Human haematopoietic stem cell development: from the embryo to the dish

Andrejs Ivanovs1,2, Stanislav Rybtsov1, Elizabeth S. Ng3,4, Edouard G. Stanley3,4,5, Andrew G. Elefanty3,4,5,* and Alexander Medvinsky1,*

ABSTRACT

Haematopoietic stem cells (HSCs) emerge during embryogenesis and give rise to the adult haematopoietic system. Understanding how early haematopoietic development occurs is of fundamental importance for basic biology and medical sciences, but our knowledge is still limited compared with what we know of adult HSCs and their microenvironment. This is particularly true for human haematopoiesis, and is reflected in our current inability to recapitulate the development of HSCs from pluripotent stem cells in vitro. In this Review, we discuss what is known of human haematopoietic development: the anatomical sites at which it occurs, the different temporal waves of haematopoiesis, the emergence of the first HSCs and the signalling landscape of the haematopoietic niche. We also discuss the extent to which in vitro differentiation of human pluripotent stem cells recapitulates bona fide human developmental haematopoiesis, and outline some future directions in the field.

KEY WORDS: HSC, Human, Embryo, hPSC, Reprogramming

Introduction

In mammals, the first transient waves of blood cells arise in the yolk sac (extra-embryonic haematopoiesis) and serve the immediate needs of the growing embryo (Silver and Palis, 1997). They include the most primitive wave, which consists mainly of large nucleated erythrocytes that is rapidly followed by a second wave of erythro-myeloid progenitors (EMPs) and lymphoid progenitors that transiently seed the foetal liver (Böiers et al., 2013; McGrath et al., 2015). The third wave, which includes self-renewing haematopoietic stem cells (HSCs) that give rise to the permanent adult haematopoietic system, emerges later inside the body of the embryo (intra-embryonic haematopoiesis) in the aorta-gonad-mesonephros (AGM) region, which is evolutionarily conserved in many vertebrates (Medvinsky et al., 1993; Medvinsky and Dzierzak, 1996). The process of HSC development in vertebrates is driven by largely conserved, although not entirely identical, molecular mechanisms (Medvinsky et al., 2011; Ciau-Uitz et al., 2016). The endothelial origin of haematopoietic progenitors and HSCs has been established by their emergence from the aortic endothelial layer (Garcia-Porrero et al., 1995; Tavian et al., 1996; Jaffredo et al., 1998), lineage tracing in vivo using genetic labelling (Zovein et al., 2008; Chen et al., 2009; Bertrand et al., 2010; Kiss and Herbomel, 2010) and observations of embryonic stem cells (ESCs) differentiating in vitro (Eilken et al., 2009; Lancerin et al., 2009). HSCs and non-self-renewing haematopoietic progenitor cells that emerge in the AGM region are organised within intra-aortic haematopoietic clusters (IAHCS). These bud predominantly from the endothelial floor of the dorsal aorta, which accordingly co-express endothelial and haematopoietic markers. The search for markers that would accurately identify developing HSCs and separate them from non-self-renewing progenitors is an important goal towards understanding the mechanisms that underlie HSC development (Rybtsov et al., 2014; Zhou et al., 2016). Important events that occur downstream of HSC emergence include colonisation and expansion of HSCs in the foetal liver and subsequent lodging in the adult bone marrow.

Our current understanding of the mechanisms that drive HSC development has come from analysis of various model organisms. We do not intend here to provide detailed overview of non-human models, as this extensive topic has been reviewed recently by us and by others (Medvinsky et al., 2011; Kim et al., 2014; Ciau-Uitz et al., 2016; Kaus et al., 2016; Robertson et al., 2016; Crisan and Dzierzak, 2016). Owing to the high accessibility of model organisms for experimentation and the genetic variability and rarity of human embryonic material, the analysis of human haematopoietic development has lagged behind these well-established experimental models and for a long time was limited to immunohistological and in vitro studies. More recently, the development of highly efficient xenograft mouse models has allowed human HSCs to be tested functionally (Shultz et al., 2012). Despite the discovery of major commonalities between mouse and human HSC development, significant differences remain that need to be characterised in order to understand human HSC development.

Attempts to recapitulate haematopoiesis development via directed differentiation of pluripotent stem cells (PSCs) – both ESCs and induced PSCs (iPSCs) – have provided a valuable source of information, although current protocols appear to replicate extra-embryonic, yolk sac-like, rather than intra-embryonic, adult-like haematopoiesis. Indeed, it may be for this reason that efforts to generate bona fide HSCs from PSCs in vitro have not been successful, undermining expectations for the generation of HSCs for clinical needs. This, in turn, has stimulated the quest for alternative methodologies, such as direct cell reprogramming (Riddell et al., 2014; Sandler et al., 2014). In this Review, we first discuss in vivo human haematopoietic development, focussing on the different anatomical sites where haematopoiesis takes place, as well as the molecular and functional characterisation of human haematopoietic stem and progenitor cells as they emerge in a spatiotemporal manner. Based on what is known regarding the development of the human haematopoietic system, we consider the
extent to which attempts to generate HSCs in vitro recapitulate in vivo development, and discuss the prospect of generating true fully functional HSCs in the dish — cells that would be suitable for transplantations in clinical settings. The generation of HSCs in vitro in the absence of genetic modifications would represent the strongest indication that the key basic mechanisms that underpin HSC development in the human embryo have become largely understood.

The anatomy of haematopoietic development

In the mouse, developmental stages can be distinguished by morphological changes occurring over less than a day. Human embryo development, however, takes a significantly longer time. Stages of human development, known as Carnegie stages (CS), are defined by external morphological characteristics and normally cover several days each (O’Rahilly and Muller, 1987). Staging according to the Carnegie classification is more accurate than using gestational age or crown-rump length, which is important when analysing human HSC development (Ivanovs et al., 2013). For consistency, where days post-conception (dpc) were used in original publications, corresponding CS are presented in parallel in this Review. Despite differences in developmental timescales, substantial parallels between mouse and human haematopoietic development can readily be observed (Fig. 1).

The yolk sac

The first wave of human haematopoiesis originates in the yolk sac, a structure that develops differently in human and mouse. Whereas in the mouse conceptus the yolk sac surrounds the body of the embryo, in the human it develops in front of the embryo body in a balloon-like form (Fig. 2A). Tight groups of ‘mesenchymal’ cells, of mesoderm origin, that are adjacent to the endoderm, differentiate into haematopoietic cells surrounded by endothelial cells, which later remodel to form the yolk sac vascular plexus. Large primitive nucleated erythrocytes represent the major haematopoietic output from the yolk sac at CS 7-8 (16-18.5 dpc), with the occasional presence of primitive macrophages and megakaryocytes (Bloom and Bartelmez, 1940; Fukuda, 1973; Luckett, 1978). By CS 10 (21-22 dpc), the first primitive erythroblasts can be observed inside the cardiac cavity, marking the onset of blood circulation, followed by appearance of the first CD45+ (PTPRC+) cells, (Tavian et al., 1999), similar to analyses in the mouse (Ghiur et al., 2008). Although a yolk sac EMP wave similar to the mouse has not been formally characterised during human development, early studies indicate that this wave may emerge in the human yolk sac during CS 13-15 (28-35 dpc) (Migliaccio et al., 1986).

The AGM region

The appearance of IAHCs on the ventral wall of the human dorsal aorta in the AGM region, reported early last century (Minot, 1912), marks the onset of the intra-embryonic, permanent adult haematopoietic wave in the vertebrate embryo. Spatiotemporal analysis revealed that IAHCs emerge at CS 13 (27 dpc), exclusively at the floor of the embryonic dorsal aorta and disappear by CS 17 (39-42 dpc) (Tavian et al., 1996, 1999). By contrast, in the mouse some IAHCs are also observed in the aortic roof (Taoudi and Medvinsky, 2007; Taylor et al., 2010; Yokomizo and Dzierzak, 2010). In the human embryo, IAHC formation covers the pre-umbilical area of the floor of the dorsal aorta and penetrates the entry to the vitelline artery (Fig. 2B,C). Early mesoderm and the endothelial compartment in early post-gastrulation embryos is marked by KDR (also known as FLK1) expression (Cortés et al., 1999). The endothelial lining of the human dorsal aorta upregulates CD34, which later marks adult HSCs, from CS 9 (19 dpc) (Tavian et al., 1996, 1999, 2001; Oberlin et al., 2002). By CS 13 (27 dpc), the first CD34+CD45+ cells emerge in the pre-umbilical region of the dorsal aorta, and their number reaches several hundred by CS 15 (33 dpc).

Although it is broadly accepted that IAHCs are formed from mesodermal precursors through endothelial intermediates, in the absence of direct tracking, various scenarios of IAHC formation can be considered (Medvinsky et al., 2011) (see also ‘The endothelial origin of human haematopoiesis’, below). Developing human IAHCs sometimes appear to penetrate through the aortic endothelial lining (Tavian et al., 1999). This could reflect a number of scenarios, including ‘excessive’ endothelial-to-haematopoietic transition (EHT) activity in situ, transendothelial cell migration towards the aortic lumen into the bloodstream (Bertrand et al., 2005; Rybsiov et al., 2011), or migration of EHT-undergoing cells in the opposite direction, towards the venous system, as described in zebrafish (Bertrand et al., 2010; Kisa and Herbomel, 2010). Expression of angiotensin converting enzyme (ACE), which labels human foetal liver HSCs (Jukubaitis et al., 2008), identifies scattered ACE+CD34+ cells beneath the dorsal aorta (Sinka et al., 2012). These have been postulated to represent IAHC precursors, which upregulate CD34 and after integration into the CD34+ endothelial lining form ACE+CD34+ IAHCs (Sinka et al., 2012). Identification of the exact migration pathways during HSC specification in the mammalian AGM region remains one of the least explored issues in the field.

The liver

The liver rudiment emerges as a diverticulum from the floor of the embryonic gut at early CS 10 (21 dpc). From late CS 10 (22 dpc), the liver rudiment contains primitive yolk sac-derived erythrocytes and CD45- cells, likely of monocytic/macrophage lineage. From CS 13 (27-29 dpc), the liver is seeded by growing numbers of CD34+CD45- cells (Tavian et al., 1999), which likely represent yolk sac-derived cells similar to mouse EMPs (McGrath et al., 2015). Although direct experimental evidence is lacking, these might be the cells that progressively replace primitive erythroblasts in the human embryo bloodstream during liver colonisation from CS 13-14 (30-33 dpc) onwards (Migliaccio et al., 1986). Concurrent emergence of IAHCs and CD34+CD45- cells in the liver at CS 13 suggests that some of these cells also represent colonisation of the foetal liver by AGM-derived cells, despite a delay in the appearance of HSCs in the liver (see ‘Spatiotemporal HSC development in the early human embryo’, below). The liver remains an important niche for haematopoietic differentiation and HSC expansion until birth. Although some analysis of the developing liver as the haematopoietic niche was conducted on the mouse, this issue in human remains a largely unexplored territory.

The placenta

In the mouse placenta, definitive HSCs are detected at the same developmental age as in the AGM region. For this reason, the placenta is considered to be one of the HSC sources in the mouse embryo (Gekas et al., 2005; Ottersbach and Dzierzak, 2005). At least from week 5-6 of development, the human placenta contains high numbers of CD34+ cells, represented by immature CD34+CD45+ cells that lack CD38 and contain high colony-forming units-culture (CFU-C) numbers, and CD34+CD45- cells committed to erythroid and myeloid differentiation containing fewer CFU-Cs (Barcena et al., 2009). At this stage, the placenta is also a site of extensive erythroid maturation: placental villi are
enriched for primitive erythrocytes that express embryonic ζ-globin in association with macrophages, which may facilitate their enucleation (Van Handel et al., 2010). Although CD34+ cells appear in the human placenta as early as at week 5 of gestation, true HSCs can be detected there only after week 9 of gestation or even later, as determined by xenotransplantation into immunocompromised mice (Robin et al., 2009; Muench et al., 2017). Although the CD34+CD45+ population that contains HSCs is embedded within a vimentin-positive stromal environment or rarely in close association with blood vessels (Muench et al., 2017), localisation of the exact HSC fraction and their regulatory microenvironment in the human placenta is currently unavailable.

**The bone marrow**
The development of the definitive bone marrow niche is closely linked to the invasion of cartilaginous bone by blood vessels and
bone ossification. Vascular invasion facilitates seeding of bone marrow with haematopoietic progenitors and HSCs. Bone marrow formation arbitrarily marks the end of the human embryonic period (CS 23; 56 dpc) (O’Rahilly and Muller, 1987). The initial CD34−CD45+ haematopoietic cells that invade the cartilaginous bone include mostly CD68+ monocytes/macrophages possibly participating in chondrolysis. This is followed by colonisation with CD34+CD45+ progenitors and HSCs (Charbord et al., 1996), although the time of onset of HSC activity in the human bone marrow remains unknown.

**Understanding haematopoietic development through in vitro approaches**

A semi-solid *in vitro* (methylcellulose) assay allows an assessment of the presence of CD34+ CFU-Cs in the early human embryo (Tavian et al., 1999). However, this assay does not reveal the potential of embryonic tissues to generate CFU-Cs or HSCs. In the mouse, this issue was addressed by pre-culturing embryonic tissues as explants to allow their further development *ex vivo* before assessing their haematopoietic potential (Cumano et al., 1996; Medvinsky and Dzierzak, 1996). This approach was adopted for human studies and revealed that the human yolk sac and para-aortic splanchnopleura (P-Sp; the precursor of the AGM region) isolated prior to the onset of circulation at CS 10 (21 dpc; to exclude cross-seeding) possess differing haematopoietic potentials (Tavian et al., 2001). The yolk sac could generate only myeloid and natural killer (NK) cells, whereas the P-Sp showed a broader spectrum of haematopoietic differentiation including lymphoid B- and T-cell lineages starting from CS 11 (24 dpc), suggesting that it could give rise to multipotent haematopoietic progenitors *in vivo* (Tavian et al., 2001). A subsequent study revealed that only the dorsal aorta and not the yolk sac contained so-called long-term cobblestone area-forming cells, which in bone marrow correlate with the presence of HSCs (Oberlin et al., 2002). Although these results are consistent with analyses in the mouse embryo, the human lympho-myeloid precursors revealed in culture may not be equal to true HSCs. Identification of the very few HSCs in the early human embryo required development of a robust *in vivo* long-term repopulation assay, which became possible only with the later advent of highly receptive xenograft mouse models.

**The endothelial origin of human haematopoiesis**

Various non-human experimental models have shown that during development haematopoietic cells emerge from the embryonic endothelium (Jaffredo et al., 2000; de Bruijn et al., 2002; Bertrand et al., 2010; Kissa and Herbomel, 2010; Chen et al., 2011). KDR+CD34− cells observed in the early human embryo mark mesodermal precursors that are likely to develop into endothelial precursors (Cumano et al., 1996; Medvinsky and Dzierzak, 1996). This approach was adopted for human studies and revealed that the human yolk sac and para-aortic splanchnopleura (P-Sp; the precursor of the AGM region) isolated prior to the onset of circulation at CS 10 (21 dpc; to exclude cross-seeding) possess differing haematopoietic potentials (Tavian et al., 2001). The yolk sac could generate only myeloid and natural killer (NK) cells, whereas the P-Sp showed a broader spectrum of haematopoietic differentiation including lymphoid B- and T-cell lineages starting from CS 11 (24 dpc), suggesting that it could give rise to multipotent haematopoietic progenitors *in vivo* (Tavian et al., 2001). A subsequent study revealed that only the dorsal aorta and not the yolk sac contained so-called long-term cobblestone area-forming cells, which in bone marrow correlate with the presence of HSCs (Oberlin et al., 2002). Although these results are consistent with analyses in the mouse embryo, the human lympho-myeloid precursors revealed in culture may not be equal to true HSCs. Identification of the very few HSCs in the early human embryo required development of a robust *in vivo* long-term repopulation assay, which became possible only with the later advent of highly receptive xenograft mouse models.

**Fig. 2. Human embryonic haematopoietic tissues.** (A) The location of the AGM region, yolk sac, liver, vitelline and umbilical arteries and placenta in a CS 15 human embryo. The location of the amnion, eye, and upper and lower limbs are also indicated for reference. Note that HSCs at this stage are localised to the AGM region and will appear in the yolk sac and liver slightly later. (B) Human IAHC development and the stage-specific spatial localisation of CD34+CD45+ cell clusters on the ventral wall of the dorsal aorta [indicated in red; modified from Tavian and Peault (2005)]. (C) VE-cadherin+CD45+RUNX1+ human IAHCs on the ventral wall of the dorsal aorta of a CS 16 human embryo. Whole-mount antibody staining (left) and sagittal confocal section of the boxed area (right). The largest clusters are usually localised close to the entry of the vitelline artery. Scale bars: 1 mm in A; 0.05 mm in C.
endothelial cells, the human AGM-derived CD34+CD45 specification (Vodyanik et al., 2006). Therefore, apart from and maternal cells, which might contaminate samples, especially in (Fig. 2A) have been assayed (Ivanovs et al., 2011). Short tandem represents a more stringent, long-term and robust assessment than transplantation into immunodeficient mice (see Box 1), which Human HSC development has been functionally evaluated through Spatiotemporal HSC development in the early human embryo

Human HSC development has been functionally evaluated through transplantation into immunodeficient mice (see Box 1), which represents a more stringent, long-term and robust assessment than in vitro assays. All haematopoietic tissues known from mouse studies, such as the AGM region, yolk sac, liver, umbilical cord and placenta (Fig. 2A) have been assayed (Ivanovs et al., 2011). Short tandem repeat analysis allowed for the discrimination between embryonic and maternal cells, which might contaminate samples, especially in placenta transplants. Examination of CS 12-17 (25-42 dpc) embryos revealed the presence of HSCs predominantly in the AGM region. Out of 11 experiments in which multipotent HSCs were detected in the AGM region, nine showed the presence of HSCs solely in that location. The first HSCs in the AGM region were detected as early as at CS 14 (30-32 dpc) and persisted until at least CS 17 (39-42 dpc). In the yolk sac, HSCs were detected only after CS 16 (35-38 dpc), approximately 5 days later than in the AGM region. In general, repopulation with yolk sac cells was significantly rarer than with AGM region cells. A robust presence of HSCs in the liver is generally detected later, from week 7-8 and the most potent subset of liver HSCs still express VE-cadherin, reflecting their endothelial origin (Oberlin et al., 2010; Ivanovs et al., 2011).

Although the appearance of HSCs in the mouse AGM region, yolk sac, umbilical cord, liver and placenta occur almost concurrently (Müller et al., 1994; de Bruijn et al., 2000; Kumaravelu et al., 2002; Gekas et al., 2005; Ottersbach and Dzierzak, 2005), their appearance in different locations in human embryo is temporally resolved, owing to a more protracted developmental period. Whereas in the mouse HSCs develop in umbilical cord and placenta prior to liver colonisation (de Bruijn et al., 2000; Gekas et al., 2005; Ottersbach and Dzierzak, 2005), human umbilical cord and placenta at analogous stages (CS 14-17; 30-42 dpc) lack HSCs (Ivanovs et al., 2011). Only from week 9 of development are HSCs reliably detected in the human placenta (Robin et al., 2009), suggesting that the human placenta is a secondary site for HSC development.

The aforementioned studies emphasise the importance of functional in vitro assays for the analysis of HSCs. Although immunohistological and in vitro studies suggested the appearance of HSCs in the CS 13 (27-29 dpc) liver (Tavian et al., 1999), transplantable HSCs appear there only after CS 17 (39-42 dpc) (Ivanovs et al., 2011). Similarly, despite the presence of cells with an HSC phenotype and CFU-Cs, week 5-6 human placenta lack definitive HSCs (Barcena et al., 2009; Robin et al., 2009; Ivanovs et al., 2011).

The exceptional regenerative potential of the first human HSCs

Transplantation of AGM cells after CS 14 (30-32 dpc) into immunodeficient mice results in long-term multilineage haematopoietic engraftment, confirming the presence of definitive HSCs (Ivanovs et al., 2011). As in the mouse, very few – normally only one – HSCs emerge in the human AGM region, as transplantation of cells from one human AGM into several recipients usually results in one haematopoietically reconstituted mouse (Fig. 3). Repopulation kinetics of human AGM region-derived HSCs are significantly slower than their mouse counterparts. By month 3 post-transplantation, haematopoietic chimerism reaches approximately 1% of total leucocyte count in the recipient blood. However, subsequent growth of human haematopoietic cells results in up to 90% chimerism by month 8 in recipient peripheral blood and bone marrow, which is accompanied by extensive generation of large numbers (more than 300) of fully functional daughter HSCs, as determined by re-transplantation into secondary recipients (Fig. 3) (Ivanovs et al., 2011). For comparison, this significant production of functional daughter HSCs exceeds the number required to prevent bone marrow failure in patients after myeloablation (Catlin et al., 2011). Computational analysis of the repopulation dynamics suggested that adult human HSCs replicate every 40 weeks (Catlin et al., 2011). However, amplification of human AGM-derived HSCs from 1 to 300 in the recipient mouse over 6 months requires

**Box 1. The history of the (NOD/SCID/Gamma) NSG mouse line for haematopoietic transplantation assays**

In the late 1980s, introduction of the antiprogestational compound mifepristone for elective termination of early pregnancies (Vervest and Haspels, 1985) made early intact human embryos available for research purposes. At about the same time, severe combined immunodeficiency (SCID) mice, which could accept human haematopoietic grafts, were generated (McCune et al., 1988). These two factors largely defined progress in functional analysis of human HSCs. As SCID mice often immunoreject xenografts, they were backcrossed to non-obese diabetic (NOD) mice, which are deficient in functional T, B and antigen-presenting cells. For a long time, NOD/SCID recipient mice (Shultz et al., 1995) were widely used to study human haematopoiesis in vivo. However, owing to residual NK cell activity, haematopoietic engraftment of human HSCs is attenuated, which especially affected experiments with low HSC numbers. This, together with the short lifespan of these mice, diminishes the utility of the model. To address this limitation, NOD/SCID/Gamma (NSG) mice were engineered by introducing an additional mutation in the IL2 receptor γ chain. These mice, which lack NK cells, live for longer and support high-level human HSC engraftment (Shultz et al., 2005). The NSG mice later permitted transplantation analysis of the first human HSCs emerging in low numbers in the AGM region (Ivanovs et al., 2014, 2011).
replication at least every 2-4 weeks, because transplanted HSCs also generate differentiated blood cells. The proliferation of the first HSCs in situ (in the embryo) may be even higher, given massive expansion of HSCs in the foetal liver.

Achieving similar levels of immunodeficient mouse engraftment with human umbilical cord blood (UCB) HSCs requires considerably higher HSC doses, with 10-20 human UCB HSCs giving approximately 40% repopulation by 8 months post-transplantation (Notta et al., 2011). Re-transplantation of bone marrow (BM) of the primary recipient was transplanted in serial dilutions into secondary recipients. Repopulation of all recipients indicated that the original AGM-derived HSC generated no fewer than 300 daughter HSCs (Ivanovs et al., 2011).

Localisation and phenotype of the first human HSCs

In mice, IAHCs are localised predominantly to the ventral domain of the dorsal aorta, correlating with the presence of functional HSCs (Taoudi and Medvinsky, 2007). Similarly, the ventral domain of the human AGM region is enriched for cells with an HSC immunophenotype and function in transplantation studies, suggesting that human HSCs also localised to IAHCs (Ivanovs et al., 2014). However, although IAHCs disappear by CS 17 (35-38 dpc), HSCs are still detectable at this stage indicating that functional HSCs in the human AGM region do not entirely correlate with the presence of IAHCs (Tavian et al., 1996, 1999). As in the mouse, the number of IAHCs significantly exceeds the low numbers of HSCs in the AGM (Ivanovs et al., 2011), suggesting that many cells in human IAHCs could be nascent HSCs, that cannot yet be detected by direct transplantation, as inferred from in vitro modelling of mouse HSC development (Taoudi et al., 2008; Rybtsov et al., 2011; Zhou et al., 2016). The development of a similar in vitro co-culture approach for recapitulating human HSC development has been unexpectedly challenging.

Xenotransplantation into immunodeficient mice revealed that the earliest human HSCs emerging in the AGM region are CD34\(^+\)VE-cadherin\(^−\)CD45\(^−\), consistent with an endothelial origin (North et al., 2002; Taoudi et al., 2005; Ivanovs et al., 2014). This population comprises 500-1000 cells and is uniformly labelled by other UCB HSC markers (Majeti et al., 2007; Notta et al., 2011), which together localise the first few definitive human HSCs to CD34\(^+\)VE-cadherin\(^+\)KIT\(^+\)THY1\(^+\)endothelin\(^+\)RUNX1\(^−\)CD38\(^−\)CD45RA\(^−\)endoglin\(^+\)RUNX1\(^−\)CD38\(^−\)cells (Ivanovs et al., 2014). Thus, compared with UCB from which HSCs can be isolated with high purity (Notta et al., 2011), identification of definitive HSCs within the AGM population is inefficient, and remains a challenge. Phenotypic and quantitative characterisation of the entire HSC hierarchy emerging in the human AGM region is hampered by the lack of an experimental system to identify immature human HSCs. Dissecting the composition of the AGM-derived cell population that contains HSCs using additional – and currently unknown – markers might provide further insights.

Human foetal liver and UCB HSCs could also be a source of useful information for the identification of immature HSCs in the AGM region (Notta et al., 2011; Prashad et al., 2015). In human, the phenotypic identity of foetal liver HSCs is less defined than in the mouse (Kim et al., 2006), although their main markers, CD34\(^+\)CD38\(^−\), are the same as those of UCB and bone marrow HSCs (Majeti et al., 2007; Doulatabad et al., 2012). GPI-80 (VNN2) marks HSCs within the CD34\(^+\)CD38\(^−\)CD90 (THY1)\(^+\)foetal liver population and could potentially be useful for enrichment of developing HSCs from earlier pre-liver stages. Similarly, it would be interesting to establish whether \(\alpha6\) integrin (also known as CD49f), a marker for UCB HSCs, can be used for this purpose (Notta et al., 2011).

The signalling landscape within the human AGM niche

HSC development is closely associated with the development of the dorsal aorta and these two developmental programmes share common signalling pathways, which have mainly been investigated using model organisms. In various species, KDR/VEGF signalling plays an important role in the formation of the dorsal aorta (Ciau-Uitz et al., 2016). Groups of KDR\(^−\)CD34\(^−\) cells can be detected at CS 10 (21 dpc) in the trunk mesoderm just before the onset of blood circulation (Cortés et al., 1999). These cells could potentially be considered as candidates for the definitive haemangioblasts described in model organisms and in hPSC cultures (Kennedy et al., 2007; Ciau-Uitz et al., 2016; Slukvin, 2016). By CS 14-15 (30-34 dpc), endothelial cells of the dorsal aorta and IAHCs co-express KDR and its ligand VEGF (Marshall et al., 1999) and upregulate transcription factors that are crucially important for haematopoietic development in IAHCs, such as RUNX1 (also known as AML1), c-myb (MYB), SCL (TAL1), GATA2 and GATA3 (Tavian et al., 1996; Labastie et al., 1998; Marshall et al., 1999; Ivanovs et al., 2014).
Ventral polarisation in IAHC and HSC development in model organisms is at least partly driven by spatial asymmetry of molecular signalling in the AGM region (Wilkinson et al., 2009; Souilhhol et al., 2016), suggesting similar mechanisms in human. From CS 13 (28 dpc), BMP4 expression is observed ventrally in a thin subendothelial mesenchymal layer of the dorsal aorta, which transiently expands and subsequently disappears by CS 16 (38 dpc). At this stage, functionally defined HSCs are still present in the AGM region (Ivanovs et al., 2011). BMP4 is a negative regulator of HSC maturation: although it is involved in the formation of the HSC niche, its direct action on HSCs is likely limited by BMP antagonists (Souilhhol et al., 2016). Notably, BMP4 can induce expression of tenascin C and fibronectin (Molloy et al., 2008), two extracellular matrix molecules that are also ventrally polarised in the human AGM region (Marshall et al., 1999). Tenascin C promotes haematoendothelial development of hESCs (Uemishii et al., 2014) and its deficiency impairs mouse bone marrow cultures, which can be rescued by fibronectin (Ohta et al., 1998). TGFβ, another well-known modulator of the extracellular matrix, is expressed in close proximity to IAHCs at the same time as BMP4 expression is downregulated. It has been proposed that TGFβ could participate in rearrangement of the niche and restriction of HSC expansion prior to foetal liver colonisation (Marshall et al., 2000).

Signalling through KIT is important for HSC maturation and its ligand, SCF (KITL in mouse; KITLG in human), is ventrally polarised in the mouse AGM region (Rybtsov et al., 2014; Souilhhol et al., 2016). KIT is expressed in IAHCs and in the first functional HSCs emerging in the human embryo (Labastie et al., 1998; Ivanovs et al., 2014). BMP4 can upregulate KIT in various tissues and, through upregulation of KIT in human aortic endothelium, it may potentially facilitate HSC initiation (Marshall et al., 2007). However, further HSC development may require suppression of BMP4 signalling, as described in the mouse (Souilhhol et al., 2016). FLT3, which marks mouse embryonic HSC precursors (Boyer et al., 2011), and its ligand are also expressed in IAHCs and surrounding endothelium, and may be involved in human HSC maturation (Marshall et al., 1999).

**Adhesion and migration**

HSC development is associated with cell migration and IAHCs are marked by a range of adhesion molecules that potentially play a role in this process. CD34 and its receptor L-selectin (CD62L; SELLP) are involved in cell adhesion processes including lymphocyte-endothelial interactions (Rosen, 2004), and are expressed in human embryonic endothelium, IAHCs and functional HSCs (Tavian et al., 1996; Ivanovs et al., 2014). The first lineage-restricted haematopoietic marker, SPN, appears during hPSC differentiation (Vodyanik et al., 2006) and can mediate adhesion through E-selectin receptor (SELE) (Matsumoto et al., 2005). CD34 and CD43 can cooperatively define adhesive behaviour of haematopoietic cells (Drew et al., 2005). Other surface molecules involved in cell adhesion – ALCAM, VE-cadherin, CD44, CD164 and VCAM1 – are also expressed in IAHCs, endothelial cells of the dorsal aorta and sometimes the mesenchymal cells underneath IAHCs (Cortés et al., 1999; Watt et al., 2000; Marshall, 2006; Ivanovs et al., 2014). Given the role of the TNF/lymphotoxin/NFκB pro-inflammatory pathway in early HSC development described in model organisms (Espin-Palazón et al., 2014; Li et al., 2014) and its impact on VCAM1, L- and E-selectin expression (Suna et al., 2008), it is conceivable that complex and finely regulated adhesion dynamics play important roles in human HSC development. However, it is likely that additional signalling pathways described in other model organisms (Ciau-Uitz et al., 2016) also play a role in human HSC development. Modelling HSC development in vitro using PSCs and other approaches could shed light on this issue.

**Re-capitulation of human haematopoietic development using pluripotent stem cells**

This section examines parallels between the in vitro haematopoietic differentiation of hPSCs and aspects of haematopoietic development in the human embryo. Rather than replicating recent reviews that comprehensively discuss protocols for the differentiation of hPSCs into HSCs or progenitors (Ditadi et al., 2016; Slukvin, 2016; Wahlster and Daley, 2016), we have structured this section to highlight the extent to which in vitro differentiation reflects human haematopoietic development (Fig. 4). We argue that ongoing efforts to generate HSCs from PSCs in vitro are both informed by and add to our understanding of the signalling events that underpin in vivo human haematopoietic development.

**Mesoderm induction and patterning**

In the mammalian embryo, ingestion of cells through the primitive streak during gastrulation generates mesoderm, which colonises the yolk sac (extra-embryonic mesoderm) and intra-embryonic sites including paraxial, intermediate and lateral plate mesoderm (Tam and Behringer, 1997). Whereas extra-embryonic mesoderm gives rise to transient yolk sac haematopoiesis, lateral plate mesoderm adjacent to the paraxial mesoderm (the splanchnopleura) forms bilateral mesodermal strips that converge to the midline where they coalesce to form the dorsal aorta. The growth factors involved in mesoderm specification, dorsal aorta formation and haematopoietic development include members of the FGF, WNT, TGFβ, retinoic acid, Hedgehog and Notch signalling pathways (Lawson et al., 2001; Gering and Patient, 2005; Chanda et al., 2013; Lizama et al., 2015; Ciau-Uitz et al., 2016).

In culture, as in vivo, mesoderm induction and patterning is mediated by the cooperative actions of highly conserved FGF, BMP4, activin and canonical WNT signals (Schier and Shen, 2000; Kimelman, 2006; Nostro et al., 2008; Woll et al., 2008; Wang and Nakayama, 2009; Bernardo et al., 2011; Yu et al., 2011). Mesoderm patterned to a haematopoietic fate is marked by the expression of primitive streak genes, such as the transcription factor genes brachyury (T), MIXL1, FOXF1, as well as cell surface receptors KDR and PDGFRα (Davis et al., 2008; Ditadi et al., 2016; Slukvin, 2016). An early study showed that the most efficient production of haematopoietic cells resulted from the exposure of the emerging mesoderm to activin and BMP4 signalling (Nostro et al., 2008), analogous to the posterior primitive streak from whence the first yolk sac haematopoietic progenitors are derived. Arguably, this might explain why the haematopoietic output also resembled that of the yolk sac and did not include long-term HSCs. It was only later that a series of studies shed light on the role of mesoderm patterning for both the suppression of primitive haematopoiesis and for the support of definitive haematopoietic lineages. Following an initial period of BMP4-based mesoderm induction, brief inhibition of activin (Kennedy et al., 2012), or provision of a WNT agonist (Gertow et al., 2013) were sufficient to inhibit the erythroid-biased primitive programme. Importantly, activin inhibition (Kennedy et al., 2012), WNT stimulation (Sturgeon et al., 2014) or a combination of the two (Ng et al., 2016) from days 2–4 of in vitro differentiation also resulted in a bias towards definitive haematopoietic lineages, as defined by the capacity to generate T lymphocytes (Kennedy et al., 2012; Sturgeon et al., 2014).
The selective expression of HOXA genes in human foetal liver and UCB-derived haematopoietic progenitor and stem cells clearly distinguished them from PSC-derived haematopoietic progenitors generated using conventional mesoderm induction protocols (Dou et al., 2016; Ng et al., 2016). HOX genes mark axial position during development and are first expressed at the primitive streak stage (Deschamps and van Nes, 2005). Therefore, it was relevant that activin inhibition and WNT stimulation during mesoderm patterning, alone and in combination, increased the expression of the CDX genes, leading to sustained expression of their target HOXA genes and more closely mimicking the HOXA expression signature seen in foetal liver and cord blood cells (Ng et al., 2016). These data support the premise that allocation of cells to extra- and intra-embryonic haematopoietic fates is initiated during mesoderm patterning, and that this is reflected in the selective expression of HOXA genes in definitively patterned mesoderm (Fig. 4).

**The generation of yolk sac-like haematopoietic lineages**

The emergence of FGF2- and VEGF-dependent blast colonies from BMP4-induced mesoderm after 2-4 days of differentiation marks the onset of haematopoiesis in human hPSC cultures (Kennedy et al., 2007; Davis et al., 2008; Vodyanik et al., 2010). As well as the surface markers KDR and PDGFRα, blast colony-forming cells (BL-CFCs) also express the apelin receptor (Vodyanik et al., 2010; Choi et al., 2012; Yu et al., 2012), and CD235α (glycoporphin A) (Sturgeon et al., 2014). As in the mouse (Huber et al., 2004), human blast colonies give rise to haematopoietic, endothelial and smooth muscle lineages (Yu et al., 2012). Haematopoietic differentiation of blast colonies is limited to primitive erythrocytes, megakaryocytes and macrophages, and, as in the mouse (Lancrin et al., 2009), probably occurs via a haematogenic endothelial cell precursor.

Further differentiation of hPSCs along a yolk sac-type path can be achieved by co-culture with mouse OP9 stromal cells (Vodyanik et al., 2006), as embryoid bodies (Pick et al., 2007; Nostro et al., 2008) or monolayer cultures (Uenishi et al., 2014) in serum-free media supplemented with growth factors, usually including VEGF, SCF and FGF2. Endothelial and haematopoietic cell surface markers are sequentially acquired, starting with CD34 after 4-6 days, and followed by CD43 at day 6-8, the expression of which first indicates that the endothelial-to-haematopoietic transition is a common mechanism for blood formation both in the embryo and in differentiating hPSCs and is not limited to a particular developmental stage.

The CD34+CD43+ haematopoietic progenitors give rise to a broad range of erythroid and myeloid lineages (Choi et al., 2012; Kennedy et al., 2012; Ng et al., 2016), reminiscent of the broad potential of the yolk sac-derived erythro-myeloid progenitor (EMP) described in the mouse that transiently seeds the foetal liver (Griffiths et al., 2015). Although EMP-type haematopoiesis has not been formally characterised in human embryos, kinetic studies report a rapid transition of clonogenic cells from the yolk sac to the foetal liver at around 5 weeks of gestation (Migliaccio et al., 1986). Given that at this time the human foetal liver does not yet possess repopulating ability (Ivanovs et al., 2011), the cells seeding the early foetal liver are likely to include yolk sac-derived EMPs, and the yolk sac-like erythroid and myeloid cells derived from hPSCs may be their human in vitro equivalent. Whether these EMP-type cells are defined as a second ‘primitive’ or a first ‘definitive’ haematopoietic wave is somewhat moot. For the moment, the descriptor ‘EMP’ may be sufficient. We do point out, however, that in contrast to haematopoietic cells derived from the AGM region, foetal liver or...
other embryonic sites, hPSC-derived EMP-like cells do not express HOXA genes (Dou et al., 2016; Ng et al., 2016).

**The generation of AGM-like haematopoietic lineages from haematogenic endothelium**

As discussed above, the appearance of IAHCs in the AGM region, which arise from a transient haematogenic endothelium in the dorsal aorta, predicts the emergence of the first stem cells that acquire repopulating activity in the vertebrate embryo. In contrast to non-haematogenic vascular cells, hPSC-derived haematogenic endothelium lacks expression of CD73 (NT5E) (Choi et al., 2012; Ditadi et al., 2015), CXCR4 and DLL4 (Ditadi et al., 2015). Also, both mouse (North et al., 1999; Thambryajah et al., 2016) and human (Choi et al., 2012; Ng et al., 2016) haematogenic endothelium are distinguished by their expression of the key transcription factors RUNX1 and GFI1, which are required for the endothelial to haematopoietic transition. Interestingly, hPSCs differentiated with a protocol that induces HOXA gene expression produce SOX17-expressing endothelial cells that bear similarity at the transcriptional level to developing dorsal aorta endothelium, and a SOX17null haematogenic endothelium that efficiently generates haematopoietic cells (Ng et al., 2016). The low level of SOX17 in haematogenic cells is consistent with literature reporting the downregulation of SOX17 during the EHT (Clarke et al., 2013; Nobuhisa et al., 2014). Similarly, single-cell analysis of hPSC-derived cells documented a downregulation of SOX17 at the EHT point (Guibentif et al., 2017).

It has not been possible to identify cell surface markers that distinguish true or ‘definitive’ AGM-derived HSCs from yolk sac-type, non-engrafting haematopoietic progenitors (Ditadi et al., 2016; Slukvin, 2016). For this reason, considerable attention has been paid to the generation of lymphocytes from hPSCs in vitro, arguing that these lineages represent definitive haematopoiesis and may indicate culture conditions that support the development of AGM-like HSCs. Many laboratories have generated T cells from hPSCs, in a Notch-dependent fashion, from endothelial precursors present at 7–9 days of differentiation (Timmermans et al., 2009; Kennedy et al., 2012; Uenishi et al., 2014; Ditadi et al., 2015; Dou et al., 2016; Ng et al., 2016). The dependence of AGM-type, but not yolk sac-type haematopoiesis on Notch signals in the mouse embryo (Kumano et al., 2003; Hadland et al., 2004) has been used to further strengthen the case that the hPSC-derived T cells are progeny of an intra-embryonic haematopoietic progenitor. An important caveat, however, is the realisation that, in the mouse, both T- and B-cell potential is present in the pre-AGM stage yolk sac (Yoshimoto et al., 2011, 2012; McGrath et al., 2015). The dependence of human yolk sac-derived EMP or lymphoid cells on Notch signals is not known.

Unlike the relative ease in which T cells can be derived, B-cell development from hPSCs has been infrequently reported. One laboratory described the generation of cells with a CD19+CD10 (MME)+ pre-B-cell phenotype, some of which matured to surface IgM+ immature B cells (French et al., 2015).

To place this in context, during mouse development there are several waves of B-cell production (Herzenberg and Herzenberg, 1989). The first innate type B cells, called B-1 cells, arise independently from haematogenic endothelia in both the yolk sac and the AGM region, prior to HSCs, and seed the foetal liver (Yoshimoto, 2015). Subsequent waves of HSC-derived B-1 and adaptive B-2 lymphocytes arise later in the foetal liver (Yoshimoto, 2015) and a wave of predominantly HSC-derived B-2 lymphocytes emerges in the bone marrow (Montecino-Rodriguez and Dorshkind, 2012). The identification of B-1 lymphocytes in humans (Rothstein and Quach, 2015), and their observation in the 10–11 week foetal liver (Bueno et al., 2016), argues for conservation of the innate immune system between species. Given that human B-1 cells differ immunophenotypically from B-2 lymphocytes (Rothstein and Quach, 2015), it should therefore be possible to discern whether B cells derived from human hPSCs include B-2 cells, and are thus unequivocally descended from HSCs.

The pattern of globin chain expression by hPSC-derived erythroid cells has also served as an indicator of developmental stage. Erythroid cells differentiated from human PSCs usually express ε- and γ-globins (Chang et al., 2006; Lu et al., 2008; Hatzistavrou et al., 2009; Dias et al., 2011), although prolonged culture in human plasma or on stromal layers has yielded predominantly γ- or γ- and β-globin-expressing cells (Qu et al., 2008; Lapillonne et al., 2010; Yang et al., 2014; Fujita et al., 2016). Importantly, the switch to β-globin synthesis occurs perinatally, and thus is not expected at the AGM or foetal liver stages of haematopoiesis. Rather, focus is on the anticipated downregulation of embryonic ε- and ζ-globins that signifies exit from primitive haematopoiesis. As a further complication, yolk sac EMP-derived erythroid precursors in mice carrying a single copy of the human β-globin locus downregulate ε-globin and express low levels of β-globin when matured in vitro (McGrath et al., 2011). As with the generation of lymphoid cells, these findings indicate that it is not easy to exclude yolk sac EMP haematopoiesis as a source of erythroid cells that appear to derive from foetal liver because they have suppressed embryonic globins.

Finally, although EMPs are traditionally associated with the yolk sac, cells with a similar restricted developmental capacity could conceivably derive from the AGM haematogenic endothelium also. Consistent with this hypothesis, clonogenic progenitors (approximately 90 colonies per embryo) are present in the VE-cadherin+CD45+ population in the human AGM region, the same fraction harbouring less frequent pre-HSCs and HSCs (Ivanovs et al., 2014). These clonogenic cells might also seed the early foetal liver, and might only be distinguishable from yolk sac-derived EMPs by their expression of HOXA genes. Although further experiments will be required as proof, there is support for this idea in a recent report of ε- to γ-globin switching and downregulation of ζ-globin in clonogenic cells derived from HOXA+ cultures (Ng et al., 2016).

Transcriptional comparisons of cells emerging from hPSC-derived HOXA+ haematogenic endothelium and AGM-derived stem progenitors has revealed a high degree of similarity across a broad range of haematopoietic cell surface markers, transcription factors and signalling molecules (Ng et al., 2016). Nevertheless, the hPSC-derived progenitors lack some essential characteristics that enable long-term haematopoietic repopulation. Although mesoderm patterning with modulation of WNT and activin signals increases HOXA expression, the pattern of expression in the hPSC-derived progenitors differs subtly from that observed in human AGM samples, with reduced expression of anterior HOXA genes and elevated levels of the most posterior HOXA genes (Ng et al., 2016).

**Alternative strategies to explore human haematopoietic development**

Teratoma formation as a model for human HSC development

*In vivo* teratoma formation may offer valuable insights into the generation of transplantable HSCs from hPSCs. Teratomas, which contain derivatives of all three germ layers, can be formed after ectopic engraftment of PSCs into a host animal. The mixture of cell
types in these tumours can preserve some developmentally important inductive interactions, leading to the formation of disorganised tissues and organs, including bone marrow (Hentze et al., 2009; Ambabe et al., 2013). Teratomas formed from hPSCs engrafted into immunodeficient mice, in some cases, generate human CD45+ cells that home to bone marrow, circulate in the recipient peripheral blood and colonise lymphoid tissues (Ambabe et al., 2013; Suzuki et al., 2013). Transplantation of total bone marrow or sorted human CD34+CD45+ bone marrow cells from primary into secondary recipients shows donor-derived multilineage peripheral blood chimerism, comparable to UCB repopulation (Ambabe et al., 2013; Suzuki et al., 2013). Teratoma-derived B and T cells appear functionally normal, and erythrocytes express adult haemoglobins, which is not usually achievable during in vitro hPSC differentiation. It is conceivable that in vivo teratoma formation might more faithfully recapitulate some of the processes that occur during embryo development and, in some cases, this could include the specification of an AGM-like region and lead to the production of HSCs. However, an extended period of observation is required to determine whether these HSC-like cells have a truly long-term haematopoietic potential (Ivanovs et al., 2011; Notta et al., 2011).

Direct cell lineage conversion: programming and reprogramming to blood

Although enforced gene expression has been shown to be a powerful tool for altering cell fate for several decades (Davis et al., 1987; Kulesza et al., 1995), the real boost to the generation of blood cells using this approach was precipitated by reprogramming studies in which fibroblasts were reverted back to pluripotency using four transcription factors (Takahashi and Yamanaka, 2006; Pereira et al., 2012). These studies spawned quests for combinations of transcription factors, overexpression of which could generate clinically relevant cell types, including HSCs. Groups chose either ESCs (Kyba et al., 2002; Elcheva et al., 2014), human endothelial cells (Sandler et al., 2014), haematopoietic cells (Doulatov et al., 2013) or fibroblasts (Pereira et al., 2013; Batta et al., 2014) as the starting material. Comparative transcriptomics of HSC and progenitor cell populations often provided lists of candidate factors to overexpress. Although many of the factors used in reprogramming studies have previously been implicated in developmental haematopoiesis, their true ‘haematopoietic potency’ becomes unveiled when they are expressed in a foreign cellular context, such as when reprogramming fibroblasts (Pereira et al., 2013; Batta et al., 2014) or undifferentiated PSCs (Elcheva et al., 2014; Vereide et al., 2014) (see Fig. 5 for details).

Using an hPSC differentiation model, the potency of two pairs of transcription factors, ETV2/GATA2 and GATA2/SCL, has been shown in activating pan-myeloid or erythroid-megakaryocytic programmes, respectively (Elcheva et al., 2014). These results reflect the importance of these factors as haematopoietic fate-determining transcription factors (Elcheva et al., 2014). Although this system failed to generate long-term repopulating cells, it is a powerful tool for the dissection of haematopoietic developmental pathways. In separate studies, a HOXA9/ERG/RORA/SOX4/MYB combination of transcription factors was identified by comparing a human HSC-enriched population with more mature progenitors from UCB (Doulatov et al., 2010; Laurenti et al., 2013). Overexpression of this combination of transcription factors in hPSC-derived haematopoietic progenitors conferred transient in vivo engraftment capacity (Doulatov et al., 2013). The data suggested that HOXA9, ERG, RORA, SOX4 and MYB belong to the same molecular regulatory network; this is supported by the observation that HOXA9 occupies Erg, Sox4 and Myb promoters (Huang et al., 2012). The interest in analysis of HOX genes reflects their strong positive impact on expansion of mouse bone marrow HSCs (Antonchuk et al., 2002; Argiropoulos and Humphries, 2007). Moreover, collinear activation of HOX genes is involved in temporal regionalisation of the mesoderm (limura and Pourquié, 2006), which may underlie functional differences in sequentially emerging haematopoietic waves. For example, HoxB4 overexpression along with adjusted Notch signalling confers multilineage long-term repopulation ability to mouse ESCs (Lu et al., 2016). Downregulation of HOXA5 and HOXA7 in human foetal liver HSCs suppresses their capacity to self-renew and repopulate NSG recipients. Although, forced expression of HOXA5, HOX A7 and HOX A9 alone or in combination in HOX A-negative hPSC-derived progenitors was insufficient to convert the progenitors to HSCs (Dou et al., 2016), a recent study reported a greater degree of success by using a larger cohort of genes (Sugimura et al., 2017). Induced expression of seven transcription factors, notably including three HOX genes, successfully converted hPSC-derived haematogenic endothelium into long-term repopulating HSCs (Sugimura et al., 2017). Finally, expression of four transcription factors (FO SB, GFI1, RUNX1 and SPI1) in human umbilical vein and dermal microvascular endothelial cells, or in adult mouse endothelial cells, were also sufficient to generate immunocompetent HSCs (Sandler et al., 2014; Lis et al., 2017). In all these studies, using transcription factors to manipulate cellular identity has provided insights into the regulation of normal haematopoietic development (Fig. 5). For example, successful reprogramming of fibroblasts into blood allowed, through obtaining a molecular signature, identification of precursors of long-term repopulating cells in the mouse placenta (Pereira et al., 2016).

Future directions and challenges

Understanding the development of the human haematopoietic system is of fundamental biological and clinical importance. Various haematopoietic diseases originate prenatally (Hunger et al., 1998; Greaves, 2005; Hong et al., 2008; Roy et al., 2012; Barrett et al., 2016), and oncogenic mutations in developing haematopoietic progenitors or stem cells can lead to pre-leukaemic changes that may not be manifested if the oncogene is activated in the adult counterparts (Barrett et al., 2016). The molecular mechanisms that regulate development and maintenance of haematopoietic progenitor and stem cells in the embryo differ significantly from those in the adult, and the functional heterogeneity that appears within the emerging HSC population during development further complicates our understanding of HSC regulation (Copley et al., 2012; Crisan et al., 2015; Beaudin et al., 2016). Although the impetus to study human haematopoietic development is clear, poor availability and heterogeneity of human embryonic material makes the analyses challenging. Direct extrapolation of results obtained from model organisms is not always appropriate owing to species-specific differences, although major genetic networks and sites of haematopoiesis are significantly conserved. Furthermore, for ethical and practical reasons, studying the period of early implantation in the human is extremely difficult. Thus, this highly relevant developmental stage remains largely inaccessible and so cannot inform protocols for derivation of HSCs from PSCs. Additionally, until recently, PSC differentiation was biased towards transitory yolk sac haematopoiesis, resulting in the failure of this in vitro model to fully inform us of the mechanisms of HSC development in vivo. Thus, deficiency in our in vivo
knowledge has not yet been compensated for by information from the in vitro system, and vice versa.

In our opinion, two approaches may be taken to address this problem. Notwithstanding their limitations, it is important to continue to use model organisms, which remain virtually unrestricted resources in which to pursue detailed analyses of HSC development. Although not free from ethical and practical considerations, research using non-human primates may also provide valuable information that will help to bridge the evolutionary gap between the mouse and human (D’Souza et al., 2016).

A second approach is the continued use of reprogramming to investigate key determinants of cell fate. Although reprogramming seems like a bold approach to generating cells of interest, it could potentially inform us about events that occur in vivo, provided that it goes beyond describing acquisition of markers and extends to functional characteristics. Research in this direction should aim to identify the best cell sources and factors for producing normal, functional HSCs. To date, the selection of genes for haematopoietic reprogramming has been based on genes known to be essential for haematopoietic development and genes that are differentially expressed in bone marrow/foetal liver/UCB HSCs compared with progenitor cells. However, the genes that drive the formation of the first exceptionally potent HSCs that emerge in the human AGM region (Ivanovs et al., 2011) remain unknown. Identification of these genes in human requires more detailed reconstruction of the developing HSC hierarchy, which is currently limited by lack of an appropriate experimental system in human, but could be facilitated by further in-depth analysis of model organisms. Although the data show a lack of definitive HSCs in human placenta and large extra-embryonic arteries, they should not be excluded from consideration as they could harbour immature HSCs.

The presence of one or more transgenes in the genome of reprogrammed human HSCs represents a clinical risk, and further makes it difficult to assess the entire scale of molecular changes in the reprogrammed cells. Reprogramming methods that do not involve transgene integration, such as non-integrating vectors and RNA, will therefore be highly desirable. Whether this approach, which has proved suitable for iPSC generation, will work for reprogramming into HSCs will depend on the activation of crucial endogenous genes that act as ‘molecular switches’ to produce a phenotypically stable cell type, long after the non-integrating factors are degraded or diluted. This is particularly important for conferring stemness to progenitor cells, and so far it is not clear whether this approach will be successful. Ultimately, the safest HSCs generated in the dish would likely be genetically unaltered cells, produced by exposure to external factors equivalent to the embryonic cues that drive HSC development. Here, the bottleneck seems to be identifying the combinations of soluble and cell-based factors and

In our opinion, two approaches may be taken to address this problem. Notwithstanding their limitations, it is important to continue to use model organisms, which remain virtually unrestricted resources in which to pursue detailed analyses of HSC development. Although not free from ethical and practical considerations, research using non-human primates may also provide valuable information that will help to bridge the evolutionary gap between the mouse and human (D’Souza et al., 2016).

A second approach is the continued use of reprogramming to investigate key determinants of cell fate. Although reprogramming seems like a bold approach to generating cells of interest, it could potentially inform us about events that occur in vivo, provided that it goes beyond describing acquisition of markers and extends to functional characteristics. Research in this direction should aim to identify the best cell sources and factors for producing normal, functional HSCs. To date, the selection of genes for haematopoietic reprogramming has been based on genes known to be essential for haematopoietic development and genes that are differentially expressed in bone marrow/foetal liver/UCB HSCs compared with progenitor cells. However, the genes that drive the formation of the first exceptionally potent HSCs that emerge in the human AGM region (Ivanovs et al., 2011) remain unknown. Identification of these genes in human requires more detailed reconstruction of the developing HSC hierarchy, which is currently limited by lack of an appropriate experimental system in human, but could be facilitated by further in-depth analysis of model organisms. Although the data show a lack of definitive HSCs in human placenta and large extra-embryonic arteries, they should not be excluded from consideration as they could harbour immature HSCs.

The presence of one or more transgenes in the genome of reprogrammed human HSCs represents a clinical risk, and further makes it difficult to assess the entire scale of molecular changes in the reprogrammed cells. Reprogramming methods that do not involve transgene integration, such as non-integrating vectors and RNA, will therefore be highly desirable. Whether this approach, which has proved suitable for iPSC generation, will work for reprogramming into HSCs will depend on the activation of crucial endogenous genes that act as ‘molecular switches’ to produce a phenotypically stable cell type, long after the non-integrating factors are degraded or diluted. This is particularly important for conferring stemness to progenitor cells, and so far it is not clear whether this approach will be successful. Ultimately, the safest HSCs generated in the dish would likely be genetically unaltered cells, produced by exposure to external factors equivalent to the embryonic cues that drive HSC development. Here, the bottleneck seems to be identifying the combinations of soluble and cell-based factors and
inhibitors that will produce bona fide long-term repopulating human HSCs and again, analysis of mouse and other model organisms as well as reprogramming studies may shed light on this problem. The generation of reporter hPSC lines will continue to be a powerful methodology for analysing human HSC development (Rafii et al., 2013; Jung et al., 2016; Ng et al., 2016). Ultimately, the depth of our understanding of human HSC development will be measured by our ability to faithfully recapitulate in vitro the embryonic pathway that leads to HSC formation. This is what will pave the way to informative in vitro modelling of childhood haematological pathologies, and to the robust generation of bona fide HSCs for possible clinical use.

Acknowledgements
The authors would like to thank Prof. I. Slukvin (Madison) for helpful comments on hPSC differentiation; Dr D. Paruzina for providing images for Fig. 2C; Dr S. Gordon-Keaylock for helpful comments and assistance during preparation of this manuscript.

Competing interests
The authors declare no competing or financial interests.

Funding
This work was supported by Bloodwise (formerly LLR) and Medical Research Council (UK). E.G.S. and A.G.E. are Research Fellows of the National Health and Medical Research Council (Australia). Work in the laboratories of E.G.S. and A.G.E. is supported by the National Health and Medical Research Council (Australia), Stem Cells Australia and the Children’s Cancer Foundation.

References

Development (2017) 144, 2323-2337 doi:10.1242/dev.134866


Rabin, C., Bollerot, K., Mendes, S., Haak, E., Criasu, M., Cerisoli, F., Lauw, I., Kaimakis, P., Jorna, R., Vermeulen, M. et al. (2009). Human placenta is a potent


Watt, S. M., Buhring, L. H., Tavazoie, S. M., Bhat, M., Buck, D., Fuchs, A., Doyonnas, R. et al. (2010). Functionally defined CD146 epitopes are expressed on CD34(+) cells throughout ontology but display distinct distribution patterns in adult hematopoietic and nonhematopoietic tissues. *Blood* 93, 3113-3121.


