GATA transcription factors in development and disease
Mathieu Tremblay*, Oraly Sanchez-Ferrás* and Maxime Bouchard‡

ABSTRACT
The GATA family of transcription factors is of crucial importance during embryonic development, playing complex and widespread roles in cell fate decisions and tissue morphogenesis. GATA proteins are essential for the development of tissues derived from all three germ layers, including the skin, brain, gonads, liver, hematopoietic, cardiovascular and urogenital systems. The crucial activity of GATA factors is underscored by the fact that inactivating mutations in most GATA members lead to embryonic lethality in mouse models and are often associated with developmental diseases in humans. In this Primer, we discuss the unique and redundant functions of GATA proteins in tissue morphogenesis, with an emphasis on their regulation of lineage specification and early organogenesis.

KEY WORDS: GATA factors, Gene redundancy, Genetics, Lineage specification, Mouse development, Transcription, Human diseases

Introduction
During vertebrate development, the GATA family of transcription factors plays pleiotropic roles in the early stages of cell differentiation and organ development across a variety of tissues. Decades of research in mouse genetics and biochemistry have revealed the crucial importance of GATA factors in tissue homeostasis and morphogenesis, and more specifically as regulators of progenitor differentiation and lineage specification. In addition, the identification of causal mutations in GATA factors has shed light on a number of human developmental disorders, such as anemia, hypoparathyroidism, deafness and infertility, as well as renal and cardiac defects.

GATA transcription factors are evolutionarily conserved among animals, plants and fungi. Vertebrates possess six paralogs, classified into two subfamilies based on their spatial and temporal expression patterns (Fig. 1A). Although originally divided as hematopoietic (GATA1/2/3) and cardiac (GATA4/5/6) GATA factors, their function and expression patterns extend well beyond these tissues. For example, GATA2 and GATA3 also have important functions in the kidney, skin, prostate, mammary gland and central nervous system (Grote et al., 2008; Kaufman et al., 2003; Kourous-Mehr et al., 2006; Lee et al., 1991; Nardelli et al., 1999). Similarly, GATA4/5/6 are crucially required in organ systems such as the lung, liver and pancreas (Molkentin, 2000). Thus, GATA factors perform key functions in a wide range of developing systems (Table 1), and understanding how each member acts in a cell-type or tissue-specific manner is crucial in understanding their role in embryonic development and their contribution to pathogenesis. Here, we briefly discuss the molecular hallmarks and mode of action of GATA proteins, review the different developmental systems in which GATA members have been implicated and link these findings to human disease.

The GATA family proteins: molecular mechanisms
GATA factors were named after the consensus DNA-binding sequence (A/T)GATA(A/G), which is recognized by the zinc-finger domains common to all family members. Of the two zinc fingers (Fig. 1A), one appears to be consistently required for consensus sequence recognition and binding (Yang and Evans, 1992). The other zinc finger either binds the GATA recognition sequence, stabilizes the interaction with certain sequences or interacts with protein partners (Bates et al., 2008; Crispino et al., 2001; Martin and Orkin, 1990; Trainor et al., 2000; Tsang et al., 1997; Wilkinson-White et al., 2015). In contrast to the highly conserved zinc-finger region, the N- and C-terminal regions, which contain transcription activation modules, diverge considerably among GATA factors (Fig. 1A) (Morrisey et al., 1997). Protein sequence conservation between all six vertebrate members identifies GATA3 as having the highest sequence similarity with both its GATA paralogs and orthologs, suggesting that it may be closest to the ancestral mammalian GATA factor (Fig. 1A,B).

GATA proteins can act as pioneer factors
The classical function of transcription factors is to bind specific DNA sequences in enhancer and promoter regions and modulate their transcriptional output. However, a subclass of transcription factors called ‘pioneer transcription factors’ has the capacity to recognize and bind heterochromatic DNA sequences, and to promote chromatin opening and the recruitment of additional transcriptional regulators. In recent years, a number of examples of pioneer activity have been reported for GATA transcription factors. Acting as a priming factor, GATA4 binds closed chromatin prior to hepatocyte specification and promotes liver-specific gene expression (Bossard and Zaret, 1998; Cirillo et al., 2002). In support of this pioneer activity, GATA4, together with another pioneer factor, FOXA3, can reprogram fibroblasts toward the hepatocyte lineage (Huang et al., 2011). Pioneer activity is also observed in epithelial tissues, where GATA proteins promote the transcription and downstream activity of multiple steroid receptor genes. In the mammary gland, GATA3 and the estrogen receptor (ERα) regulate each other and, along with FOXA1, can nucleate a remodeling complex at heterochromatic enhancer regions of ERα target genes, leading to the opening and epigenetic marking of sites for active transcription (Eckhoute et al., 2007; Kong et al., 2011). Alone, FOXA1 or ERα are not sufficient to fully open the chromatin, supporting a bona fide pioneer activity for GATA3 (Eckhoute et al., 2007; Kong et al., 2011). Similarly, in the prostate, GATA2 enters in a mutual regulatory loop with the androgen receptor (AR) and favors chromatin opening of AR target genes in conjunction with FOXA1 (Böhm et al., 2009; He et al., 2014; Wang et al., 2007; Wu et al., 2014; Xiao et al., 2016). GATA2 also regulates progesterone receptor (Pgr) gene expression in the
uterus, and acts as a co-factor of PGR to activate its target genes (Jeong et al., 2005; Magklara and Smith, 2009; Rubel et al., 2012, 2016; Zhang et al., 2013). The integrated activity of GATA factors with steroid hormone receptors is intriguing and may reflect an ancestral role of GATA factors in hormonal responses. Together, these findings highlight that, through their pioneer activity (Fig. 2), GATA proteins act as primary regulators of lineage decisions and cell fate transitions.

GATA transcriptional complexes
The pioneer activity of GATA factors and their subsequent role as classical transcriptional regulators are achieved largely via their interaction with co-regulators to assemble a transcriptional complex and recruit chromatin remodeling proteins (Katsumoto et al., 2004; Lowry and Mackay, 2006; Takemoto et al., 2002; Zhou and Ouyang, 2003). GATA factors typically regulate gene expression in a combinatorial way, acting together with other tissue-specific transcription factors (Fig. 2). During hematopoietic stem cell differentiation, they can form transcriptional complexes with FOG1/2 (friend of GATA), LMO1/2, SCL (TAL1), E-proteins (E2A, HEB, E2-2), LYL1 and LDB1/2 to perform elaborate regulatory activity (Krivega et al., 2014; Love et al., 2014; Meier et al., 2006; Tripic et al., 2009; Wadman et al., 1997; Wilson et al., 2010). A similar complex involving GATA2, LMO4, SCL and NLI exists in the embryonic central nervous system (CNS) and controls the binary cell fate choice between GABAergic (V2b) and glutamatergic (V2a) interneuron development (Joshi et al., 2009). Other molecular interactions between GATA factors and ETO2, RUNX1, ERG or FLI-1 have also been described (Goardon et al., 2006; Meier et al., 2006; Schuh et al., 2005; Tripic et al., 2009; Wilson et al., 2010). These multiprotein complexes can function both as activators or repressors of target genes (Wadman et al., 1997).
<table>
<thead>
<tr>
<th>Gene</th>
<th>Tissue</th>
<th>Cells targeted</th>
<th>Mouse models</th>
<th>Associated diseases in humans</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gata1</td>
<td>Blood</td>
<td>Megakaryocytes</td>
<td>Pf4Cre, TxCre and hemizygote Gata1AHS/Y</td>
<td>Transient myeloproliferative disorder and acute megakaryoblastic leukemia, inherited macrocytic anemia and neutropenia, thrombocytopenia with α-talassemia, macrothrombocytopenia with severe anemia, anemia, and congenital erythropoietic porphyria</td>
<td>Shivdasani et al., 1997; Gutierrez et al., 2008; Meinders et al., 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Erythrocytes</td>
<td>HbbCre (β-globin), MxCre, TxCre, Rosa26Cre and Gata1DNT rescued germline knockout</td>
<td></td>
<td>Gutierrez et al., 2004; Gutierrez et al., 2008; Mancini et al., 2012; Kaneko et al., 2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mast cells</td>
<td>Rosa26Cre and CD11cCre</td>
<td></td>
<td>Ohneda et al., 2014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dendritic cells</td>
<td>Rosa26Cre and CD11cCre</td>
<td></td>
<td>Gutierrez et al., 2007; Scheenstra et al., 2016</td>
</tr>
<tr>
<td></td>
<td>Testis</td>
<td>Sertoli cells</td>
<td>DhhCre</td>
<td>Emberger syndrome (lymphedema), myelodysplastic syndrome, chronic myeloid leukemia, acute myeloid leukemia and immunodeficiency 21</td>
<td>Lindeboom et al., 2003; Kaimakis et al., 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hematopoietic stem cells</td>
<td>VecCre</td>
<td></td>
<td>Kitajima et al., 2006; Huang et al., 2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Megakaryocytes</td>
<td>Gata2 overexpression</td>
<td></td>
<td>Li et al., 2015</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mast and basophil cells</td>
<td>Rosa26Cre</td>
<td></td>
<td>Tsai et al., 1994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Erythrocytes</td>
<td>Germline knockout</td>
<td></td>
<td>Onodera et al., 2016; Zhou et al., 1998; Rubel et al., 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dendritic cells</td>
<td>CD11cCre</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gonad</td>
<td>Epithelial cells and progenitors (vagina, uterus, seminal vesicle and testis)</td>
<td>PgrCre and Gata2YAC rescued germline knockout</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kidney</td>
<td>Trophoblasts</td>
<td>UBCCre</td>
<td></td>
<td>Home et al., 2017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nephric duct cells and renal tubular cells</td>
<td>Pax8GATA-TELCre and Gata2YAC rescued germline knockout</td>
<td></td>
<td>Zhou et al., 1998; Yu et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Prostate</td>
<td>Epithelial cells</td>
<td>PrbCre (Probasin)</td>
<td></td>
<td>Xiao et al., 2016</td>
</tr>
<tr>
<td></td>
<td>Pilutary and thyroid gland</td>
<td>Gonadotropes</td>
<td>Gata2Cre (aGSU), PgrCre and Pit1-Gata2tg</td>
<td></td>
<td>Dassen et al., 1999; Charles et al., 2006; Rubel et al., 2016; Kala et al., 2009; Haugas et al., 2016; Jager et al., 2016; Lahtii et al., 2016</td>
</tr>
<tr>
<td></td>
<td>Nervous system</td>
<td>GABAergic neurons</td>
<td>En1Cre</td>
<td>Hypothyroidism</td>
<td>Craven et al., 2004; Kala et al., 2009; Zhou et al., 2000</td>
</tr>
<tr>
<td></td>
<td>Ear</td>
<td>Serotonergic neurons</td>
<td>Germline knockout and En1Cre</td>
<td></td>
<td>Haugas et al., 2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V2 interneurons</td>
<td>Germline knockout</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Endothelium</td>
<td>Embryonic vasculature</td>
<td>Gata2-VECre (vascular enhancer)</td>
<td></td>
<td>Lim et al., 2012</td>
</tr>
<tr>
<td></td>
<td>Bone</td>
<td>Lympathic vasculature</td>
<td>Prox1Cre</td>
<td></td>
<td>Kazenwadel et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Adipose tissue</td>
<td>Osteoblasts</td>
<td>Prx1Cre and osteocalcinCre</td>
<td></td>
<td>Li et al., 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mensenchymal stem cells and adipocytes</td>
<td>Prx1Cre and adiponectinCre</td>
<td></td>
<td>Li et al., 2016</td>
</tr>
<tr>
<td>Gata2</td>
<td>Blood</td>
<td>T cells (lymphoid)</td>
<td>CD4Cre, RorcCre, Tnfrsf4Cre, LckCre and OX40Cre</td>
<td>T cell acute lymphoid leukemia</td>
<td>Pai et al., 2003; Yagi et al., 2010; Pai et al., 2004; Skapenko et al., 2004; Zhu et al., 2004; Kanhere et al., 2012; Zhong et al., 2016; Yu et al., 2015; Wohlfert et al., 2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Treg cells (lymphoid)</td>
<td>Foxp3Cre</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>Natural killer cells</td>
<td>NKP46Cre</td>
<td>DiGeorge syndrome</td>
<td>Raid et al., 2009</td>
</tr>
<tr>
<td></td>
<td>Placenta and uterus</td>
<td>Cardiomyocytes</td>
<td>Germline knockout</td>
<td></td>
<td>Grote et al., 2008; Home et al., 2017</td>
</tr>
<tr>
<td></td>
<td>Kidney</td>
<td>Nephric duct cells</td>
<td>Germline knockout and HoxB7Cre</td>
<td>HDR syndrome (renal defects)</td>
<td>Grote et al., 2006; Grote et al., 2008; Xiao et al., 2016; Nguyen et al., 2013; Shafer et al., 2017</td>
</tr>
<tr>
<td></td>
<td>Prostate</td>
<td>Prostate epithelial cells</td>
<td>PrbCre and Nkx3.1Cre</td>
<td></td>
<td>Kourous-Mehr et al., 2006; Asselin-Labat et al., 2007; Oliver et al., 2012</td>
</tr>
<tr>
<td></td>
<td>Breast</td>
<td>Mammary epithelial cells</td>
<td>K14Cre, BLG/GBL and MMTVCre</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adrenal and parathyroid gland</td>
<td>Parathyroid cells</td>
<td>Germline knockout</td>
<td>HDR syndrome (hypoparathyroidism)</td>
<td>Grorionieva et al., 2010</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Gene</th>
<th>Tissue</th>
<th>Cells targeted</th>
<th>Mouse models</th>
<th>Associated diseases in humans</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin and hair follicles</td>
<td>Epidermal cells, and hair follicle cells and progenitors</td>
<td>K14Cre and germline knockout</td>
<td></td>
<td></td>
<td>Kaufman et al., 2003; de Guzman Strong et al., 2006; Kurek et al., 2007</td>
</tr>
<tr>
<td>Ear</td>
<td>Neurosensory cells (inner ear)</td>
<td>Foxg1Cre</td>
<td></td>
<td></td>
<td>Duncan et al., 2011; Haugas et al., 2012; Duncan and Fritzsch, 2013</td>
</tr>
<tr>
<td></td>
<td>Spiral ganglion neurons (inner ear)</td>
<td>Bhlhb5Cre, Bhlhe22Cre and Pax2Cre</td>
<td></td>
<td></td>
<td>Appler et al., 2013; Yu et al., 2013; Duncan and Fritzsch, 2013; Luo et al., 2013</td>
</tr>
<tr>
<td>Lens</td>
<td>Noradrenergic cells</td>
<td>DbhCre</td>
<td></td>
<td></td>
<td>Tsarovina et al., 2010</td>
</tr>
<tr>
<td></td>
<td>Lens fiber cells</td>
<td>Dbh-Gata3Cre-rescued germline knockout</td>
<td></td>
<td></td>
<td>Maeda et al., 2009</td>
</tr>
<tr>
<td>Nervous system</td>
<td>Sympathetic neurons and adrenal medullary chromaffin cells</td>
<td>Dbh-Gata3Cre-rescued germline knockout</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neurons</td>
<td>Germline knockout</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adipose tissue</td>
<td>Serotonergic neurons</td>
<td>PerCre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gata4</td>
<td>Heart</td>
<td>Cardiomyocytes</td>
<td>Germline knockout, hypomorph, knock-in Gata4^{V217G/V217G}, Nlx2.5Cre, Mef2c-AHFCre, Is11Cre, aMHCCre, bMHCCre and Tnnt2Cre (cTNT)</td>
<td>Congenital heart disease</td>
<td>Kuo et al., 1997; Molkentin et al., 1997; Crispino et al., 2001; Pu et al., 2004; Zeisberg et al., 2005; Rojas et al., 2008; Ma et al., 2008; Oka et al., 2006; Zhao et al., 2008; van Berlo et al., 2010, 2013; He et al., 2014; Malek Mohammadi et al., 2017</td>
</tr>
<tr>
<td></td>
<td>Adipocytes</td>
<td>Knockout embryonic stem cells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gonad</td>
<td>Epithelial cells (genital ridge)</td>
<td>Wnt1Cre, Wt1Cre and Osr1Cre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sertoli cells (testis)</td>
<td>Knock-in, Gata4^{V217G/V217G}, AtnrCre and Nra1Cre (Sf1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Granulosa cells (ovary)</td>
<td>Amhr2Cre, Wt1Cre, Cyp19Cre, Nra1Cre (Sf1) and PRCre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adrenal gland</td>
<td>Steroidogenic cells</td>
<td>Nr5a1Cre (Sf1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intestine and stomach</td>
<td>Epithelial cells</td>
<td>VillinCre and knock-in Rosa26Gata4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Midgut endoderm and intestine cells</td>
<td>ShhCre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pancreas</td>
<td>Foregut endoderm and pancreatic β cells</td>
<td>Foxa3Cre and Pdx1Cre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>Mesenchymal cells</td>
<td>Knockout embryonic stem cells chimera, germline heterozygote and knock-in Gata4^{V238G/V238G}</td>
<td>Pancreatic agenesis and diabetes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Liver sinusoidal endothelial cells</td>
<td>Stab2Cre, Lyve1Cre and AlbCre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liver</td>
<td>Hepatic mesenchymal cells</td>
<td>Gata4Cre (CR2 enhancer)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone</td>
<td>Neural crest cells, craniofacial cells and dental papilla mesenchymal cells</td>
<td>Wnt5Cre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Osteoblasts</td>
<td>Col1a1Cre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gata5</td>
<td>Heart</td>
<td>Endocardial cells and cardiomyocytes</td>
<td>Germline knockout and Tie2Cre</td>
<td>Congenital heart disease</td>
<td>Laforest and Nemer, 2011; Messaoudi et al., 2015</td>
</tr>
</tbody>
</table>

Continued
Table 1. Continued

<table>
<thead>
<tr>
<th>Genes</th>
<th>Tissues</th>
<th>Cells targeted</th>
<th>Mouse models</th>
<th>Associated diseases in humans</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gonad</td>
<td>Blood</td>
<td>Macrophages</td>
<td>LyZ2Cre and LysMCre</td>
<td></td>
<td>Gauthier et al., 2014; Okabe and Medzhitov, 2014</td>
</tr>
<tr>
<td>Heart</td>
<td>Cardiomyocytes</td>
<td></td>
<td>SM22aCre and Wnt11Cre</td>
<td>Congenital heart disease</td>
<td>Zhao et al., 2005; Zhao et al., 2008; van Berlo et al., 2013</td>
</tr>
<tr>
<td>Limb</td>
<td>Steroidogenic cells</td>
<td></td>
<td>Nr5a1Cre</td>
<td></td>
<td>Tevosian et al., 2015; Padua et al., 2015</td>
</tr>
<tr>
<td>Intestine and colon</td>
<td>Epithelial cells (intestine and colon)</td>
<td>VillinCre</td>
<td></td>
<td></td>
<td>Sodhi et al., 2006; Beuling et al., 2012</td>
</tr>
<tr>
<td>Pancreas</td>
<td>β Cells, hepatocyte progenitors and foregut endoderm (pancreas)</td>
<td>Pdx1Cre, Foxa3Cre and Ptf1aCre</td>
<td>Pancreatic agenesis and diabetes</td>
<td></td>
<td>Xuan et al., 2012; Carrasco et al., 2012; Martellini et al., 2013; Villamayor et al., 2018</td>
</tr>
<tr>
<td>Lung</td>
<td>Epithelial cells and airway stem cells</td>
<td>SftpcCre and knock-in Gata4V238G/V238G</td>
<td></td>
<td></td>
<td>Ackerman et al., 2007; Zhang et al., 2008; Tian et al., 2011; Zhao et al., 2005</td>
</tr>
</tbody>
</table>

Studies using mouse models have helped unravel the function of GATA factors in different tissues. Shown are the genetic mouse models generated [germline knockout, knock-in, transgenic overexpression (tg) or Cre-mediated conditional inactivation] and the human diseases found to harbor mutations in GATA factors. Owing to space limitations, not all models are described in the text.

complexes, FOG1/2 proteins are key to recruiting epigenetic regulators, including CtBP and NuRD, to induce locus-wide histone demethylation and deacetylation, respectively (Fox et al., 1999; Hong et al., 2005; Rodriguez et al., 2005). The histone acetyl transferase p300 also acts as a positive mediator of different GATA-dependent transcriptional programs in several systems, and is modulated by the polycomb complex (PRC) (Blobel et al., 1998; Dai and Markham, 2001; Flodby et al., 2017; He et al., 2012; Hosokawa et al., 2013a; Kakita et al., 1999; Lamonica et al., 2011; Wada et al., 2000; Yamashita et al., 2002). Additional post-translational modifications, such as phosphorylation and SUMOylation, fine-tune GATA protein activity (Chun et al., 2003; Collavin et al., 2004; Partington and Patient, 1999; Towatari et al., 1995).

As part of their function in transcriptional regulation, GATA factors can contribute to 3-dimensional chromatin reorganization. Together with FOG, LDB1, MED1 and BRG1, GATA factors can promote chromatin looping by bringing together distant enhancers and promoter elements (Fig. 2) (Kim et al., 2009; Song et al., 2007; Stumpf et al., 2006). A direct role in bridging distant regulatory elements has been demonstrated for GATA1 at the β-globin locus, for GATA2 at the Kitl locus and for GATA3 at Th2 cytokine loci (Jing et al., 2008; Spilianakis and Flavell, 2004; Vakoc et al., 2005). Overall, the great variety of GATA-interacting protein partners is indicative of highly regulated and context-specific functions.

The GATA switch
The differential recruitment of GATA factors and their respective regulatory complexes can also be achieved by changes in their expression pattern. They can influence each other’s expression such that they function consecutively during lineage commitment. This sequential activity of GATA factors on their target genes is referred to as the ‘GATA switch’ (Fig. 2). This switch happens for multiple shared target genes leading to distinct transcriptional outputs (Bresnick et al., 2010; Huang et al., 2016; Im et al., 2005). For example, the displacement of GATA2 by GATA1 upon erythroid differentiation is tightly coupled to the repression of Gata2 transcription through recruitment of EKLF and the SWI/SNF chromatin remodeling complex. This causes removal of the histone acetylase CBP and a change in chromatin looping conformation (Grass et al., 2006; Martowicz et al., 2005). The switch from GATA2 to GATA1 also affects the chromatin landscape of differentiating erythrocytes through a shift in histone methyltransferase EZH2 to EZH1 (Xu et al., 2015). A similar switch has been identified in the placenta between GATA3 and GATA2, where the Gata2 locus is repressed by GATA3 in progenitors, and is expressed only as differentiation occurs (Ray et al., 2009). A GATA switch also occurs at the Wdr77 locus during both erythroid maturation and lung development, leading to proliferation inhibition and terminal differentiation in both cell types (Yu et al., 2016).

Together, the identification of GATA co-factors and their associated complexes reveals a sophisticated transcriptional activity for GATA factors that extends from pioneer factors to fine transcriptional activators and repressors, depending on the associated molecular complexes (Fig. 2). The transcriptional response elicited by GATA proteins is intimately associated with cellular context and leads to a specific molecular output for each tissue.
GATA factors in differentiation and organogenesis

The crucial role of GATA factors as transcriptional regulators was first identified in hematopoietic and cardiac development. Since then, germline and conditional gene inactivation models in mice have been used as powerful tools to decipher the diverse functions of GATA factors in a variety of developmental systems. This effort has led to a better understanding of the activity of GATA proteins in embryo patterning and organogenesis and, later, was informative for understanding human disorders through the identification of disease-causing mutations in GATA genes (Table 1). The following sections detail the developmental systems in which GATA factors play a prominent role, beginning with the hematopoietic and cardiac systems – two systems in which GATA factors have been most extensively studied – and then briefly discussing other tissues.

The hematopoietic system

As previously alluded to, the first and best-documented role for GATA transcription factors is in hematopoietic system development, which is initiated in the embryo but persists during postnatal and adult stages. Blood cell differentiation is divided into primitive and definitive hematopoiesis (Ivanovs et al., 2017). Primitive hematopoiesis takes place in the embryonic yolk sac from embryonic day (E) 7.25 to E9.0 and produces only erythrocytes and macrophages. In contrast, definitive hematopoiesis, which progresses from the aorta-gonad-mesonephros (AGM) region to the fetal liver between E10.5 and E12.5, and to the bone marrow and spleen around birth, is dependent on the hierarchical differentiation of multipotent hematopoietic stem cells (HSCs). The differentiation of HSCs generates lineage-committed progenitor cells that give rise to all blood cell lineages. These progenitors differentiate into either myeloid cells (monocytes, macrophages, granulocytes, erythrocytes, megakaryocytes/platelets and dendritic cells) or lymphoid lineages (T cells, B cells, NK cells and innate lymphoid cells). A number of studies have revealed that this occurs via the widespread activity of GATA1/2/3 (Fig. 3).

In the mouse, GATA1 and GATA2 regulate both primitive and definitive erythropoiesis. Accordingly, knockouts for either gene die around E10-E11 from severe hematopoietic defects (Fujisawa et al., 1996; Ling et al., 2004; Pepvy et al., 1995; Tsai et al., 1994; Tsai and Orkin, 1997). During primitive hematopoiesis, Gata1 and Gata2 act redundantly and only double-knockout embryos show a complete block in erythropoiesis (Fujisawa et al., 2004). Upon activation of definitive hematopoiesis, Gata1 and Gata2 are primarily required for stem/progenitor cell maintenance and act as commitment factors for the different lineages (Fig. 3). The crucial importance of Gata2 in stem/progenitor cells was shown by gain- and loss-of-function experiments, which both result in a reduction in HSC numbers and a block of their differentiation (Heyworth et al., 1999; Persons et al., 1999; Tipping et al., 2009; Tsai and Orkin, 1997). Gata3 also plays a role in the long-term self-renewal of HSCs through the control of cell cycle entry (Frelin et al., 2013; Ku et al., 2012).

At later stages of hematopoietic development, as lineage specification proceeds from progenitor cells, Gata1 acts as a primary regulator of erythropoietic cell differentiation (Weiss and Orkin, 1995). This discovery, made in the mouse, was later supported by the presence of inactivating Gata1 mutations in dyserythropoietic anemia in humans (Nichols et al., 2000). The mutually antagonistic function between Gata1 and the transcription factor Pu.1 forms one of the earliest binary lineage decisions, specifying HSCs into erythro-megakaryocytic or myelo-lymphoid lineages, respectively (Hoppe et al., 2016; Laiosa et al., 2006). At the molecular level, Pu.1 displaces Gata1 from DNA, promoting myeloid differentiation, whereas Gata1 inhibits the interaction between Pu.1 and its co-activator JUN (Zhang et al., 2000), thereby allowing erythroid cell specification. A role for Gata1 was additionally found in the differentiation of megakaryocytes, eosinophils and mast cells (Gutierrez et al., 2007; Harigae et al., 1998; Hirawase et al., 2002; Shivdasani et al., 1997; Vyas et al., 1999), as well as in dendritic cell differentiation, survival and function, where it acts primarily at the lineage bifurcation point between monocytes and dendritic cells (Gutierrez et al., 2007).

Gata2 similarly acts at multiple decision points during myeloid lineage differentiation. Gata2 is necessary for the differentiation of myeloid cells, such as dendritic cells, from a common myeloid
progenitor (CMP) and prevents the aberrant expression of a T-cell-associated gene signature (Onodera et al., 2016). During mast cell specification from granulocyte-monocyte progenitors (GMPs), GATA2 must replace GATA1 at key loci to promote lineage commitment (Cantor et al., 2008). Another example of a GATA-regulated lineage decision is found in megakaryocyte-erythrocyte progenitors (MEPs), in which GATA2 promotes megakaryocyte differentiation at the expense of the erythrocyte lineage (Ikonomi et al., 2000) (Fig. 3). In support of a crucial activity in myeloid cell differentiation, GATA2 mutations have been identified in patients with immunodeficiency 21 (IMD21) (Ostergaard et al., 2011) and in lymphedema with myelodysplasia (Hsu et al., 2011).

Hence, together GATA1 and GATA2 activate the transcriptional program of specific cell fates, while repressing regulators of alternative cell fates in the myeloid and erythroid lineages. Interestingly, a similar lineage decision event occurs between the myeloid and lymphoid lineages, where the balance is controlled through mutual repression between the lymphoid regulator GATA3 and the myeloid regulator GATA2 (Nandakumar et al., 2015) (Fig. 3). Following this lymphoid lineage decision, GATA3 is crucially required both for T-cell lineage specification and survival (Pai et al., 2003), and to antagonize B-cell and myeloid developmental fates, which is performed partly by repression of PU.1 (Scripture-Adams et al., 2014). Down the T-cell differentiation path, GATA3 informs the choice between helper and cytotoxic lineages. In the absence of Gata3, specification to the T helper lineage is impaired, whereas overexpression of GATA3 inhibits cytotoxic T cell differentiation (Hernández-Hoyos et al., 2003; Pai et al., 2003). GATA3 is next required for the decision between Th1 and Th2 lineages through a cross-antagonizing mechanism between GATA3 and the transcription factor T-BET (TBX21) (Ouyang et al., 1998; Zhu et al., 2004). This lineage specification event involves a GATA3/CHD4/p300 activation complex at Th2-specific gene loci and a GATA3/CHD4/NURD repression complex at T-bet and Th1-specific gene loci (Hosokawa et al., 2016, 2013b). By repressing the alternative cell fate, GATA3 acts as a master regulator of Th2 differentiation, while T-BET promotes the Th1 lineage (Zhu et al., 2004).

The innate counterparts of cytotoxic T cells are the natural killer (NK) cells. Although Gata3 is dispensable for early development of NK cells, it is required for thymic NK cell maturation (Vosshenrich et al., 2006). Gata3 deletion in NK cells affects both maturation and homing to peripheral organs (Ali et al., 2016). Hence, GATA3 is a crucial regulator of both adaptive and innate lymphoid cells, and exerts its influence by controlling the development and function of T cells, B cells and thymic NK cells.

Overall, the exhaustive analysis of GATA factor activity in the hematopoietic system has been instrumental in understanding the concept of binary lineage decisions, both at the cellular and molecular levels. As we discuss below, binary lineage decisions extend beyond the hematopoietic system and, accordingly, numerous roles for GATA factors in the development of other organs/tissues have been reported.

The cardiovascular system
Heart development begins with the fusion of the cardiac crescent, which generates the heart tube. Through a looping process, the heart tube forms the different cardiac chambers and arterial trunk
The formation of the heart requires the coordination of several processes, including cell specification, cell differentiation and tissue patterning. This leads to tissue compartmentalization and commitment of the cardiovascular progenitors to different cell lineages: cardiomyocytes, endothelial, smooth muscle and epicardial cells (Lescoart et al., 2018).

GATA4/5/6 are central regulators of cardiac development (Clowes et al., 2014; Peterkin et al., 2005). Among them, GATA4 has the earliest unique function in the heart, as its inactivation in the mouse leads to embryonic lethality between E8.5 and E10.5 due to severe defects in ventral body patterning and a lack of pericardial cavity and heart tube (Kuo et al., 1997; Molkentin et al., 1997). Analysis of chimeric embryos in which the visceral endoderm was composed of wild-type cells revealed important activities for Gata4 in proepicardium development, heart tube looping, chamber formation and in the generation of the septum transversum mesenchyme (Pu et al., 2004; Watt et al., 2004; Zeisberg et al., 2005) (Fig. 4B). GATA4 has also been shown to regulate endocardial cushion formation by promoting epithelial-to-mesenchymal transition (EMT) (Rivera-Feliciano, 2006). As seen in other systems, the interaction of GATA factors with co-factors reinforces their activity and specificity. For example, GATA4 interacts with FOG2 to regulate heart muscle vasculature formation (Chlon and Crispino, 2012; Crispino et al., 2001; Tevosian et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000) and further regulates myocardial gene expression by interacting with different transcription factors, such as retinoic acid receptors, NKX2.5 or Nfatc4 and Mef2 (Clabby et al., 2000). This precise balance of co-factor interactions is essential for the confinement of chamber genes and AV canal genes in their respective cardiac region.

GATA5, another member of the ‘cardiac’ group, is expressed in the myocardium as well as in the endocardium and derived endocardial cushions. Conditional inactivation of Gata5 in endocardial cells leads to hypoplastic hearts and partially penetrated bicuspid aortic valve formation (valves with two leaflets rather than three), suggesting an autonomous role for GATA5 in endocardial cushion formation and cardiac valve development (Laforest and Nemer, 2011) (Fig. 4C).

As Gata6-deficient embryos die at an early stage due to defects in extra-embryonic endoderm differentiation (Koutsourakis et al., 1999; Morrissey et al., 1998), conditional inactivation approaches have necessary to decipher its role in heart morphogenesis. The specific inactivation of Gata6 in neural crest-derived vascular smooth muscle cells leads to perinatal lethality and cardiovascular defects, including interrupted aortic arch and persistent truncus arteriosus (Lepore et al., 2006) (Fig. 4D). This suggests that GATA6 contributes to the morphogenetic patterning of the aortic arch and cardiac outflow tracts by controlling the proper migration of cardiac neural crest cells to the developing cardiac outflow tract. On the other hand, GATA6 in cardiac progenitors is uniquely required for ventricular septal morphogenesis (Tian et al., 2010).

In addition to their specific roles, functional redundancy is a hallmark of GATA4/5/6 in heart development. GATA4 and GATA6 act redundantly in the regulation of cardiac lineage specification as the inactivation of both genes leads to a complete failure to initiate the cardiac morphogenetic program (Zhao et al., 2008) (Fig. 4E). In addition, genetic cooperation between GATA4/5/6 has been demonstrated in compound heterozygote mice for Gata3 and either Gata4 or Gata6, which exhibit pronounced heart defects and lethality (Laforest and Nemer, 2011) (Fig. 4F). As expected for such crucial regulators of heart development, inactivating mutations in GATA4/5/6 have been identified in humans with congenital heart defects (Garg et al., 2003; Kassab et al., 2016; Kodo et al., 2009; Van Der Bom et al., 2011; Wang et al., 2014; Zhang et al., 2015).

Although GATA4/5/6 are recognized as the primary cardiac GATA factors, Gata3 is also expressed in various developing cardiac structures. GATA3 has a significant role in the formation of the cardiac outflow tract, as revealed by catecholamine-based rescue of Gata3 mutant lethality (Raids et al., 2009) (Fig. 4G). These defects include a shorter outflow tract and reduced rotation of truncus arteriosus during the looping stages. Interestingly, individuals with DiGeorge syndrome with a deletion in the 10p region that contains the GATA3 gene present with multiple problems, including cardiac defects (Epstein and Buck, 2000; Wilson et al., 1993).

Other mesoderm-derived tissues

The urogenital system

The urogenital system comprises the kidneys, gonads and urogenital tracts, which develop from the intermediate mesoderm (Stewart and Bouchard, 2014) (Fig. 5A). Both Gata2 and Gata3 are expressed in the developing urogenital system (Ainoya et al., 2012; Grote et al., 2006; Hasegawa et al., 2007). In the renal primordium, Gata3 is largely restricted to the epithelial nephric duct, whereas Gata2 is expressed in the intermediate mesoderm surrounding the duct. At a later stage, they are co-expressed in the developing collecting ducts of the kidney. In the mouse, germline inactivation of Gata3 leads to a complete absence of kidneys and urinary tracts due to severe defects in elongation and guidance of the nephric duct (Grote et al., 2008, 2006) (Fig. 5B). GATA3 occupies, together with the transcription factors Pax2 and Lmx1, a strategic position at the base of the gene regulatory network driving kidney development (Bouafia et al., 2013; Chia et al., 2011; Grote et al., 2008, 2006). GATA3 is later necessary for nephric duct insertion into the cloaca (primordium of the bladder and urethra), and for branching morphogenesis of the kidney, in part through the regulation of the receptor tyrosine kinase Ret (Chia et al., 2011; Grote et al., 2008, 2006) (Fig. 5C). GATA2 also plays a key role in renal development, as evidenced by studies using transgenic Gata2 expression to rescue Gata2-deficient embryos. These experiments identified a crucial requirement for GATA2 in regulating ureter positioning in the renal mesenchyme, likely acting through Bmp4 signaling (Ainoya et al., 2012; Zhou et al., 1998). Interestingly, both Gata2- and Gata3-related malformations seen in mice reproduce defects from individuals with congenital anomalies of the kidney and urinary tract (CAKUT) (Hoshino et al., 2008; Uetani and Bouchard, 2009; Yosypiv, 2012). In addition, inactivating GATA3 mutations in humans lead to renal defects as part of HDR (hypoparathyroidism, sensorineural deafness and renal disease) syndrome (Van Esch et al., 2000).

GATA4/5/6 contribute to other aspects of urogenital system development, acting mostly in the gonads and lower urinary tract. GATA4 is primarily associated with sex determination and gonadal development (Viger et al., 1998). The conditional deletion of Gata4 (using Wnt1Cre) in the coelomic epithelium results in a complete
Fig. 4. Specific and redundant roles of GATA factors during heart morphogenesis. (A) Schematic representation of heart development in the mouse. (B-G) The impact of Gata3, Gata4, Gata5 and Gata6 inactivation on heart development. (B) Gata4−/− deficient hearts present an early lethality due to a failure in heart looping. (C) Both germline and conditional Gata5 inactivation (using Tie2Cre) lead to aortic valve abnormalities. (D) Conditional inactivation of Gata6 in neural crest-derived vascular smooth muscle cells (using Wnt1Cre) prevents cardiac neural crest cell colonization and leads to aortic and ventricular septal defects. (E) Concomitant loss of both Gata4 and Gata6 (via tetraploid embryo complementation) leads to acardia. (F) Compound heterozygous mice for Gata5 and either Gata4 or Gata6 exhibit multiple heart defects, leading to death. (G) Gata3 germline knockout embryos rescued for noradrenaline deficiency show severe heart malformations, such as ventricular septal defects, double-outlet of right ventricle, anomalies of aortic arch and persistent truncus arteriosus.

A, atrium; IT, inflow tract; LA, left atrium; LV, left ventricle; NCC, neural crest cells; OT, outflow tract; RA, right atrium; RV, right ventricle; SV, sinus venosus; V, ventricular region.
Strikingly, adrenal development (Kiiveri et al., 2002; Pihlajoki et al., 2013). Of the gonad differentiation program, supporting a close relationship between these two lineages (Pihlajoki et al., 2013). Although Gata4 alone is dispensable for adrenal gland development, the loss of both Gata4 and Gata6 leads to adrenogenital syndrome due to loss of expression of the master regulator Sry. Together, these results emphasize once again the functional redundancy between GATA4 and GATA6, and their unique requirement at different stages of organ development (Tevosian et al., 2015).

The sexual dimorphic role observed for some GATA factors extends to Gata5, as the inactivation of this gene leads to urogenital tract abnormalities only in females (Molkentin et al., 2000). These include malformations in the urethra and vaginal opening, and are linked to a disruption of early morphogenetic movements. This phenotype mimics a condition of proximal hypospadias in women (Molkentin et al., 2000). Together, these results reveal a pleiotropic role for GATA4/5/6 during genital system development, ranging from early specification to terminal differentiation.

Endothelial vascular cells

The developmental functions of GATA factors also include a role in the vascular system. Loss of Gata2 specifically in the embryonic endothelium or in the lymphatic vasculature has revealed important function for GATA2 in both the morphogenesis and maintenance of lymphatic vessel valves and vasculature (Kazenwadel et al., 2015; Lim et al., 2012). This requirement is elicited by VEGF and involves the early activation of a lineage-specific program, including the regulation of Sox18, Etn2 and Tall (Kanki et al., 2017). In this respect, GATA2 respects the propensity of GATA factors to promote lineage specification and early stage morphogenesis. In line with this role for GATA2 in the lymphatic system, inactivating GATA2 mutations have been identified in individuals with Emberger syndrome, a disease characterized by lymphedema and a susceptibility to the development of myelodysplastic syndrome/acute myeloid leukemia (MDS/AML) (Ostergaard et al., 2011).
Mesenchymal cell differentiation

A requirement for GATA2 and GATA4 has also been reported in mesoderm-derived mesenchymal stem cells and their differentiation into adipocytes and osteoblasts. In this system, GATA2 primarily promotes mesenchymal stem cell proliferation and prevents premature differentiation to both adipocyte and osteoblast cell lineages, while GATA4 plays a positive role in osteoblast cell differentiation (Güemes et al., 2014; Li et al., 2016; Zhou et al., 2017; Kamata et al., 2014; Patankar et al., 2011; Tsai et al., 2005). Their function in these tissues is therefore compatible with the general observation that GATA proteins act as regulators of lineage specification and tissue homeostasis.

Ectoderm-derived tissues

Ectoderm-derived epithelial tissues such as the mammary gland and the skin exhibit a distinct stem cell hierarchy. In these tissues, GATA factors govern the differentiation programs driving lineage specification, but are also central to the maintenance of epithelial tissue architecture.

Mammary gland

In the mammary gland, Gata3 is expressed in the terminal end buds, which host the stem/progenitor cell pool during development (Koulos-Mehr et al., 2008). As the gland develops, GATA3 regulates duct elongation and branching morphogenesis, as well as normal differentiation of luminal epithelial cells (Koulos-Mehr et al., 2008, 2006). Post puberty, the absence of Gata3 causes lactation failure and pup lethality owing to a failure in cell differentiation, which results in an accumulation of progenitor cells in both the alveolar and ductal lineages (Asselin-Labat et al., 2007). In the duct, GATA3 acts as a pioneer factor with FOXA1 to open the chromatin at ER-responsive genes (Kong et al., 2011). A different situation is observed in the alveolar secretory epithelium, where GATA3 is also required for epithelial cell differentiation but acts independently of ERα and FOXA1, suggesting that it drives an independent transcriptional program in this lineage. Hence, GATA3 directs luminal differentiation of progenitor cells, while also being necessary for the maintenance of luminal epithelial cell integrity and function (Asselin-Labat et al., 2007; Koulos-Mehr et al., 2006).

The epidermis

GATA3 additionally plays unique roles in skin development. In this system, multipotent stem cells give rise to the epidermis and hair follicle lineages (Gonzales and Fuchs, 2017). Inactivation of Gata3 in basal cells leads to a loss of the epidermal barrier and causes postnatal death due to dehydration through the epidermis (de Guzman Strong et al., 2006). Gata3 deletion in the skin also causes an expansion of inner root sheath (IRS) progenitors that fail to differentiate and instead progress toward a hair shaft fate, creating an aberrant hair structure (Kauffman et al., 2003; Lim et al., 2000). The specific deletion of Gata3 in the epidermis and hair follicles causes a delay in hair growth and maintenance, abnormal hair follicle organization and defects in skin differentiation (de Guzman Strong et al., 2006; Kurek et al., 2007). At the molecular level, GATA3 integrates morphogenetic signals such as WNT, NOTCH and the BMP pathway at the crossroads of hair follicle versus epidermal fates and in IRS versus hair shaft cell fate decisions (Genander et al., 2014; Kaufman et al., 2003; Kurek et al., 2007). In the skin, GATA3 therefore controls several lineage decisions as well as terminal epithelial integrity (Chikh et al., 2007; Masse et al., 2014).

The inner ear and lens

Both the inner ear and lens develop from epithelial placodes undergoing invagination and subsequent differentiation. In the inner ear, Gata3 and Gata2 are expressed in a sequential manner, the former regulating the expression of the latter. In line with an early expression, Gata3 germline deficiency leads to aberrant otic placode invagination accompanied by otic epithelial closure and detachment defects, which affect the outgrowth of the semicircular and cochlear ducts (Karj et al., 2001; Lilleväli et al., 2006, 2004). These defects are associated with the reduced expression of genes involved in cell adhesion and motility (Lilleväli et al., 2006). As expected for a downstream regulator, loss of Gata2 in the otic region gives rise to a milder phenotype, causing growth defects in only the semicircular ducts (Haugas et al., 2010).

Following otic vesicle invagination and patterning, neuronal progenitor cells delaminate from the vesicle to form the vestibulo-acoustic ganglia, while the remaining cells form the neurosensory epithelium. In the absence of Gata3, prosensory cell formation is prevented, leading to defects in spiral ganglion neurons and cochlear neurosensory epithelia (Duncan et al., 2011; Haugas et al., 2012). The inactivation of Gata3 specifically in the neurosensory epithelial system of the cochlea leads to aberrant tectorial membrane and microvilli development. GATA3 therefore ensures both the specification and correct wiring of efferent inner ear neurons (Appler et al., 2013) through upregulation of pro-neurosensor and sensory epithelial gene networks (Lilleväli et al., 2006). Interestingly, mice heterozygous for Gata3 show progressive hearing loss resulting from hair cell degeneration (van der Wees et al., 2004). This phenotype parallels the systemic deafness observed in individuals with HDR syndrome, which results from inactivating mutations in GATA3 (Van Esch et al., 2000). Together, these findings highlight the importance of GATA3 in inner ear development and hearing/deafness.

In contrast to its role in the otic placode, GATA3 does not affect the invagination of lens placodes. However, it is essential for subsequent lens differentiation, as shown by the downregulation of γ-crystallin genes and failure to degrade fiber cell nuclei in Gata3-deficient embryos. GATA3 further controls cell cycle exit of lens progenitor cells through the upregulation of Cdkn1b (p27) and Cdkn1c (p57), and downregulation of Ccnd2 (cyclin D2) expression (Maeda et al., 2009). Together, these results identify GATA3 as an important regulator of lens cell fate progression.

Other ectoderm-derived tissues

In addition to the skin, lens, otic vesicle and mammary gland, GATA2 and GATA3 are required for progenitor cell maintenance and lineage specification in other ectoderm-derived tissues. For example, GATA2 and GATA3 perform important regulatory functions at different stages of neuronal development. In the hindbrain, GATA2 and GATA3 act redundantly during neuronal specification, but also play unique roles as selectors in subsequent stages of neuronal differentiation toward the serotonergic and glutamatergic lineages (Craven et al., 2004; Haugas et al., 2016; Lahti et al., 2016). A unique role for GATA2 has also been identified for the development of sympathetic (Tsarovina et al., 2004) and GABAergic neurons (Kala et al., 2009; Willett and Greene, 2011). In sympathetic neurons, GATA3 is essential for the expression of tyrosine hydroxylase (TH), a key enzyme of the catecholamine biosynthetic pathway (e.g. dopamine and noradrenaline). Accordingly, inactivation of Gata3 in the mouse germ line causes noradrenaline deficiency and embryonic lethality.
around E10.5, which can be rescued to birth by the administration of downstream catecholamine intermediates (Lim et al., 2000).

In the pituitary gland, the interaction between GATA2 and PIT1 controls cell fate decisions. GATA2 acts synergistically with PIT1 to upregulate thyrotropic and repress gonadotropin genes, whereas the expression of GATA2 in the absence of PIT1 leads to the expression of gonadotrope-specific genes (Dasen et al., 1999). Here again, the combinatorial interaction of GATA factors with specific transcription factors is an important mechanism of action in the control of lineage specification.

Endoderm-derived tissues
During development, the endoderm compartment gives rise to the gastrointestinal tract and associated organs, and is divided into foregut, midgut and hindgut. The liver, lungs and pancreas all emerge from the foregut endoderm, whereas the small intestine develops from the midgut endoderm and the colon from the hindgut. The prostate gland is also endoderm derived, arising from the primitive urogenital sinus located at the junction between bladder and urethra. In recent years, roles for GATA factors in the development of multiple endoderm-derived organs have been described.

The liver
During liver development, GATA4 acts as a pioneer factor in hepatocyte cell differentiation by rendering chromatin accessible to other transcription factors while establishing liver-specific gene expression (Watt et al., 2007; Zheng et al., 2013). GATA4 additionally controls the specification and function of liver sinusoidal cells (fenestrated blood vessels). The specific deletion of Gata4 in these cells switches discontinuous liver sinusoids into continuous capillaries (Géraud et al., 2017), and mutant mice die from anemia due to liver hypoplasia, fibrosis and impaired colonization by hematopoietic progenitor cells. Thus, GATA4 is not only important as a master regulator of hepatocyte differentiation but also for hepatic microvascular endothelium specification. The analysis of Gata4 or Gata6 mutant embryos rescued for extra-embryonic defects with tetraploid embryonic stem cells (ESCs) additionally identified defects in liver bud growth and commitment of the endoderm to a hepatic cell fate (Watt et al., 2007; Zhao et al., 2005). GATA4 and GATA6 possibly compensate for each other during the specification stage, as the initial hepatic fate specification is not affected in single mutants. These results suggest crucial roles in different aspects of liver morphogenesis.

The lungs
In the lungs, GATA6 is essential for branching morphogenesis and late epithelial cell differentiation, but not for endodermal specification. This was demonstrated in chimeras derived from Gata6-deficient embryonic stem cells (ESCs), which show branching defects of the pulmonary endoderm as well as a block in bronchial epithelial cell differentiation (Koutsourakis et al., 2001). GATA6 controls the balance between bronchioalveolar stem cell expansion and epithelial differentiation (Zhang et al., 2008). In addition, GATA4 and FOG2 are required for normal lung development and diaphragm function, acting via the regulation of lung mesenchymal cells (Ackerman et al., 2007; Jay et al., 2007).

The pancreas
Although both Gata4 and Gata6 are expressed in pancreatic progenitors, their expression becomes mutually exclusive and restricted to endocrine and exocrine lineages, respectively (Decker et al., 2006). In mice, only the loss of both Gata4 and Gata6 leads to complete pancreas agenesis, which is indicative of functional redundancy between both genes at an early stage (Carrasco et al., 2012; Xuan et al., 2012). In the absence of both factors, the epithelium fails to expand properly and the specification of both endocrine and exocrine lineages is considerably affected. These defects partly result from an upregulation of the sonic hedgehog (SHH) pathway, which leads to the conversion of pancreatic progenitor cells toward the stomach or intestinal cell lineages (Xuan and Sussel, 2016). GATA4 and GATA6 are thus essential to maintain pancreas identity through suppression of the SHH pathway. Interestingly, in vitro differentiation assays using human pluripotent stem cells (hPSCs) have revealed that the formation of pancreatic progenitors is highly sensitive to Gata6 and Gata4 gene dosage (Shi et al., 2017). Whereas a reduced number of progenitors is seen in Gata4-deficient as well as in Gata4/6 compound heterozygote cells, pancreatic progenitor cells are absent in embryos derived from Gata6-deficient hPSCs. In accordance with this, mutations in both Gata4 and Gata6 have been linked to diabetes and pancreatic agenesis in humans (Allen et al., 2011; Chao et al., 2015; De Franco et al., 2013; Smith et al., 2014; Stanescu et al., 2015). Individuals with Gata6 mutations also present heart, gut, thyroid and intra-uterine growth retardation problems, highlighting the important role played by Gata6 in multiple tissues (Bonnefond et al., 2012; Chao et al., 2015; De Franco et al., 2013).

The intestine
Gata4 and Gata6 are expressed in different parts of the intestine. Whereas Gata6 is expressed in the distal part of the small intestine (the ileum) and in the colon, Gata4 is only expressed in the proximal duodenum and jejunum (Battle et al., 2008; Bosse et al., 2006; Fang et al., 2006). Loss of Gata4 in the jejunal epithelium leads to severe defects in fat and cholesterol absorption, resulting from a switch from jejunal to ileal gene expression profiles (Battle et al., 2008). Gata4 is co-expressed with Fog1 in a proximal-distal gradient along the small intestine (Beuling et al., 2008). With the help of a mouse model specifically preventing the interaction between Gata4 and Fog1, it was found that the Gata4:Fog1 complex is required to represses ileal fate specification in the proximal intestine, whereas Gata4 acts independently of Fog1 to activate proximal intestine-specific genes (Beuling et al., 2008). More recently, Gata4 was revealed as being sufficient to induce jejunal enterocyte identity when ectopically expressed in the ileum (Thompson et al., 2017), indicative of a strong jejunal lineage specification potential. Deletion of Gata6 in the small intestine alters ileal epithelial cell populations by reducing enteroendocrine and Paneth cell number and by increasing goblet cell number. It also causes changes in the ileal enterocyte-specific gene expression pattern, shifting it towards a colon-like pattern (Beuling et al., 2011). The absence of both Gata4 and Gata6 leads to a loss of enterocytes, an increase in goblet cell number and, ultimately, neonatal death (Walker et al., 2014). This phenotype is in part due to deregulation of the NOTCH pathway, which is important for the choice between enterocyte and goblet cell fate (Walker et al., 2014). More caudally, GATA6 is required for the expression of goblet cell and colonocyte-specific genes, as well as for the control of proliferation, migration and lineage maturation in the colon epithelium (Beuling et al., 2012). Hence, the intestine is another developmental system in which the GATA4-GATA6 pair plays an essential and redundant role in cell differentiation.
The prostate
The prostate epithelium architecture is analogous to the mammary epithelium and is mainly composed of luminal and basal cells, interspersed with rare neuromodocrine cells. As is the case for the mammary gland, GATA3 plays a key role in prostate development. However, in contrast to its luminal-restricted expression observed in the mammary gland, Gata3 is expressed in both basal and luminal lineages of the prostate (Shafer et al., 2017). Loss of Gata3 in the developing prostate (using Nkx3.1\textsuperscript{Cre}) does not prevent luminal cell differentiation but causes an accumulation of intermediate progenitor cells associated with loss of polarity, tissue hyperplasia and defective branching morphogenesis (Fig. 6). These malformations stem from defects in lineage commitment through oriented cell division in progenitor cells (Shafer et al., 2017). In the adult prostate, Gata2 and Gata3 act redundantly, as only double knockouts (using Pb\textsuperscript{lox/lox}) show altered expression of AR and an expansion of the basal cell compartment (Xiao et al., 2016). Interestingly, a negative-feedback regulatory loop exists between GATA2 and AR, whereby GATA2 directly controls Ar gene expression, while androgen signaling downregulates Gata2 expression (Xiao et al., 2016). Hence, GATA factors are essential for both the development and the maintenance of prostate epithelial homeostasis.

Pre-implantation embryo development
In addition to their roles in the various germ layer-derived tissues, GATA factors play important roles during pre-implantation development. In mouse blastocysts (E5.5), Gata2 and Gata3 are specifically expressed in the trophectoderm lineage, which gives rise to the placenta, but are not expressed in the inner cell mass from which the embryo proper develops (Home et al., 2017). Both factors act redundantly in trophoblast and placenta formation, where they control the self-renewal of stem/progenitor cells and their progressive differentiation (Home et al., 2017). Here again, GATA3 acts as a pioneer factor to activate lineage-specific factors, including Gata2 (Ray et al., 2009). GATA6 is important in the inner cell mass where it modulates the FGF/ERK pathway to control the commitment between primitive endoderm and epiblast (Bessonard et al., 2014; Schrodé et al., 2014). Ectopic expression of GATA6 in ESCs is sufficient to drive them toward the primitive endoderm lineage (Artus et al., 2011; Fujikura et al., 2002; Shimosato et al., 2007), whereas its inactivation blocks primitive endoderm formation and expands the epiblast lineage (Cai et al., 2008).

Conclusions
Since their discovery in the early 1990s, GATA transcription factors have been shown to play essential roles in development, and their study has contributed significantly to our understanding of the basic principles of tissue development and morphogenesis. Their prominent and pleiotropic roles in lineage specification from stem/progenitor cells has been invaluable for deciphering the cellular and molecular mechanisms by which cell diversification and subsequent differentiation occur. The activation of cell type-specific genes by GATA factors is often coupled with a blockade of alternative gene programs, through either activation of lineage-restricted transcription factors, feed-forward loops or mutual antagonisms between drivers of alternative cell fates. This leads to a cascade of regulatory changes that locks in the fate of progenitor progeny and allows the morphogenetic program to proceed. Besides lineage specification, GATA factors also play crucial roles in the early stages of tissue morphogenesis. As expected for such important developmental regulators, most have been associated with deleterious genetic diseases in human.

Moving forward, the greatest challenges will be to integrate the specific activity of GATA factors into the more complex regulatory networks they are part of. It is unclear, for example, what properties and circuitry associated with GATA factors make them suitable for driving specific lineage decisions at the expense of others. Although many cellular defects resulting from GATA perturbation have been identified, few of the underlying gene networks have been characterized in great detail. Are there commonalities between the networks underlying hematopoietic, trophoblast and mammary lineage specification? What differentiates those networks from those in which GATA factors are not directly involved in lineage specification? The close association between GATA factors and hormone signaling (e.g. estrogen, progesterone and androgen signaling) in different organ systems is also intriguing. We can now build on a large amount of data on the reported roles of GATA factors in different developmental systems to tackle the challenging task of integrating phenotypic observations into comprehensive molecular and cellular systems.

Acknowledgements
We apologize to authors whose work could not be cited due to space limitations. Special thanks to the members of the Bouchard laboratory for critical reading of the manuscript.

Competing interests
The authors declare no competing or financial interests.

Funding
M.B. is supported by a grant from the Canadian Institutes of Health Research (MOP-130431) and holds a Senior Research Scholar Award from the Fonds de la Recherche du Québec-Santé. O.S.-F. is supported by a KRESCENT fellowship from the Kidney Foundation of Canada.

References


Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). Nat. Genet. 43, 929-931.

Ouyang, W., Ranganath, S. H., Weindel, K., Bhattacharya, D., Murphy, T. L., Sha, W. C. and Murphy, K. M. (1998). Inhibition of Th1 development mediated by GATA-3 through an IL-4-independent mechanism. Immunity 9, 745-755.


