Role of GATA factors in development, differentiation, and homeostasis of the small intestinal epithelium

Boaz E. Aronson,1,2 Kelly A. Stapleton,1 and Stephen D. Krasinski1,3

1Division of Gastroenterology and Nutrition, Department of Medicine, Children’s Hospital Boston, and Harvard Medical School, Boston, Massachusetts; 2Department of Pediatrics, Emma Children’s Hospital, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; and 3Gerald J. and Dorothy R. Friedman School of Nutrition Science and Policy, Tufts University, Boston, Massachusetts

Submitted 11 April 2013; accepted in final form 7 January 2014

Aronson BE, Stapleton KA, Krasinski SD. Role of GATA factors in development, differentiation, and homeostasis of the small intestinal epithelium. Am J Physiol Gastrointest Liver Physiol 306: G474–G490, 2014. First published January 16, 2014; doi:10.1152/ajpgi.00119.2013.—The small intestinal epithelium develops from embryonic endoderm into a highly specialized layer of cells perfectly suited for the digestion and absorption of nutrients. The development, differentiation, and regeneration of the small intestinal epithelium require complex gene regulatory networks involving multiple context-specific transcription factors. The evolutionarily conserved GATA family of transcription factors, well known for its role in hematopoiesis, is essential for the development of endoderm during embryogenesis and the renewal of the differentiated epithelium in the mature gut. We review the role of GATA factors in the evolution and development of endoderm and summarize our current understanding of the function of GATA factors in the mature small intestine. We offer perspective on the application of epigenetics approaches to define the mechanisms underlying context-specific GATA gene regulation during intestinal development.

endoderm; GATA; intestinal development; intestinal differentiation; intestinal epithelium

GATA FACTORS ARE AN ANCIENT family of transcription factors that mediate processes of development, differentiation, and gene expression in multiple tissues and cell types. Six GATA factors are conserved in all vertebrates. On the basis of mutational analyses of GATA1 (70) and GATA4 (84), vertebrate GATA factors contain transcriptional activation domains localized to the NH2-terminal region, two highly conserved zinc fingers and adjacent basic regions, a nuclear localization signal sequence, and a COOH-terminal domain of unknown function (Fig. 1A). GATA zinc fingers comprise four cysteine residues (type IV), Cys-X2-Cys-X17-Cys-X2-Cys, that coordinate a divalent zinc ion (Fig. 1B) and, together with the adjacent basic regions, direct binding to the nucleotide sequence element WGATAR (W = A or T, R = A or G) and also interact with other regulatory proteins to modulate GATA transcriptional control of target genes.

Vertebrate GATA factors are separated into two subfamilies: 1) GATA1, GATA2, and GATA3 and 2) GATA4, GATA5, and GATA6. Members of the GATA1/2/3 subfamily are expressed in developing blood cells, where they regulate differentiation-specific gene expression in hematopoiesis (reviewed in Ref. 93), GATA2 also functions in urogenital development, pituitary cell fate specification, and generation of V2 interneuronal cells in the spinal cord (reviewed in Ref. 21). GATA3 is also expressed in the adrenal glands, kidneys, central and peripheral nervous systems, inner ear, hair follicles, skin, and breast tissue (reviewed in Ref. 51). Members of the GATA4/5/6 subfamily are expressed in various mesoderm- and endoderm-derived tissues, such as heart, liver, pancreas, lung, gonad, and gut, where they play critical roles in regulating tissue-specific gene expression (reviewed in Ref. 78). Here we review the role of GATA factors in the evolution, development, and function of the small intestinal epithelium.

GATA Factors Are Expressed in Endoderm of Ancient Metazoa

Gut formation is one of the earliest outcomes of multicellularity (reviewed in Refs. 20 and 107). It occurs by gastrulation, a process in which invagination of the blastocyst lining at the blastopore leads to the formation of a gastrovascular cavity, called the archenteron, that is lined by two epithelial sheets, the outer ectoderm and the inner endoderm (reviewed in Refs. 72 and 110). In the transition from protost to metazoan, thought to occur during the Cambrian explosion 542 million years ago, selective pressures forced the division of labor among cells: the ectoderm became specialized for protection, locomotion, and sensing, and the endoderm became specialized for digestion and absorption of nutrients. Diploblastic animals possessing two germ layers, ectoderm and endoderm, exhibit radial (rather than bilateral) symmetry and have a single opening that serves as both mouth and anus. They are found exclusively in the phylum Cnidaria, which includes corals, sea anemones, jelly-
fish, and hydra. Mesoderm, or the middle layer, thought to be a derivative of endoderm, arose ~40 million years after the emergence of ectoderm and endoderm. Triploblastic animals possessing three germ layers are linked to the evolution of axis formation and bilateral symmetry. In contrast to diploblastic animals, triploblastic bilaterians (with the exception of a few highly reduced forms) have complete digestive tracts with separate mouth and anus. The through-gut system enabled the directional movement of nutrients, facilitating the diversification of cells along the digestive tract to control the processing of food at different stages of digestion. This, in turn, allowed for the utilization of a wider assortment of food types, enabling an expanded range of mobility. Thus the through-gut digestive system was one of the earliest and key innovations in the evolution of metazoans.

Triploblastic bilaterians are further subdivided into protostomes and deuterostomes on the basis of how mouth and anus develop. During gastrulation, invagination occurs at the blastopore, and the two opposing poles of the blastopore margin close, forming a slitlike structure that defines the ventral midline. In protostomes (including arthropods, nematodes, and annelids), the openings at each end become mouth or anus, whereas in deuterostomes (including echinoderms and chordates), the formation of the mouth occurs by a secondary breach of the archenteron.

Although multiple GATA factors have been identified in plants, fungi, and molds (66), no GATA factors have been reported for members of the phylum Porifera (sponges), the earliest members of the animal kingdom, or their hypothesized single-cell evolutionary predecessors the choanoflagellates. However, Martindale et al. (73) cloned and characterized the first report of a GATA factor in the starlet sea anemone Nematostella vectensis, the only GATA factor in the cnidarian diploblast. NvGata, with homology to the vertebrate GATA1/2/3 and GATA4/5/6 subfamilies (47), is the first detected in individual cells throughout the blastula, which coalesce at the blastopore and ingress into the blastocoel (cavity in blastocyst). By the late gastrula stage, all the gastrodermal cells lining the gastrovascular cavity (called the coelenteron in diploblasts) express NvGata, and expression continues into the polyp stage. This report demonstrates a pattern of expression of an ancient GATA factor that intersects with the evolutionary emergence of gastrulation and endodermal specification (Fig. 2).

Expression of GATA factors has also been reported for developing bilaterian protostome flatworms (17, 18, 46, 71), all of which express at least two GATA factors, generally, a single, conserved GATA1/2/3 ortholog that is expressed in endoderm and at least one, but usually multiple, GATA4/5/6 orthologs that are expressed in mesoderm and endoderm derivatives. These findings suggest that the two vertebrate GATA factor classes originated from distinct ancestral orthologs and that the GATA gene duplication and functional divergence that led to these two ancestral GATA factors occurred after the split from diploblasts (Fig. 2). These studies also aligned the vertebrate GATA4/5/6 subfamily with endodermal differentiation.

Members of the phylum Echinodermata, the second-largest grouping of deuterostomes, after chordates, express two Gata...
genes (50, 63). Gatac and Gatae, which are orthologous to vertebrate Gata1/2/3 and Gata4/5/6, respectively. The Gatae mRNA is first detected in a ring around the vegetal pole of the blastula. During gastrulation, transcripts are detected surrounding the blastopore, in the posterior archenteron, and in the anterior mesoderm of the archenteron. In the late gastrula and early larval stages, expression is localized to the midgut and hindgut and to the developing coelomic pouches.

Using a combination of large-scale perturbation analyses, computational methodologies, genomic data, cis-regulatory analyses, and molecular embryology, Davidson et al. (30) derived a mesoderm/endoderm gene regulatory network (GRN) for the sea urchin embryo. In the 24-h blastula, all the genes known to execute mesoderm and endoderm fate are active. Early inputs are provided by the B lymphocyte-induced maturation protein-1 (blimp1, formerly called krox1) and orthodenticle homolog-1 (otx1) genes, which regulate virtually all the genes that specify mesendoderm, a transient precursor cell population that subsequently differentiates into mesoderm and endoderm. The principal target is gatae, which contributes to activation of six other transcriptional regulatory genes of the mesendoderm and also feeds back to the otx1 cis-regulatory system, permanently locking down the mesendodermal specification state.

In summary, expression of GATA factors in metazoans can be traced back to the evolution of gastrulation and the development of endoderm (Fig. 2). In a diploblast, a single GATA factor that may be ancestral to the vertebrate GATA1/2/3 and GATA4/5/6 subfamilies is expressed at the blastopore prior to gastrulation and then in all endodermal cells after gastrulation. In triploblasts, at least two GATA orthologs are expressed, generally corresponding to the vertebrate GATA1/2/3 and GATA4/5/6 subfamilies. Triploblastic protostomes generally have a single GATA1/2/3 ortholog, but multiple GATA4/5/6 orthologs, whereas early triploblastic deuterostomes have a single member of each subfamily that eventually expanded into three members of each GATA subfamily in vertebrates. In all cases, the GATA4/5/6 ortholog is expressed at the blastopore prior to gastrulation, in ingressing cells during gastrulation, in all endodermal cells after gastrulation, and in cells along the digestive tract at maturity. GRN analysis in the sea urchin indicates a pivotal role for GATA factors in mesoderm and endoderm specification and gut development.

Divergent GATA Factors Act Sequentially in Endoderm Formation and Gut Development in Invertebrates

The function of GATA factors and their placement in the regulatory cascade of endoderm formation and gut development in the vertebrate protostomes Caenorhabditis elegans and Drosophila melanogaster (Drosophila) have been well reviewed (reviewed in Refs. 55, 87) and are updated here.

Seven GATA orthologs specify endoderm fate and terminal differentiation of the C. elegans intestine. C. elegans expresses no less than 11 orthologous GATA factors, of which 7 have been implicated in endoderm and/or gut formation (reviewed in Ref. 55). The sequences of these GATA factors display significant divergence, with only the elt-1/Gata1/2/3 ortholog containing complete dual zinc finger domains. The remaining C. elegans GATA factors, all lacking the first zinc finger, group with the GATA4/5/6 subgroup and, like lower Metazoa, display a biased expansion of this subgroup of GATA factors (47).

Embryogenesis in C. elegans (reviewed in Ref. 99) begins when the asymmetric first cleavage produces a large anterior daughter called AB and a smaller posterior daughter called P1. P1 then divides to produce the anterior germline and mesodermal precursor cell P2 and the posterior mesodermal precursor cell EMS (Fig. 3A). The anterior daughter of EMS, called MS, gives rise to many mesodermal cell types, while the posterior daughter of EMS, the E blastomere, is the progenitor of the sole endoderm-derived organ, the intestine. Early embryonic patterning events are essentially completed by the 28-cell stage, when gastrulation begins with the ingestion of the intestinal precursor cells. The mature intestine comprises the midgut of the C. elegans alimentary tract, connecting to the pharynx (or foregut, derived partly from MS) and rectum (or hindgut, derived from P2).

Tissue specification of the EMS blastomere is controlled by the maternal bzip homeodomain transcription factor SKN-1, the translation of which is confined to EMS by maternally provided mRNA only to the posterior embryo. SKN-1 directly activates transcription of the genes encoding MED-1 and MED-2, two nearly identical and functionally redundant GATA-like transcription factors (Fig. 3A). The med-1 and med-2 genes are unique in that they lack introns and express GATA-like factors that contain a single type IV zinc finger that differs from classical GATA zinc fingers by a single amino acid. This alteration results in recognition of the noncanonical DNA sequence RAGTATAE, suggesting that med-1 and med-2 have diverged in function and may be a nematode invention (67). MED-1 and MED-2 play essential roles in mesendoderm specification by activating the T-box gene tbx-35 in the MS blastomere (19) but play only a minor role in endoderm specification and gut development (25, 68).

Specification of the E blastomere fate is determined by expression of the genes encoding the redundant pair of GATA factors, END-1 and END-3, which are activated by SKN-1 and, to a lesser extent, MED-1/MED-2 (Fig. 3A). While individual end-1 or end-3 knockouts develop normally, double knockouts of end-1 and end-3 do not form an intestine (94). Ectopic expression of either factor converts other embryonic cells to an endodermal fate (69, 126). End-1 and end-3 appear to be the result of a recent duplication, with each diverging to activate two distinct, but overlapping, E lineage regulatory pathways, perhaps through differences in their DNA-binding domain specificities (11). End-1 and end-3 are also expressed differently; end-3 is expressed prior to that of end-1 (and may activate end-1 gene expression), and single end-3 knockouts reveal a delay in E lineage activation that is not apparent in single end-1 knockouts. Expression of both end-1 and end-3 is extinguished prior to terminal gut differentiation.

The major target of END-1/END-3 in the early intestine (2E) is the gene encoding yet another GATA factor, ELT-2 (43) (Fig. 3A). Deletion of elt-2 results in a lethal arrest at birth; the newly hatched larvae have a malformed, but clearly specified, intestine (43, 74, 75). Promoters of genes that are exclusive to, or highly enriched in, the intestine contain GATA binding sites (compared with <5% for control promoters), and most of these genes are downregulated in elt-2 null worms. ELT-2 may regulate these target genes in cooperation with Notch signaling by physically interacting with the Notch-dependent effector
LAG-1/CSL and, together, selecting target genes for endodermal expression (89). Thus ELT-2 plays a central role in the establishment and maintenance of most aspects of terminal C. elegans intestinal physiology.

Although the GATA orthologs ELT-4 and ELT-7 are also expressed in the mature C. elegans intestine, deletion of either elt-4 or elt-7 alone results in essentially wild-type intestine (5, 42, 104). ELT-4 is a very small GATA factor, barely the size of a single zinc finger, and has no discernible function as determined by deletion and overexpression experiments (42). Deletion of elt-4 in the context of other elt deletions, however, has not been reported; thus redundant functions with ELT-2 and/or ELT-7 remain a possibility. Simultaneous deletion of elt-2 and elt-7, but not of either alone, eliminates all morphological features of differentiation in patches of the gut, suggesting cross-regulatory functions between ELT-2 and ELT-7 in gut development (104). Elt-7, like elt-2, is activated by END-1/END-3, is expressed before elt-2, and activates its own expression and expression of elt-2, constituting an apparent positive-feedback system that locks down gut differentiation. In summary, up to seven GATA4/5/6 orthologs play cooperative and sequential roles in guiding cell fate specification, endoderm commitment, and terminal differentiation of the nematode intestine (Fig. 3A).

Three GATA orthologs are necessary for endoderm development in Drosophila melanogaster. Five GATA genes are found in the Drosophila genome, only one of which (dGATAc, grain) is orthologous to the GATA1/2/3 subgroup (47). Three GATA orthologs, serpent (srp, dGATAb), dGATAc (grain), and dGATAe, are related to endoderm development (reviewed in Ref. 87). With the exception of srp, which lacks an NH2-terminal zinc finger due to a splicing isoform, all the Drosophila GATA factors have two zinc fingers associated with basic domains.

In Drosophila, the digestive tract develops by invaginations of prospective endoderm at the anterior and posterior regions of the early blastoderm (stage 7–8) that migrate bilaterally along the yolk toward the trunk, encounter each other in the middle of the trunk (stage 9–10), and spread over the yolk to form a continuous epithelial tube that encloses the yolk (stage 12–13), eventually giving rise to the epithelial lining of the midgut (stage 15–16) (reviewed in Ref. 87). The prospective endodermal regions of the early blastoderm (stage 5–6) are specified by graded, maternally derived signaling from the transmembrane receptor tyrosine kinase, Torso (Tor) (Fig. 3B). Tor signaling activates the two earliest zygotic gap genes tailless (tl) and huckebein (hkb), which are expressed in a nested pattern, with the Hkb regions included in Tll-positive domains. While the Tll-positive, Hkb-negative domains specify ectodermal foregut and hindgut, Hkb-positive, Tll-positive regions specify endoderm, although Tll itself is not necessary for this process.

In anterior and posterior terminal regions, Hkb activates srp, the earliest GATA gene expressed in Drosophila endoderm (95, 98) (Fig. 3B). Expression of srp begins at an early blastoderm stage in the prospective endodermal regions but disappears in the endoderm around stage 10–11, long before terminal differentiation of the gut occurs. Loss of Srp activity results in a transformation of prospective endoderm into ectodermal foregut and hindgut (98), while overexpression of srp ectopically
induces endodermal genes (87). Thus srp, like end-1/end-3 in C. elegans, is necessary for endoderm specification but is not required later to maintain terminal gut differentiation.

Recently, it was shown that Srp induces an endodermal epithelial-mesenchymal transition (EMT) (22). During formation of the Drosophila endoderm, epithelial cells adopt a mesenchymal state and migrate through the embryo and later reepithelialize to give rise to a large portion of the intestinal tract. In srp mutants, endodermal cells retain epithelial shape and polarity and do not migrate, and dE-cadherin, a junctional protein that mobilizes in EMT, remains tightly localized to the apical domain of cells. Thus Srp confers critical migratory capabilities on epithelial cells during intestinal development.

dGATA4c is expressed in the head, posterior spiracles, and central nervous system, as well as in developing endoderm beginning at stage 12, and continues to be expressed in anterior regions of midgut at least to stage 15 (reviewed in Ref. 87). In nongut tissue, dGATA4c affects organ shape by locally controlling cell rearrangement and morphogenetic movement, but neither deficiency nor overexpression of dGATA4c affects midgut morphology or endodermal gene expression. Thus the function of dGATA4c in endoderm development remains unknown.

dGATA6 is expressed specifically in the anterior and posterior endoderm at stage 8, inversely correlating with weakening srp expression, and continues to be expressed in the larval and adult midgut (91) (Fig. 3B). dGATA6 is not expressed in srp mutants and is induced when srp is overexpressed, indicating that srp is required for dGATA6 expression. Loss of dGATA6 causes a general downregulation of genes associated with terminal differentiation of the midgut, while dGATA6 overexpression induces ectopic expression of these genes (91). However, a subset of terminal differentiation genes activated by Srp, but not dGATA6, suggests that dGATA6-independent pathways that are downstream of Srp regulate some intestinal genes (92). In embryos deficient in srp or dGATA6, a master gene of ectodermal hindgut, brachycenteron (byn), is ectopically expressed in innate endodermal regions. Taken together, these data indicate that Srp specifies endodermal fate and, together with its target dGATA6, activates the terminal gut differentiation program while suppressing an ectodermal fate.

In summary, sequential expression of discrete GATA4/5/6 orthologs mediates endoderm specification and intestinal development in C. elegans and Drosophila. In both cases, a maternally derived process asymmetrically activates the early and transient expression of a GATA factor (end-1/end-3 in C. elegans and srp in Drosophila) in prospective endoderm that specifies endodermal fate. This GATA factor, in turn, activates a later GATA factor (elt-2/elt-7 in C. elegans and dGATA6 in Drosophila) that triggers a terminal gut differentiation program. Thus the sequential functions of distinct GATA factors in the gene regulatory pathways of endoderm development are well conserved in the protosome invertebrates C. elegans and Drosophila.

Conserved GATA4/5/6 Orthologs Are Necessary for Endoderm Development in Vertebrates

Deuterostome vertebrates possess a remarkably conserved molecular program in which the Nodal signaling pathway is necessary and sufficient to initiate mesendoderm development (reviewed in Refs. 102, 106, 127, and 128). In general, high levels of Nodal signaling promote endoderm development, whereas lower levels specify mesoderm development. Nodal signaling stimulates the downstream expression of a conserved group of transcription factors, including Mix-like homeomain proteins, Sry-related high-mobility group box (SOX) factors, forkhead box (FOX) factors, and GATA factors. Generally, these factors together function to segregate endoderm from mesoderm and to lock down endodermal fate. In this section, we update the specific functions of GATA factors in endoderm specification and intestinal development in Xenopus laevis (Xenopus) and Danio rerio (zebrafish).

Gata4/5/6 regulate endodermal gene expression in X. laevis. The early Xenopus embryo forms two distinct halves, the pigmented animal pole and the yolky vegetal pole. Fertilization occurs at the animal pole and triggers a displacement of egg cytoplasm, establishing an asymmetric distribution of morphogenetic factors. In the early blastula, the yolky vegetal cells along the blastocoel floor are fated to become endoderm. In the midblastula stage (~4,000 cells), the so-called midblastula transition, in which the cell cycle lengthens and zygotic transcription begins, occurs. A blastopore forms by invagination of local endodermal cells on the anterior-dorsal side of the embryo, marking the beginning of gastrulation. As gastrulation proceeds, the endoderm and mesoderm become internalized, and the leading edge of the dorsal-anterior endoderm migrates to the position of the ventral foregut. In the postgastrula Xenopus embryo, the majority of the endoderm mass is located ventrally, and the presumptive organ domains are arranged along the anterior-posterior axis, consistent with their final position in the gut. The mass of endodermal tissue elongates and eventually cavitates to form a primitive gut tube. The digestive tract then undergoes extensive remodeling from larva to adult to adapt from an aquatic herbivorous to a terrestrial carnivorous life.

Nodal signaling in Xenopus is induced in the late blastula stage by the maternally inherited T-box transcription factor VegT, which is localized to the vegetal region, where it directly activates the transcription of Nodal-related genes (Fig. 4A). At this stage, gata4, gata5, and gata6 are induced in the yolky vegetal cells of the prospective endoderm (1, 29, 118). Induction of gata gene expression is abrogated by VegT loss of function (120), indicating that gata expression is dependent on Nodal signaling. Knockdown of Gata6 activity using morpholinos leads to a decrease in endodermal gene expression and defects in gut morphology, while forced ectopic expression of gata4, gata5, or gata6 in ectoderm induces endodermal markers such as sox17, hepatocyte nuclear factor-1B (hnf1β), and foxa2 (1, 118). These data suggest that Nodal signaling acts through gata4/5/6 to establish and maintain endodermal gene expression.

Active migration of the leading edge of mesendoderm across the blastocoel roof of the Xenopus embryo is necessary for the development of tissues such as head mesoderm, heart, blood, and liver. Gata4 and gata6 are expressed in this migratory tissue during gastrulation. Gain-of-function experiments with these Gata factors reveal an induction of cell spreading and migration in nonmotile cells of presumptive ectoderm, while expression of a dominant-negative form of Gata6 severely impairs the ability of the dorsal leading edge of the mesendoderm to spread (40). These studies implicate a critical role for
indicated as inducers of homologs have been identified in mouse. However, injection of mouse otx2 and promotes the expression of Gata5 (96, 111). Both loss- and gain-of-function experiments stages, Nodal activates zebrafish 128) (Fig. 4B). In endodermal progenitors during late blastula bryonic yolk syncytial layer, is thought to activate Nodal maternally inherited VegT directly stimulates tran- scription of Nodal genes in Xenopus, an as-yet-unidentified maternal signal, likely originating from the maternal extraem- bryonic yolk syncytial layer, is thought to activate Nodal signaling in zebrafish (reviewed in Refs. 102, 106, 127, and 128) (Fig. 4B). In endodermal progenitors during late blastula stages, Nodal activates zebrafish faust (fau), which encodes Gata5 (96, 111). Both loss- and gain-of-function experiments reveal that Gata5 regulates the amount of endoderm formed and promotes the expression of sox17 and foxa2, markers of specified endoderm (96, 97). Combinatorial mutant analysis demonstrates that Gata5 cooperates with Bon, a Mix-like homeobox transcription factor, downstream of Nodal signaling and upstream of casanova (a sox17-related gene) to regulate endoderm formation (97). Gata5 and Bon may synergize with eomesodermin, a localized maternal determinant required for endoderm induction, to induce cas, possibly by facilitating the assembly of a transcriptional activation complex on the cas promoter (10).

The lateral-to-medial migration of endoderm during gastrulation, thought to be analogous to the movement of mammalian rostral endoderm as it folds to form the foregut (119), is impaired by selective inhibition of gata5 translation in the yolk syncytial layer, the tissue ortholog of the mammalian extraembryonic visceral endoderm (115). Since extraembryonic GATA4, and not GATA5, is required for ventral folding morphogenesis in mice (59, 79), zebrafish Fau/Gata5 has been suggested to be a functional ortholog of mammalian GATA4 (115). In contrast to fau/gata5 mutants in which endoderm specification is impaired, zebrafish gata6 mutants show normal endoderm specification, differentiation, and morphogenetic movement, but subsequent organogenesis fails, in that the intestine lacks normal epithelial folds, and liver and exocrine pancreatic tissues fail to form (52).

Zebrafish gata6 morphants express early endoderm markers, indicating that the endoderm is specified, but die prior to gut development. Titration of the gata6 morpholino to allow low-level expression of Gata6 and embryonic survival reveals a specific defect in liver development (52), similar to that in mice (125). Interestingly, loss of microRNA-145, which targets gata6 in zebrafish gut smooth muscle, results in defects in smooth muscle function as well as gut maturation (124), implicating a role of mature smooth muscle in promoting epithelial differentiation, and an importance for nonepithelial Gata factors in gut homeostasis, an area that has not been well studied.

Morpholino knockdown, chromatin immunoprecipitation, and GRN analyses revealed that Otx2 is necessary for endoderm specification in zebrafish and likely directly activates gata5 and gata6 (111). Gata5 and gata6 then activate each other, forming an autoregulatory positive-feedback system that locks down endoderm specification. Thus the Otx-Gata regulatory loop found in sea urchin (30) is conserved in fish and possibly frogs.

GATA4 and GATA6 Are Required for Extraembryonic Endoderm Development in Mice

Amniotes, including reptiles, birds, and mammals, are vertebrates whose embryos are totally enclosed in a fluid-filled sac, an evolutionary advantage that enabled breeding away from water. During embryogenesis, amniotes develop extra-embryonic structures, called extraembryonic endoderm, that protect and nourish the developing embryo. Extraembryonic endoderm is distinguished from definitive endoderm, the germ layer precursor of the gut and its accessory organs, although both share many developmental features and molecular markers. Since most of the GATA work has been conducted on mice, this section focuses on the role of GATA factors in murine endoderm development.

Although Nodal signaling is essential for specification of the definitive endoderm in the mouse embryo (101) and GATA...
transcription factors are necessary downstream Nodal effectors of endoderm development in *Xenopus* and zebrafish, it remains uncertain whether GATA factors have an analogous role in the development of the definitive endoderm in mice. However, it is clear that GATA factors are critical for extraembryonic endoderm specification and mouse embryo viability.

After fertilization and zygotic cleavage of the early mouse embryo, a growing mass of undifferentiated cells, the morula, is formed by the second embryonic (E) day (reviewed in Refs. 100 and 109). Each of the eight cells of the morula, called a blastomere, increases its surface contact with its neighbor through a process called compaction, resulting in a polarization of cells. By E3.5, a blastocyst of ~32 cells is established in which ~12 internal cells comprise the new inner cell mass (ICM) and the remaining 20 cells comprise the surrounding trophectoderm (TE), the first cell lineage to be specified in the developing mouse embryo. After implantation at E4, the blastocyst elongates, and the placenta develops around the egg cylinder and begins to provide maternal blood to the embryo. At this stage, the primitive endoderm (PrE) lining the blastocoel cavity is specified from the ICM. The TE and PrE contribute mainly to extraembryonic tissues that form the placenta and yolk sac, respectively, whereas the pluripotent cells of the ICM, called the epiblast (Epi), give rise to all the cell lineages of the future embryo, including the definitive endoderm.

Prior to gastrulation, PrE differentiates into visceral endoderm (VE), which forms from cells that remain in contact with the Epi, and parietal endoderm (PE), which forms from cells that migrate over the inner surface of the TE. VE cells have microvilli and synthesize proteins involved in nutrient transport and also function in patterning the early embryo by participating in primitive streak formation and embryonic mesoderm induction by antagonizing Nodal signaling. Distinct from the VE, the PE cells produce a large amount of extracellular matrix that is deposited along the inner surface of the TE to form Reichert’s membrane, a thick basement membrane-like structure that separates the yolk cavity from the maternal tissues. Extraembryonic endoderm thus provides critical patterning cues and essential maternal-fetal connections required to sustain and protect the embryo after implantation. Disruption of extraembryonic endoderm development results in arrest at gastrulation (34), demonstrating the importance of extraembryonic endoderm for fetal development.

The definitive endoderm, from which the gut and its accessory organs are formed, arises from the Epi at gastrulation (beginning at E6.5). Cells in the Epi undergo EMt, migrate through the primitive streak, and differentiate into mesoderm and endoderm. The presumptive definitive endoderm cells invade and displace cells in the VE, although lineage-tracing experiments suggest that the definitive endoderm at least partly comprises VE cells (60). Morphogenesis begins at E7.5, when the epithelial sheet folds over at the anterior and posterior ends, forming foregut and hindgut pockets, termed the anterior intestinal portal and caudal intestinal portal. The folding continues caudally and rostrally until the folds meet, completing tubulogenesis. During somite stages, the mass of endodermal tissue elongates and cavitates to form a primitive gut tube (E8.5). At E9.5, endodermal organ buds form, and from E9.5 to E13.5, the gut tube lengthens and the circumference increases. From E14.5 to E17.5, the stratified epithelium transitions to a columnar epithelium, cytodifferentiation occurs, and villi form. Crypts develop during the first 2 wk after birth, in contrast to crypt development in humans, which occurs before birth (reviewed in Ref. 81). The small intestine undergoes "re"-differentiation during the weaning transition, when expression of enzymes and transporters shifts to accommodate the change from a milk-based to a solid-food diet.

In contrast to *Xenopus* and zebrafish, in which GATA5 is essential for endoderm and gut development (96, 97, 118), GATA5 is not expressed in mouse embryonic stem (ES) cells (24) or primitive or definitive endoderm (83). Gata5 null mice are sterile and show abnormalities in the female urogenital systems (80), but defects in endoderm or gut development have not been reported. However, Gata5 is upregulated in Gata4 null (Gata4−/−) or Gata6 null (Gata6−/−) ES cells (24), suggesting that GATA5 is redundant with GATA4 and GATA6 during early development. Resolution of the function of GATA5 in endoderm development must await combination knockout studies.

Gata6−/− mice die at E5.5–E6.5 due to defects in PrE formation and subsequent extraembryonic endoderm development (56, 86). Gata6−/− ES cells fail to develop extraembryonic endoderm in vivo and in vitro, and the expression of early and late markers of VE is significantly downregulated or abolished (56, 86), while forced overexpression of Gata6 in ES cells promotes differentiation toward extraembryonic endoderm (41). These data demonstrate that GATA6 plays a critical role in differentiation of the PrE and survival of the embryo past the early primitive streak stage.

A model is emerging for the role of GATA6 in PrE specification (reviewed in Refs. 100 and 109) (Fig. 5). At the eight-cell stage (E2.0), all blastomeres retain the potential to form all cell lineages, but following compaction, segregated expression of Pou domain, class 5, transcription factor 1 (Pou5f1, formerly known as Oct4) and caudal type homeobox 2 (Cdxb2) accompanies specification of the ICM and TE, respectively. At E3.5, the cells of the ICM in the preimplantation blastocyst, although morphologically indistinct, are already heterogeneous. At this stage, Nanog, which encodes a homeodomain protein necessary to maintain pluripotency in ES cells in vitro (77), and Gata6 are expressed in the ICM in a random, mutually exclusive “salt-and-pepper” pattern (26) (Fig. 5A). Lineage-tracing experiments showed that Nanog-positive cells specify Epi, while Gata6-positive cells become PrE. By E4.5, overt segregation of PrE and Epi has occurred in the ICM, in which Gata6-positive PrE cells localize to the surface of the ICM facing the blastocoel cavity, while the Nanog-positive cells of the Epi are confined inside the ICM (Fig. 5A).

The segregation of Epi and PrE lineages in the ICM is regulated by FGF signaling (61), FGF4, FGF receptor (FGFR) 2, and SH2/SH3 adaptor (GR2B), which together activate the MAPK signaling pathway, are necessary for PrE formation (Fig. 5B). In embryos lacking Grb2, Gata6 expression is lost, all ICM cells become Nanog-positive, and no PrE is formed (26), indicating that all ICM cells have adopted the Epi fate at the expense of PrE. ES cells overexpressing a dominant-negative FGFR cannot differentiate into PrE, but this phenotype can be rescued by overexpression of GATA factors (64), suggesting that activation of Gata factor expression is the key PrE-promoting event downstream of FGF signaling. The precise mechanism that underlies Gata6/Nanog segregation is...
unknown, although an inverse correlation in expression of Fgf4 and Fgfr2 in the early ICM (48) suggests a mechanism in which Fgf4 expression from presumably Epi-fated cells generates a lateral inhibitory signal to Fgfr2-expressing, PrE-fated cells (Fig. 5B), resulting in mutually exclusive specification of Epi and PrE.

The sorting mechanism responsible for the relocation of randomly positioned PrE progenitors to the surface of the ICM facing the blastocoel remains unknown. While overexpression of Gata6 in deeper ICM cells did not significantly change their position, a combination of Gata6 and Wnt9A, which encodes a Wnt ligand known to be expressed in the surface ICM, facilitates repositioning of Gata6-expressing cells (76). One of the GATA6-induced genes is disabled homolog-2 (Dab2), which encodes a phosphoprotein involved in endocytosis. Dab2 is detectable in PrE of the implanting blastocyst at E4.5 and is downregulated in Gata6 mutants (85). Homozygous Dab2-deficient mutants are embryonic lethal by E6.5 due to defective relocation of the PrE progenitors (121). Thus DAB2 is an essential downstream effector of GATA6 that may function in PrE progenitor positioning in the blastocyst. The relationship of DAB2 to WNT9A-facilitated sorting remains to be determined.

Although targeted deletion of Gata4 is embryonic lethal at E9.5 (see below), long after the PrE has been specified, GATA4 may still function in PrE development. Cultured Gata4−/− mouse ES cells show a defective formation of PrE (88, 105), while forced overexpression of Gata4 in ES cells is sufficient to induce the proper differentiation program toward extraembryonic endoderm (41). Gata6−/− ES cells reveal a decrease in Gata4 expression (56, 86), while Gata4−/− embryos show an increase in Gata6 expression (59, 79). These data suggest a linear pathway in which GATA4 is downstream of GATA6 in PrE differentiation.

Although GATA6 and, possibly, GATA4 function in PrE specification and, thus, VE and PE formation, little is known about the specific roles of GATA4 or GATA6 in the further differentiation or maintenance of VE or PE. Using embryoid bodies derived from mouse ES cells in which both copies of the Gata4 gene were disrupted, Soudais et al. (105) found that, in contrast to Gata4-sufficient embryoid bodies, which were covered by a layer of VE, Gata4−/− embryoid bodies formed no VE. Gata4−/− ES cells also showed defects in VE development but can be induced to undergo epithelial differentiation by retinoic acid (23). Since Gata4−/− ES cells have the capacity to differentiate along other lineages, it was concluded that Gata4 was specifically required for VE formation (105).

PE produces large amounts of basement membrane components, such as laminin-1 and collagen IV. Mouse F9 embryonal carcinoma cells can be induced to differentiate into PE-like cells, including expression of basement membrane components, by treatment with retinoic acid and dibutyryl cAMP. Silencing of Sox7 or combined silencing of Gata4 and Gata6 results in suppression of PE differentiation. Although overexpression of Sox7 alone was insufficient to induce PE differentiation, overexpression of Gata4 or Gata6 in Sox7-silenced F9 cells restored the differentiation into PE (44). These data suggest that Sox7 is required for the induction of Gata4 and Gata6, and the interplay among these transcription factors plays a crucial role in PE differentiation.

Gata4−/− mice die between E7.5 and E9.5, with a migration and folding defect in ventral morphogenesis that results in a
failure to form a primitive heart tube and foregut (59, 79). At E8.0, Gata4 is expressed in cardiogenic splanchnic mesoderm and in definitive and extraembryonic endoderm. To define the cause of the Gata4^+/- defect, chimeric mice were produced in which Gata4^+/+ cells were restricted to extraembryonic endoderm and small portions of foregut and hindgut definitive endoderm (88) but deleted in all other cells of the developing embryo. Heart and foregut developed normally, indicating that expression of GATA4 in extraembryonic or definitive endoderm, rather than mesoderm, is required for ventral morphogenesis. Tetraploid embryo complementation, in which tetraploid embryos contribute cells to the extraembryonic endoderm, while the fetus, including definitive endoderm, is derived solely from ES cells, was used to produce Gata4^+/- embryos with wild-type extraembryonic endoderm (116). Despite an absence of GATA4 in all other cells of the developing embryo including definitive endoderm, ventral folding morphogenesis and heart tube formation occurred normally, indicating that the GATA4 in extraembryonic endoderm, rather than mesoderm or definitive endoderm, is required for ventral morphogenesis and heart development. Thus, although GATA4 may have overlapping functions with GATA6 in PrE formation at ~E6, it is specifically required later in extraembryonic endoderm for definitive endoderm migration and ventral folding morphogenesis at around E8.

Tetraploid embryo complementation was also used to determine the importance of GATA factors in extraembryonic endoderm for later endoderm organ development. Gata4^+/- embryos with wild-type extraembryonic endoderm form the dorsal pancreas normally but show defects in ventral pancreatic development (117). Gata6^+/- embryos with wild-type extraembryonic endoderm specify hepatic cells normally, but the specified cells fail to differentiate, and the liver bud does not expand (125). In vivo footprinting showed that a DNA-binding site for GATA factors is occupied on a liver-specific, transcrip-tional enhancer of the serum albumin gene long before albumin expression and tissue specification in endoderm development (12). GATA4 was one of a select few factors capable of binding to sites in compacted maternatin (28), suggesting that GATA factors may act as pioneering factors in liver development.

Conditional deletion of Gata factors in the intestinal epithelium has been established using villin-Cre transgenes in which Cre is expressed in the epithelial cells of the small and large intestine, including stem cells (37). Using a tamoxifen-inducible model in which Gata4 was deleted prior to cytodifferentiation and villus formation and during the suckling period, only modest alterations in gene expression, but no overt changes in intestinal development, were noted (13).

In summary, Nodal signaling is essential for endoderm and gut development in vertebrates. In Xenopus and zebrafish, Nodal signaling stimulates the downstream expression of members of the Gata4/S6 subfamily (96, 111, 120), which, in turn, contribute to endoderm development (1, 96, 97, 111, 118). In mice, Gata5 is not necessary for mouse endoderm development (80), while Gata6 is essential for PrE development in the blastocyst (56, 86), possibly for mediating migration of committed PrE cells from the ICM to the blastocoeal surface (reviewed in Refs. 100 and 109). GATA4 likely also functions in PrE development prior to gastrulation, perhaps downstream of GATA6 (56, 59, 79, 86, 100, 109), but its expression in extraembryonic endoderm is essential for ventral folding morphogenesis and primitive heart and gut tube formation after gastrulation (59, 79, 88). The precise role of GATA factors in definitive endoderm remains to be determined. Further investigation using combination Gata deletion and/or alternative Cre drivers is necessary to define the specific and overlapping functions of GATA factors in early intestinal development and their function in definitive endoderm.

GATA4 and GATA6 Show Specific and Redundant Functions in the Mature Mouse Small Intestine

The mature mammalian small intestine is lined by a highly specialized epithelium that exhibits a wide-ranging, yet tightly regulated, functional diversity along its cephalo-caudal axis (Fig. 6A), resulting in an exquisite efficiency in the absorption of dietary nutrients. The functional diversity is linked to a continuous renewal process in which stem cells located at or near the base of the crypts of Lieberkühn produce proliferating transit-amplifying progenitor cells that ultimately differentiate into five principal postmitotic cell types comprising one type of absorptive cell (absorptive enterocytes) and four types of secretory cells (enteroendocrine, goblet, Paneth, and tuft cells) (reviewed in Ref. 90) (Fig. 6B). Absorptive enterocytes, goblet cells, enteroendocrine cells, and tuft cells migrate up the crypt to populate the villi, fingerlike structures that protrude into the lumen to increase absorptive surface area, whereas Paneth cells migrate to the base of crypts. The differentiated cells are eventually removed by apoptosis and/or extrusion. Cells of the villus epithelium turn over in 3–4 days, whereas Paneth cells at the base of crypts turn over in 3–6 wk.

Current models of intestinal epithelial differentiation support the existence of two populations of stem cells, rapidly dividing crypt base columnar cells marked by leucine-rich repeat-containing G-protein coupled receptor 5 (LGR5), and slow-cycling, quiescent cells that reside at the so-called +4 position (reviewed in Refs. 90 and 122). The stem cells produce proliferating transit-amplifying cells that undergo a series of transitions, ultimately giving rise to the differentiated cell lineages of the small intestinal epithelium (Fig. 6C). The Wnt and bone morphogenetic protein (BMP) signaling pathways are implicated in crypt morphogenesis and proliferation, whereas the Notch signaling pathway plays a critical role in determining epithelial cell fate by regulating the balance of absorptive vs. secretory cells (reviewed in Refs. 90 and 122). Notch signaling activates the expression of hairy and enhancer of split 1 (Hes1), which encodes a transcription factor that specifies the absorptive enterocyte lineage. Progenitor cells that elude Notch signaling express atonal homolog 1 (Atoh1, formerly called Mathl), which encodes a transcription factor that designates the secretory lineages. Within secretory progenitors, growth factor-independent 1 (Gfi1) distinguishes enteroendocrine from goblet/Paneth progenitors; neurogenin 3 (Neurog3) specifies the enteroendocrine lineage; sam pointed domain-containing Ets transcription factor (Spdef), a GFI1 target, promotes goblet differentiation; and the Wnt targets SRY box-containing gene 9 (Sox9) and ephrin type B receptor 3 (Ephb3) are necessary for the differentiation of Paneth cells and their localization to the crypt base, respectively. The role
GATA factors and intestinal epithelial homeostasis

of GATA factors in the topographic patterning and regulation of intestinal proliferation, lineage commitment, differentiation, and gene expression in the mature mouse small intestine is the focus of this section.

GATA4/5/6 are expressed in distinct and overlapping patterns in the mature mouse small intestine. Gata4, Gata5, and Gata6 mRNAs are expressed in mature mouse small intestine (2, 62, 82, 83). Along the proximal-distal gradient, Gata4 mRNA is generally constant but declines sharply in the distal ileum, while Gata5 mRNA shows a generally increasing proximal-distal pattern and Gata6 mRNA is generally constant throughout the small intestine (7, 14, 35, 112). The distal decline in Gata4 expression was confirmed for GATA4 protein in mice (14, 112) and humans (14, 49). A tracing approach, in which Cre excision resulted in expression of alkaline phosphatase under the control of the endogenous Gata4 promoter, was used to determine that Gata4 is expressed in the proximal 85% of small intestine and is sharply downregulated in the distal 15% (9), coincident with the distal ileum.

Immunohistochemical analysis of GATA protein expression in crypt and villus compartments, and in specific lineages, is highly variable, probably owing to differences in the specific antibodies used. GATA4 is expressed in the nuclei of epithelial cells in crypts and on villi (14, 16, 32, 49, 112), and containing with proliferation or lineage markers indicates that GATA4 is expressed in proliferating cells in crypts and in absorptive enterocytes on villi, but not in goblet or enteroendocrine cells; it remains unclear whether GATA4 is expressed in differentiated Paneth cells (14, 36, 49); expression of GATA4 in intestinal stem cells or tuft cells has not been reported. GATA5 has been suggested to be specific for secretory cells (36) and has also been reported to be transiently expressed in the nuclei of absorptive enterocytes before weaning in mice (~2 wk of age) (33), but further work is necessary to confirm these patterns of expression. GATA6 is expressed in the nuclei of cells in crypt and villus epithelium in mice (7, 14, 32, 103) and humans (49). GATA6 is expressed in proliferating crypt cells and in all differentiated lineages tested in mice (7). GATA6 is expressed in all but enteroendocrine cells in humans (49), suggesting a possible species difference in intestinal GATA6 expression. It should be noted that the field is hampered by antibodies for GATA5 and GATA6 that cross-react with other proteins and that, unlike GATA4, immunostaining for GATA5 or GATA6 has not been verified using conditional knockout mice as controls.

GATA factors bind and activate intestinal genes. The realization that GATA factors were expressed in the mature small intestine led to a search for putative GATA target genes. Using EMSAs and transient cotransfection assays in cultured cell lines, multiple groups showed that GATA4, GATA5, and/or GATA6 bind and activate the promoters of multiple intestine-specific genes (16, 31, 32, 38, 39, 45, 54, 58). Mutations in these GATA motifs generally abrogated binding and transactivation. Dusing et al. (35) showed that a transgene constructed from the intestine-specific enhancer in the adenosine deaminase (Ada) gene that contains consensus GATA binding sites had high expression levels in murine duodenum, but a homologous transgene with selective mutations in all three enhancer
GATA binding sites had markedly reduced duodenal expression and alterations in the proximal-distal and cell-specific expression patterns, demonstrating the importance of GATA motifs for transcriptional activation in vivo.

**GATA4 mediates regional identities in absorptive enterocyte gene expression and function.** GATA4 activates a subset of absorptive enterocyte genes in the proximal small intestine. Gata4 is expressed in the proximal 85% of small intestine but is not expressed in the distal ileum (9), coinciding with transitions in the expression of subsets of absorptive enterocyte genes (Fig. 7) and suggesting a role in defining the topographic patterning of gene expression in the small intestine. GATA4 was the preferred GATA factor in nuclear extracts from murine jejunum that bound the GATA motifs in the regulatory regions of sucrase-isomaltase (Si) (16) and lactase (Lct) (112), two absorptive enterocyte genes. In isolated Gata4−/− cells in the small intestines of chimeric mice, fatty acid-binding protein 1 (Fabp1) was downregulated (31). Conditional inactivation (14) or deletion (4) of Gata4 using villin-Cre transgenes did not reveal major changes in structure, proliferation, or lineage specification in the small intestine but showed a general downregulation in jejunum of a specific subset of absorptive enterocyte genes, including Lct and Fabp1. Interestingly, Si and fatty acid-binding protein 2 (Fabp2) gene expression was unaffected by conditional Gata4 deletion. Closer scrutiny revealed that while Lct and Fabp1 are not expressed in distal ileum, Si and Fabp2 are expressed in this region of small intestine (57, 108) and, thus, are discordant with Gata4 expression (Fig. 7), suggesting that Si and Fabp2 are not GATA4 targets. Comparison of the global gene expression profiles of control jejunum, control ileum, and Gata4−/− jejunum revealed that Gata4−/− jejunum lost expression of 53% of the “jejunal-specific” genes (i.e., genes normally expressed in jejunum, but not distal ileum) (4). Gata5 and Gata6 were expressed normally when Gata4 was deleted (4, 14, 32), indicating that neither GATA5 nor GATA6 is able to compensate for GATA4 loss.

Multiple studies support a mechanism in which GATA4 cooperates with HNF1α to activate specific absorptive enterocyte genes. In transient cotransfection experiments, GATA4 cooperates with HNF1α (112, 113), indicating that neither GATA5 nor GATA6 is able to compensate for GATA4 deletion. Conditional deletion of Gata4 also revealed an upregulation in jejunum of absorptive enterocyte genes, the expression of which is normally restricted to the distal ileum (4, 9, 14), including solute carrier family 1 member A2 (Slc10a2), which encodes the apical sodium-dependent bile acid transporter (Asbt) and fatty acid-binding protein 6 (Fabp6), which encodes the ileal lipid-binding protein (Ilbp). ASBT was induced on villi within 24 h of induction of Gata4 deletion, indicating that GATA4 represses Slc10a2 gene expression within differentiated absorptive enterocytes, rather than by a process that originates in progenitor cells in crypts (8). Zinc finger protein, multitype 1 (Zfpml1, also known as friend of GATA type 1 [Fog1]) is coexpressed with Gata4 in absorptive enterocytes and, like Gata4, shows a declining proximal-distal pattern of expression (8). Conditional expression of a knockin Gata4 mutant that is unable to bind FOG cofactors, but otherwise functions normally, leads to a significant upregulation of Slc10a2 expression (8), suggesting that FOG1 may participate in the repressive effect of GATA4 on certain target genes. Comparison of the global gene expression profiles of control jejunum, control ileum, and Gata4−/− jejunum revealed that Gata4−/− jejunum gained expression of 47% of the ileal-specific genes (genes normally expressed in distal ileum, but not jejunum) (4).

Together, these findings indicate that GATA4 has a specific function, distinct from other intestinal GATA factors (GATA4-specific pathway), in which it both activates and represses subsets of absorptive enterocyte genes in the proximal intestine (Fig. 8). GATA4 activates a discrete subset of genes that is normally not expressed in the distal ileum (“jejunal” gene set), likely in cooperation with HNF1α, and represses a specific subset of genes in the proximal intestine that are normally expressed in the distal ileum (“ileal” gene set), possibly in cooperation with HNF1α. Although it remains to be determined how GATA4 regulates its target genes, whether by directly binding their promoters and/or enhancers or by indirect mechanisms, GATA4 nonetheless is responsible for maintaining the regional identity in absorptive enterocyte gene expression between the proximal intestine and the distal ileum.

**GATA4 maintains regional function in the mature small intestine.** Conditional Gata4 deletion also results in alterations in intestinal physiology. Using a conditional, noninducible knockout approach in which Gata4 is inactivated early in intestinal development (E12.5), Battle et al. (4) showed that the intestinal deletion of Gata4 resulted in smaller mice that weighed consistently less than their gender-matched littermate
controls. Lipid metabolism was identified as the molecular/cellular function most affected by loss of GATA4 in the intestinal epithelium. Plasma levels of glucose were not different, but the levels of cholesterol and phospholipids were significantly reduced in the mice lacking GATA4 in their intestine. These mice also showed a 21% reduction of fat absorption and a 93% reduction of cholesterol absorption. Global profiling analysis of jejunum showed a downregulation of genes associated with lipid transport and an upregulation of genes critical for bile acid uptake, including Fgf15, a key bile acid signaling molecule. Hepatic cholesterol 7 alpha-hydroxylase (Cyp7a1) expression was downregulated in liver, supporting a loss of jejunal-specific functions: 1) absorptive enterocyte gene expression by activating lipid-binding proteins and apolipoproteins ("lithic gene set") and repressing specific colonic genes ("colonic gene set"), 2) crypt cell proliferation, likely in cooperation with CDX2, 3) enteroendocrine cell commitment by promoting Neurog3 expression, and 4) Paneth cell differentiation by promoting a terminal Paneth program while inhibiting a goblet program in committed (SOX9-, EPHB3-positive) Paneth progenitors. GATA factors are indicated in yellow and non-GATA factors in blue.

Using a conditional, inducible model, we showed that proximal bile acid absorption was indeed induced by intestinal Gata4 deletion (9). Although total bile acid excretion remained unchanged, taurocholate transport from the mucosal to the serosal fluid in everted gut sacs was increased in proximal segments and bile acid concentration in luminal contents was depleted in distal segments when intestinal Gata4 was deleted, confirming a proximal induction of bile acid absorption. Bile acid pool size was not different, but, owing to a relative increase in absorption of specific bile acids along the length of the small intestine, the makeup of the pool was shifted to a more hydrophobic composition. These data were the first to show that bile acid absorption can be induced in absorptive enterocytes that normally do not exhibit this characteristic and that the restriction of Slc10a2 expression and active bile acid absorption to the distal small intestine by GATA4 plays a key role in determining the composition, but not the size, of the bile acid pool.

We also tested the hypothesis that an induction of Slc10a2 gene expression and bile acid absorption in proximal small intestine by conditional mutation or deletion of Gata4 is sufficient to correct bile acid malabsorption resulting from ileocecal resection (ICR) (9). We utilized two Gata4 mutant models, each producing a moderate (22% of ileal levels) or high (69% of ileal levels) proximal induction of Slc10a2 mRNA expression. ICR in wild-type mice resulted in an anticipated increase in bile acid excretion, reduction of the bile acid pool size, and compensatory increase in transcription of hepatic bile acid biosynthetic enzymes compared with sham-operated controls, reflecting bile acid malabsorption. ICR in the Gata4 mutant mice showed an induction of bile acid absorption in the proximal small intestine that was sufficient to reduce bile acid excretion, maintain the bile acid pool size, and reduce the compensatory upregulation of hepatic bile acid biosynthetic enzymes. Lipid absorption was not examined in this study. This is the first example of a nontransplant intervention capable of restoring intestinal bile acid absorptive function following ileectomy and establishes the concept that reducing GATA4 activity may be clinically useful for restoring lost ileal function due to disease or resection.

GATA4 and GATA6 modulate crypt cell proliferation, secretory cell differentiation, and absorptive enteroocyte gene expression. To define the function of GATA6 in the small intestine, we compared the intestinal phenotype produced among single and double Gata4/Gata6 conditional knockout mice using the villin-CreERT2 driver (7). In the distal ileum, where Gata6 is expressed but Gata4 is not (Fig. 7), conditional deletion of Gata6 resulted in a decrease in crypt cell proliferation, decreases in differentiated enteroendocrine and Paneth cells, and alterations in absorptive enterocyte gene expression. This altered phenotype was not present in the proximal intestine, where Gata4 is expressed. Instead, Paneth cells were increased, perhaps as a compensatory response to the loss of Paneth cells in distal ileum. When Gata4 and Gata6 (herein, Gata4/Gata6) were conditionally deleted, the proximal intestine showed all the changes in proliferation, differentiation, and gene expression found in the distal ileum of single Gata6 conditional knockout mice. These data indicate that, in addition to the GATA4-specific functions in maintaining jejunalodeal distinctions in absorptive enterocyte gene expression, GATA4 and GATA6 share common functions (GATA4/GATA6-redundant pathway) in regulating proliferation, differentiation, and gene expression in the mature small intestine (Fig. 8). Cellular proliferation in intestinal crypts is necessary for the continuous renewal of the intestinal epithelium. Gata4/Gata6
deletion results in a reduction of the number of Ki67- and bromodeoxyuridine-positive cells in crypts, as well as a decrease in villus height and epithelial cell number, demonstrating a requirement for GATA factors in supporting intestinal renewal and the maintenance of the absorptive surface area (7). In Caco-2 cells, an intestinal cell line that proliferates rapidly until reaching confluence and then undergoes differentiation, CDX2, a key transcriptional regulator of intestinal genes, occupies a specific set of gene loci in proliferating cells but a different set in differentiated cells (114). GATA motifs were the most highly represented motif (other than CDX motifs) in the CDX2 occupancy sites in proliferating cells, and GATA6 was found to preferentially co-occupy these sites, suggesting that CDX2 co-regulates intestinal proliferation with GATA4/GATA6 (Fig. 8).

Conditional deletion of Gata4/Gata6 also showed an altered secretory cell phenotype. Enteroendocrine cell number and Neurog3 mRNA abundance were decreased, suggesting a GATA requirement for enteroendocrine cell specification in ATOH1-secretory progenitors. The number and morphology of goblet cells on villi were normal, but Paneth cells were re-ATOH1-secretory progenitors. The number and morphology of GATA requirement for enteroendocrine cell specification in Neurog3 secretory cell phenotype. Enteroendocrine cell number and GATA6 (Fig. 8).

GATA4/GATA6 also activates and represses specific absorptive enterocyte genes (Fig. 8). Many of the genes downregulated by Gata4/Gata6 deletion encode lipid transporters and apolipoproteins (“lipid” gene set) but are distinct from the GATA4-specific lipid transporters and carriers discussed above. Many of the genes upregulated by conditional Gata4/ Gata6 deletion are normally more highly expressed in colon than small intestine (“colonic” gene set). Thus the GATA4-specific and GATA4/GATA6-redundant pathways both function in the proximal-distal patterning of absorptive enterocyte gene expression.

In summary, these studies outline two fundamental pathways of GATA regulation in the mature mouse small intestine, a GATA4-specific pathway and a GATA4/GATA6-redundant pathway (Fig. 8). In the GATA4-specific pathway, GATA4 regulates the proximal-distal transcriptome in the gastrointestinal tract by distinguishing jejunal vs. ileal gene expression and function (4, 9, 14). GATA6 is unable to compensate for this pathway, supporting a mechanism in which GATA4 is specifically selected by its target genes for regulation. In the GATA4/GATA6-redundant pathway, GATA4 or GATA6 is capable of regulating multiple processes, including proliferation, lineage specification, and terminal differentiation (7). These findings indicate that GATA target genes in the intestinal epithelium are regulated by mechanisms that involve differential recruitment of specific GATA factors.

**GATA6 Regulates Colonic Epithelial Homeostasis**

Although the focus of this review is on the role of GATA factors in small intestinal homeostasis, we recently reported that GATA6 is necessary for normal colonic epithelial differentiation (6). The colonic epithelium, also a derivative of endoderm, has crypts and a surface epithelium, but no villi. It is maintained through a process of continuous cellular renewal in which crypt stem cells give rise to three differentiated cell lineages, colonocytes, goblet cells, and enteroendocrine cells. Although neither GATA4 nor GATA5 is expressed in colon, GATA6 is expressed in colonic epithelium of humans (49, 53) and mice (6). Using improved GATA6 antibodies, validated using conditional Gata6 knockout mice, we showed that GATA6 is expressed in all proliferating and differentiated cells in the mature mouse colon (6). Conditional, inducible deletion of Gata6 in the small intestinal and colonic epithelium using the villin-CreERT2 driver resulted in an altered crypt structure, a decrease in crypt proliferation, and a delayed crypt-to-surface epithelium migration rate (6). Specific colonocyte genes were significantly downregulated, goblet cells were immature, and the mRNAs for specific hormones produced by endocrine cells were altered. Gata6 knockdown in human LS174T cells, a colonic adenocarcinoma cell line, resulted in the downregulation of a similar subset of colonocyte genes. GATA6 occupied active loci in enhancers and promoters of some of these genes, suggesting that they are direct targets of GATA6. These data demonstrate that GATA6 is necessary for proliferation, migration, lineage maturation, and gene expression in the mature mouse colonic epithelium. A role in supporting colonic epithelial proliferation is consistent with the suggestion that GATA6 may promote colorectal cancer (reviewed in Ref. 3).

**Summary and Perspectives**

In this review, we show that GATA factors are essential for the earliest designation of presumptive endoderm and the establishment and maintenance of homeostasis in the mature small intestine. Invertebrates demonstrate sequential expression of divergent GATA factors in which early orthologs essential for endoderm development are extinguished, giving way to other orthologs critical for later intestinal development. Vertebrates, on the other hand, show a continuum of expression of the same members of the Gata4/5/6 subfamily throughout development, suggesting an evolutionary advance in the ability of the same GATA factor to regulate different functions dependent on the developmental time frame. GATA6 is essential for the earliest designation of PrE, possibly by regulating the migration of PrE-committed cells to the blastocoel surface in the early blastocyst. GATA4 may also participate in PrE development, along with GATA6, but is essential for ventral folding morphogenesis after gastrulation. In vertebrates and invertebrates, cell migration of presumptive or established endoderm is emerging as a key function of GATA factors in embryogenesis. The role of GATA factors in the differentiation of PrE into VE and PE (23, 44, 105) and in the regulation of cytodifferentiation and villus formation (13) has received some attention, but additional work using combination knockouts with different Cre drivers is necessary to establish the individ-
ual and overlapping functions at these stages of development. Conditional knockout approaches in adult mice have revealed two fundamental pathways of GATA regulation: 1) a GATA4-specific pathway that functions to demarcate the differences in absorptive enterocyte gene expression between proximal intestine and distal ileum (4, 9, 14) and 2) a GATA4/GATA6-redundant pathway that regulates crypt cell proliferation, secretory cell differentiation, and absorptive enterocyte gene expression (7).

While the general function of GATA factors at various stages of development can be gleaned from the downstream effect of Gata deletion, we do not yet know the first-line direct targets of individual GATA factors that put these downstream outcomes in play or how GATA factors endow embryonic germ layers to gain competence to differentiate into distinct cell types or how they modulate chromatin in vivo to control transcriptional activation or repression. Almost 15 years ago, Zaret utilized in vivo footprinting to show that GATA factors occupy GATA motifs in chromatin in genes silent in endoderm, providing these genes with the potential to be activated later in development, a phenomenon called genetic potentiation (reviewed in Ref. 123). Using chromatin pull-down assays and global epigenetics approaches in vivo, we now have the ability to determine the occupancy sites of GATA factors within chromatin in vivo and map these sites to genes that change in gene knockout studies, providing a reasonable account of first-line direct targets of specific GATA factors at various stages of development. We can also characterize histone modifications at GATA occupancy sites and the effect of GATA deletion on these modifications, establishing a fundamental basis of how GATA factors modulate chromatin to effect transcriptional potentiation, activation, or repression. Finally, characterization of DNA motifs and occupancy of other transcriptional regulators at GATA-occupancy sites, as well as identification of partner proteins that physically interact with GATA factors using immunoprecipitation/proteomics techniques, will allow us to define distinct pathways of GATA regulation. Application of these techniques at different developmental time frames will establish the fundamental underpinnings of GATA regulation throughout the continuum of intestinal development.

ACKNOWLEDGMENTS
We thank Drs. E. Beuling, T. Bosse, M. P. Verzi, and R. J. Grand for reading the manuscript and providing helpful suggestions.

GRANTS
This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases Grant RO1-DK-061382 to S. D. Krasinski and grants from the Nutricia Research Foundation, KWF Kankerbestrijding, Prins Bernhard Cultuurfonds (The Netherlands), and the European Society for Pediatric Research (Switzerland) to B. E. Aronson.

DISCLOSURES
No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS
B.E.A. and S.D.K. drafted the manuscript; K.A.S. and S.D.K. prepared the figures; B.E.A., K.A.S., and S.D.K. edited and revised the manuscript and approved the final version of the manuscript.

REFERENCES
11. Boeck ME, Boyle T, Bao Z, Murray J, Mericle B, Waterston R. Conditional knockout approaches in adult mice have revealed two fundamental pathways of GATA regulation. Application of these techniques at different developmental time frames will establish the fundamental underpinnings of GATA regulation throughout the continuum of intestinal development.

Downloaded from http://ajpgi.physiology.org/ by 10.220.32.246 on June 9, 2017
Review

G488 GATA FACTORS AND INTESTINAL EPITHELIAL HOMEOSTASIS


