promoting stem cells as well as normal cell proliferation. In addition, TEAD is likely to be involved in both the stemness and proliferation function of YAP/TAZ. However, depending on cell context, YAP/TAZ must induce expression of different genes between stem cells and differentiated cells. It is also possible that besides TEAD, different transcription factors may be used by YAP/TAZ in stem cells compared to differentiated cells to induce downstream target gene expression, thereby promoting and maintaining stem cells.
Liver: Liver progenitor cells and tumorigenesis

The liver is the most important metabolic organ and has a high regenerative capacity and is able to regenerate after more than 70% hepatectomy. Hepatocytes are the predominant cell type in the adult liver and are mitotically quiescent. The regenerative capacity of the liver depends on hepatocyte proliferation, although the liver also contains oval cells (OCs) which are capable of generating a transit precursor compartment. Liver regeneration has been known for many years, although the underlying mechanisms and how the liver senses when it has reached its original size are still poorly understood [90].

Previous studies in Drosophila have implicated the Hippo pathway as a central mechanism that restricts tissue overgrowth during development and it derails under pathological conditions contributes to tumorigenesis [91]. The Hippo pathway impinges on the transcriptional coactivator Yki to regulate the transcription of target genes involved in cell growth, proliferation, and survival. Conservation of mammalian homologs for all the known components of the Drosophila Hippo pathway has facilitated investigation of the physiological roles of Hippo signaling in mammals. While it was already suggested based on the Drosophila data that the Hippo pathway is involved in mammalian tumorigenesis, Dong et al [12] provided functional evidence that the mammalian Hippo pathway is a potent regulator of organ size and that its dysregulation leads to tumorigenesis in the liver. Induction of YAP overexpression using a conditional YAP transgenic mouse resulted in massive hepatomegaly via an increase in the number, but not the size, of the liver cells. Interestingly, the YAP-induced enlarged livers reverted back to their original size without any gross abnormalities when the expression of transgenic YAP was repressed. These data clearly establish a predominant role of YAP in organ size control in mice [92,93]. However, when YAP overexpression was maintained for an extended period of time, the transgenic mice develop liver tumors similar to hepatocellular carcinoma, suggesting a role of hyper-YAP activation in cancer development.

More recently, other components of the Hippo pathway have been shown to repress proliferation and restrict liver growth. Deletion of both MST1 and MST2 results in embryonic lethality [94–96]. However, a single copy of MST1 or MST2 (mice with genotype of either MST1+/−, MST2+/− or MST1+/-, MST2−/-) is sufficient to support normal embryonic development. During later stages of development, loss of MST1 or MST2 promotes proliferation of liver stem cells/progenitor cells such as oval cells. Proliferation of liver progenitor cells gives rise to both hepatocytes and cholangiocytes (biliary epithelial cells, BECs), which are the prominent epithelial cells of the bile duct. This eventually leads to the development of liver tumors due to the loss of heterozygosity of the remaining MST1 or MST2. These mice display characteristics of hepatocellular carcinoma (HCC) and cholangiocarcinoma (CC) with expansion and transformation of a mixed population of tumor-associated liver progenitors [94]. Interestingly, YAP protein levels were increased, while YAP phosphorylation and LATS1/2 phosphorylation were significantly reduced, relative to wild-type, in the absence of MST1/2, indicating that YAP is a downstream effector of MST1/2 in the liver. In contrast, TAZ protein levels and phosphorylation status are decreased in the MST1/2-knockout liver and tumors, suggesting that TAZ, a potential oncogene, is unlikely to play a major role in overproliferation and tumorigenesis in this model [95].

Intriguingly, liver-specific deletion of Savi enhanced proliferation and expansion of hepatic progenitor cells (OCs) and these mice eventually developed liver tumors with a mixed HCC and CC phenotype, distinct from HCC which originates from the aberrant proliferation of hepatocytes only. However, the levels of phosphorylated YAP and phosphorylated LATS1/2 were not affected in the Savi KO livers, suggesting that Sav is likely to play an essential role in OC expansion and tumorigenesis in this model but, surprisingly, acts independently of LATS1/2 and YAP [96,97].

Studies of NF2 conditional knockout mice also support a role for YAP in liver tumorigenesis [93,98,99]. Inactivation of NF2 results in hepatocyte and BEC proliferation, widespread hepatocellular carcinoma, and bile duct hamartomas comprising cytokeratin-positive biliary epithelial cells. Zhang et al [98] reported that NF2 and YAP act antagonistically to each other in the Hippo pathway to regulate liver development and physiology. Deletion of only one copy of YAP was sufficient to reverse the expansion of liver progenitor cells and tumorigenesis driven by the loss of NF2. Consistent with this finding, the NF2-deficient liver showed reduced phosphorylation of YAP and LATS1/2 and increased YAP nuclear localization, providing functional evidence that the main tumor suppressive mechanism of NF2 is mediated through inactivating YAP. On the other hand, EGFR signaling has also been implicated in NF2 deletion-induced tumorigenesis. Pharmacologic inhibitors of EGFR blocked OC expansion and tumorigenesis triggered by NF2 deletion [99]. Benhamouche et al also showed that liver-specific deletion of NF2 leads to an early and dramatic expansion of progenitor cells without any detectable alteration in YAP localization and phosphorylation, arguing against a role for YAP in NF2 KO-induced tumorigenesis. Future studies are necessary to clarify the discrepancy of these two reports regarding NF2 deletion-induced YAP activation [98,99]. However, the general consensus is that NF2 acts upstream of YAP and that other downstream effectors of NF2 may also contribute to tumorigenesis.

Collectively, these data suggest that Hippo pathway components may play an important role in maintaining hepatocyte quiescence and regulating organ size in mammals, yet their dysregulation can lead to stem cell expansion, overgrowth, and tumorigenesis through multiple mechanisms. There are differences in the phenotypes observed in the conditional knockout mouse models of various Hippo pathway components (Fig 3). Thus, further studies are needed to fully elucidate the roles of these Hippo pathway components and their mechanisms of action in regulating of liver progenitor cells.

Skin: Epidermal progenitor cells

The skin (epidermal tissue) in the human body undergoes constant replenishing, completely replacing itself every 2 weeks throughout an individual’s life [100]. The epidermal stem cells are located within the basal layer and have a high proliferative capacity to continuously produce new epidermis while still maintaining structural integrity. During development, the basal epidermal cells generate proliferative progenitor cells, which can only divide for a limited number of cycles; these cells then leave the basal layer, migrate...
Recent findings have implicated the importance of the Hippo pathway in epidermal development and homeostasis. It has been shown that inactivation of SAv1 (WW45) alleles leads to early embryonic lethality, and histological examination displayed a thickening of the epidermal skin layer in the embryos. WW45-null primary keratinocytes show hyperproliferation, progenitor expansion, decreased apoptosis, and inhibition of terminal differentiation.

Figure 3. Schematic illustration of Hippo pathway in liver.
Endodermal progenitors generate hepatocytes and cholangiocytes that surround the bile duct system in adult liver. The Hering canal cells give rise to bipotential oval cells, which are capable of generating both hepatocytes and cholangiocytes. Hepatocyte regeneration is responsible for liver growth after partial hepatectomy. The exposure of the adult liver to hepatotoxins induces the proliferation of oval cell, but hepatocytes are slow to respond or do not respond at all to toxic injury. (A) In the hepatocytes, MST1/2 is activated by proteolytic cleavage that resulted in the loss of the Sav1 interacting SABAH domain. Cleaved MST1/2 is required to phosphorylate Mob1, but Sav1 is not required for Hippo pathway activity. This facilitates YAP phosphorylation, resulting in cytoplasmic retention by 14-3-3 binding and degradation by ubiquitin-proteasome-dependent manner. Lats1/2 activity is unaffected by MST1/2 inactivation in hepatocytes. But loss of NF2 decreases Lats1/2 and YAP phosphorylation, suggesting that the existence of unknown kinase other than MST1/2. Additionally, RASSF family proteins seem to have a role in MST1/2 regulation. Lats1/2 might indirectly inhibit YAP by activating unidentified kinase distinct from MST1/2. (B) In the oval cells, MST1/2-regulated phosphorylation on YAP Ser127 is unaffected by Sav1 inactivation. However, Sav1 regulates YAP protein level and localization via as yet defined mechanisms. There are no clear links between MST1/2 and Lats1/2 activation in oval cells. The mechanism underlying the inactivation of YAP inhibits oval cell proliferation. However, the role of oval cells in liver regeneration remains controversial.

toward the skin’s surface as they terminally differentiate, eventually leading to constant skin remodeling. When the skin is injured, wound healing greatly accelerates this regenerating process by which these inner progenitor cells migrate outwards. Epidermal growth must be carefully balanced, because inadequate proliferation results in the thinning of skin and loss of protection, whereas excessive growth leads to hyperproliferative disorders.
These observations suggest that the Hippo pathway restricts the pool of these progenitor cells. Through molecular and genetic studies, two groups have independently shown that YAP overexpression results in expansion of the epidermal stem cells and progenitor cells in the epidermis [42,102]. Mice carrying the YAP transgene reveal epidermal thickening, hyperkeratosis, and squamous cell-like carcinoma in skin grafts. Conversely, deletion of YAP in the epidermis or disruption of the YAP–TEAD interaction during epidermal development resulted in epidermal hypoplasia and loss of keratinocyte proliferation. This phenotype was attributed to the gradual loss of the epidermal stem/progenitor cells and the progenitor cells' limited capacity for self-renewal.

Surprisingly, deletion of MST1/2 did not lead to epidermal hypoplasia, indicating that YAP is regulated through an alternative mechanism that is not dependent on canonical Hippo pathway components MST1/2 in the skin [41]. Consistent with an MST1/2-independent regulation of YAP, recent studies have shown that MST1/2 is not required for YAP activation by G-protein-coupled receptor (GPCR) signaling. Cell adhesion and α-catenin have also been implicated in YAP regulation. Interestingly, skin-specific deletion of α-catenin, a component of adherens junctions and an important tumor suppressor in epithelia, resulted in keratinocyte hyperproliferation and squamous cell carcinoma that resemble the phenotypes observed in YAP transgenic mice [41]. α-Catenin is considered a critical sensor for cell density and provides the cell with neighborhood information through the formation of density-dependent cell–cell junctions (adherens junctions). Similar to α-catenin, the Hippo signaling pathway has been implicated in cell contact inhibition of proliferation as well as tissue growth control [103]. Notably, α-catenin can directly interact with YAP and suppress YAP function, possibly by sequestering YAP at the plasma membrane and preventing it from entering the nucleus [41]. These findings provide a mechanistic explanation for how α-catenin modulates YAP activity by translating context-dependent information to regulate stem cell proliferation and tissue expansion. It should be noted that there is strong evidence supporting that angiomotin mediates cell–cell contact and tight junction signals to inhibit YAP function by both increasing YAP phosphorylation and physical binding [78,104].

Nervous system: neural progenitor cells

YAP and TEAD2 are highly expressed in neural stem cells (NSCs), which are multipotent progenitors present in the nervous system. NSCs are capable of self-renewing and produce multiple neural lineages which ultimately compose the central nervous system (CNS) [85,105]. In the vertebrate’s developing neural tube, YAP is expressed by ventricular zone progenitor cells and co-localizes with Sox2, a neural progenitor marker [106,107]. Overexpression of either YAP or a transcriptionally active form of TEAD in the neural tube leads to reduced neural differentiation and a marked increase in neural progenitor cell numbers due to accelerated cell cycle progression and recurring cell cycle exit. These effects are associated with the induction of cyclin D1 and the down-regulation of NeuroM. Conversely, loss of YAP triggers cell death and promotes premature neuronal differentiation in the chick neural tube [106]. Both YAP gain-of-function and loss-of-function studies in Xenopus demonstrate that YAP is required for expansion of Sox2+ neural plate progenitors and Pax3+ neural crest progenitors at the neural plate border and for maintaining these progenitor cells in an undifferentiated state. The effects of YAP on Pax3+ neural crest progenitors are through the direct regulation of Pax3 transcription. YAP acts through TEAD to stimulate Pax3 expression. Previous studies have also suggested that mouse TEAD is responsible for activating the Pax3 promoter and neural crest expression in the mouse as well [108]. It is well documented that the expansion of mouse neural progenitors is mediated by the activation of the Notch pathway; however, in the frog embryo, YAP’s ability to repress neural differentiation is likely independent of Notch signaling [107]. It has been shown that YAP is amplified or up-regulated in human Shh-dependent medulloblastoma, a brain tumor in children. Similarly, it was observed that YAP and its target transcription factor TEAD1 are highly expressed in mouse Shh-dependent medulloblastomas [109]. In addition, YAP is a target of Shh signaling in the developing cerebellum. YAP expression and nuclear localization are induced in proliferating cerebellar granule neural precursors, which are thought to be the cells of origin for certain medulloblastomas. Additionally, it has been suggested that mutation of Patched1 (PTCH1), which encodes an inhibitor of hedgehog pathway, leads to the activation of YAP in a non-cell-autonomous manner and alters hedgehog pathway in medulloblastoma cells and tissue samples [110]. These studies show a critical role for YAP and TEAD in neuronal progenitor cells and medulloblastoma development.

Large-scale RNAi screens reveal that Fat1 cadherin, the closest homolog of Drosophila dFat, is spatially restricted to the intermediate regions of the neural tube and acts though YAP to regulate the number of neural progenitor cell pools within the dp4-vpl domain [111]. Loss of NF2 also caused an overexpansion of the neocortical progenitor pool by increasing YAP/TAZ protein levels, enhancing nuclear localization of both these proteins, and up-regulating their target genes in the mammalian dorsal telencephalon [112]. In addition, Hippo signaling had previously been implicated in Fts/Ds signaling through its regulation of cell proliferation and differentiation in Drosophila, although there was no direct evidence to implicate Fts/Ds signaling in regulating the vertebrate Hippo pathway [113]. Cappello et al recently suggested a connection of FAT4/DCHS1 and YAP in mammals. They reported that knockdown of FAT4 or DCHS1 promotes neural progenitor cell proliferation and malpositioning of cells in the developing cerebral cortex [114]. These mouse data demonstrate that reduced levels of FAT4 and DCHS1 increase the activity of unphosphorylated YAP and a YAP-responsive transcriptional reporter. Together, these findings reveal a novel function of NF2 and FAT4 signaling in inhibiting neural progenitor expansion during brain development and establish YAP/TAZ as key effectors.

To date, the proposed model is that YAP promotes NSC proliferation by serving as an effector of the Shh pathway in the brain. A full understanding of the role of the Hippo pathway in NSC requires future studies to examine crosstalk between Hippo and other signaling pathways such as the MAPK, Ephrin, Wnt, and Notch pathways that are also thought to control brain development.
Cardiac progenitor cells and muscle progenitor cells

The fetal heart grows through the proliferation of cardiomyocytes, and following birth, postnatal cardiomyocytes undergo hypertrophy to reach an optimal size. Although it was traditionally believed that the adult human heart lacks adequate myocardium regenerative potential for repair, recent studies have identified endogenous stem cells with the regenerative capacity to repair lost or damaged heart tissue during the late cardiac development of the adult heart [115].

Unlike other tissues such as the liver, the role of Hippo signaling in the heart is less well understood. It has recently been shown that a cardiac-specific deletion of Taz or overexpression of a constitutively active YAP mutant in embryos results in embryos within cardiomegaly due to increased cardiomyocyte proliferation. Ablation of either the MST1/2 or LATS1/2 kinases, the upstream inhibitory kinases of YAP, causes an upregulation of cell death resulting from an overgrowth phenotype due to elevated cardiomyocyte proliferation, similar to the Sav cKO heart [116,117]. Genetic interaction studies have shown that nuclear YAP interacts with β-catenin in cardiomyocytes, directly activating β-catenin target genes to promote Wnt signaling, which has already been implicated in cardiac repair and cell reprogramming. Loss of β-catenin in the Sav cKO hearts suppressed the overgrowth phenotype caused by Hippo pathway inactivation, suggesting that the Hippo pathway restrains cardiomyocyte proliferation and heart size by inhibiting Wnt signaling [118]. Another recent study showed that YAP activates the IFG pathway during heart development, resulting in the inactivation of GSK3β, which inhibits β-catenin degradation [119]. More recently, Xin et al have reported that expression of constitutively active YAP promotes proliferation of adult cardiomyocytes and enhances adult heart regeneration in response to injury. YAP-expressing cardiomyocytes behave similar to embryonic cells with regard to their regenerative potential [120].

Conversely, loss of YAP leads to embryonic lethality through myocardial hypoplasia, due to reduced cardiomyocyte proliferation in the embryonic heart [119,121]. Thus, YAP connects Hippo signaling and other growth-promoting pathways, such as IGF and Wnt signaling, to regulate embryonic and neonatal cardiomyocyte proliferation. This is mediated at least in part by its interaction with β-catenin, directly promoting a stemness gene expression program [117–119,121].

A role for the Hippo pathway in skeletal muscle is beginning to be delineated. YAP overexpression in C2C12 myoblasts and primary mouse muscle stem cells blocks the progression of myoblasts through the myogenic program and preserves the progenitor-like and proliferative properties [122,123]. High YAP expression and activity expands the pool of activated satellite cells, the resident stem cells in skeletal muscle, and prevents the differentiation of this cell population. Interestingly, overexpression of TAZ increases myogenic gene expression in a MyoD-dependent manner, thereby promoting myogenic differentiation [124]. Despite the high level of sequence identity between YAP and TAZ, their opposite effects on muscle progenitor fate is a nice illustration of the complexity and context specificity associated with Hippo pathway activation or inhibition and the resulting transcriptional response. Obviously, further studies need to be carried out in vivo to conclusively determine the role of Hippo signaling, particularly the opposing functions of YAP and TAZ, in cardiac and skeletal muscle biology.

Intestine: Intestinal stem cells

Intestinal stem cells (ISCs) are responsible for the constant renewal and repair of the intestinal epithelium to maintain tissue homeostasis [125,126]. Recent studies have highlighted the role of the Hippo pathway and its effectors YAP and Yki in intestinal regeneration following tissue injury in both mice and Drosophila, respectively. In general, the loss of Hippo signaling and/or the elevated YAP activity is associated with stem cell expansion in various organs [125,127]. However, in the intestine, there are contradictory reports regarding the role of YAP in ISC expansion and intestinal regeneration across different species and experimental settings.

The function of the Hippo pathway and YAP in ISCs has mostly been studied in the context of intestinal regeneration following tissue injury in transgenic animal models (Fig 4). In the DSS-induced colonic regeneration model by Cai et al, YAP protein levels are elevated following tissue injury. In addition, the specific deletion of YAP in the intestinal epithelium prevented DSS-induced intestinal regeneration, suggesting that YAP is required for these processes [128]. Correlating with the function of the Hippo pathway to suppress YAP activity, loss of Hippo signaling in Sav1-deficient crypts displayed accelerated regeneration upon DSS-induced injury in a YAP-dependent manner [128]. Similarly, Zhou et al [129] showed that deletion of the core Hippo kinase MST1/2 in the intestinal epithelium resulted in a marked expansion of the ISC compartments due to YAP hyperactivation. Ubiquitous overexpression of YAP-S127A, which lacks the phosphorylation site required for inactivation by the Hippo pathway, also resulted in the loss of differentiation markers and expansion of an undifferentiated cell population in the mouse intestine [92].

On the other hand, Barry et al [130] reported that specific expression of YAP in the intestinal epithelium suppresses intestinal renewal and reduces the ISC population by restricting Wnt/β-catenin signaling. Intestinal regeneration after irradiation is characterized by hyperactivation of Wnt/β-catenin signaling. Consistently, deletion of YAP resulted in Wnt hypersensitivity and led to ISC expansion and crypt hyperplasia after injury by irradiation. These results are at odds with the role of YAP in the DSS-induced colonic regeneration model.

Another inconsistency is in the crosstalk between YAP and Wnt/β-catenin signaling and their role in intestinal regeneration. The Sav1-deficient mouse colons developed polyps after DSS-induced regeneration, which showed nuclear accumulation of YAP, but not β-catenin [128]. This is consistent with the observation by Barry et al [130] that YAP-S127A expression restricts Wnt/β-catenin signaling during intestinal regeneration. In contrast, Zhou et al [129] reported that in the MST1/2-deficient intestinal epithelium, nuclear accumulation of YAP correlates with β-catenin activation. Uncontrolled tissue regeneration after injury can become oncogenic, like in colon cancer. In this context, Barry et al underscored that YAP is silenced in a subset of highly aggressive human colorectal carcinomas, whereas Zhou and co-workers showed a striking prevalence of YAP overexpression in 95% of colonic cancer specimens [129,130]. The complex nature of YAP in the context of ISC expansion, intestinal regeneration, and its role to Wnt/β-catenin signaling certainly requires further investigation. Nevertheless, these studies point to a role of YAP in ISC, either positively by
directly promoting ISC or negatively by indirectly inhibiting Wnt signaling.

In the *Drosophila* midgut, the Hippo pathway and Yki facilitate intestinal regeneration after tissue injury [131–133]. Perturbation of Hippo signaling or overexpression of a constitutively active Yki mutant (Yki-S168A) induced the expression of the Upd (Out-stretched), which is a cytokine that stimulates expansion of ISC through the JAK/STAT pathway. However, further investigation is required to address whether Upd acts in an autocrine fashion via Hippo-Yki signaling in the enterocytes [132].

**Hippo signaling and cancer stem cells**

As discussed above, the Hippo pathway plays a key role in regulating organ size and tumorigenesis by inhibiting cell proliferation, promoting apoptosis, and regulating stem/progenitor cell expansion [134,135]. Phosphorylated YAP/TAZ localizes to the cytosol, decreasing tumor growth, whereas unphosphorylated YAP/TAZ is localized mainly in the nucleus and promotes cell and tumor growth. Indeed, there is considerable evidence that abnormal Hippo signaling is associated with tumor progression. As expected, elevated expression and activity of YAP/TAZ correlates with various human cancers [103,136–138].

**Figure 4. The context-dependent role of YAP in intestinal stem cell expansion.**

In the intestinal stem cells (ISC), the Hippo pathway inhibits YAP activity by phosphorylation and cytosolic retention of YAP. The cytosolic YAP directly binds to β-catenin and subsequently inhibits the canonical Wnt signaling. In Mst1/2−/− intestinal epithelium, loss of Hippo pathway regulation promotes dephosphorylation and nuclear translocation of YAP/β-catenin and induces their target gene expression. Activation of YAP/β-catenin results in the expansion of ISC. However, a controversial role of YAP has been demonstrated in the context of Wnt-induced intestinal regeneration. In YAP−/− intestinal epithelia, hyperactivation of Wnt/β-catenin signaling results in ISC expansion, whereas YAP overexpression represses Wnt/β-catenin signaling, which leads to the loss of ISC and epithelial self-renewal. In this context, YAP functions to inhibit the nuclear translocation of disheveled (Dvl).
Moreover, TAZ has been shown to be a key regulator of cancer stem cells (CSCs) in breast cancer [139]. In addition, YAP and TAZ are highly expressed in CSCs of medulloblastomas [109]. Increasing evidence has suggested that tumor growth is dependent on CSCs, which represent a small subset of cells within a tumor but have the ability to self-renew, differentiate into other tumor cell types, and initiate tumor formation. CSCs are also thought to be resistant to chemotherapeutic agents and are responsible for cancer recurrence and metastasis. High-grade tumors are characterized by a higher population of CSCs within the tumor. Microarray analysis of 993 primary human breast tumors has identified a list of genes highly expressed in G3 (tumors that poorly differentiated tumors) compared to G1 (benign tumors) [139]. Interestingly, elevated YAP/TAZ activity is observed in G3 tumors, which are also characterized by the expression of embryonic and normal mammary stem cell genes. Using a model for tumor progression, Cordenonsi et al demonstrated a role for TAZ in breast cancer cells [139]. Upon injection in mice, MII cells, which are Ras-transformed MCF10A-T1K, generate low-grade tumors. On the other hand, MIV cells, which are malignant MCF10A-CA1a cells derived from the in vivo spontaneous evolution of MII cells, readily formed tumors resembling G3 tumors. TAZ was highly expressed in the MIV cells, but not the MII cells, whereas YAP levels were comparable across both cell lines. Overexpression of active TAZ increases MCF10A proliferation and the formation of invasive carcinomas. These observations support an important role of TAZ in breast cancer stem cells.

Other studies have shown that nuclear TAZ is highly expressed in high-grade glioblastomas. Ectopic expression of TAZ leads to increased invasion, self-renewal, and tumor initiating capacity to generate properties similar to mesenchymal-like stem cells [140]. Conversely, knockdown of TAZ expression in mesenchymal-like stem cells decreases their mesenchymal properties and limits their capacity to self-renewal and initiate in glioma. Collectively, it is clear that TAZ enhances the self-renewal capacity and tumorigenic potential contributing to both the initiation and progression of breast cancer and glioma. Therefore, TAZ could be a potential molecular target for treating aggressive tumors that have uncontrolled TAZ activation.

Dysregulation of the Hippo pathway has been identified in a broad range of human cancers, including liver, lung, colorectal, ovarian, and prostate [12,103,137]. Studies have shown that YAP activity is increased as a result of increased expression and nuclear localization in human tumor samples. This is consistent with inactivation of the Hippo pathway which is known to inhibit YAP and TAZ activity mainly by promoting these transcriptional coactivators’ cytoplasmic localization and ubiquitin-mediated degradation. In addition, YAP gene amplification (somatic mutation) has been reported in various human and murine tumor models [136,141]. Collectively, these data suggest that unrestrained YAP activity can counteract classical tumor suppressor checkpoints.

Compared with other well-known oncogenic signaling pathways, only few cancers are known to be associated with a direct mutation of a Hippo pathway component. Of note, Lats2 is mutated in approximately 40% of mesothelioma cases [142]. Interestingly, Mst1/2 and Lats1 are tumor suppressors in mice, and although mutation in these genes have not been identified in human cancer, silencing of these genes have been reported to data, suggesting that these genes may be inactivated by non-mutational mechanisms [94–96,129,143–146]. NF2 is a potent upstream regulator of the Hippo pathway, and an inactivating mutation in NF2 is associated with several human cancers including acoustic neuromas, meningiomas of the brain, and schwannomas of the dorsal roots of the spinal cord [146,147]. A high frequency of NF2 mutations has also been reported in mesothelioma [148,149]. Recently, a TAZ and calmodulin-binding transcription activator 1 (CAMTA1) fusion gene has been reported in epithelioid hemangioendothelioma, a rare form of sarcoma [150,151]. The role and mechanism of this fusion protein in cancer progression is still unclear, but may relate to the transcriptional regulatory functions ascribed to both TAZ and CAMTA1.

Many studies have reported a high frequency of mutations in various GPCRs (GPR98, GRM3, ACTRI1, LPHN3, and BAII3) and G-proteins (GNAS, GNAQ, and GNAO1) across a wide range of cancers, particularly in melanoma [152–154]. Notably, activating mutations in GNAQ and GNAI1 have been observed in approximately 50% of uveal melanomas. And in these uveal melanomas with activating mutations in GNAq or GNA11, we found that YAP is constitutively activated and its activation is pathologically critically important.

Collectively, extensive studies have established a critical role for the Hippo pathway in human tumorigenesis. Inhibiting YAP/TAZ may be a new therapeutic area for treating cancers with a dysregulated Hippo pathway.

Conclusions

Although most of the Hippo pathway components were initially identified in Drosophila, much research has recently been done in mammalian cells and animal models, revealing this pathway’s important contribution to tissue homeostasis, organ size control, cancer development, and stem cell biology. As the key downstream effectors of the Hippo pathway, YAP/TAZ is involved in embryonic stem cells as well as tissue-specific stem cell self-renewal, and tissue regeneration and homeostasis of the liver, intestine, pancreas, heart, skin, and central nervous system. Moreover, compelling evidence supports a role for YAP/TAZ in cancer stem cells. Therefore, components of the Hippo pathway may be good therapeutic targets in diseases such as degeneration and cancer.

Since the discovery of the Wrts kinase in Drosophila in 1995, for the first decade research in the Hippo pathway was largely limited to Drosophila. However, rapid progress, especially in the last several years, has been made regarding the identification of upstream components, signals, and mechanisms of regulation in both Drosophila and mammalian systems. Cell polarity, adhesion, mechanotransduction, as well as diffusible signals acting through GPCRs, have all been identified as regulators of Hippo pathway activity. However, many key questions remain to be addressed. The function of YAP/TAZ has been investigated in only a few cell types. Further studies to uncover the physiological roles of YAP/TAZ in a broad range of tissue-specific stem cells and various types of cancer stem cells will likely expand our knowledge of the Hippo pathway in regulating tissue homeostasis during development and adulthood as well as cancer initiation and metastasis. Organ size regulation is a fundamental question in biology, though the signals critical for sensing organ size control, with each organ presumably having its own specific signals, are unknown. Research into the molecular signals controlling organ size will be of paramount importance not
only for the Hippo pathway but also for the field of developmental biology. Because the Hippo pathway is regulated by a wide range of signals, both physical and chemical, how the Hippo pathway integrates all of these inputs from multiple signaling pathways to generate a concerted cellular response remains a question of high interest. Understanding the molecular mechanisms by which the Hippo pathway controls development, regeneration, tissue homeostasis, and injury/repair will require the input of researchers across multiple disciplines, including genetic, genomic, developmental, systems biology, cell biology, biochemistry, and cancer biology. Given the increasing research interest in this pathway, continued rapid progress is eagerly anticipated.

Acknowledgments
We apologize for the many important contributions to the Hippo pathway field that we could not cite owing to space constraints. The authors would like to thank Steven W Plouffe and Dr. Carsten G. Hansen for critical reading of this manuscript. This work is supported by grants from the National Institutes of Health to K.L.G. (CA132809 and EYO226116).

Conflict of interest
The authors declare that they have no conflict of interest.

References
We apologize for the many important contributions to the Hippo pathway and the role of the Hippo pathway in cell proliferation and differentiation. Development 127: 1315–1324


Hippo pathway in stem cells and cancer


Hippo pathway in stem cells and cancer  Jung-Soon Mo et al
Heterotrimeric G proteins control stem cell proliferation through CLAVATA signaling in Arabidopsis

Takashi Ishida1,†, Ryo Tabata1,‡, Masashi Yamada2,3,†, Mitsuhiro Aida4, Kanako Mitsumasu1, Masayuki Fujiwara5, Katsushi Yamaguchi5, Shuji Shigenobu5, Masayuki Higuchi4, Hiroyuki Tsuji4, Ko Shimamoto4, Mitsuyasu Hasebe6,7, Hiroo Fukuda2 & Shinichiro Sawa1,*

Abstract

Cell-to-cell communication is a fundamental mechanism for coordinating developmental and physiological events in multicellular organisms. Heterotrimeric G proteins are key molecules that transmit extracellular signals; similarly, CLAVATA signaling is a crucial regulator in plant development. Here, we show that Arabidopsis thaliana Gβ mutants exhibit an enlarged stem cell region, which is similar to that of clavata mutants. Our genetic and cell biological analyses suggest that the G protein beta-subunit1 AGB1 and RPK2, one of the major CLV3 peptide hormone receptors, work synergistically in stem cell homeostasis through their physical interactions. We propose that AGB1 and RPK2 compose a signaling module to facilitate meristem development.

Keywords Arabidopsis thaliana; heterotrimeric G protein; peptide hormone; RECEPTOR-LIKE PROTEIN KINASE 2; stem cell homeostasis

Introduction

Coordinated cell proliferation and cell differentiation are essential processes in multicellular organisms. To achieve these functions, organisms have developed scrupulously designed cell-to-cell communication systems over the course of evolution. Plants have established unique ligand-receptor-based signaling modules, such as the CLAVATA (CLV) pathway, which compromises the CLV3 peptide hormone and the extracellular leucine-rich repeat (LRR) domain-containing receptors CLV1, CLV2-CORYNE (CRN)/SUPPRESSOR OF LLP1 2 (SOL2), and RECEPTOR-LIKE PROTEIN KINASE 2 (RPK2) [1,2]. In the shoot apical meristem (SAM) of Arabidopsis thaliana (Arabidopsis), CLV signaling restricts the expression of the homeobox-containing transcription factor WUSCHEL (WUS) [3–5]. Conversely, WUS promotes the expression of CLV3, forming a negative feedback loop that controls the number of stem cells [3–5]. While the peptide-binding plasma membrane components have been well studied, the molecules that mediate intracellular signaling by these receptors are largely unknown. The protein phosphatase KAPP and a Rho GTPase-related protein have been shown to physically interact with the CLV1 receptor [6], and the protein phosphatase 2Cs POLTERGEIST (POL) and POL-LIKE 1 (PLI) are also known to be signaling mediators [7]. However, further analyses are needed to trace the signaling pathway from the receptor to cellular processes.

On the other hand, heterotrimeric G proteins, composed of alpha (Gα), beta (Gβ), and gamma (Gγ) subunits, are important signaling molecules that link extracellular signals to intracellular mechanisms in eukaryotes [8,9]. The basic components and mechanisms of G protein signaling have been studied extensively in mammalian cells: G protein-coupled receptors (GPCRs) sense extracellular ligands and stimulate G proteins, whereupon Gα and Gβγ dissociate and provoke variable cellular events [8]. In resting cells, GDP-bound Gα associates with Gβγ, and ligand-stimulated GPCRs promote the exchange of GDP for GTP, causing Gα and Gβγ to dissociate. Although land plants express similar G protein components, they are controlled by slightly different systems compared with canonical G proteins. In plants, Gα can spontaneously exchange GDP for GTP, while the seven-pass transmembrane domain-containing protein RGS1 inhibits G signaling through the formation of an inactive
complex [10–12]. Extracellular ligands bound to RGS1 stimulate the release of Gα to activate signaling. This self-activating ability helps explain the absence of clear GPCR homologues in plant genomes [13]. Several transmembrane proteins have been annotated as plant GPCRs based on their sequence; however, clear evidence that these candidates function as GPCRs has not been reported [9,13]. These facts provide possibilities for the mode of G signaling which is the presence of alternative, non-canonical GPCRs and GPCR-independent function of G proteins. Despite their unique regulatory mechanisms, plant G proteins are involved in various aspects of morphological and physiological processes, much like their mammalian counterparts [12,14–18]. Recently, Bommert et al [19] reported genetic evidence that maize Gα modulates CLV signaling in the control of shoot meristem size. However, the biochemical and cell biological processes underlying the cross-talk between the CLV pathway and G proteins remain unclear, as these extraordinary phenotypes have only been reported for Gα mutants in maize [19]. Here, we show that the Arabidopsis G protein beta-subunit1 (agb1) mutant exhibits an enlarged SAM, similar to that of clv mutants. Genetic analysis suggests that AGB1 works together with RPK2, a leucine-repeat-receptor-like kinase (LRR-RLK), in stem cell homeostasis. Bimolecular fluorescence complementation (BiFC) assays and co-immunoprecipitation (co-IP) analyses indicate that AGB1 associates with RPK2. These results establish the involvement of AGB1 in meristem development in the RPK2-dependent signaling pathway and indicate the diversity of CLV signaling in plants.

Results and discussion

Identification of mutations in a gene encoding a heterotrimeric G protein β subunit in clv2 enhancer 1 mutants

To decipher the molecular mechanisms underlying the CLV signaling pathway, we conducted a genetic screen to search for mutations that enhance the phenotypes of clv2 mutants. As a result, we isolated 48 mutants with obviously enlarged SAMs, which have been designated clv2 enhancer (clen) mutants. From these mutants, clv2 enhancer 1 (clen1) was selected for further study. clv2-101, a null allele of clv2 mutation, displays approximately twofold (21.25 μm) larger SAMs than wild-type plants, in which SAM height is 10.18 μm on average, and the clv2 alleles can1 double mutant exhibits sevenfold (71.73 μm) larger SAMs (Fig 1A–C and G). Similarly, the pistils of wild-type flowers have 2 carpels, whereas clv2-101 and double-mutant plants present 2.5 and 3.8 carpels on average, respectively (Fig 1H–J and N). The enlarged SAMs and increased carpel numbers observed in clv2 clen1 mutants relative to wild-type or clv2 mutant plants suggest decreased CLV3 signaling activity [3]. Using a positional cloning approach, the clen1 mutation was roughly mapped to near the nga1139 marker (33/34 chromosomes) on chromosome 4, and the genomic DNA sequence was analyzed via the SOLiD system to identify mutations [20]. Thus, we detected a nucleotide substitution in the AGB1 gene that converts a Trp residue into a stop codon (Supplementary Fig S1). In addition to the clen1 mutant, we have identified this mutation in 8 clen mutants and have found other four point mutations in five additional clen mutants (Supplementary Fig S1).

The nonsense mutation in the AGB1 gene is expected to be responsible for the observed clen1 phenotypes. We therefore examined SAM size and carpel number in the previously isolated agb1-2 and clv2 agb1-2 mutants [17]. agb1-2 produced 1.4-fold larger SAMs than wild-type plants, while the carpel number was still 2 (Fig 1D, G, K and N), suggesting that the single mutation in AGB1 is sufficient to affect SAM height. Moreover, the clv2 agb1-2 double mutant exhibited a clv2 clen1-like phenotype, showing 6.5-fold larger SAMs and 1.9 times the number of carpels compared with wild-type plants (Fig 1E, G, L and N). Although overexpression of AGB1 did not affect plant architecture in the wild type (Supplementary Fig S2), it suppressed the enhanced abnormalities of the clv2 clen1 mutant, resulting in a clv2-like phenotype (Fig 1F, G, M and N). These plants also resembled a clv2 clen1 mutant that harbors a genomic fragment of the AGB1 gene (Supplementary Fig S3). These results show that a mutation in AGB1 enhances the abnormalities of the clv2 mutant and suggest that AGB1 regulates SAM height and carpel number.

AGB1 is involved in SAM maintenance through the CLV3 signaling pathway

CLV3 restricts cell proliferation in the SAM, and synthetic CLV3 peptide treatment induces SAM consumption due to diminished cell proliferation [5,21]. To investigate whether the enlarged SAM phenotype observed in agb1 is a consequence of a disturbance of CLV3 signaling, we examined the sensitivity of the agb1-2 mutant to the CLV3 peptide (Fig 2). Wild-type seedlings grown on MS media containing 5 μM CLV3 did not develop stems under these conditions, even at 20 days after germination (Fig 2E, M and N). Conversely, 10% of agb1-2 mutants developed a stem at the same stage (Fig 2G, O and Q). Furthermore, clv2 agb1-2 double mutants showed strong resistance compared with clv2 or agb1-2 mutants (Fig 2H, P and Q).

Next, we examined the genetic relationship between AGB1 and WUS, which is known to function downstream of CLV signaling [4,5]. Similar to the wus-101 single mutant, the SAM was terminated in the wus-101 agb1-2 double mutant (Supplementary Fig S6), indicating that WUS is epistatic to AGB1. Taken together, these findings suggest that AGB1 regulates the SAM activities through a CLV3-related pathway.

Heterotrimeric G proteins are expressed in the inflorescence meristem

Given that AGB1 is a heterotrimeric G protein subunit, the involvement of other G protein components was predicted. To examine whether G proteins are expressed in SAMs, we performed in situ mRNA hybridization experiments. Expression of both GPA1 and AGB1 was observed in the inflorescence meristem, floral meristem, and floral organ primordium (Supplementary Fig S7A and B). Conversely, weak AGG1 expression signals were detected, whereas AGG2 was not (Supplementary Fig S7C and D). The expression of these genes in vegetative SAMs and inflorescences was supported by the Arabidopsis eFP Browser database (Supplementary Fig S7E–G). Thus, the fact that not only AGB1 but also other G protein components were expressed in SAMs highlighted the possibility that the G protein signaling complex is involved in the CLV signaling pathways.
The heterotrimeric G protein γ subunit, but not Gα, is also involved in CLV3 signaling

To investigate the possibility that the Gα and Gγ subunits are involved in CLV signaling, we examined both SAM height and CLV3 peptide sensitivity in Gα and Gγ null mutants, designated gpa1-4 and agg1-1c agg2-1, respectively [22,23]. A recent report showed that a maize Gα mutant exhibited a very large meristem phenotype [19]. In contrast, gpa1-4 plants did not show obvious SAM or carpel abnormalities, and the mutation did not affect the degree of resistance to CLV3 compared with the wild type (Supplementary Table S1, Supplementary Fig S8A). Furthermore, the additional mutation of GPA1 did not affect the clv2 mutant phenotype (Supplementary Table S1). Conversely, the agg1-1c agg2-1 double mutant produced 1.4-fold larger SAMs than the wild type as well as 2-carpel siliques, and 10% of the Gγ mutants maintained a SAM even on CLV3-containing media (Supplementary Table S1, Supplementary Fig S8A), similar to what was observed in the agg1-2 mutants. Furthermore, clv2 agg1-1c agg2-1 triple mutants showed enhanced abnormalities and SAMs were maintained in the triple mutant at the similar frequency as in the clv2 agg1 double mutants when treated with CLV3 (Supplementary Table S1, Supplementary Fig S8A).

In mammals and plants, Gβ and Gγ are known to form a heterodimer [8,9,24]. Arabidopsis Gγ appears to act with Gβ during CLV signaling to regulate SAM height and carpel number. This idea is also supported by the results of the examination of carpel number phenotypes at higher temperatures. Morphological abnormalities in flowers are occasionally strengthened at higher temperatures [25]. Accordingly, the agg1-2 and agg1-1c agg2-1 mutants both exhibited 3-carpel pistils, whereas wild-type and gpa1-4 plants all presented 2...
Gbeta and CLAVATA control stem cells in plants

Takashi Ishida et al

2014

ª

Figure 2. AGB1 is involved in SAM maintenance through CLV3 signaling.

A–P Eighteen-day-old seedlings of wild-type (A, E, I, and M), clv2 (B, F, J, and N), agb1-2 (C, G, K, and O), and clv2 agb1-2 (D, H, L, and P) plants. The plants were grown on agar medium with (E–F, M–P) or without (A–D, I–L) 5 µM CLV3 peptide. (I–P) represent closer views of (A–H). Scale bars = 1 cm in (A–H), 1 mm in (I–P).

Q Quantification of the seedlings showing terminated SAMs observed 20 days after germination.
carpels (Supplementary Fig S8B). The difference between the G\alpha mutant phenotypes and those of the G\beta and G\gamma mutants is not unexpected, as G\alpha mutations are often reported as the exception among G protein mutants [13].

**G\beta\gamma controls SAM maintenance in the RPK2 pathway**

The fact that the clo2-101 agb1-2 double mutant showed detectable additive phenotypes suggests that AGB1 mediates CLV3 signaling in a CLV2-independent manner. In contrast to clo2 agb1-2 mutants, which exhibit 3.1-fold larger SAMs than clo2 mutants, the clo1 agb1-2 and rpk2 agb1-2 mutants exhibited 2-fold and 1.5-fold larger SAMs than the corresponding single mutants (Fig 3A). Furthermore, comparing rpk2 and rpk2 agb1-2 plants revealed similar carpel numbers, whereas the clo2 agb1-2 and clo1 agb1-2 mutants showed a significantly increased carpel number relative to the single mutants (Fig 3B). All of the double mutants showed a significantly enhanced phenotype with the corresponding single mutants. However, the degree of enhancement was smaller in rpk2 agb1 for SAM height and only the rpk2 agb1 produced a similar number of carpels when compared with the rpk2 single mutant. These results lead us to hypothesize that AGB1 is at least partially involved in the RPK2-dependent CLV signaling pathway.

**AGB1 associates with RPK2 in planta**

Based on the results of the genetic analyses, AGB1 is expected to interact with CLV signaling components either directly or indirectly. AGB1 has been observed to localize to the plasma membrane, nucleus, cytoplasm, and Golgi apparatus [24,26–28]. Our expression analysis confirmed the presence of AGB1-GFP signals at the plasma membrane (Supplementary Fig S9). Because LRR-containing receptor complexes also localize to the plasma membrane, any complex that these proteins form is likely to be present here.

We next tested whether AGB1 associates with CLV1, RPK2, and CLV2 using BiFC assays. Protoplasts transformed with AGB1 and RPK2 exhibited a positive BiFC signal when a CLV3-expressing vector was co-transformed, and the signal was localized to the plasma membrane (Fig 4A and C). However, we did not detect an interaction between AGB1 and either CLV1 or CLV2 (Fig 4A and C). AGB1 is therefore predicted to receive CLV signals through RPK2, though it is unclear how CLV3 facilitates this interaction. Furthermore, we performed a co-IP assay to confirm the physical interaction between AGB1 and RPK2. FLAG-tagged AGB1 was pulled down with Venus-tagged RPK2c, which contains the C-terminal intracellular domain, whereas most of the AGB1-FLAG disappeared after IP when expressed alone or with mCherry-Venus (Fig 4D). These results suggest that RPK2 is capable of interacting with AGB1. Taking these results together with the genetic data, we propose that AGB1 functions preferentially with RPK2 on the CLV signaling to regulate cell proliferation activities in SAMs.

**Conclusion**

Heterotrimeric G proteins are evolutionarily conserved signaling molecules that mediate the transduction of extracellular cues into intracellular signals, in combination with transmembrane GPCRs. In plants, several transmembrane proteins have been reported as GPCRs. In this study, we have shown that an LRR-RLK receptor, RPK2, is able to interact with G proteins. Surprisingly, among the examined G protein mutants, only the G\alpha mutant did not exhibit any abnormalities in CLV signaling-related processes, suggesting that mutations in GPA1 did not disrupt CLV signaling. However, we cannot exclude the possibility that G\alpha or related proteins serve as a bridge between RPK2 and G\beta\gamma dimers, as in canonical GPCR and G protein interactions. The *Arabidopsis* genome encodes three extra-large G proteins (XLGs), which contain a G\alpha-like domain.
As any potential overlapping functions of XLGs were not addressed in this study, further analyses will be needed to evaluate the biological relevance of XLGs not only in CLV signaling but also in G protein function. In fact, a recent report showed that maize Gα mutants exhibit enlarged meristem phenotype leading the authors to infer a function of Gα in SAM maintenance [19]. Despite clear evidence of the involvement of Gα in the maize CLV-like pathway, further research is required before any generalizations can be made because severe phenotypes, such as those observed in the maize Gα mutant, have not been reported in either Arabidopsis or rice Gα mutants [13, 19, 30]. The critical amino acids for Gα function have been reported. In particular, the Thr residue in the switch I region of GPA1 is important for the interaction between regulatory proteins for activation [31]. Although Arabidopsis, rice, and other species, including dicots, gymnosperms, and animals, harbor the conserved amino acid in the Gα subunit, maize and some other monocots do not exhibit this residue [12]. This evolutionarily distinct background of heterotrimeric G proteins could be another explanation for the differing meristem phenotypes observed in maize compared with Arabidopsis (Fig 1 and [12, 19]).

Taken together, our results suggest the hypothesis that CLV3-RPK2 signaling activates a heterotrimeric G protein through an interaction, at least in Arabidopsis. Thus, these results support the notion that LRR-RLK-type receptor RPK2 acts as an alternative GPCR, similar to canonical GPCR-G protein systems. This situation contrasts with that in maize, where FEA2, an LRR type receptor, mediates CLV-like signaling and G proteins [19]. Although

---

**Figure 4. AGB1 is able to interact with RPK2.**

A. Protoplasts expressing the indicated proteins tagged with the N- or C-terminal halves of Venus.
B. Positive and negative controls for the BiFC analysis. BiFC signals (upper), mCherry fluorescence (middle), and merged (bottom) images are shown.
C. Quantification of the BiFC assays. The results for the positive control and the experiments for CLV1, CLV2, RPK2, and AGB1 are shown. BiFC signals were measured as described in Materials and Methods. The percentages of cells with BiFC signals are indicated by yellow bars (n = 20).
D. Co-IP assay showing the physical interaction between AGB1 and RPK2. AGB1-FLAG alone or with mCherry-Venus or RPK2-Venus was transiently co-expressed in protoplasts. Total protein extracts were subjected to IP with an anti-GFP antibody. The presence of AGB1-FLAG (upper) and Venus-tagged proteins (bottom) was determined by Western blotting. Note that the AGB1-FLAG co-expressed with RPK2c-Venus was condensed after IP. The co-IP experiments were repeated three times, with similar results.

Data information: Scale bars = 10 μm in (A) and (B).

---

[9, 29]. As any potential overlapping functions of XLGs were not addressed in this study, further analyses will be needed to evaluate the biological relevance of XLGs not only in CLV signaling but also in G protein function. In fact, a recent report showed that maize Gα mutants exhibit enlarged meristem phenotype leading the authors to infer a function of Gα in SAM maintenance [19].
Arabidopsis and maize utilize common G proteins for meristem maintenance, different systems consisting of various combinations of receptors and G protein subunits are employed. Therefore, these differences might contribute to the diversity of the signaling pathways in plant development.

Materials and Methods

Plant materials

The following Arabidopsis wild-type and mutant lines were obtained: wild-type Columbia-0 (Col-0); clv1-101 (CS858348) in the Col-2 background; clv3-8 ER in an unknown background (CS3604) [32]; rpki-2 [5], clv2-101 (GK686A09), gpa1-4 (SALK_001846), agt1-2 (CS6536), and wus-101 (GK870H12) in the Col-0 background; and agg2-1 agg3-1 (kindly provided by Jimmy Botella) in a mix of Col-0 and Wассilewskija (Ws).

SAM measurement

Seven-day-old seedlings were fixed with 70% ethanol, cleared in a mixture of chloral hydrate, glycerol, and water (8:1:2; w/v/v), and observed using a microscope (ZEISS AXIO Imager M1) that was equipped with Nomarski optics. The base of the SAM was defined as the location of the leaf primordium, and the height was measured between the top and base of the SAM, as described [5].

Peptide assay

The CLV3 peptide was synthesized as described previously [33]. Seedlings were grown on MS plates containing CLV3 peptides until 18–20 days after germination.

Protoplast transformation

Arabidopsis leaf mesophyll protoplast transformation was performed as described previously [34]. True leaves of 3-week-old seedlings were collected and chopped in an enzyme solution containing 0.6% Cellulase ‘ONOZUKA’ RS (Yakult Pharmaceutical Industry) and 0.6% Macerozyme R10 (Yakult Pharmaceutical Industry). Isolated protoplasts were washed and re-suspended at a concentration of 2 × 10^7 protoplasts per ml for polyethylene glycol (PEG)-mediated transformation. Vectors for transient expression were mixed with the protoplasts in transformation buffer [0.4 M mannitol, 0.1 M Ca(NO₃)₂, and 40% PEG (w/v) (Sigma)]. After washing, the protoplasts were incubated in liquid culture medium containing 0.4 M mannitol for 12–24 h at 23°C.

Further experimental details are provided in Supplementary Methods.

Supplementary information for this article is available online: http://embor.embopress.org

Acknowledgements

We thank A Miyawaki (RIKEN), S Takayama (NAIST), and M. Kakita (Nagoya University) for the BiFC vectors and N Inada (NAIST) for confocal microscopy. This work was supported by grants from KAKENHI (22150002, 23119517, 23012034, 24114001, 24114009, 24370024, 24657035, and 24658032 to SS; 25119713 and 25440134 to TI) and the NIBB cooperative research program (12-103) to SS.

Author contributions

MH, HT, KS, MH, HF, and SS conceived or designed the experiments. TI, RT, MY, MA, KM, MF, KY, and SS performed the experiments. KY, SS, and MH analyzed the data. TI, RT, MY, MA, and SS wrote the manuscript. KS was deceased on September 28, 2013.

Conflict of interest

The authors declare that they have no conflict of interest.

References

3. Clark SE (1995) CLAVATA3 is a specific regulator of shoot and floral meristem development affecting the same processes as CLAVATA1. Development 121: 2057


Targeted gene therapy and cell reprogramming in Fanconi anemia

Paula Rio1,2,†, Rocío Baños1,2,†, Angelo Lombardo3,†, Oscar Quintana-Bustamante1,2, Lara Alvarez1,2, Zita Garate1,2, Pietro Genovese3, Elena Almarza1,2, Antonio Valeri1,2, Begoña Díez1,2, Susan Navarro1,2, Yaima Torres4, Juan P Trujillo2,5, Rodolfo Murillas6, Jose C Segovia1,2, Enrique Samper4, Jordi Surralles5, Philip D Gregory7, Michael C Holmes7, Luigi Naldini3,8,** & Juan A Bueren1,2,*

Abstract

Gene targeting is progressively becoming a realistic therapeutic alternative in clinics. It is unknown, however, whether this technology will be suitable for the treatment of DNA repair deficiency syndromes such as Fanconi anemia (FA), with defects in homology-directed DNA repair. In this study, we used zinc finger nucleases and integrase-defective lentiviral vectors to demonstrate for the first time that Fanca can be efficiently and specifically targeted into the AAVS1 safe harbor locus in fibroblasts from FA-A patients. Strikingly, up to 40% of FA fibroblasts showed gene targeting 42 days after gene editing. Given the low number of hematopoietic precursors in the bone marrow of FA patients, gene-edited FA fibroblasts were then reprogrammed and re-differentiated toward the hematopoietic lineage. Analyses of gene-edited FA-iPSCs confirmed the specific integration of Fanca in the AAVS1 locus in all tested clones. Moreover, the hematopoietic differentiation of these iPSCs efficiently generated disease-free hematopoietic progenitors. Taken together, our results demonstrate for the first time the feasibility of correcting the phenotype of a DNA repair deficiency syndrome using gene-targeting and cell reprogramming strategies.

Keywords: cell reprogramming; Fanconi anemia; gene-targeting; iPSCs; zinc finger nucleases

Introduction

The progressive development of engineered nucleases has markedly improved the efficacy and specificity of targeted gene therapy, opening new possibilities for the treatment of inherited and acquired diseases in the clinics (Tebas et al., 2014). In contrast to conventional gene therapy with integrative vectors, targeted gene therapy enables the insertion of foreign sequences (i.e., therapeutic genes or small oligonucleotides) in specific sites of the cell genome. Thus, depending on the genetic etiology of the disease, the gene-targeting approach may pursue the correction of a specific mutation or, alternatively, the insertion of the therapeutic transgene into safe loci of the genome, often referred to as ‘safe harbors’ (Naldini, 2011).

In spite of the advances in the field, the question of whether or not targeted gene therapy will be applicable to diseases where homology-directed repair (HDR) is affected has never been explored. Taking into account that Fanconi anemia (FA) proteins participate in HDR (Taniguchi et al., 2002; Yamamoto et al., 2003; Niedzwiedz et al., 2004; Yang et al., 2005; Nakanishi et al., 2011) and coordinate the action of multiple DNA repair processes, including the action of different nucleases and homologous recombination (see reviews in Kee & D’Andrea, 2010; Kottemann & Smogorzewska, 2013; Moldovan & D’Andrea, 2009), we aimed to investigate for the first time the possibility of conducting a targeted gene therapy strategy in FA cells.

Genetically, FA is a complex disease where mutations in sixteen different genes (FANCA, -B, -C, -D1/BRCA2, -D2, -E, -F, -G, -I, -J/BRIP1, -L, -M, -N/PALB2, -O/RAD51C, -P/SLX4; -Q/ERCC4/XPF) have been reported (Bogliolo et al., 2013). Among all these genes, mutations in FANCA account for about 60% of total FA patients (Casado et al., 2007; Auerbach, 2009). Importantly, while few recurrent mutations (i.e., truncation of exon 4 in Spanish gypsies or mutations...
in exons 13, 36, and 38) have been observed in FA-A patients. FANCA mutations are generally private mutations, which include point mutations, microinsertions, microdeletions, splicing mutations and large intragenic deletions (Castella et al., 2011). Thus, considering the large number of genes and mutations that can account for the FA disease, the insertion of a functional FA gene in a ‘safe harbor’ locus would lead to the generation of a targeted gene addition platform with a broad application in FA, regardless of the complementation group and mutation type of each patient.

Recent studies by our group and others aiming at the identification of ‘safe harbor sites’ in the human genome have shown robust and stable expression of transgenes integrated in the human PPP1R12C gene, a locus also known as AAWS1, across different cell types (Smith et al., 2008; Lombardo et al., 2011). Additionally, no detectable transcriptional perturbations of the PPP1R12C and its flanking genes were observed after integration of transgenes in this locus, indicating that AAWS1 may represent a safe landing path for therapeutic transgene insertion in the human genome (Lombardo et al., 2011). These observations, together with the development of artificial zinc finger nucleases (ZFNs) that efficiently and selectively target the AAWS1 locus, have facilitated gene editing strategies aiming at inserting therapeutic transgenes in this locus, not only in immortalized cell lines but also in several primary human cell types, including induced pluripotent stem cells (hiPSCs; Hockemeyer et al., 2009; DeKelver et al., 2010; Lombardo et al., 2011; Zou et al., 2011b; Chang & Boughassira, 2012).

Because a defective FA pathway not only predisposes FA patients to cancer (Rosenberg et al., 2008) but also to the early development of bone marrow failure due to the progressive extinction of the HSCs (Larghero et al., 2002; Jacome et al., 2006), our final aim in these studies was the generation of gene-edited, disease-free FA-HSCs, obtained from non-hematopoietic tissues of the patient. Thus, in our current studies, we firstly pursued the specific insertion of the therapeutic FANCA gene in the AAWS1 locus of FA-A patients’ fibroblasts. Thereafter, gene-edited FA cells were reprogrammed to generate self-renewing disease-free iPSCs and finally re-differentiated toward the hematopoietic lineage, as previously described with FA cells corrected by conventional LV-mediated gene therapy (Raya et al., 2009).

Our goal of conducting a combined approach of gene editing and cell reprogramming in FA cells was particularly challenging taking into account the relevance of the FA pathway both in HDR (Taniguchi et al., 2002; Yamamoto et al., 2003; Niedzwiedz et al., 2004; Yang et al., 2005; Moldovan & D’Andrea, 2009; Kee & D’Andrea, 2010; Naknishi et al., 2011; Kottemann & Smogorzewska, 2013) and cell reprogramming (Raya et al., 2009; Muller et al., 2012; Yung et al., 2013). In spite of these hurdles, the strong selective growth advantage characteristic of corrected FA cells allowed us to establish a new approach for the efficient generation of FA HPCs harboring specific integrations of the therapeutic FANCA gene in a safe harbor locus.

Results

Efficient gene-targeting-mediated complementation of fibroblasts from FA-A patients

To promote insertion of a FANCA expression cassette into the AAWS1 locus, an integrate-defective lentiviral vector (IDLV) harboring the EGFP and FANCA transgenes flanked by AAWS1 homology arms (donor IDLV) was generated (Fig 1A top). In this donor IDLV, FANCA is under the transcriptional control of the human PGK promoter. In addition, a promoterless EGFP cDNA preceded by a splice acceptor (SA) site and a translational self-cleaving 2A sequence was also included upstream of the FANCA cassette. Upon targeted-mediated insertion into AAWS1, the EGFP cassette will be placed under the transcriptional control of the promoter of the ubiquitously expressed PPP1R12C gene, thus allowing the FACSorting of gene-targeted cells (Fig 1A). Besides the donor IDLV, an adenoviral vector expressing a ZFN pair (AdV5/35-ZFN), designed to induce a DNA double-strand break in the AAWS1 locus, was used to enhance the efficiency of gene targeting in this locus (Hockemeyer et al., 2009).

To investigate the feasibility of performing gene targeting in FA-A cells, skin fibroblasts from four FA-A patients with different mutations in AAWS1 were transduced either with the donor IDLV alone, or with the donor IDLV and the AdV5/35-ZFNs simultaneously. Fourteen days after transduction, cells were analyzed by flow cytometry to measure the proportion of EGFP+ fibroblasts. While <0.05% of the cells transduced with the donor IDLV alone were positive for EGFP, 0.2–1.1% of FA fibroblasts that had been co-transduced with the donor IDLV and the ZFNs-AdV were EGFP+ (see Fig 1B and representative analyses in Supplementary Fig S1). Strikingly, the percentage of EGFP+ cells markedly increased during the in vitro culture of these cells, reaching levels between 5.5 and 13.4% (Fig 1B), showing the proliferation advantage of gene-edited FA-A fibroblasts.

Because the prolonged in vitro culture of FA fibroblasts results in increased rates of cell senescence (Muller et al., 2012), in a new set of experiments, fibroblasts from three FA patients (FA-S2, FA-123 and FA-644) were transduced with an excisable hTERT-expressing LV (Salmon et al., 2000) prior to performing the gene-targeting procedure. Transduction of FA fibroblasts with hTERT-LVs resulted in a marked increase in telomerase activity (see representative data in Supplementary Fig S2). Significantly, the proportion of EGFP+ cells was markedly increased (3–4-fold) in hTERT-transduced versus untransduced FA fibroblasts from FA patients (Fig 1C), indicating that hTERT improved the efficacy of gene targeting in FA-A fibroblasts. Consistent with data obtained with non-immortalized fibroblasts, when immortalized gene-edited FA fibroblasts were maintained in culture, a progressive increase in the proportion of EGFP+ cells was also observed (see data from geFA-52T in Fig 1D). Strikingly, around 40% of treated FA-A fibroblasts were EGFP+ after 42 days in culture in the absence of any selectable drug (Fig 1D).

PCR analyses with two pairs of primers that amplify, respectively, the 5’ and the 3’ integration junctions between the EGFP/FANCA cassette and the endogenous AAWS1 locus evidenced the insertion of the EGFP/FANCA cassette into the AAWS1 locus of sorted EGFP+ geFA-52T fibroblasts (Fig 1E). In these gene-edited FA fibroblasts, the activity of hTERT was also confirmed (Supplementary Fig S2).

To investigate whether the insertion of the therapeutic hFANCA cassette in the AAWS1 locus of FA-A fibroblasts corrected the cellular phenotype of the disease, the functionality of the FA pathway in FA-52T fibroblasts was tested both before (negative control) and after the gene-targeting procedure. As a positive control, healthy
Figure 1. Efficacy of gene targeting of FANCA in the AAVS1 locus of primary hFA-A fibroblasts.

A Top: schematic representation of the donor integrase-defective lentiviral vector (IDLV) used to promote insertion of the EGFP/FANCA cassette into the AAVS1 locus. Middle: AAVS1 locus with the zinc finger nucleases (ZFNs) target site. Bottom: AAVS1 locus upon ZFN-mediated targeted insertion of the EGFP/PGK-FANCA cassette. Black arrow shows transcription of the EGFP from the endogenous PPR12C promoter; HA, homology arm; SD, splice donor; SA, splice acceptor; BGHpA, bovine growth hormone polyadenylation signal; SV40pA, simian virus 40 polyadenylation signal. Constituents of the LTR (US-R-AU3) are also indicated.

B Proliferation advantage of targeted Fanconi anemia (FA) fibroblasts (EGFP+) previously transduced with hTERT (FA-52T fibroblasts).

C Comparative analysis of gene targeting in FA-A fibroblasts, untransduced or transduced with a lentiviral vector expressing hTERT. Analyses were performed 14 days after gene targeting.

D In vitro proliferation advantage of targeted FA fibroblasts (EGFP+) previously transduced with hTERT (FA-52T fibroblasts).

E Targeted integration analysis of the EGFP/PGK-FANCA cassette into the AAVS1 site by PCR using primers specific for the 5’ or 3’ integration junctions (red arrows in the top schematic) defined as 5’T I or 3’T I, respectively.
donor fibroblasts (H.D. Fib) were analyzed in parallel. The presence of nuclear FANCD2 foci, fully dependent on the expression of all the FA core complex proteins, including FANCA (Garcia-Higuera et al., 2001), was determined in these samples after DNA damage induced by mitomycin C (MMC). In contrast to uncorrected FA-52T fibroblasts (FA-52T Fib.), which did not generate FANCD2 foci even after MMC exposure, a significant proportion of the geFA-52T fibroblasts generated FANCD2 foci, mainly after treatment with MMC, thus mimicking the response of H.D. fibroblasts (Fig 2A). Because the main characteristic of FA cells is the increased chromosomal instability upon exposure to DNA inter-strand cross-linking (ICL) drugs, we also investigated the response of both uncorrected and gene-edited FA-A fibroblasts to diepoxybutane (DEB). While in FA-52T fibroblasts DEB induced a significant increase in the number of chromosomal aberrations per cell (from 0.05 ± 0.05 to 1.7 ± 0.46 aberrations/cell)—including chromatid breaks and

---

**Figure 2. Phenotypic correction of the gene-edited FA-A fibroblasts.**

A. Top: histogram showing the percentage of FA-A fibroblasts, untransduced or co-transduced with the donor integrase-defective lentiviral vector (IDLV) and the AdV5/35-ZFNs (geFA-52T Fib.), showing FANCD2 foci in the absence or the presence of mitomycin C (MMC). Bottom: representative images of FANCD2 foci (red) in cells shown in the top histogram, after MMC treatment.

B. Chromosomal instability induced by diepoxybutane (DEB) in untreated (FA-52T) and gene-edited FA fibroblasts (geFA-52T Fib). Left: representative FISH analysis was performed by staining telomeres (in green), centromeres (in pink) and chromosomes (in blue). Right: histogram showing the number of chromosomal aberrations per cell.

Data information: Values are shown as mean ± s.e. from three independent experiments (A) or analysis of twenty different metaphases per group (B). All P-values were calculated using two-tailed unpaired Student’s t-test.
radial chromosomes, typically found in FA patients’ cells—the same DEB treatment did not induce any increase in the number of chromosomal aberrations in gene-targeted FA fibroblasts (Fig 2B).

Taken together, these results show the feasibility of correcting the phenotype of FA cells using gene targeting strategies, in particular by promoting the insertion and expression of FANCA in the AAAS1 safe harbor locus of fibroblasts from FA-A patients.

Efficient generation of disease-free iPSCs from FA fibroblasts corrected by gene targeting

To generate disease-free FA-iPSCs, FA fibroblasts subjected to gene editing (geFA-123, geFA-S2 and geFA-S2T) were first enriched for EGFP+ cells by cell sorting and then reprogrammed using a poly-cistronic excisable LV expressing the human SOX2, OCT4, KLF4, and cMYC transgenes from the EF1A promoter (STEMCCA vector; Somers et al, 2010). Consistent with previous observations (Raya et al, 2009), uncorrected FA fibroblasts did not generate iPSCs after reprogramming, even after transduction with the TERT-LV (data not shown). Although several iPSC-like colonies were generated from gene-edited FA-123 fibroblasts (115 AP+ cells/100,000 fibroblasts), no stable iPSC lines could be generated from FA fibroblasts simply subjected to gene editing, most probably because of the pro-senescent nature of these cells. In marked contrast to these observations, the reprogramming of FA fibroblasts that were first transduced with the hTERT-LV and then subjected to gene editing generated 230 iPSC-like clones, most of which could be maintained after serial in vitro passages (Supplementary Fig S3). Twelve iPSC clones generated from geFA-S2T fibroblasts were further expanded and differentiated into fibroblasts to perform additional studies to confirm the integration site of the EGFP/FANCA construct. First, qPCR analyses were conducted to determine the mean copy number per cell of the EGFP/FANCA cassette. As shown in Supplementary Table S1, all iPSC clones that were positive for integration of the cassette were also positive for the PCR band corresponding to the specific insertion in the AAAS1 locus.

Three geFA-iPSC clones (clones 16, 26 and 31) were selected for further characterization. The pluripotency of these gene-corrected clones was first analyzed by both alkaline phosphatase (AP) staining and immunohistochemistry staining of different pluripotency genes. Representative pictures in Fig 3A and Supplementary Fig S4A showed that all tested geFA-iPSC clones were highly positive for AP, NANOG, TRA-1-60, OCT4, and SSEA-4 expression. RT-qPCR analyses of the expression of endogenous pluripotency genes NANOG, OCT4, SOX2, KLF4, and cMYC were consistent with the pluripotent nature of these clones (Supplementary Fig S4B). In all cases, a very low expression of the ectopic reprogramming transgenes was found, indicating substantial inactivation of the EF1A promoter present in the reprogramming vector. As expected for bona fide iPSC clones, OCT4 and NANOG promoters were hypomethylated in gene-corrected FA-iPSC clones, in clear contrast to the high level of methylation observed in H.D. fibroblasts (Supplementary Fig S4C). To further demonstrate the pluripotency of geFA-iPSCs, in vivo, cells were subcutaneously inoculated in NSG mice. Characteristic teratomas containing complex structures representing the three embryonic germ layers were observed 8–10 weeks after implantation. Immunofluorescence staining confirmed the expression of definitive endoderm markers (Fox2A), neural structures that expressed neuroectodermal markers (β-III-tubulin) and the generation of mesoderm (Brachyury) and mesoderm derivatives tissue such as muscle (α-SMA; Fig 3B).

To confirm the insertion of the FANCA cassette into the AAAS1 locus in the gene-corrected FA-iPSC clones, Southern blot analyses were performed on genomic DNA extracted from gene-edited FA-iPSC clones 16, 26, and 31. Blots hybridized with probes for the exogenous EGFP and the endogenous AAAS1 genes confirmed the monoallelic integration of the EGFP/FANCA cassette into the AAAS1 locus and the absence of random integration in any of the three tested clones (Fig 3C,D).

Once demonstrated the generation of bona fide gene-edited FA-iPSCs, in the next set of experiments, we aimed to verify whether these geFA-iPSCs were disease free, as shown for their parental gene-edited FA fibroblasts (Fig 2). First, we verified by qRT-PCR that hFANCA mRNA levels corresponding to the three tested geFA-iPSC clones were similar to levels observed in the control ES cell line and markedly higher when compared to uncorrected FA-S2T fibroblasts (Fig 4A). Western blot analysis confirmed the expression of FANCA in all the three tested clones (Fig 4B). Even more, since FANCA is necessary for the relocation of FANCD2 to damaged DNA sites, we investigated the presence of nuclear FANCD2 foci in three geFA-iPSC clones exposed to MMC. As shown in Fig 4C, these analyses further confirmed the expression and functionality of FANCA in the three tested geFA-iPSC clones. Consistent with the restored FA pathway of gene-edited FA-iPSCs, DEB did not induce a significant increase in the number of chromosomal aberrations in FA-corrected cells. Remarkably, the number of chromosomal aberrations in geFA-iPSCs (0.2 ± 0.1 aberrations/cell; Fig 4D) was ten times lower to the number observed in their parental uncorrected fibroblasts (see Fig 2B).

To assure the identity of the different geFA-iPSC clones, the presence of the original pathogenic mutations described in patient FA-52 (c.710-5T>C and c.3558insG) was investigated by Sanger sequencing both on FA-S2T fibroblasts and geFA-iPSC clones 16, 26, and 31 (Supplementary Fig S5). The confirmation of both pathogenic mutations in the three tested geFA-iPSCs, together with our observations showing that all stable iPSC clones contained the AAAS1-targeted FANCA gene (Supplementary Table S1) and had a functional FA pathway, demonstrates that the disease-free nature of gene-edited FA-iPSCs is a consequence of the functional insertion of FANCA within the AAAS1 safe harbor site of these reprogrammed FA cells.

Aiming to excise the STEMCCA vector from the genome of geFA-iPSCs, cells from clone 16 were transduced with an IDLV co-expressing the Cre recombinase and the Cherry fluorescence marker (Papapetrou et al, 2011). Thereafter, individual colonies were isolated to select those clones with a lower number of copies of the STEMCCA provirus. Two clones were selected: Excised clones
In clone 16.2, the excision of the hTERT provirus was also confirmed (<0.05 copies as deduced from q-PCR analyses). RT-qPCR analysis performed in these two subclones showed the persistent expression of endogenous pluripotency genes (SOX2, OCT4, KLF4, NANOG, and cMYC) and the absence of ectopic transgenes expression (Supplementary Fig S6A). As expected from bona fide pluripotent iPSC clones, these two clones generated teratomas with structures characteristic of the three germ layers (Supplementary Fig S6B).

Figure 4. Disease-free Fanconi anemia phenotype of corrected geFA-iPSCs.
A Histogram showing the levels of hFANCA expression in gene-edited FA-iPSC clones and human ES (H9FA-52T) relative to untreated FA-52T fibroblasts. Data are shown as mean ± s.e. of three different analyses.
B Western blot analysis showing FANCA expression in geFA-iPSC clones in comparison with fibroblasts from HD and a FA-A patient.
C Representative immunofluorescence analysis of FANCD2 foci in geFA-iPSCs after DNA damage with mitomycin C (MMC). D Chromosomal instability induced by diepoxybutane (DEB) was also tested in geFA-iPSC16. FISH analysis was performed using probes to detect telomeres (green), centromeres (pink) and chromosomes (blue). Right: histogram showing the number of chromosomal aberrations per cell. Data information: Data are shown as mean ± s.e. from three different experiments (A) or analysis of twenty different metaphases per group (D). All P-values were calculated using two-tailed unpaired Student’s t-test.

Figure 3. Pluripotency characterization and insertion site analyses of gene-edited FA-A iPSCs.
A Expression of TRA1-60, SSEA-4, OCT4, and NANOG pluripotency markers by immunofluorescence staining of gene-edited FA-iPSCs (geFA-iPSCs; clone 16).
B Immunofluorescence analysis of ectoderm (β-III-tubulin), endoderm (Fox2A), and mesoderm (α-SMA and Brachyury) in teratomas generated from geFA-iPSCs (clone 16).
C Southern blot analysis of genomic DNA extracted from the indicated gene-corrected FA iPSC clones (geFA-iPSCs) and from parental fibroblasts, either unmanipulated (FA) or after gene editing (ge-FA iPSCs, clones 16, 26, and 31). Genomic DNA was digested with BglI and hybridized with a probe for PPP1R12C. The band of 9.6 kb corresponds to the targeted integration in PPP1R12C, while the 3.3 kb correspond to the untargeted allele.
D Southern blot analysis of samples shown in (C) digested with BstXI and hybridized with a probe (P) for EGFP. One single band of 5.1 kb is expected for specific integrations in PPP1R12C.
16.1 and 16.2, with a number of 0.35 ± 0.10 and <0.05 copies/cell, respectively. In clone 16.2, the excision of the \( \text{hTERT} \) provirus was also confirmed (<0.05 copies as deduced from q-PCR analyses). RT-qPCR analysis performed in these two subclones showed the persistent expression of endogenous pluripotency genes (\( \text{SOX2}, \text{OCT4}, \text{KLF4}, \text{NANOG}, \) and \( \text{cMYC} \)) and the absence of ectopic transgenes expression (Supplementary Fig S6A). As expected from bona fide pluripotent iPSC clones, these two clones generated teratomas with structures characteristics of the three germ layers (Supplementary Fig S6B).

Figure 4. Disease-free Fanconi anemia phenotype of corrected geFA-iPSCs.

A Histogram showing the levels of \( \text{hFANCA} \) expression in gene-edited FA-iPSC clones and human ES (H9) relative to untreated FA-52T fibroblasts. Data are shown as mean ± s.e. of three different analyses.

B Western blot analysis showing \( \text{FANCA} \) expression in geFA-iPSC clones in comparison with fibroblasts from HD and a FA-A patient.

C Representative immunofluorescence analysis of \( \text{FANCD2} \) foci in geFA-iPSCs after DNA damage with mitomycin C (MMC).

D Chromosomal instability induced by diepoxybutane (DEB) was also tested in geFA-iPSC 16. FISH analysis was performed using probes to detect telomeres (green), centromeres (pink) and chromosomes (blue). Right: histogram showing the number of chromosomal aberrations per cell.
Analysis of the genetic stability of gene-edited FA fibroblasts and iPSCs

Because of the chromosomal instability of FA cells, we investigated by means of karyotype analyses and aCGH analyses whether the different manipulations of FA-52 fibroblasts and their corresponding iPSCs induced chromosomal instability. As shown in Table 1, no evident karyotype or aCGH abnormalities were observed in expanded FA-52 parental fibroblasts when compared with a reference human DNA sample. Even more, the transduction with hTERT-LV and the gene-editing process did not induce evident chromosomal abnormalities in these cells. Reprogrammed geFA-52 iPSCs also had a normal karyotype, although a deletion in the 16p12.2p12.1 locus was noted in the aCGH analysis. After excision with the Cre recombinase, in addition to the 16p deletion, a mosaic trisomy in chromosome 5 was observed (See Table 1 and Supplementary Fig S7).

Generation of disease-free hematopoietic progenitors from gene-edited FA-A iPSCs

In experiments corresponding to Fig 5 and Supplementary Figs S8 and S9, we investigated whether hematopoietic progenitor cells derived from gene-edited FA-iPSCs were disease-free. To conduct these experiments, embryoid bodies from geFA-iPSCs were incubated with hematopoietic cytokines as described in Materials and methods. As shown in representative analyses from Supplementary Fig S8A, the hematopoietic differentiation of geFA-iPSCs after 21 days of in vitro stimulation was demonstrated by the presence of hematopoietic precursors (CD43+/CD44+), committed hematopoietic progenitors (CD34+CD45+) and also mature hematopoietic cells (CD45+CD34-). When the hematopoietic differentiation of excised and non-excised iPSC clones was compared, the proportion of CD45+ and CD34+CD45+ was consistently increased in the case of the excised vs the non-excised clones (see data from two independent experiments in Fig 5A and Supplementary Fig S8). Consistent with the flow cytometry data, granulo-macrophage and erythroid colonies were generated by geFA-iPSC-differentiated cells in methylcellulose. As it was observed in the flow cytometry studies, higher numbers of hematopoietic progenitors were generated by excised versus non-excised geFA-iPSC (Fig 5B). In all instances, colonies derived from geFA-iPSC were almost as resistant to MMC as healthy cord blood progenitor cells, in contrast to the MMC hypersensitivity observed in BM progenitors from FA patients (Fig 5C).

Finally, to investigate whether gene-edited FA-iPSCs were also able to differentiate toward the hematopoietic lineage in vivo, one of the teratomas generated by the excised geFA-iPSCs (clone 16.2) was analyzed for the presence of human hematopoietic markers. As shown in Supplementary Fig S9, 3% of the cells present in this teratoma consisted on hCD45+mcCD45- cells. Within this population, 3.5% corresponded to hCD45+ cells, thus revealing the in vivo differentiation potential of this clone.

Table 1. aCGH analysis in FA-52 fibroblasts prior to and after gene editing and in gene-edited iPSCs-derived clones

<table>
<thead>
<tr>
<th>Cells</th>
<th>aCGH result</th>
<th>OMIM GENES</th>
<th>Karyotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>FA-52 fibroblasts*</td>
<td>–</td>
<td>–</td>
<td>46 XY</td>
</tr>
<tr>
<td>geFA-52T fibr.</td>
<td>–</td>
<td>–</td>
<td>46 XY</td>
</tr>
<tr>
<td>geFA-52T iPSC clones</td>
<td>–</td>
<td>–</td>
<td>46 XY</td>
</tr>
<tr>
<td>Clone 16′</td>
<td>Deletion</td>
<td>16p12.2p121</td>
<td>*</td>
</tr>
<tr>
<td>Clone 16 Ex′</td>
<td>Deletion</td>
<td>16p12.2p121</td>
<td>*</td>
</tr>
</tbody>
</table>

*Comparison analyses between expanded fibroblasts from patient FA-52 (FA-52 fibroblasts) and a reference male DNA sample.
†Comparison analyses between expanded, TERT-transduced, and gene-edited FA-52 fibroblasts (geFA-52T fibr) with respect to FA-52 fibroblasts.
‡Comparison analyses between geFA-52T iPSCs clone 16 and clone 16 Ex (after excision of the reprogramming cassette) and FA-52 fibroblasts.

Discussion

Thanks to the development of artificial nucleases capable of generating DNA double-strand breaks (DSBs) in pre-determined sequences of the genome (Porteus & Baltimore, 2003; Urnov et al, 2010; Cong et al, 2013; Joung & Sander, 2013), targeted gene therapy is entering into the clinics (Tebas et al, 2014). Whether these approaches will be amenable to the treatment of DNA repair deficiency syndromes such as FA is, however, uncertain. In this respect, it is currently known that FA proteins participate in maintaining the genomic stability of the cell and coordinate the actions of multiple repair processes, including HDR (Kottemann & Smogorzewska, 2013), making these cells particularly appropriate for investigating the feasibility of performing targeted gene therapy in syndromes associated with DNA repair defects and genome instability. Although the mechanisms explaining how the FA pathway promotes HDR are still unclear, most evidence suggests that the monoubiquitination of FANCD2—which is critically dependent on the presence of all the FA core complex proteins, including FANCA—is essential for the recruitment of several HDR factors (such as BRCA1, BRCA2, and RAD51) to damaged chromatin (see review in Kee & D’Andrea, 2010).

To investigate whether gene targeting was feasible in FA cells we focused on the most frequent FA complementation group, FA-A (Casado et al, 2007; Auerbach, 2009), and investigated the possibility of inserting the therapeutic transgene in a safe harbor locus of the human genome—the AAVS1 locus (Lombardo et al, 2011). Strikingly, our first results in Fig 1 clearly demonstrate the feasibility of performing gene targeting in FA-A cells with significant efficacies (up to 4%), comparable with efficacies reported in primary cells competent for DNA repair (DeKelver et al, 2010; Lombardo et al, 2011; Sebastiani et al, 2011; Soldner et al, 2011; Zou et al, 2011a). The feasibility of performing gene targeting in FA-A cells could be explained by different hypotheses. First, as previously described in other systems (Matrai et al, 2011; Peluzzo et al, 2013), a transient though early expression of FANCA may be induced by the donor IDLV, thus facilitating the insertion of the exogenous therapeutic cassette through a HDR process. Besides
this hypothesis, we should contemplate the possibility that the limited HDR activity of FA-A cells (Nakanishi et al., 2005, 2011) could be sufficient to facilitate the ZFN-mediated integration of our donor IDLV in the AAVS1 site. Finally, although the integration of the therapeutic cassette in the AAVS1 locus might have occurred through an HDR-independent process, as reported in other models (Anguela et al., 2014), PCR and Southern blot analyses showed the expected amplicons and band length for targeted
integration of the cassette, strongly suggesting that AAVS1 targeting took place through a HDR mechanism. In this respect, while the specificity of gene targeting might be reduced in FA cells, our data clearly show that all the FA-iPSC clones harbored one single copy of FANCA specifically integrated in the PPRR12C target gene (Table 1). Consequently, this result further supports the efficacy and the specificity of our gene targeting approach.

With the main objective of preventing the predisposition to senescence of FA cells (Muller et al., 2012), the transduction of hTERT-LV in FA-A fibroblasts induced an unexpected effect in these cells, which consisted of a significant increase in the efficacy of gene editing (Fig 1). Whether or not this effect is specific for FA cells or whether it is simply mediated by the enhanced proliferation rate of TERT-transduced FA cells is currently unknown. Nevertheless, to the best of our knowledge, the improved gene targeting mediated by hTERT observed in our experiments constitutes a new finding that has not been previously reported in any other experimental model. The observation that transduction with hTERT also facilitates the generation of gene-edited FA-iPSCs is consistent with previous data showing the relevance of hTERT in cell reprogramming (Batista et al., 2011; Pomp et al., 2011; Winkler et al., 2013). In safety terms, even though the hTERT provirus could be efficiently excised from transduced cells with the Cre recombinase, further approaches based on the transient expression of hTERT during gene editing and/or cell reprogramming would constitute safer approaches to limit potential genomic insults during the ex vivo manipulation of the samples.

Interestingly, EGFP analyses in gene-edited FA fibroblasts showed that in the absence of any artificial selection process, a progressive increase in the proportion of targeted cells (up to 40% after 42 days in culture) was observed, mimicking the improved growth proliferation properties of FA precursor cells in mosaic patients (Waisfisz et al., 1999; Gregory et al., 2001; Gross et al., 2002) or in experimental models of FA gene therapy (Rio et al., 2008). Consistent with previous observations in FA cells corrected by LV-mediated gene therapy (Raya et al., 2009), this proliferation competence of FA-corrected cells was particularly remarkable when samples were subjected to cell reprogramming, confirming the relevance of the FA pathway during the process of iPSC generation. Similar conclusions were obtained in two additional studies (Muller et al., 2012; Yong et al., 2013), although these studies showed that reprogramming of FA cells can occur, albeit with a very low efficiency compared to gene-complemented FA cells.

Studies in Figs 2 and 4 showing the generation of nuclear FANCD2 foci and the chromosomal stability of gene-edited FA fibroblasts and iPSCs upon exposure to ICL drugs demonstrate that the specific targeting of FANCA in the AAVS1 locus has completely corrected the phenotype of FA-A fibroblasts and bona fide iPSCs. Although transduction of FA fibroblasts with the hTERT-LV might have had consequences upon the genetic instability of FA cells, our karyotype and aCGH studies indicate that neither the expansion nor the transduction with hTERT-LV or the gene-editing processes induced evident chromosomal abnormalities in FA fibroblasts. In contrast to these results, data in Table 1 and Supplementary Fig S7 showed the presence of chromosomal abnormalities in reprogrammed and excised geFA-iPSCs. Importantly, different genetic defects have also been reported in non-FA-iPSCs (Mayshar et al., 2010; Gore et al., 2011; Laurent et al., 2011; Cheng et al., 2012; Ruiz et al., 2013) that were associated with the generation of the iPSCs (Mayshar et al., 2010; Gore et al., 2011; Hussein et al., 2011; Laurent et al., 2011) and/or with mutations that pre-existed in the somatic population of origin (Young et al., 2012). This indicates that the presence of chromosomal abnormalities in our iPSCs is not exclusive of their FA genetic background and that the different mechanisms accounting for mutations in non-FA-iPSCs would be applicable to our geFA-iPSCs.

Consistent with the previous study showing the generation of disease-free FA-iPSCs through conventional gene therapy approaches (Raya et al., 2009; Muller et al., 2012), our new study shows the efficient hematopoietic differentiation of gene-edited FA-iPSCs. Moreover in the current study, we observed the generation of increased numbers of hematopoietic progenitors from geFA-iPSCs subjected to excision of the reprogramming cassette, confirming previous observations showing that the residual expression of reprogramming genes limits the iPSC differentiation potential (Ramos-Mejia et al., 2012). The hematopoietic differentiation observed in these experiments and the robust expression of FANCA targeted into the safe harbor AAVS1 locus should account for the generation of a high number of hematopoietic progenitors with normalized response to MMC.

In summary, our study demonstrates for the first time the possibility of conducting efficient and precise targeted-mediated gene therapy in HDR-deficient cells. Moreover, we show the feasibility of reprogramming these cells to generate iPSC-derived gene-edited hematopoietic progenitors characterized by a disease-free phenotype. Our approach thus constitutes a new proof-of-concept with a potential future clinical impact to optimize the generation of gene-corrected HSCs from non-hematopoietic tissues of patients with inherited diseases, including DNA repair deficiency and genetic instability syndromes, like FA.

Materials and Methods

Cell lines and primary fibroblasts from FA-A patients

293T and HT1080 cells (ATCC: CRL-11268 and ATCC: CCL-121) were used for the production and titration of the LVs, respectively. Cells were grown in Dulbecco’s modified medium GlutaMAX™ (DMEM; Gibco) supplemented with 10% fetal bovine serum (FBS, BioWhitaker) and 0.5% penicillin/streptomycin solution (Gibco). Skin fibroblasts were obtained from FA-A, FA-5, FA-664, and FA-52 patients and were maintained in DMEM (Invitrogen) supplemented with 20% FBS (BioWhitaker) and 1% penicillin/streptomycin solution (Gibco) at 37°C under hypoxic conditions (5% of O2) and 5% of CO2. Patients were classified as FA-A patients as previously described (Casado et al., 2007). The ES4 and H9 (NIH Human Embryonic Stem Cell Registry, http://stemcells.nih.gov/research/registry/) lines of hES cells were maintained as originally described (Raya et al., 2008). FA patients and healthy donors were encoded to protect their confidentiality, and informed consents were obtained in all cases according to Institutional regulations of the CIEMAT. All studies conformed the principles set out in the World Medical Association Declaration of Helsinki.
Vectors

pCCL.sin.CPT.AAVS1.loxp.PA.2A.GFP.pA.loxp.PGK.FANCA.pA.Wpre donor transfer LV (donor IDLV) was generated using elements from the backbones pCCL.PGK.FANCA.Wpre* (Gonzalez-Murillo et al., 2010) and pCCLsin.CPT.AAVS1.2A.GFP.pA (Lombardo et al., 2011). The integrase-defective third-generation packaging plasmid pMD.Lg/pRRE.D64Vint was used to produce IDLV particles (Lombardo et al., 2007). pLM.CMV.Cherry.2A.Cre (Papapetrou et al., 2011) and pLox.TERT.ires.TK vectors (Salmon et al., 2000) were provided by Addgene. For reprogramming experiments, the EF1α STEMCCA lentiviral vector kindly provided by Dr Mostoslavsky was used (Sommer et al., 2010). This vector contains the cDNAs for OCT4, SOX2, c-MYC, and KLFL flanked by loxP sequences for their subsequent excision. ZFNs targeting intron 1 of the PPPIR12C gene were expressed from an Adenoviral Vector (AdV5/35) under the control of the CMV promoter (Lombardo et al., 2011).

Cell transduction

For gene editing experiments, fibroblasts from FA-A patients were transduced either with donor IDLV alone (150 ng HIV Gag p24/mL) or together with AdV5/35-ZFNs (multiplicity of infection (MOI) 200). Fourteen days post-transduction, the proportion of EGFP+ cells was determined by flow cytometry (BD LSRFortessa cell analyzer, Becton Dickenson Pharmingen). To immortalize fibroblasts from FA-52 to FA-123 patients, 105 cells were transduced at MOI 1 with the pLox.TERT.ires.TK LV (Salmon et al., 2000) for 24 h. To excise the reprogramming cassette and hTERT from established hiPSCs, single cell suspensions were generated by incubation with acutase (Gibco) and transduced for 10 h with the IDLV pLM.CMV.Cherry.2A.Cre. Immediately after transduction, 2 × 105 cells/10 cm2 dish, expressing Cherry protein, were sorted and new subclones of the parental geFA-iPSCs were generated.

Hematopoietic differentiation

iPSC colonies were detached using collagenase type IV (Gibco) for 30 min at 37°C, washed and centrifuged at 200× g, resuspended in differentiation media composed by KO-DMEM (Gibco) supplemented with 20% non-heat-inactivated FBS (Biowhitaker), 1% NEAA (Lonza; Biowhitaker), L-Glu (1 mM; Invitrogen), β-mercaptoethanol (0.1 mM; Gibco) and hrBMP4 (0.5 ng/mL; Prepotech) and plated in ultra-low attachment plates (Costar). After 2 days, media were replaced by Stempro 34 (Invitrogen) supplemented with 0.5% pen/streptomycin, L-Glu (2 mM; Invitrogen), MTG (40 mM; Sigma), ascorbic acid (50 μg/mL; Invitrogen), hrSCF, hrFlt3 ligand and TPO (100 ng/mL; Bio-Rad), hrIL3 (10 ng/mL; Biosource), hrIL6 (10 ng/mL; Prepotech), hrBMP4 (50 ng/mL; Prepotech), Wnt11 (200 ng/mL; R&D), and rhVEGF (5 ng/mL; Prepotech). Media were changed every 3–4 days. At day 7, media were replaced by fresh media where rhWnt-11 was substituted by rhWnt-3a (200 ng/mL; R&D). Media were changed every 3–4 days. At day 14 and 21, immunophenotypic analysis of the differentiated cells was performed by flow cytometry, and colony-forming unit assays were conducted (See Supplementary Methods).

Flow cytometry

Transduction with the AdV5/35-ZFNs and the donor IDLV, was analyzed by flow cytometry analysis (FACSCalibur; Becton Dickinson Pharmingen). Immunophenotypic analysis of the hematopoietic differentiated cells was performed using the following antibodies according to the manufacturer's instructions: phycoerythrin (PE)-Cy7-conjugated anti-human CD34 (BD Pharmingen), PE-conjugated anti-human CD31 (eBiosciences), allophycocyanin (APC)-conjugated anti-human CD45 (BD), and fluorescein isothiocyanate (FITC)-conjugated anti-human CD43 (BD). Fluorochrome-matched isotypes were used as controls. 4',6-Diamidino-2-phenylindole (DAPI; Roche)-positive cells were excluded from the analysis. Analysis was performed using FlowJo software.

Immunofluorescence and Western blot of Fanconi anemia proteins

Analyses of FANCD2 foci were performed by immunofluorescence of primary fibroblasts or iPSCs treated for 16 h with 200 nM of MMC. After MMC treatment, cells were stained with rabbit polyclonal anti-FANC D2 (Abcam, ab2187-50) as previously described (Hotta & Ellis, 2008; Raya et al., 2009). Cells with more than ten foci were scored as positive. FANCA expression was analyzed by Western blot (Raya et al., 2009) using the following antibodies: hFANCA (ab5063 Abcam) and anti-beta Actin to mouse antibody (ab6276, Abcam) as control. Goat polyclonal antibody to rabbit IgG (HRP, ab6721-1; Abcam) and sheep polyclonal antibody to mouse IgG—H&L (HRP, ab 6808, Abcam) were used as secondary antibodies. Protein quantification was done with Image J software.

FANCA expression by qRT-PCR

The expression of human FANCA mRNA was analyzed in the different clones of geFA-iPSCs by real-time quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR; Gonzalez-Murillo et al., 2010) using primers described in Supplementary Methods. Parental fibroblasts from FA-52 and ES H9 were used as controls.

Gene targeting analysis: PCR and Southern blots

For PCR analysis, genomic DNA was extracted with DNeasy Blood & Tissue Kit (Qiagen). To detect the targeted integration of the HDR cassette in the AAVS1 locus, two different pair of primers for the 3’ or the 5’ integration junction (5’TI and 3’TI, respectively) were used (Supplementary Table S2). PCR was conducted as follows: 2 min at 94°C, 40 cycles of 30 s at 94°C, 30 s at 58°C (5’TI) and 59°C (3’TI), 1 min at 72°C and one final step for 5 min at 72°C. The proper target integration amplified a 1195 bp amplicon for the 5’TI and a 1314 bp fragment for the 3’TI that were resolved in agarose gel at 2%. For Southern blot analyses, genomic DNA was extracted and digested either with BstXI enzyme or with BglI (both from New England Biolabs). Matched DNA amounts were separated on 0.8% agarose gel, transferred to a nylon membrane (Hybond XL, GE Healthcare) and probed either with the 32P-radio labeled sequence of a fragment of EGFP to detect specific (5.1 kb) and non-specific integrations or with a probe of AAVS1 gene located outside of the...
The paper explained

Problem
Gene targeting is becoming a true alternative to conventional gene therapy with integrative gammarretoviral or lentiviral vectors. It is however unknown whether these approaches would be applicable to inherited syndromes like FA, characterized by homology-directed DNA repair (HDR) defects. Additionally, the existence of 16 different FA genes, each of them with multiple mutations potentially accounting for the disease, would imply the necessity of developing individualized targeted gene therapy strategies in FA patients.

Results
We have demonstrated for the first time an efficient and specific targeting of FANCA in the AAV5 safe harbor locus of FA-A patients’ fibroblasts. This approach allowed us to develop a gene-editing platform applicable to all FA subtypes and FA gene mutations based on the insertion of the therapeutic FA gene in a safe harbor locus. Moreover, gene-edited FA-A fibroblasts were reprogrammed to generate disease-free iPSCs, which could be re-differentiated toward the hematopoietic lineage in a process that resulted in the generation of gene-edited, disease-free, hematopoietic progenitor cells.

Impact
Our data showing that gene targeting is feasible in FA opens the possibility of using similar strategies in different inherited syndromes characterized by defects in HDR and genome instability. The generation of disease-free HSCs through the specific insertion of therapeutic transgenes in a safe harbor locus of non-hematopoietic cell tissues, additionally constitutes an implemented approach to overcome HSC defects characteristic of many DNA repair deficiency syndromes, like Fanconi anemia.

Supplementary information for this article is available online: http://embomolmed.embopress.org

Acknowledgements
The authors would like to thank Prof. Juan C. Izpisua-Belmonte and Dr Guillermosto Guenechea for helpful discussions; Laura Cerrato for technical assistance with iPSCs; and Aurora de la Cal for coordination with the FA Network. We are also indebted to the FA patients, families, and clinicians from the FA network. This work was supported by grants to J.A.B. from the European Union (FP7 GA 222878 PERSIST), Spanish Ministry of Economy and Competitiveness (International Cooperation on Stem Cell Research Plan E; Ref PLE 2009/0100; SAF 2009-07164 and SAF 2012-39893), Fondo de Investigaciones Sanitarias, Instituto de Salud Carlos III (RETICS-RD 2012/013, SAF2010/BMD-2420), and La Fundación Privada La Marató de TV3, 121430/31/32, to J.S. from the Generalitat de Catalunya (SGRO489-2009), the ICAC-Academia Foundation, the Marató de TV3 (464/C/2012), the Spanish Ministry of Science and Innovation (SAF2012-31883), the European Commission (HEALTH-F5-2012-305421), and the European Regional Development FEDER Funds; to L.N. from Telethon (TICET grant D2), European Union (FP7 GA 222878 PERSIST, ERC Advanced Grant 249845 TARGETINGGENETHERAPEUTY) and the Italian Ministry of Health. The authors also thank the Fundación Marcelino Botín for promoting translational research at the Hematopoietic Innovative Therapies Division of the CIERBER. CIBERER is an initiative of the Instituto de Salud Carlos III, Spain.

Author contributions
Contribution: PR, RB, AL, LN, and JAB conceived and designed the experiments PR, RB, AL, OQ-B, LA, ZG, PG, EA, AV, BD, SN, YT, JPT, and RM conducted experiments. JCS, ES, JS, PDG, and MCH provided reagents, tools, and ideas. PR, RB, AL, LN, and JAB wrote the paper.

Conflict of interest
P.D.G. and M.C.H. are current or former employees of Sangamo BioSciences, Inc. The rest of the authors declare that they have no conflict of interest.

For more information
Fanconi Anemia Research Foundation: www.fanconi.org

References
Paula Rio et al. Targeted gene therapy in Fanconi anemia

EMBO Molecular Medicine


differentially expressed genes after transduction with improved lentiviral vectors. Blood 96: 3392 – 3398

License: This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
Intercellular network structure and regulatory motifs in the human hematopoietic system

Wenlian Qiao², Weijia Wang³, Elisa Laurenti²,³, Andrei L Turinsky⁴, Shoshana J Wodak⁴,⁵, Gary D Bader³,⁶,⁷, John E Dick²,³ & Peter W Zandstra¹,⁷,⁸,⁹,¹⁰,*

Abstract

The hematopoietic system is a distributed tissue that consists of functionally distinct cell types continuously produced through hematopoietic stem cell (HSC) differentiation. Combining genomic and phenotypic data with high-content experiments, we have built a directional cell–cell communication network between 12 cell types isolated from human umbilical cord blood. Network structure analysis revealed that ligand production is cell type dependent, whereas ligand binding is promiscuous. Consequently, additional control strategies such as cell frequency modulation and compartmentalization were needed to achieve specificity in HSC fate regulation. Incorporating the in vitro effects (quiescence, self-renewal, proliferation, or differentiation) of 27 HSC binding ligands into the topology of the cell–cell communication network allowed coding of cell type-dependent feedback regulation of HSC fate. Pathway enrichment analysis identified intracellular regulatory motifs enriched in these cell type- and ligand-coupled responses. This study uncovers cellular mechanisms of hematopoietic cell feedback in HSC fate regulation, provides insight into the design principles of the human hematopoietic system, and serves as a foundation for the analysis of intercellular regulation in multicellular systems.

Keywords feedback regulation; hematopoietic stem cell; intercellular signaling

Subject Categories Network Biology; Stem Cells

DOI 10.15252/msb.20145141 | Received 20 January 2014 | Revised 9 June 2014 | Accepted 17 June 2014

Introduction

The hematopoietic system is a distributed tissue consisting of multiple phenotypically and functionally distinct cell types. Hematopoietic stem cells (HSCs), at the apex of the hematopoietic developmental hierarchy, populate and sustain the system through highly coordinated self-renewal and differentiation processes. Increasing evidence suggests that HSC fate decisions are regulated in part via feedback mechanisms including HSC autocrine signaling and paracrine signaling from differentiated hematopoietic cells (Csaszar et al, 2012; Heazlewood et al, 2013). However, the key signaling molecules and cell types involved and how multiple often competing feedback signals act to regulate HSC fate in a coordinated manner are poorly understood.

We previously used mathematical modeling and bioinformatic strategies to systematically characterize the role of feedback signaling in regulating human umbilical cord blood (UCB) HSC fate in vitro (Kirouac et al, 2009, 2010). We identified lineage-dependent stimulatory and inhibitory signals that constitute a dynamic and complex feedback signaling network for hematopoietic stem and progenitor cell (HSPC) proliferation. This led to the development of an effective culture system capable of expanding human UCB HSC by globally diluting inhibitory feedback signals (Csaszar et al, 2012), pointing to the relevance of the network that our modeling approach uncovered. However, how the feedback signaling network is organized and how HSCs sense and interpret the signals produced by different cell types remains to be elucidated.

Network analysis is a powerful approach to detect the design principles of many types of distributed systems. This strategy has been used to interpret ecological (Olesen et al, 2007), social (Apicella et al, 2012), financial (Vitali et al, 2011), and molecular (Jeong et al, 2001) systems, but has never been applied to cell–cell communication (CCC) networks. We hypothesized that mapping the hierarchical hematopoietic signaling network would provide insight into its regulatory structure and function, in particular how feedback mechanisms control HSC fate decisions. From a network structure perspective, we were particularly interested in understanding how network structures including modular (network division into
sub-networks) and promiscuous (overlapping connectivity and sub-specialization of network components) strategies impact hematopoietic system behavior and HSC fate regulation.

Existing hematopoietic intercellular signaling networks have been constructed based on theoretical interactions between cells (Frankenstein et al., 2006) or curation of ligand–receptor interactions in heterogeneous cell populations (Kirouac et al., 2010). By taking advantage of high-resolution sorting of hematopoietic cells and transcriptome profiling, we created a CCC network to represent intercellular signaling between 12 highly resolved and phenotypically defined populations of stem, progenitor, and mature cell types from uncultured human UCB samples. We computationally analyzed the properties of the system and validated predictions using in vitro HSC fate responses to network-predicted HSC-targeting ligands. Our results support a model whereby differentiated hematopoietic cells influence HSC fates by regulating key intracellular regulatory nodes through cell type-dependent feedback signals. Control parameters such as relative cell frequency and local compartmentalization (niches) are opportunities to impose specificity in HSC fate regulation. Overall, our findings provide insight into the design principles of the human hematopoietic system focusing on the mechanisms of CCC in the feedback regulation of HSC fate. Further, our approach provides a fundamentally new strategy for analyzing intercellular regulation in multicellular systems.

Results

A hematopoietic cell–cell communication network is constructed from transcriptomic data

Our strategy for constructing and analyzing hematopoietic CCC networks is shown in Fig 1 that we will refer to throughout the manuscript. Transcriptomic data (Novershern et al., 2011; Laurenti et al., 2013) of 12 phenotypically defined, highly enriched hematopoietic cell types (Fig 2A) were the resource for network construction (Fig 1; step 1a). The data captured the intuitive biological properties of corresponding cell types as defined by gene ontology (Fig 2B; see also Supplementary Table S1 and Materials and Methods). For example, stem and progenitor cells (hereafter collectively referred to as the primitive cells), except for megakaryocyte-erythroid progenitors (MEP), over-expressed HSC proliferation and differentiation genes; MEP and erythroblasts (EryB) over-expressed erythrocYTE and megakaryocyte (Mega) differentiation genes; monocytes (Mono) over-expressed genes related to leukocyte and neutrophil (Neut) biological properties; and precursor B cells (PreB) over-expressed genes related to PreB differentiation.

To construct the CCC network, we compiled a database (Supplementary Table S2) of 341 receptors (or receptor genes) and their cognate ligands equivalent to 253 ligands (or ligand genes) (Materials and Methods). Hierarchical clustering of the receptor and ligand gene expression values recapitulated the developmental relationship (primitive cell compartment versus mature cell compartment) between the 12 cell types (Fig 2C), indicating similar expression of ligand and receptor genes in cells of the same developmental stage. Specifically, the primitive cells exhibited correlated receptor expression at higher confidence (average $P = 0.005$) and correlated ligand expression at lower confidence (average $P = 0.175$) than the mature cells in which average $P$-values for receptor expression and ligand expression were 0.0900 and 0.0570, respectively. Thus, we suspected changes in the ligand and receptor expression in blood cells during progression through differentiation.

In the construction of CCC networks, we assumed that the differentially over-expressed genes of each cell type are predictive of the cell type’s protein expression (Schwanhauser et al., 2011), and representative of the cell type’s biological properties. To determine an appropriate false discovery rate (FDR) to define differential over-expression, we tested FDRs of 1%, 5%, 10%, 20% and 25% and then compared the set of receptors identified at each threshold to a benchmark of known cell type-associated receptors (see Materials and Methods). A FDR of 10% detected the known cell type-associated receptors with the optimal combination of sensitivity and specificity (Supplementary Fig S1), and thus the ligands (Supplementary Table S3A) and receptors (Supplementary Table S3B) differentially over-expressed according to this threshold were used in the subsequent analyses (Fig 1; step 1b).

A CCC network is a directional bipartite graph (Fig 2D) composed of connections between differentially over-expressed ligand and receptor genes of the cell types of interest, based on 933 ligand–receptor interaction pairs (Supplementary Table S2) involving the 341 receptors and 253 ligands in Fig 2C (Materials and Methods for network construction). Sixteen class-I cytokines including CNTF, CSF2, CTF1, IL2, IL3, IL4, IL5, IL6, IL7, IL9, IL11, IL13, IL15, IL21, LIF, and OSM require interaction with heteromultimeric receptors to initiate intracellular signaling cascades (Robb, 2007). Given that our network was constructed from gene expression data, from a modeling perspective, we assumed that the greater the number of receptor species that a cell expresses for a ligand, the higher the probability that the ligand binds to the cell. We considered the interactions of each ligand and its cognate receptors independently; this practice did not affect our conclusions on network structures as shown below. Some differentially over-expressed ligands and receptors did not have interaction partners in the analyzed cell types. For example, KIT expressed on HSC-enriched cells (HSCe: human UCB Lin−CD34−CD38−CD45RA−CD90+/−) binds to SCF, a ligand produced by perivascular cells in the bone marrow niche (Ding et al., 2012), which our system did not have information about. Such ligands or receptors were connected to a hypothetical “Others” population representing an unknown number of additional cell types that potentially impact hematopoiesis. Based on these rules, a CCC network containing 1,344 ligand production-binding relationships between 249 ligand nodes and 13 cell nodes was constructed (Supplementary Table S4), of which 178 ligands mediated the connection between the 12 cell nodes of interest and 117 ligands targeting HSCe (Fig 1; step 1c). This CCC network paves a new way of depicting the hematopoietic hierarchy, and we next sought to analyze its properties.

As a starting point for our analysis, we separated the CCC network into two networks representing ligand production and ligand binding, respectively. The cell types were ranked in different orders based on the number of their interacting ligands in the two processes (Fig 2E). Distribution of the cell types based on the numbers of their produced ligands was approximated by a linear
1. Construction

A. Hematopoiesis hierarchy

B. Protein-protein interaction databases

C. Cell-cell communication network

2. Analysis

A. Analysis of ligand production network

B. Analysis of ligand binding network

3. Validation

A. In vitro screens for HSCe-targeting ligands

B. Cell-cell communication network for HSCe fate regulation

Figure 1. Computational and experimental workflow of the study.

The study is divided into network construction, analysis, and validation stages. Successive steps within each stage were alphabetically labeled. HSCe: human UCB HSC-enriched (Lin-CD34+CD38-CD45RA-CD49f-CD90+/+) cells.
Figure 2.
Interaction between blood cells and ligands in the ligand production process is modular

A cell-to-ligand interaction, $A_{ij}$, in the ligand production network was defined if cell $i$ produced ligand $j$. Simultaneously, clustering the cell types and the ligands suggested that groups of ligands were associated with subsets of cells in the network (Fig 3A). Silhouette widths (Rousseeuw, 1987) measuring the relatedness of the cell types’ ligand production supported the existence of 4 ligand–cell modules (Fig 3B, Supplementary Fig S2): the primitive cell module (HSCe + MLP + CMP + MEP + GMP), neutrophil–monocyte module (Neut + Mono), erythroid module (EryB), and a module of all the other cell types (Boso + Eos + Mega + PreB) (Fig 1; step 2a). A priori biological processes of 190 ligands (Supplementary Table S5) suggested that each blood cell module produced ligands with biased biological functions. For instance, ligands of the neutrophil–monocyte module enriched in exogenous signals that inhibit cell survival (HG $Z$-scores were 1.63 and 2.98 for Mono and Neut, respectively) and signals that mediate cell survival via NF-$\kappa$B (HG $Z$-scores were 2.15 and 1.43 for Mono and Neut, respectively); ligands of Baso, Eos, and PreB within the (Boso + Eos + Mega + PreB) module enriched in signals that direct differentiation cell fates of T helper cells (HG $Z$-scores were 1.17, 2.65, and 3.18 for Baso, Eos, and PreB, respectively); and ligands of EryB enriched in signals that regulate G1-S cell cycle transition (HG $Z$-score = 1.41) (Fig 3C). See Supplementary Table S6 for the other HG enrichment $Z$-scores.

In summary, our analysis suggested that blood cell ligand production is peculiar to blood cell identities, and a modular interaction structure exists in the ligand production network. This conclusion is robust to the choice of FDR threshold for differential gene over-expression (Supplementary Fig S2B) and the incorporation of hetero-multimeric receptor expression in network construction (Supplementary Fig S2C). Furthermore, ligand production of hematopoietic cell modules indicated characteristic biological properties. Considering HSC feedback regulation, this raised the possibility of HSC feedback control by cell module- or cell type-specific signaling.

Interaction between ligands and blood cells in the ligand binding process is promiscuous

We next sought to determine whether the ligand binding network had a similar structure to the ligand production network. A ligand-to-cell interaction, $B_{ij}$, in the ligand binding network was defined if cell $i$ expressed receptor(s) for ligand $j$. Interrogation of the network (Fig 4A) using spectral co-clustering (Dhillon, 2001) suggested a significantly less modular interaction structure than in the ligand production network (Fig 3A) ($t$-test $P < 0.001$), with ubiquitously shared ligand binding among the 12 cell types due to non-specific ligand–receptor interactions (Supplementary Fig S3A). The promiscuous network structure is robust to the choice of FDR threshold for differential gene over-expression (Supplementary Fig S3B) and the incorporation of hetero-multimeric receptor expression in network construction (Supplementary Fig S3C). Interestingly, HSCe which normally reside in the bone marrow niche with progenitor and maturing cells (Fig 4B) interacted with ligands of the greatest diversity. This raised the question of how HSCe fate can be specifically regulated in response to physiological demand. We hypothesized two different mechanisms: relative cell frequency that allows more abundant cell types skew the ligand species and resources available to HSCe, and cell compartmentalization that limits the access of HSCe to locally available ligands. We then explored, computationally, the effects of the two mechanisms on the quantity and identity of HSCe-targeting ligands (Fig 1; step 2b).

To explore the role of cell frequency in skewing HSCe–targeting ligands, we compared ligand binding in two scenarios by assuming that the probability of binding a ligand is a function of cell frequency given non-regulated receptor ligand affinities. In the first scenario, we modeled ligand binding in the system of mono-nucleated cells (MNC) isolated from fresh human UCB samples. Based on flow cytometry analysis, Neut was the most abundant cell type in the system (Fig 4C) according to the phenotypic definition we used; consequently, the cell type was the major ligand sink that significantly influenced ligand accessibility of the other cell types (Fig 4Cii). In contrast, HSCe, a quantitatively underrepresented cell type in the MNC system, had negligible ligand access despite the large number ligands targeting the cell type (Fig 4A). In the second scenario, we modeled ligand binding using cell frequencies from progenitor cell-enriched UCB samples (Fig 4D), in which cell composition is reminiscent of the progenitor enrichment seen during development or in the bone marrow niche (Nombela-Arrieta et al, 2013). Increased frequency of HSCe elevated their access to the available ligand resources (Fig 4Dii). This analysis indicates that controlling...
hematopoietic cell relative frequency can modulate ligand exposure to HSCe.

Then, we explored the role of cell compartmentalization. While an increasing number of hematopoietic cell types such as erythroblasts (Soni et al., 2008), megakaryocytes (Huang & Cantor, 2009), monocytes (Chow et al., 2011), and B cell progenitors (Nagasawa, 2007) are found in the stem cell niche within the bone marrow environment, the exact location and direct feedback role of these cell types on HSC fate decisions is not clear. We used OR gates to model the feedback effect of these cell types on HSCe as a function of their localization based on the extant knowledge of 190 ligands (Supplementary Table S5). The model consisted of four compartments to represent cells of different developmental stages: HSCe themselves, progenitor cells (PC = CMP + GMP + MEP + MLP), mature cells in the stem cell niche (MCN = EryB + Mega + Mono + PreB), and granulocytic mature cells in the peripheral blood or tissues (MCP = Baso + Eos + Neut) (Fig 4E). The spatial relationship between each compartment and HSCe was modeled by the probability of the ligands produced by the compartment reaching HSCe (Materials and Methods). Specifically, we assumed that (i) there is no diffusion for HSCe autocrine ligands, so the probability of HSCe autocrine binding $P_{HSCe}$ is 1; (ii) PC reside close to HSCe, so $P_{PC}$ is 0.8; (iii) MCN reside further away from HSCe than PC, so $P_{MCN}$ is 0.7; (iv) physical barriers between the stem cell niche and the peripheral tissues prevent MCP ligands from reaching HSCe, so $P_{MCP}$ is 0.1. We found that HSCe expressed a broad spectrum of biological function-associated ligands by each cell module (Fig 4B). Asterisks (*) indicate the enriched ligand sets defined as HG $Z$-score $> 1.15$.

**Figure 3. Modular ligand–cell interaction structure in the ligand production network.**

A Hierarchical clustering based on Jaccard distances identifies four cell modules separated by the blue lines.

B Silhouette widths for the four cell modules in (A).

C Expression of a priori biological function-associated ligands by each cell module in (B). Asterisks (*) indicate the enriched ligand sets defined as HG $Z$-score $> 1.15$.

See also Supplementary Table S5 and Supplementary Figure S2.
The model consisted of four compartments to represent cells of hematopoietic cell relative frequency can modulate ligand expo-

A hierarchical clustering based on Jaccard distances identifies four cell modules separated by the blue lines.

**Figure 4.**

**A** Mono-nucleated cell compartment

**B** Stem and progenitor cell compartment

**C** Stem cell niche

**D** Peripheral tissues

**E** Insignificant signal strength

**F** Insufficient signal strength

**G** Communication to HSCe, PC, MCN, MCP

Figure 4.
This cell movement potentially alters the HSC microenvironment. We next sought to predict the spatial effect of Mono, Mega, EryB, and PreB on HSCe feedback regulation. Our simulation results (Fig 4G) revealed the importance of Mega-produced HSCe-targeting ligands in innate inflammatory response terms and the importance of Mono-produced HSCe-targeting ligands in regulating angiogenesis-associated terms. Strikingly, it was evident that EryB-produced HSCe-targeting ligands are associated with regulating cell cycle progression, cell survival and proliferation, which warrants future experimental validation. This analysis indicates that regulation of cell identities in HSCe microenvironment or niche can modulate ligand exposure to HSCe.

In summary, our analysis uncovered promiscuous ligand-to-cell interactions in the ligand binding network. HSCe were found to express receptors for a broad range of ligands, implying the existence of physical parameters such as relative cell frequency and compartmentalization in HSC fate regulation. Our subsequent simulation revealed a potential importance of Mega, Mono, and EryB ligands in HSC fate regulation. To explore how hematopoietic cell type-dependent signals feedback to HSCe, we next performed high-content in vitro experiments for HSCe-targeting ligands.

Validation of HSCe-targeting ligands using a high-content in vitro phenotypic assay

High-content in vitro experiments were performed by following the protocol in Fig 5A. HSCe-targeting ligands in the CCC network (Supplementary Table S4) were ranked according to the molecular interaction confidence scores (Ceol et al, 2010) for ligand–receptor interactions (Supplementary Table S2) and the receptor gene expression levels in HSCe from the Transcriptomic data. Thirty-three ligands were prioritized for experimental tests (Materials and Methods, Supplementary Table S7). We examined the phenotypic impact of each ligand on 40 HSC-enriched cells (HSC-e: Lin–CD34+/CD133+/CD90+ and 1:20 for CD49f+CD90–RholowCD38+CD45RA–CD49f–CD34+ CD133+CD90+) isolated from human UCB samples; this population contains approximately one NOD-scid-IL2Rgcf–/– repopulating cell per 13 cells (combination of 1:10 for CD49f+CD90+ and 1:20 for CD49f–CD90– HSC-enriched cells) (Notta et al, 2011). Each ligand was tested in a short-term assay at three doses in the presence of three basal cytokines (BC)—SCF, THPO, and FLT3LG (Petzer et al, 1996; Madlambayan et al, 2005; Csaaszar et al, 2012). On day 7, the numbers of CD34+/CD133+CD90+ cells (defined as HSC-enriched cells) (Mayani & Lansdorp, 1994; Dorrell et al, 2000; Danet et al, 2001; Ito et al, 2010), CD34+ cells that were CD133– or CD90– (defined as progenitor cells; see Supplementary Fig S5 for functional quantification using the colony-forming cell assay), and CD34+ cells (defined as mature cells) were quantified. The BC cocktail-supplemented culture output 704 ± 425 (mean ± s.d. from 33 biological replicates) consisted of 6.35 ± 3.21% HSC-enriched cells, 27.75 ± 6.86% progenitor cells, and 65.90 ± 10.04% mature cells. This established a reference for detecting the effects of test ligands on HSC-e fate decisions (Supplementary Fig S6). In addition to the BC cocktail, TGFB1 (10 ng/ml) (Batard et al, 2000) and StemRegenin 1 (SR1, 0.75 μM) (Boitano et al, 2010) were used as the negative and positive control for HSC-e expansion, respectively (Fig 5B).

In vitro effect of the 33 ligands was quantified by signed one-tail P-values from the nested ANOVA detailed in the Materials and Methods (Supplementary Fig S7A). P-values of the 35 ligands (including TGFB1 and SR1) at their most effective dose on human UCB HSC-e are shown in Fig 5C. For ligands that did not have any significant effect, results of the highest working concentrations were reported. See Supplementary Fig S8 for cell number comparison between the tested conditions and the BC control. See Supplementary Tables S8 and S9 for results of all the testing conditions. These in vitro data allowed us to examine the impact of the cell types of interest on HSC fate regulation in the CCC network.

Provisional feedback signaling networks for cell type-associable HSC fate modulation

Measurement of the in vitro effect of the 33 ligands on HSC-e allowed creation of a directional CCC network. First, we categorized each ligand into one of the five functional categories [inducing quiescence, inducing self-renewal, inducing differentiation, inducing proliferation (self-renewal + differentiation), and inhibiting proliferation] in terms of their manipulation in HSC-e fate decisions using the P-values in Supplementary Table S9 and the classifier in Table 1. A representative ligand is given for each category in Supplementary Fig S7B. The ligands, at the working concentrations reported. See Supplementary Fig S8 for cell number comparison between the tested conditions and the BC control. See Supplementary Tables S8 and S9 for results of all the testing conditions. These in vitro data allowed us to examine the impact of the cell types of interest on HSC fate regulation in the CCC network.
found to direct HSC-e fate decisions (Fig 1; step 3a), indicating a significant enrichment of prediction capacity in this analysis (Binomial $P = 0.0001$, Materials and Methods).

Intriguingly, dose-dependent HSC-e fate regulation was observed for some ligands. For example, TNFSF10, at a working concentration of 1 ng/ml, did not affect the number of HSC-enriched cells, progenitor cells, or mature cells (ANOVA $P$-values were 0.2747, 0.2642, and 0.3721, respectively). When the ligand was used at 10 ng/ml, it led to an increase in the number of HSC-enriched cells (ANOVA $P = 0.0036$), so it induced HSC-e self-renewal. At a working concentration of 100 ng/ml, however, the ligand led to a significant decrease in the number of HSC-enriched cells (ANOVA $P = 0.0007$), progenitor cells (ANOVA $P = 0.0094$), and mature cells (ANOVA $P = 0.0207$) (Supplementary Fig S6bii), so it inhibited HSC-e proliferation, which may be due to the pro-apoptotic effect of the ligand (Zamai et al., 2000). Dose-dependent HSC-e fate regulation was also observed for FGF1, FGF2, IL11, and TNFSF12 (Supplementary Table S9). This result is reminiscent of differential

---

**Figure 5. HSC-e respond to exogenously added HSCe-targeting ligands.**

A The experimental and analytical protocol. HSC-e: human UCB Lin+Rho-dlowCD34+CD90+CD45RA+CD49f+. BC basal cocktail consisted of 100 ng/ml SCF, 50 ng/ml THPO, and 100 ng/ml FLT3LG.

B Fold changes between the results of (i) negative control (TGFB1) and (ii) positive control (SR1) and that of the cell culture supplemented with BC only. HSC-enriched cells: CD34+CD133+CD90-; progenitors: CD34+ cells that are CD133+ or CD90+, mature cells: CD34-. Data are from 33 biological replicates.

C Signed one-tail $P$-values from the nested ANOVA when comparing the cell counts of testing conditions to the BC control. Positive $P$-values indicate that effect of a test ligand was greater than that of the BC control, and negative $P$-values indicate the effect of a test ligand was less than that of the BC control. Ligand concentration is in ng/ml, except for SR1 that is in µM.

See also Supplementary Figs S4, S5 and S6.

© 2014 The Authors

Molecular Systems Biology 10 741 | 2014 9
activation of pathways that are involved in diverse biological processes (Kale, 2004). Furthermore, categorization of some ligands such as FGF2 (working concentration, WC = 50 ng/ml) and BMP6 (WC = 100 ng/ml) was sensitive to the statistical significance threshold, suggesting their indeterminate role in regulating HSC-e fate decisions may be context dependent. The ligands were excluded accordingly in the subsequent analyses.

We explored how ligands produced by different cell types influenced HSC-e fate decisions by performing a functional enrichment analysis for the ligands expressed by each of the 12 cell types in the CCC network using the ligand function categorization (Fig 6A) as a reference. To ensure that there were sufficient data to draw qualitative conclusions, the analysis was performed based on the categorization at the intermediate confidence level while excluding BMP6 in which categorization was indeterminate at that confidence level. Assuming each ligand acts independently in HSC-e fate regulation, this analysis allowed us, for the first time, to predict the role of each cell type in the HSC-e feedback regulation. As shown in Fig 6B, progenitor cells such as CMP, MEP, GMP, and MLP predominantly expressed ligands that induced HSC-e quiescence and self-renewal; EryB expressed ligands of diverse functions as expected from the results shown in Fig 3C. In contrast to a majority of the cell types, which expressed at most three types of directive signals for HSC-e fate decisions, HSCe expressed ligands inducing self-renewal, quiescence, and differentiation, and inhibiting proliferation. This is reminiscent of self-sufficient autocrine signaling of HSC (Kirit w et al, 2005) possibly to compensate for their disadvantage in accessing exogenous signals.

<table>
<thead>
<tr>
<th>Function categories</th>
<th>HSC-enriched cells</th>
<th>Progenitor cells</th>
<th>Mature cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Quiescence induction</td>
<td>–</td>
<td>–</td>
<td>↓</td>
</tr>
<tr>
<td>Self-renewal induction</td>
<td>↑</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Differentiation induction</td>
<td>–</td>
<td>↑</td>
<td>–</td>
</tr>
<tr>
<td>Proliferation induction</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Proliferation inhibition</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
</tr>
</tbody>
</table>

Dash “−” indicates no change from the basal cytokine control.

Table 1. Functional definition of ligands for HSC-e fate regulation based on a cell number comparison between the conditions having the ligands of interest and the basal cytokine control.

Figure 6. In vitro experiments lead to functional categorization of HSCe-targeting ligands.

A. Functional categorization for the 35 HSCe-targeting ligands, including the negative control TGFβ1 and the positive control SR1. The ligands were categorized at different confidence levels: High, intermediate, and low confidence levels refer to ANOVA p-value significance thresholds 0.01, 0.02, and 0.05, respectively. See Table 1 for definition of the functional categories.

B. Functional enrichment was performed for the HSCe-targeting ligands produced by each cell type. The color scale indicates the HG enrichment Z-scores.
Hematopoietic cell communication network

Molecular Systems Biology

due to low cell frequency (Fig 4C). Collectively, we propose that both the progenitor cells and the mature cells regulate HSC-e fate decisions via feedback signaling yet through different mechanisms—the progenitor cells feed back HSC-e self-renewal and quiescence signals, whereas the more mature cells feed back HSC-e predominantly proliferation and differentiation signals (Fig 1; step 3b).

**Pathway enrichment analysis suggested intracellular regulatory motifs for HSC-e fate decision-making**

The association between HSCe-targeting ligands and different cell types allowed us to construct a qualitative CCC network focusing on HSC-e fate regulation (Fig 7A). A database survey on the intracellular signaling pathways of the HSC-targeting ligands suggested that intracellular regulatory motifs are associative with the ligands responsible for directive effects on HSC-e cell fate decisions in vitro (Fig 7B, Supplementary Fig S9, Materials and Methods). Specifically, signaling activity of the HSC-e quiescence-inducing ligands (such as BMP6 and INHBA), self-renewal-inducing ligands (such as ANGPT1, ANGPT2, NGF, and TNFSF12), proliferation-inducing ligands (such as TGFβ1, TNFSF10, and TNF) were attributable to SMAD (permutation $P = 0.044$, Supplementary Fig S9A), NF-κB (permutation $P = 0.122$, Supplementary Fig S9C), STAT (permutation $P = 0.04$, Supplementary Fig S9C), and caspase cascade (permutation $P = 0.059$, Supplementary Fig S9D) pathways, respectively.

Our qualitative CCC network can be depicted in three ways: a directional network weighted by receptor frequency (Fig 7C), a directional network weighted by cell frequencies in the MNC compartment (Fig 7D), and a weighted directional network with compartmentalization (Fig 7E) overlaid to illustrate the roles of cellular dynamics and spatial distribution in HSC fate regulation through feedback signaling. For example, Neut was the largest cell population in the MNC isolated from human UCB (Fig 4C), so TNFSF10 and TNF from Neut were potentially the major signals to inhibit HSC-e proliferation. However, the stem cell niche-peripheral barrier would typically protect HSC-e from the inhibitory signals.

In summary, we combined the topology of the CCC network, the in vitro effect of 33 ligands on HSC-e fate decisions, and pathway information of the ligands. Our results support a model whereby hematopoietic cells influence HSC toward certain cell fates by regulating the key intracellular regulatory motifs through cell type-specific feedback signals.

**Discussion**

While it is accepted that feedback regulation of HSC fate decisions is important to stable hematopoiesis (Csaszar et al, 2012; Heazlewood et al, 2013), it has been unclear how the feedback system operates. Extensive effort has been made to understand how stromal cells in the bone marrow microenvironment regulate HSC fate decisions (Zhang et al, 2003; Nakamura et al, 2010; Kunisaki et al, 2013). In addition, we propose a hematopoietic cell-driven feedback system that regulates HSC fate decisions through intercellular signaling.

We constructed a bipartite graph to represent the CCC network between 12 hematopoietic cell types isolated from human UCB (and orphan signals entering the network). We found that the CCC network can be depicted in two formats based on signal directionality—ligand production and ligand binding, and each format was analyzed as an individual network. The high degree of modularity in the ligand production network pointed to cell type-specific production of ligands for HSC-e cell fate regulation. In contrast, the ligand-to-cell interactions in the ligand binding network were promiscuous, and HSCe were one of the cell types that bound the most ligands, suggesting that HSCe have broad environment sensing capacity (Takizawa et al, 2012). Our analysis raised important questions about how feedback specificity is achieved in HSC fate regulation. In silico simulation posed the hypothesis that additional control mechanisms including those observed in vivo (cell type frequency control and HSC niche localization or compartmentalization) are required to confer specificity in hematopoietic cell-mediated feedback regulation of HSC fate decisions. To test the hypothesis, we prioritized 33 HSC-targeted ligands in the CCC network for in vitro experiments. We anticipated the roles of the 33 ligands in directing HSC-e fate decision using a cell surface marker expression-based phenotypic assay. The in vitro data allowed us to uncover what signals each of the 12 cell types feeds back to HSC-e. For instance, the mature cells, particularly Mono and granulocytes (Neut, Baso, and Eos), were found to express mainly inhibitory signals for HSC-e proliferation and inducing signals for HSC-e differentiation, which in combination can exhaust the HSC population because of the extensive cell cycling and division involved in the proliferation and differentiation processes (Hock et al, 2004; Zhang et al, 2006). However, under a normal in vivo condition, monocytes and granulocytes mainly circulate in the peripheral tissues; their secreted ligands have limited access to HSC in the bone marrow compartment because of the blood–bone marrow barrier. The identified importance of cell compartmentalization in protecting HSC from exogenous signals is consistent with our observation that global media dilution enhances in vitro HSC production when physical barriers between HSC and the mature cells are absent (Csaszar et al, 2012). We also found that progenitor cell types—CMP, MEP, GMP, and MLP—that typically co-localized with HSC in the bone marrow niche tend to function as a unit, enriched for ligands for HSC maintenance by inducing HSC quiescence and self-renewal. This finding supports the use of periodic primitive cell selection to increase in vitro HSC production (Madlambayan et al, 2005) and suggests technologies that target the HSC niche composition to control HSC fate in vivo.

The pathway enrichment analysis pointed to specific intracellular regulatory motifs associated with ligands of different in vitro effects on HSC-e fate. Specifically, HSC-e quiescence-inducing ligands such as BMP6 (Holien et al, 2012) and INHBA (Burrette et al, 2005) regulate the expression of SMADs to arrest cell growth. The HSC-e self-renewal-inducing ligands such as angiopoietins (Hughes et al, 2003), NGF (Descamps et al, 2001), and TNFSF12 (Kawakita et al, 2004) were found to regulate the activity of NF-κB in which deletion in the mouse hematopoietic system compromised the self-renewal and long-term hematopoietic repopulation ability of HSC (Zhao et al, 2012; Stein & Baldwin, 2013). The HSC-e proliferation-inducing ligands such as CSF2 (Carter, 2001; Gu et al, 2007), CSF3 (Harel-Bellan & Farrar, 1987), and IL11 (Yoshizaki et al, 2006) were found to induce the expression of STATs for cell proliferation. Finally, the HSC-e proliferation inhibitory ligands such as TGFβ1 (Shima et al, 1999), TNF (Mallick et al, 2012), and TNFSF10 (Kischkel et al,
A Cell-cell communication network for HSC-e fate regulation

**Figure 7.** HSC-e feedback signaling network points to intracellular regulatory motifs for HSC-e fate regulation.

A Cell-cell communication network for HSC-e fate regulation: The hematopoietic cell-driven network for HSC-e fate regulation. The positive and negative feedback signals are in respect to in vitro expansion of CD34+CD33+CD90+ cells.

B Intracellular regulatory motifs associated with ligands of different directive effects on in vitro HSC-e fate.

C Interactions between ligand-producing cells and ligands are weighted by the number of corresponding receptors (in terms of species) expressed in HSC-e. The thicker the edge, the higher the weight.

D Interactions in (C) are weighted by cell frequencies obtained from fresh human UCB mono-nucleated cell samples shown in Fig 4G.

E Interactions in (D) are weighted by spatial compartmentalization, where 10% of the ligands from peripheral compartment (Baso, Eos and Neut) reach HSC-e. The expressed ligands of the “Others” population, such as BMP2, LTA CSF2, and IL17, are not shown due to the lack of cell frequency information.

See also Supplementary Fig S9.

2000) initiated caspase cascade to cause cell death. Although many connections between exogenous ligand stimulation, pathway node activity, and cell phenotype changes were established in cancer cell lines, these connections led us to the anticipation that exogenous ligands direct HSC-e toward different cell fate by regulating the activity of specific cell fate decision-associated intracellular regulatory motifs, which opens opportunities for future study.

In summary, our results demonstrate the importance of cell-to-cell communication in human UCB stem cell fate control. Hematopoietic cells influence HSC toward certain cell fates by regulating...
the key intracellular regulatory motifs through cell type-specific feedback signals. Further, control parameters such as relative cell frequency and spatial compartmentalization (niches) are opportunities to impose specificity in HSC fate regulation. A particularly interesting extension of our current work is to analyze how defects in HSC niche composition and physical structure or defects in HSC intracellular regulatory motifs affect feedback regulation of HSC fate decisions in vivo and consequently causes hematopoietic disorders such as leukemogenesis (Schepers et al., 2013).

One limitation of this study is that we used only transcriptomic data rather than proteomic data to construct the CCC network. Although there is a general agreement between mRNA and protein expression levels of ligands and receptors in mammalian cells (De Haan et al., 2003; Madlambayan et al., 2005; Schwanhausser et al., 2011), gaining better understanding of the dynamics of mRNA expression and the corresponding protein expression can be important in understanding context-specific network structures and their dynamic evolution. The newly developed mass cytometry (Bendall et al., 2011) offers a novel single cell proteomic approach to achieve this goal. A second limitation of this study is that we defined the exogenous effects of 33 ligands on HSC-e fate decision according to in vitro measurements of a cell surface marker expression-based phenotypic assay. Discrepancy between our observation about the in vitro effects of the tested ligands and their documented effects in literature may be attributable to the differences in experimenting cell populations and culture conditions. Further functional validation of the surface markers to cell function fidelity would certainly strengthen our analysis of network directionality; ultimately, our network should guide the selection of potentially novel HSC-e-regulating cell types, ligands, and their key intracellular signaling nodes for in-depth in vivo characterization. A final limitation of this study is that we used a static (human UCB) network to predict potentially dynamic feedback relationships between HSC-e and the other cell types. Exploring how the network connections change during culture evolution (Qiao et al., 2012) is an important next step. The assumption of our static network is direct (as opposed to indirect) feedback from each cell type to HSC-e. Although our in vitro study was specifically designed to enrich for direct effects of ligands on HSC-e by using the HSCe receptor expression information as a criterion for selecting test ligands and using a short culture time (7 days) (Caszar et al., 2014), further analysis of multi-step and adaptive feedback is needed to strengthen links to in vivo hematopoiesis.

Collectively, cell–cell communication is fundamental to biologic tissues. However, it has not been extensively explored as a network because a large number of underpinning variables need to be considered. Here, we provide a framework to systematically depict cell–cell communication as a network while exploring the roles of cell frequency and spatial distribution in the system. As a next step, connecting the CCC network with more widely studied protein–protein interaction (Kirouac et al., 2010) and gene regulatory (McKinney-Freeman et al., 2012) networks through mechanistic models of intracellular signaling activity and the resulting cellular responses (Janes et al., 2005) will allow us to understand how HSCs integrate exogenous signals to make fateful decisions. The outcome will not only contribute to the development of more effective methods for HSC production, but also further our knowledge about HSC (niche) biology and cell–cell communication as a layer of biological regulation.

Materials and Methods

Microarray datasets

Illumina data of primitive cells and progenitor B cells (ProB: CD34+CD10+CD19+; three biological replicates) were obtained from the authors of Laurenti et al. (2013). The primitive cells are HSCe (Lin-CD34-CD38-CD49f+CD45RA-CD90-CD10-; 10 biological replicates), CMP (Lin-CD34-CD38-CD135-CD45RA-CD7-CD10-; five biological replicates), MLP (Lin-CD34-CD38-CD90-CD45RA-; five biological replicates), MEP (Lin-CD34-CD38-CD135-CD45RA-CD7-CD10-; five biological replicates), and GMP (Lin-CD34-CD38-CD135-CD45RA-CD7-CD10-; five biological replicates). The data are accessible at Gene Expression Omnibus (GEO) (Edgar et al., 2002) through accession number GSE42414. Quantile signals of the Illumina data were calculated using the normalizeQuantile() function in the limma package (v3.16.3) of BioConductor.

Affymetrix CEL files of mature cells and ProB (CD34+CD10-CD19+; five biological replicates) were downloaded from GEO (accession number GSE47595 (Novershtern et al., 2011), accessed on 2011-11-20). The mature cells are Mega (CD34+CD41+CD61+CD45; six biological replicates), Ery (CD34+CD71+GlyA+; six biological replicates), Neut (FScIlSSC5CD16+CD11b+; four biological replicates), Baso (FScIlSSC5CD22+CD123+CD3+/-CD45+/; six biological replicates), Eos (FScIlSSC5CD123+CD33+/; five biological replicates), Mono (FScIlSSC5CD14+CD45+/; five biological replicates), and PreB (CD34+CD10+CD19+; three biological replicates). Quality of the Affymetrix arrays was assessed using the simpleaffy (v2.32.0) and AnnotationDbi (v1.18.4) packages of BioConductor. The arrays with average background more than 2 s.d. from the mean background level of all arrays and the arrays with present percent is less than 1.5 s.d. from the mean present% of all arrays were not used for this study. Robust multi-array average (RMA) signals of the selected arrays were computed using the justRMA() function in the limma package (v3.16.3) of BioConductor. Affymetrix annotation for GeneChip U133AAv2 (GEO accession number: GPL4686) was used.

To combine the Illumina and the Affymetrix datasets, each dataset was normalized by the averaged gene expression signal of the respective ProB arrays. An averaged signal was calculated for probes of the same gene according to Entrez gene identifiers. The post-processed datasets were merged by Entrez gene identifiers.

Ligand functional enrichment analysis

For the gene set enrichment analysis (GSEA) in Fig 2B, 13 hematopoietic gene sets (Supplementary Table S1) were compiled from the GeneGO database on 2012-11-15. GSEA was performed using the GSEA software (v2, http://www.broadinstitute.org) with the minimum gene set size equal to 1, and the other settings as defaults. See Supplementary Table S1 for GSEA Z-scores.

For the biological process enrichment analysis in Figs 3C and 4F, gene sets in Supplementary Table S5 were curated from the MetaCore pathway database (http://thomsonreuters.com/metacore/, accessed on 2014-03-05). The material is reproduced under a license from Thomson Reuters; it may not be copied or re-distributed in
whole or in part without the written consent of the scientific business of Thomson Reuter.

**Ligand/receptor database**

Using gene ontology terms “cytokine activity,” “growth factor activity,” “hormone activity,” and “receptor activity,” 417 genes with ligand activity and 1,723 genes with receptor activity were compiled from BioMart (Kasprzyk, 2011) (accessed on 2012-02-29). Ligand–receptor interaction pairs documented in public databases were compiled using the iRefWeb (Turner et al, 2010) resource (accessed on 2012-03-05). Additional 38 ligand–receptor interaction pairs from literatures (as on 2013-02-04) were included. See Supplementary Table S2 for the resulting 933 ligand–receptor interaction pairs in Supplementary Table S2. If “Cell A” expresses a receptor for ligand x and “Cell B” expresses ligand x, an arrow is drawn from “Cell B” to “Cell A.” Networks were built in R (v2.15.1) and visualized in Cytoscape (v2.8.3). The R code is available upon request.

**Hierarchical clustering**

The hierarchical clusters in Fig 2C were obtained using the Ward agglomeration method with the Manhattan distance matrix. Confidence of the clusters was quantified by approximately unbiased (AU) P-values (Shimodaira, 2002, 2004), a type of bootstrap P-values, computed using the pvclust package (v1.2-2) in R (v3.0.0).

**Identification of differentially over-expressed genes**

For the cell type of interest, one-way pairwise Wilcoxon test (R, v2.15.1) was performed between the gene expression profiles of the interested cell type and the profiles of each of the other cell types. P-values were adjusted using the Benjamini & Hochberg method (or false discovery rates, FDR). At a given threshold, the ligand and receptor genes that differentially over-expressed comparing to six other cell types (the threshold was set arbitrarily) were defined as the differentially over-expressed ligands and receptors of the cell type. The identified receptors of each cell type were compared to hematopoietic cell type-specific receptors using receiver operating characteristic (Supplementary Fig S1). The cell type-specific receptors are (1) ACVR1L1 (for TGFBI), ENG (for TGFBI), EPOR (for KIT), FBKIP1A (for TGFBI), IL12R (for IL7), IL7R (for IL7), ITGAV (for TGFBI), ITGB6 (for TGFBI), ITG9N (for TGFBI), KIT (for KITLG), LTBPI (for TGFBI), LTBP4 (for TGFBI), MPL (for THPO), TGFBR1 (for TGFBI), TGFBR2 (for TGFBI), TGFBR3 (for TGFBI), VTN (for TGFBI), CD34 and ITGA6 (CD49) for HSCe; (2) IL3RA (for IL3), CSF2RA (for CSF2), CSF2RB (for CSF2), CSF3R (for CSF2), EPOR (for KIT), KIT (for KIT), MPL (for THPO), CD34, CD38, FLT3 (CD135) for CMP; (3) MPL (for THPO), EPOR (for EPO), CD34 and CD38 for MEP; (4) CSF3R (for CSF3), CD34, CD38, FLT3, PTPRC (CD45RA) for GMP; (4) IL2RG (for IL7), IL7R (for IL7), CD34, PTPRC (CD45RA) for MLP; (5) MPL (for THPO), ITGA2B (CD41), ITGB3 (CD61) for Mega; (6) EPOR (for EPO), GYPA (CD235a) for EryB; (7) CD14 for Mono; (8) CD22 and IL3RA (CD123) for Basso; (9) IL3RA (CD123) for Eos; (10) FGCR3A (CD16) and ITGAM (CD11b) for Neut; and (11) IL2RG (for IL4), IL4R (for IL4), IL13RA1 (for IL4), MME (CD10), and CD19 for PreB.

**Network construction**

Directionality of the CCC network was defined by the expression of ligand and receptor genes on the cell types of interest, and the ligand–receptor pairs in Supplementary Table S2. To assess the effects of each test ligand (in addition to SCF, THPO, and “Cell A” expresses a receptor for ligand x and “Cell B” expresses ligand x, an arrow is drawn from “Cell B” to “Cell A.” Networks were built in R (v2.15.1) and visualized in Cytoscape (v2.8.3). The R code is available upon request.

**Bipartite network analysis**

Clustering for the ligand production networks was performed based on Jaccard distances appropriate for binary graph adjacency matrices (Gower & Legendre, 1986). Clustering for the ligand binding networks was performed using the spectral co-clustering algorithm (downloaded from http://adios.tau.ac.il/SpectralCoClustering/ on 2013-06-01) appropriate for weighted graph adjacency matrices (Dhillon, 2001).

Potential of apparent competition ( Muller et al, 1999) of cell type i to cell type j, $P_{ij}$, was computed as

$$P_{ij} = \sum_k \left( \frac{f_i R_k}{\sum_l f_i R_k} \cdot \frac{f_j R_k}{\sum_l f_j R_k} \right)$$

where $f_i$ is the normalized cell frequency of cell type i by the total cell frequency of the analyzed cell types, thus $f_i$ is between 0 and 1; $R_k$ is the number of receptors that cell type i expressed for ligand k; $K$ is the total number of ligands that cell type i binds; $I$ is the total number of ligands that cell type i binds; and M is the total number of cell types that ligand k binds. The figures were drawn by modifying the plotPAC() function in the bipartite package (v1.18) in R (v3.0.0).

**Network comparison**

To compare interaction patterns between the network of ligand source and the network of ligand sink, for each network, the numbers of overlapped ligands between one module and the other modules were obtained. The overlap of ligands between modules in the network of ligand source $S = \{9, 13, 10, 12, 17\}$, and the overlap of ligands between modules in the network of ligand sink $T = \{75, 75, 69\}$ Two-sample t-test was performed for $S$ and $T$ in R (v3.0.0).

**Flow cytometry analysis**

Human UCB samples were collected from consenting donors according to ethically approved procedures at Mt. Sinai Hospital (Toronto, ON, Canada). Mono-nucleated cells were obtained by depleting red blood cells (RBC) using RBC lysis buffer (0.15 M NH4Cl, 0.01 M KHC03, 0.1 mM EDTA) as previously described (Kirouac et al, 2009). Lineage-negative (Lin−) cells were isolated from the mono-nucleated cell fraction using the StemSep system or the EasySep system for human progenitor cell enrichment (StemCell Technologies, Inc., Vancouver, BC, Canada), according to the manufacturer’s protocol. Cell frequencies shown in Fig 4C and 4D were obtained from mono-nucleated cells of fresh UCB samples and thawed Lin− cell samples, respectively. The cells were stained using the following antibodies in 1:100 unless stated otherwise: CD90 (FITC, 1:50), CD38 (PE, PECy5, APC), CD45RA (1:50, APC), CD34
(PE-Cy7), CD49f (PE-Cy5, 1:50), CD7 (FITC), CD10 (FITC), CD135 (1:50, PE), CD45RA (1:50, APC), CD71 (FITC), CD235a (PE), CD61 (FITC), CD41 (PE), CD45 (PE-Cy7), CD14 (PE), CD16 (PE), CD11b (PE-Cy7), CD22 (FITC), CD33 (PE), CD123 (PE-Cy5), CD19 (FITC), and CD10 (PE). All the antibodies were from BD Biosciences, Missis-sauga, ON, Canada.

Logic modeling

The effect of cell localization on the identity of HSCe-targeting ligands $M_{\text{HSCe}}$ was simulated using an OR gate model:

$$M_{\text{HSCe}} = (x_{\text{HSCe}} \cdot x_{\text{LinPC}}) \cup (x_{\text{PC}} \cdot x_{\text{MCN}}) \cup (x_{\text{MCPC}} \cdot x_{\text{MMP}}),$$

where $x_{\text{HSCe}}$, $x_{\text{PC}}$, $x_{\text{MCN}}$, and $x_{\text{MMP}}$ are the differentially over-expressed ligands of HSCe, progenitor cells (CMP, GMP, MEP, and MLP), mature cells in the stem cell niche (MCN), and mature cells in the peripheral tissues (MCP). Randomly generated logic vectors $x_{\text{HSCe}}$, $x_{\text{PC}}$, $x_{\text{MCN}}$, and $x_{\text{MMP}}$ represented the probability ($P_{\text{HSCe}}$, $P_{\text{PC}}$, $P_{\text{MCN}}$, and $P_{\text{MMP}}$) of the ligands of each compartment to reach HSCe. Enrichment ($E$) of HSCe-targeting ligands $M_{\text{HSCe}}$ in a biological process mediated by ligand set $B$ was quantified as following:

$$E = \frac{n(M_{\text{HSCe}} \land B)}{n(B)},$$

where $n(M_{\text{HSCe}} \land B)$ is the number of HSCe-targeting ligands in biological process $B$, and $n(B)$ is the number of ligands in biological process $B$. For each test condition (i.e., combination of $P_{\text{HSCe}}$, $P_{\text{PC}}$, $P_{\text{MCN}}$, and $P_{\text{MMP}}$), enriched scores from 500 simulations were averaged. Content of 11 manually curated ligand sets of biological processes are tabulated in Supplementary Table S5.

In vitro experiments

Human Lin$^{-}$ cells were isolated from UCB samples collected from consenting donors according to ethically approved procedures at Mt. Sinai Hospital (Toronto, ON, Canada). Forty Lin$^{-}$ Rho$^{-}$CD34$^{-}$CD38$^{-}$CD45RA$^{-}$CD49f$^{-}$ cells were sorted and dispensed per well in a 96-well V-bottom plate with a MoFloXDP flow cytometer (Beckman Coulter). The cells were cultured in a serum-free condition supplemented with 100 ng/ml SCF, 100 ng/ml FLT3LG, 50 ng/ml THPO, and a test ligand at specific concentration. On day 7, cells were stained. Total cell counts ($N_{\text{Total}}$), CD34$^{-}$CD133$^{-}$CD90$^{-}$ cell counts ($N_{\text{HSCe-enriched}}$), and CD34$^{-}$ cell counts ($N_{\text{Mature}}$) were obtained using an LS RFortessa flow cytometer (BD Bioscience). Progenitor cell counts were calculated as $N_{\text{Total}} - N_{\text{HSCe-enriched}} - N_{\text{Mature}}$. See also “optimization of in vitro experiments” in the Supplementary Information S1.

Statistical analysis

To assess the effects of each test ligand (in addition to SCF, THPO and FLT3LG) on in vitro HSC-e fate decisions, a mixed-linear model was constructed with the experiment identifier as the random effect to account for the variability from experiment to experiment. The analysis was performed using the lme() function of the nlme package (v3.1-113) in R (v2.15.1). The source code is provided as Supplementary Information S1.

Since we were mostly concerned with not missing any effective ligands (type II error) that will inform future research, nominal $P$-values of the mixed model were reported without correction for multiple tests. The ligands were categorized using definition in Table 1. Ligand categorization was performed for significance $P$-value thresholds of 0.01, 0.02 and 0.05 (Supplementary Table S9). See also “statistical analysis for in vitro experiments” in the Supplementary Information S1.

At the $P$-value threshold of 0.02, 5 ligands were found to be neutral to HSC-e and 27 were categorized into five functional categories (inducing HSC-e quiescence, self-renewal, differentiation and proliferation, and inhibiting HSC-e proliferation). Assuming the probability that a selected ligand is functional is 0.5 and that the effectiveness of test ligands were independent from each other, the ligand selection process was modeled as a binomial process with distribution $X \sim B(33, 0.5)$, where 33 is the number of test ligands. The expected number of effective ligands was 33$\cdot$$0.5 \approx 16$. The probability of having 27 effective ligands is

$$P(X = 27) = \binom{33}{27}0.5^{27}(1 - 0.5)^6 \approx 0.0001$$

Prior to the in vitro experiments for testing the activity of HSCe-targeting ligands on HSC-e, we sought to prioritize ligands for experiments. To do that, we performed a literature survey on ligands that had been used in in vitro cell culture of human cord blood-derived cells; 11 ligands fell in this category (Supplementary Table S7). Ligands such as ANGPT1, ANGPT2, ANGPTL3, and BMP2 had been used in mice or human bone marrow cells (Supplementary Table S7), so they were also prioritized for experiments in our study. Excluding these ligands from our analysis, 15 ligands out of 18 tested ligands were effective. The corresponding probability is

$$P(X = 27) = \binom{18}{15}0.5^{15}(1 - 0.5)^3 \approx 0.003$$

To dictate the respective regulatory effects of HSCe, CMP, GMP, MEP, MLP, Mega, EryB, Mono, Neut, Eos, Baso, PreB, and Others on HSC-e cell fates, the tested ligands of each cell type were extracted from the CCC network in Supplementary Table S4. Functional enrichment analysis was performed for each cell type using hypergeometric Z-scores,

$$Z = \frac{k - n \frac{m}{N}}{\sqrt{n \frac{m}{N} \left(1 - \frac{m}{N}\right) \left(\frac{N - m}{N - 1}\right)}}$$

where $N = 117$ is the number of HSCe-targeting ligands expressed by the 13 cell types, $m$ is the number of ligands in a given function group, $n$ is the number of expressed ligands of the cell type of interest, and $k$ is the number of expressed ligands in the function group of interest.

Functional HSC-e feedback signaling network

In Fig 7C, strength of the produced signals of function group $k$ from cell type $i$ to HSC-e was modeled as
Molecular Systems Biology

\[ S_{i,k} = f_i \sum_{n=1}^{N} R_{n,k} \alpha_{k,n} \]

where \( f_i \) is the frequency of cell type \( i \), \( n \) is the number of expressed ligands of function group \( k \) by cell type \( i \), and \( R \) is the expression level of receptor gene \( n \). Cell frequencies are from Fig 4Cl.

Pathway analysis

Intracellular regulatory factors downstream of 16 out of the 19 ligands shown in Fig 7A are available in the MetaCore database (http://thomsonreuters.com/metacore/, accessed on 2014-04-01). The regulatory factors of each ligand were compiled and compared to the regulatory factors of the other ligands of the same functional group. Enrichment of ligands of the same functional group to each regulatory factor was calculated by a permutation test. For each regulatory factor, random functional categorization (quiescence induction, self-renewal induction, proliferation induction, and proliferation inhibition) was performed for 100,000 times. The ratio between the number of times that a regulatory factor randomly fell in a functional category more frequent than the actual categorization and 100,000 is defined as the permutation \( P \)-value. The results of pathway analysis for HSC-e differentiation-inducing ligands are not presented because pathway information was only found for one differentiation-inducing ligand BMP4, and the data are not sufficient for an enrichment analysis. The material from the MetaCore pathway database is reproduced under a license from Thomson Reuters.

Supplementary information for this article is available online:
http://msb.embopress.org

Acknowledgements

WQ was supported by Ontario Graduate Scholarships and a National Science and Engineering Research Council postgraduate scholarship. WW was supported by an Ontario Stem Cell Initiative post-doctoral fellowship. EL was supported by the Swiss National Science Foundation, Roche, and the FSMB (Fondation Suisse pour les Bourses en Médecine et Biologie). This work was supported by grants to SJW from the Canadian Institutes of Health Research, the Ontario Research Fund, and the SickKids Foundation; grants to GDB from NRNB (U.S. National Institutes of Health, National Center for Research Resources Grant Number P41 GM103504); grants to JED from Genome Canada through the Ontario Genomics Institute, Ontario Institute for Cancer Research, and a Summit Award with funds from the province of Ontario, the Canadian Institutes for Health Research, a Canada Research Chair, the Princess Margaret Hospital Foundation, the Terry Fox Research Institute, Canadian Cancer Society Research Institute, and in part by the Ontario Ministry of Health and Long Term Care (OMOHLTC, the views expressed do not necessarily reflect those of the OMOHLTC), and grants to PWZ from the Human Frontier Science Program, the Leukemia and Lymphoma Society of Canada, the Canadian Stem Cell Network, and the Ministry of Research and Innovation of Ontario. PWZ is the Canada Research Chair in Stem Cell Bioengineering. The authors would like to thank the members of the PWZ laboratory and Dr. Daniel Kirouac for their helpful discussion.

Author contributions

WQ and PWZ conceived and designed the study and wrote the manuscript. WQ performed in silico studies and analyzed in vitro data. WW performed in vitro experiments and contributed to drafting the manuscript. EL, ALT, SJW, GB and JED contributed reagents/materials/analysis tools. All the authors reviewed the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Conference on Knowledge Discovery and Data Mining, pp 269 – 274. New York, NY, USA: ACM


Hematopoietic cell–cell communication network

Wenlian Qiao et al

License: This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.