CALL FOR PAPERS | Comparative Genomics

How form follows functional genomics: gene expression profiling gastric epithelial cells with a particular discourse on the parietal cell

Benjamin J. Capoccia,1 Won Jae Huh,1 and Jason C. Mills1,2

Departments of 1Pathology and Immunology and 2Developmental Biology, Washington University School of Medicine, St. Louis, Missouri

Submitted 30 December 2008; accepted in final form 8 February 2009

Capoccia BJ, Huh WJ, Mills JC. How form follows functional genomics: gene expression profiling gastric epithelial cells with a particular discourse on the parietal cell. Physiol Genomics 37: 67–78, 2009. First published February 10, 2009; doi:10.1152/physiolgenomics.90408.2008.—The cellular composition and morphology of the stomach epithelium have been described in detail; however, the molecular mechanisms that regulate the differentiation of the various cell lineages as well as the function of mature gastric cells are far less clear. Recently, dissection of the molecular anatomy of the stomach has been boosted by the advent of functional genomics, which allows investigators to determine patterns of gene expression across virtually the entire cellular transcriptome. In this review, we discuss the impact of functional genomic studies on the understanding of gastric epithelial physiology. We show how functional genomic studies have uncovered genes that are useful as new cell lineage-specific markers of differentiation and provide new insights into cell physiology. For example, vascular endothelial growth factor B (Vegfb) has been identified as a parietal cell-specific marker that may allow parietal cells to regulate the mucosal vascular network. We also discuss how functional genomics has identified aberrantly expressed genes in disease states. Human epididymis 4 (HE4), for example, was recently identified as a metaplasia-induced gene product in mice based on microarray analysis. Finally, we will examine how analysis of higher-order patterns of gene expression can go beyond simply identifying individual genes to show how cells work as integrated systems. Specifically, we show how application of a Gene Ontology (GO) analysis of gene expression patterns from multiple tissues identifies the gastric parietal cell as an outlier, unlike other differentiated cell lineages in the stomach or elsewhere in the body.

stomach; cancer; stem cell

THE GENOMIC ERA HAS USHERED in a new understanding of the morphological and functional properties of stem cells and their lineages. Where the anatomist describes the heterogeneity of cell lineages based upon observation of tissue architecture and organelle distribution within the cell, the genomicist seeks to define the “molecular anatomy” of a tissue. As these two fields converge, the characteristic gene expression patterns of different cell lineages can be used as a molecular roadmap to discover new physiological functions of differentiated cells and to uncover the sequence of differentiation steps a stem cell takes to become a mature cell. In this review, we will discuss how the field of functional genomics has contributed specifically to our understanding of stomach physiology.

The microanatomy of the corpus of the stomach has been well charted. It comprises a network of glandular invaginations known as gastric units, which function to secrete acid, digestive enzymes, and mucus into the gastric lumen. Pioneering studies by Karam and Leblond (33–37) have shown that the epithelium lining the murine gastric unit is constantly being renewed throughout life. This renewal is driven by a multipotent gastric stem cell that resides within the middle of the unit in an area called the isthmus (Fig. 1). Tritiated thymidine autoradiography and electron microscopy studies have shown that these highly proliferative, granule-free stem cells give rise to three separate daughter cells: the presurface mucous cells (aka prepit cells), pre-neck, and preparietal cells (34–36). These early daughter cells maintain several stem cell features including a high nucleus-to-cytoplasm ratio, small mitochondria, low secretory activity, and a relatively high level of free ribosomes. However, as these committed daughter cells differentiate, they begin to develop a more mature phenotype indicated by increased cytoplasm-to-nucleus ratio, larger mitochondria, and increased secretory activity indicated by an expanding trans Golgi network, increased rough ER, and larger, more abundant secretory granules.
Unique to each cell lineage are the secretory proteins they produce and the direction in which they migrate from the isthmus. The surface mucous cell precursor migrates apically toward the gastric lumen to form an epithelial layer of mucus-secreting cells ("surface-associated mucous cells"; see Fig. 1) (35, 81). On the other hand, the neck cell precursor migrates toward the base of the unit to help form a region of mucus-secreting neck cells called the neck zone. As neck cells descend further toward the base of the unit, they mature into digestive enzyme-secreting zymogenic (chief) cells that constitute the primary cell type in the base zone of the gastric unit (7, 36, 68). Unlike the pit and neck cell lineages, parietal cells mature within the isthmus to become acid-secreting parietal cells; the vast majority of which then migrate away from the surface into the neck, although in rodents, some can also be found nearer the surface (see Figs. 1, 2) (33). Additionally, several types of endocrine cells can be found scattered throughout the different regions of the gastric unit, providing feedback signals in response to changes in the tissue environment (31, 37, 44). Each cell lineage must execute a specific molecular program to allow for the proper anatomy and function of the gastric unit.

We here define functional genomics as the method whereby large-scale, unbiased analyses of gene expression are used to explain the function of biological systems. One goal of functional genomics is to generate searchable databases of genes expressed in normal cells so that global patterns of biological function can be inferred; in other words, molecular anatomy can provide clues about physiology. Additionally, such data can be used to detect and define disordered cellular function (pathophysiology) in disease states, with the expectation that such information will lead to earlier and more accurate diagnoses, to more specific therapies, and to more precise monitoring of therapeutic responses. In general, the purpose of functional genomic studies of the normal stomach mucosa has been to determine: 1) the global changes in gene expression that occur as each lineage progresses toward its fate and 2) the rare regulatory genes responsible for inducing these lineage-specific global expression changes. Additional studies have tried to uncover molecular pathways expressed during aberrant patterns of differentiation (i.e., metaplasia, cancer) that occur in the setting of inflammation and injury. In this review, we will follow the impact of functional genomics on our molecular understanding of the stomach. We will show how identification of new patterns of gene expression in different lineages has provided not only novel cell lineage-specific markers but also new insights into gastric physiology and pathophysiology. Finally, we will show how recent bioinformatic approaches allow us to analyze multiple functional genomic analyses at once and draw conclusions not just about changes in individual gene expression but about the higher-order patterns and networks of gene expression that define different cell lineages.

CELL LINEAGE MARKERS AND THEIR FUNCTION

In this section we will examine how functional genomics has helped to identify novel molecular markers of each gastric epithelial lineage within the corpus of the stomach and how identifying these genes contributes to our understanding of the overall function of the stomach. It should be noted that there have not been published functional genomic studies of specific lineages within the adult antrum of the stomach. This is somewhat surprising, because 1) most human tumors and most tumors arising in mouse models of gastric neoplasia appear to originate in the antrum or at the junction between antrum and corpus and 2) one common
form of metaplastic, preneoplastic transformation of the corpus epithelial cell differentiation pattern is characterized by an antrum-like pattern (see below).

Several methods have been taken to generate the gene expression profiles that we will review in this section (Table 1). Two basic approaches have been used to purify cell populations to profile: 1) tissue was mechanically and enzymatically dispersed and the cell lineage of interest isolated using differential sedimentation properties or known surface lineage-specific markers, and 2) cells were laser-capture microdissected (LCM) from their niches, taking advantage of the well-ordered layout of cell lineages in the gastric unit. The first approach has the advantage that higher purity of target cells can be achieved, but it depends on having existing cell lineage-specific markers and involves considerable disruption of tissue architecture with concomitant stress on the isolated cells that would be expected to lead to confounding changes in gene expression (26). The laser-capture approach, on the other hand, because it involves assaying expression of genes in cells in their nearly native state, generates assays of gene expression that are much more faithful to the in vivo state; however, cells cannot be isolated to the same degree of purity, and, because cell yields are lower, many lower abundance transcripts cannot be detected.

Parietal Cells

Parietal cells (PCs) are a unique cell lineage, not only with respect to the other lineages in the stomach but even with respect to the body as a whole (Figs. 1–3). As opposed to the presurface mucous and preneck progenitor cells, preparietal cells fully differentiate within the isthmus and then migrate as mature PCs predominantly into the neck zone, though PCs can also be found in the base and, in mice, above the isthmus among upwardly migrating surface mucous cells (7, 33). Historically, PCs have been best characterized for the role they play in acid secretion; however, it has become clear that they also play a critical role in maintaining normal patterns of cell differentiation and progenitor cell proliferation (5, 7, 28–30, 33, 50, 76, 77). Genetically engineered ablation of PCs in transgenic mice blocks terminal differentiation of zymogenic cells (ZC), causes premature differentiation of neck cells, and increases proliferation of the multipotent stem cell and its committed daughters (7, 50, 51, 77). Furthermore, PC loss has been shown to be strongly linked to the development of gastric cancer (61, 63, 64, 80). Given the importance of acid secretion in physiology and pathophysiology and given the importance of these cells as governors of differentiation of other gastric cell lineages, the PC has, not surprisingly, been the most studied of the gastric cell lineages and the first cell to be selected for functional genomic analysis.

The first functional genomic analysis of PCs used a cell sorting, magnetic-bead-based approach for purification, and gene expression was determined using early generation Affymetrix GeneChip technology. The resulting gene expression profile showed that PCs preferentially expressed many genes that correlated with their known physiological role. For PCs to pump protons into the stomach lumen, they require an abundant energy source. Accordingly, a good portion of the PC transcriptome is devoted to cellular energy metabolism and mitochondria maintenance (55). Furthermore, genes regulating vesicular trafficking (ADP-ribosylation factor 1, Arf1), free radical scavenging (mitochondrial superoxide dismutase 2, Sod2), cytoskeleton remodeling (Mu-protocadherin, Mucedhi), lipid metabolism (fatty acid desaturase, Fads1), and calcium signaling (calmodulin-2, Calif2), all characteristic of a highly metabolically active secretory cell with abundant membrane remodeling requirements, were preferentially upregulated in PCs (55).

More recently, Lambrecht and colleagues (43) performed a more extensive gene expression profile using a 22K rat oligonucleotide expression microarray of H+/K+-ATPase α-subunit expressing PCs isolated by fluorescence-activated cell sorting (FACS). Their results confirmed that PCs devote considerable transcriptional energy to proton transport. They also added to our understanding of PC physiology by identifying substantial PC enrichment for transcripts of the K+ channel, Kcnq1 (43). They propose that Kcnq1, with its regulatory subunit, Kcne2, might serve as the principal luminal K+ efflux channel based on follow-up studies using antibodies showing localization of Kcnq1 protein to the PC canalicular secretory membrane. These findings jibe with other reports that Kcne2 and Kcnq1 protein are expressed in human and rat PCs and in human PCs (18, 22, 41). Perhaps because of its role in PC homeostasis and in turn the role of PCs in regulating overall mucosal homeostasis, deletion mutants of Kcnq1 in both mice and rats show altered zymogenic lineage differentiation and hyperproliferation of the gastric progenitor cells, as evinced by mice and rats with an intragenic Kcnq1 deletion (23, 41, 47).

In addition to helping us understand better how PCs secrete acid, PC gene expression profiling has also led to a better understanding of how PCs might regulate differentiation within gastric units. Jain and coworkers (29) isolated H+/K+-ATPase α-subunit expressing PCs from gastrin-deficient mice (GAS-KO) by FACS and profiled their gene expression using the U74AV2 Affymetrix chip. They have shown that in addition to regulating acid production, gastrin regulates expression and secretion of growth factors from PCs including heparin-binding epidermal growth factor (Hbegf). Their work also confirms the earlier study by Gordon and co-workers (55), showing that PCs preferentially express insulin-like growth factor binding protein-2 (Igfbp2) and PTH-like peptide (Pthlh), indicating possible mechanisms by which PCs could directly control cell division and differentiation. In addition, the GAS-KO study indicated an increased expression of several Wnt and Myc target genes, suggesting that GAS-KO PCs are functionally immature and that gastrin signaling is necessary for their maturation (29). Interestingly, Vegfb was enriched in PCs as well, suggesting that, in addition to providing a direct signal for cell proliferation, these cells also have the ability to provide or maintain blood supply to the gastric unit (7, 31, 55). Figure 2 shows how Vegfb protein and a known marker for PCs, H+/K+-ATPase B-subunit, both stain specifically PCs within the gastric unit of mice; a similar pattern is seen in human mucosa (unpublished observations by the authors). Collectively, the studies show how molecular profiles of the PC can elucidate not only additional facets of their known physiological function in ability to facilitate gastric acid production but also provide novel insight into the molecular pathways by which PCs mediate the growth and differentiation of gastric epithelial progenitor (GEP) cells, a PC function that, as we will see again, has a significant impact in disease.
Table 1. Functional genomic studies of the gastric epithelium

<table>
<thead>
<tr>
<th>Method of Isolation</th>
<th>Species</th>
<th>Gene Expression Platform</th>
<th>Example Molecules Characterized</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parietal Cells</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mills et al., PNAS, 2001 (55)</td>
<td>mouse</td>
<td>Mullk Affymetrix GeneChip</td>
<td>Igfbp2, Pthlh, Vegfb, Aqp4, Kcnq1</td>
</tr>
<tr>
<td>Lambrecht et al., Physiol Genomics, 2005 (43)</td>
<td>rat</td>
<td>22K rat oligonucleotide expression array</td>
<td>Kcnq1, Apelin (Apln)</td>
</tr>
<tr>
<td>Jain et al., Physiol Genomics, 2006 (29)</td>
<td>rat</td>
<td>U74AV2 Affymetrix GeneChip</td>
<td>Kcnq1, Aqp4, Crb, Pthlh</td>
</tr>
<tr>
<td>Mueller et al., Gastroenterology, 2004 (58)</td>
<td>rat</td>
<td>spotted mouse cDNA microarray</td>
<td>Atp5b, Vegfb</td>
</tr>
<tr>
<td><strong>Gastric Epithelial Progenitors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mills et al., PNAS, 2002 (53)</td>
<td>mouse</td>
<td>Mullk Affymetrix GeneChip</td>
<td>Igfl, Magoh, Akt</td>
</tr>
<tr>
<td><strong>Endocrine Cells</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andersson et al., Biochem Biophys Res Commun, 2005 (3)</td>
<td>mouse</td>
<td>mouse 430 Affymetrix GeneChip</td>
<td>chromogranin A, histidine decarboxylase, Vmat2</td>
</tr>
<tr>
<td>Lambrecht et al., Physiol Genomics, 2006 (44)</td>
<td>rat</td>
<td>22K rat oligonucleotide expression array</td>
<td>Apelin receptor (Aplnr)</td>
</tr>
<tr>
<td>Jain et al., Physiol Genomics, 2007 (31)</td>
<td>rat</td>
<td>44K whole rat genome expression oligo array</td>
<td>Vmat2, Hdc</td>
</tr>
<tr>
<td>Lambrecht et al., Physiol Genomics, 2007 (42)</td>
<td>rat</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Zymogenic Cells</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramsey et al., Development, 2007 (68)</td>
<td>mouse</td>
<td>mouse 430 2.0 Affymetrix GeneChip</td>
<td>Mistl (Bhlhb8)</td>
</tr>
<tr>
<td>Mueller et al., Gastroenterology, 2004 (58)</td>
<td>rat</td>
<td>spotted mouse cDNA microarray</td>
<td>pepsinogen C, phospholipase A2 group 1B</td>
</tr>
<tr>
<td>Nozaki et al., Gastroenterology, 2008 (61)</td>
<td>rat</td>
<td>mouse 430 2.0 Affymetrix GeneChip</td>
<td>HE4 (Wfdc2)<em>, Mcm3</em></td>
</tr>
<tr>
<td><strong>Early Development</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verzi et al., Gastroenterology, 2008 (81)</td>
<td>mouse</td>
<td>mouse 430 2.0 Affymetrix GeneChip</td>
<td>Foxq1</td>
</tr>
<tr>
<td>Choi et al., Mol Cell Biol, 2008 (12)</td>
<td>mouse</td>
<td>SAGE (Serial Analysis of Gene Expression)</td>
<td>Nkx6-3</td>
</tr>
<tr>
<td>Choi et al., Development, 2006 (11)</td>
<td>mouse</td>
<td>targeted, transcription factor-specific RT-PCR</td>
<td>Mistl, MathI</td>
</tr>
<tr>
<td><strong>Surface Mucous Pit Cells</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramsey et al., Development, 2007 (68)</td>
<td>mouse</td>
<td>Mouse 430 2.0 Affymetrix GeneChip</td>
<td>Foxq1, Cdhl, Grb-2</td>
</tr>
<tr>
<td>Mueller et al., Gastroenterology, 2004 (58)</td>
<td>mouse</td>
<td>spotted mouse cDNA microarray</td>
<td></td>
</tr>
</tbody>
</table>

*Genes upregulated in zymogenic cell metaplasia.
GEPs constitute the total epithelial population, and mRNA processing and cytoplasmic localization (Mago-m, splicingosome components, nuclear pore components, microtubule-associated proteins). They interpreted this pattern of gene expression to indicate that epithelial cell progenitors rely heavily on maintaining graded states of differentiation where the protein components required for the function of the mature cell are held in a state of “readiness” at the level of the mRNA until a signal is given for their translation into protein (53). It will be interesting to determine whether GEPs regulate translation through storage of untranslated transcripts that await translation as they commit to terminal cell lineages. Another intriguing clue about molecular regulation of gastric epithelial differentiation generated from the GEP gene expression profiles is as follows. Given that expression profiling indicated that PCs elaborate a factor (Igfbp2) that would decrease delivery of Igf to the gastric unit (55) and that GEPs elaborate both Igf and the receptor for Igf, one could hypothesize that regulation of growth within the gastric unit could depend on communication between the PC and GEP through Igf signaling.

Endocrine Cells

In any secretory tissue, the ability to adapt cell function to changing environmental stress is essential for maintaining homeostasis. In the stomach, the endocrine signaling pathways that regulate acid secretion from PCs have been extensively characterized. The interplay between G cells in the distal portion of the stomach and both parietal and enterochromaffin-like-(ECL) cells in the corpus is mediated by the gastric hormone gastrin. Gastrin, secreted by G cells, can directly stimulate acid secretion by binding to gastrin/cholecystokinin (CCK)-B receptors on PCs and indirectly by stimulating histamine release from ECL cells (9).

The ECL cell is the cell that most proximately regulates acid secretion in the stomach, given that, in response to gastrin, it releases histamine, which is the most potent activator of PC acid secretion (2, 9, 24, 30, 52, 74, 86). One of the first studies to present a functional analysis of ECL cells was performed by Andersson et al. (3). In this study, ECL cells were activated by treatment of mice with omeprazole for 7 days, and the resulting gene expression of the whole corpus mucosa was determined using mouse expression set 430 Affymetrix GeneChip. ECL cells were shown to preferentially express genes involved in secretory granule biogenesis [chromogranin A (Chga), synaptophysin (Sypr)], histamine synthesis (histidine decarboxylase, Hdc), membrane trafficking (secretory carrier membrane protein 1, Scamp1), and in the general secretory process (Vmat2) (3). Interestingly, these studies also indicate Pthlh as an important paracrine signal expressed in ECL cells (29, 31, 55), though, as mentioned, other groups have found this gene to be preferentially expressed in PCs, suggesting that this may have been due to PC contamination in the ECL study.

More recently, Lambrecht and co-workers (44) carried out a detailed functional analysis of rat ECL cells that were isolated by density gradient and then sorted based upon their ability to accumulate the fluorescent cationic dye acridine orange by FACs. Gene expression analysis was determined using a 22K rat oligonucleotide expression array and compared with expression profiles of all gastric epithelial cells as well as a pure population of PCs. From this analysis, six functional groups

---

**Fig. 2.** Vegfb expression in PCs. Neck zones representative of a single mouse stomach are depicted. A: Vegfb expression in green, determined using goat-anti-Vegfb from Santa Cruz Biotechnology. The experiments were performed according to protocols approved by the Washington University School of Medicine Animal Studies Committee. B: merge of Vegfb in PCs with anti-H-K-ATPase B-subunit (red).
were represented among the list of genes preferentially expressed in ECL cells: 1) neuroendocrine peptide synthesis and secretion, 2) regulation of local inflammation, 3) regulation of epithelial regeneration by secretion of growth factors, 4) wound healing, 5) hormonal regulation of food intake, and 6) control of circadian rhythm (44). Interestingly, the receptor for the PC specific neuropeptide apelin (Aplnr) was shown to be specifically expressed in ECL cells. Previous work from Lambrecht and colleagues (43) identifying apelin as a PC-specific protein coupled with their ECL expression profile indicated a novel pathway by which PCs and ECL cells communicate. The investigators confirmed expression of both apelin and Aplnr at the protein level and present evidence that the apelin/Aplnr pathway works as a negative feedback loop whereby PCs can regulate their own acid secretion in a gastrin-independent process by inhibiting histamine release from ECL cells (44). These studies epitomize how functional genomics can lead to testable hypotheses about the physiological relationships among cells.

To study the effect of gastrin on gene expression in ECL cells, work from the labs of both Samuelson and Sachs has created expression profiles of ECL cells in animal models of gastrin deficiency (31, 42). Jain and Samuelson (31) compared expression profiles of stomachs from gastrin-deficient mice (GAS-KO) and wild-type mice using Affymetrix 430A GeneChips. These results indicated a role for gastrin signaling in regulating histamine synthesis and transport (Hdc, Vmat2), maintenance of secretory granules (Chga), and regulated secretion (Cckbr) in ECL cells reaffirming the importance of gastrin in regulating ECL function. Similar results were shown by Lambrecht and co-workers (42) in rats. Collectively, these data highlight the importance of endocrine signaling in the maintenance of the stomach mucosa and suggest several pathways in which G cells, ECL cells, and PCs can communicate to regulate the functional and proliferative properties of the surrounding tissue.

Neck Cell to ZC Transition

As neck cells migrate toward the base of the gastric unit, they transition into ZCs. ZCs are characterized by large, electron-dense granules, a well-defined apical and basal cytoplasm, and abundant, highly structured rER (Fig. 1). The first functional genomic analysis of ZCs used cells purified by tissue disruption followed by differential sedimentation. Gene expression was analyzed on Mu11K Affymetrix GeneChips (54). Analysis of ZC global gene expression indicated the expected increased levels of known murine zymogenic secretory products including pepsinogen and intrinsic factor. Consistent with the physiology of a highly polarized secretory cell, ZCs showed increased expression of transcripts involved in translocation of nascent polypeptides into the ER (Tram1), protein modification in the ER and Golgi (Ero1β), and vesicular trafficking (GABA receptor-associated protein, Gabarap) (54). Interestingly, these studies also showed an increased expression in ZCs of platelet-derived growth factor (Pdgf), which in concert with Vegfb secreted by PCs (Fig. 2), may regulate angiogenesis in the gastric mucosa (54). The results underline the tremendous cell biological changes that must occur as the mucous-secreting neck cell, with its own apparenantly distinct physiological function, transitions into the elaborate yet functionally distinct serous secretory ZC (7, 20, 68).

It was in large part to identify the genes that could regulate the neck to ZC transition that led to generation of another profile of ZCs, this time using LCM to harvest cells from their niche-specific context. This analysis on Affymetrix MOE430v2 GeneChips showed substantial overlap of ZC-enriched genes with the genes in the earlier profile list but in addition identified several transcription factors. Two of these, Xbp1 and Mist1 (Bhlhb8), were also expressed in acinar salivary cells. Xbp1 has been shown to be involved in the differentiation of highly secretory cells (46, 70, 71), and Mist1 had been shown to be important for serous secretory cell development in multiple organs (65, 66, 73). Mist1−/− mice showed increased census of cells expressing equal parts neck and ZC markers (transition cells). Mist1−/− ZCs also exhibit defects in cell shape, organelle localization, and zymogenic granule homeostasis (68). Mueller and colleagues (58), in functional genomics studies to determine the effects of Helicobacter pylori infection on gastric epithelial cells, also identified Mist1 as specific for the ZC lineage. Collectively, these data implicate Mist1 as a key regulator of ZC physiology, although, interestingly, Mist1−/− mice have no significant change in the number of ZCs, indicating Mist1 is not required for cells to select the ZC fate.

Early Stomach Development

In the mouse, embryonic development of the gastrointestinal tract relies on a conserved sequence of signals that regulate the fate of epithelial cell differentiation. Studies from Shivdasani and coworkers (11, 81) have defined numerous transcription factors that regulate the early development of the stomach using large-scale genomic analysis including Foxq1. Foxq1 is able to influence the differentiation of surface mucous cells. Gene expression arrays of stomach tissue taken from a radiation-induced mutant strain of mice that is homozygous for the Foxq1 allele and wild-type mice were generated using MOE430Av2 Affymetrix Mouse Expression Arrays. Foxq1-deficient mice had a complete absence of Muc5ac, the dominant stomach mucin produced by surface mucous cells (81). These data are the first to implicate a single transcription factor in the development of surface mucous cells.

Other studies from the Shivdasani group have identified the transcription factors Barx1 and Nkx6.3 as important mediators of stomach development (12, 39). Using Serial Analysis of Gene Expression to compare gene profiles of E12 mouse stomach and intestine, they identified the homeodomain protein Barx1 as expressed preferentially in the fetal stomach. Further analysis of the function of Barx1 confirmed its role in the organization of the stomach epithelium as an inhibitor of the Wnt signaling pathway (39). Nkx6.3 was identified as a novel transcription factor that is expressed in the gastric antrum that regulates the maturation of gastrin-producing G cells (12).

All together, the functional genomic analyses of the normal gastric lineages have been key in identifying not only additional nuances in cell physiology but also, for the first time, have helped us determine some of the genes that regulate differentiation. Thus, as the result of functional genomics, we now know that Mist1 is key in developing ZC physiology, that Foxq1 is required for surface mucous cells to secrete mucins and that Nkx6.3 underlies G cell formation. Work is still
needed to identify additional such factors that might be important for development of PCs and neck cells and the other endocrine cell populations.

The data presented in this section paint a still incomplete picture of the dynamics of stomach physiology. However, there is reasonable concordance in cell-lineage-specific genes identified from study to study. Specific examples include Kcnq1 [identified in Mills et al. (55) in mice and by Lambrecht et al. in rat (43)], CD36 and Vegfb [identified by Mills et al.(55) and by Mueller et al. (58)]. Pthlh, identified by Mills et al. (55), was confirmed by Jain et al. (29). Mist1, identified by Ramsey et al. (68) was also identified by Mueller et al. (58) as being a ZC gene. There has been similar overlap among the multiple ECL transcriptional profiles.

We have also performed comparisons of higher-order gene expression patterns among transcriptional profiles generated by multiple labs and across multiple functional genomic platforms. We always find that there is cell lineage concordance: in other words, PC expression profiles cluster with other PC expression profiles, no matter the lab or platform (Fig. 3 and unpublished observations; see also Refs. 13, 14).

In the end, the functional genomic expression profiling studies described in this section do provide overlapping evidence for conserved molecular pathways that regulate epithelial growth, function, and communication. For example, the fact that Vegfb has been identified at the transcript level as PC-specific by multiple groups and at the protein level (7) (Fig. 2) suggests that it may play a key role in recruiting or modeling the vast capillary network needed for PCs to fuel their enormous needs for oxygen and to remove the prodigious bicarbonate tide that is the residue of acid secretion. Other PC-derived growth factors, such as Pthlh, may be key in PC paracrine signaling. And the role of the apelin-apelin receptor axis has emerged in large part due to cell lineage-based functional genomic studies.

To date, studies directed toward defining lineage-specific gastric epithelial cell biomarkers have been performed in rodent models, whereas profiles of human gastric mucosa have been performed as controls for studies of disease processes and have not involved assaying normal, lineage-specific gene expression patterns. However, in published accounts, gene expression patterns in rodents have so far correlated well with human gene expression patterns, wherever they have been investigated. For example Vegfb, identified on mouse microarrays, is also PC specific in humans (Fig. 2 and data not shown). As mentioned, Kcnq1, identified in rat PCs, is also expressed in human PCs. Obviously, one key difference between rodent and human lineage expression patterns is that the cellular source of intrinsic factor in rodents is overwhelmingly the ZC and in humans, it is the PC. Further studies of normal lineage expression patterns in human, likely using LCM, will be needed to determine overall rates of concordance between humans and common model species.

Despite the advances in our understanding of the molecular pathways taken between GEPs and their mature lineages, relatively little is known of the mechanisms of carcinogenesis in the stomach. While there is some hope that the studies performed in animal models will coalesce to provide clues into the molecular origins of neoplasia, the substantially divergent patterns of gene transcription profiles so far observed have

![Fig. 3. PCs represent a unique cell population among all cells in the body. A: tree diagram where Gene Ontology (GO) term-based profiles/signatures of the different cell populations cluster stem cells from various tissues together and differentiated cells from various tissues together. Note how PC populations are outliers. Sources for the stem and differentiated cell expression profiles are described in Ref. 10. B: each point represents a gene expression profile from a given cell lineage, with stem, differentiated, and PC profiles from A all plotted in terms of their relative representation of each GO term. GO terms depicted are among the most highly represented where there was statistically significant difference between stem and differentiated cells. Note that for most of these, PCs are similar to other differentiated cells; however, for GO terms like “nucleus,” PCs show a sort of extreme differentiated pattern (see text). C: similar to B, but depicted are the 3 most common GO terms that show no difference between stem and differentiated cells. Note how in each case the PC expression profiles are outliers in having these GO terms particularly well represented.](http://physiolgenomics.physiology.org/Downloadedfrom)
been problematic for the study of human gastric cancers. In the next section we will discuss what functional genomics has told us about the transformation from normal to metaplastic and neoplastic growth and its implications in human neoplasia.

CANCER

Gastric cancer is one of the leading causes of cancer-related death worldwide (67). The transformation from normal to metaplastic growth on the cellular level has been described in great detail; however, the molecular mechanisms of gastric carcinogenesis remain unclear. Early diagnosis continues to be the best correlate for long-term survival. Despite recent advances in diagnostic and treatment technologies, most gastric cancers are not detected until they have reached an advanced stage of growth (72). With the advent of global gene expression analysis, the recognition of altered patterns of molecular differentiation may help identify disease-specific genes and, accordingly, to allow early diagnosis, to better determine prognosis, to follow effectiveness of therapy, and, eventually, to develop rapid and exact therapies. In this section we will discuss the relevance of microarray-based studies in the elucidation of the molecular pathways of gastric cancer.

Neoplasia

Unfortunately, so far, attempts to identify a cohort of gastric cancer-related genes that aid in diagnosis, prognosis, or monitoring have not been particularly fruitful. Multiple studies have been performed using functional genomics to profile gene expression in gastric adenocarcinoma, the most common form of gastric cancer [recently reviewed elsewhere (16, 85)]. Investigators in these studies have used a variety of techniques for identifying and isolating tumor cells (e.g., LCM) and a variety of profiling approaches (e.g., cDNA microarrays). The “take-home message” to date is that there is little consensus around any individual, specific genes identified by functional genomics that can aid in diagnosis or prognosis. There are, however, general patterns that are similar to those seen in multiple cancer types and simply reflect the fact that tumor cells grow more rapidly and must remodel the extracellular matrix to be invasive.

Microarray analyses of primary gastric tumors have shown that the stomach mucosa is susceptible to the same functional changes that take place in most cancers and that the severity of the diagnosis can be correlated with the frequency of expression changes in genes involved in metastasis (ECM remodeling, cell motility, cell adhesion) (16, 25, 27, 32, 84). Inoue and colleagues (27) used the IntelliGene Human Cancer chip cDNA microarray to analyze gastric tissue from 43 patients at various stages of metastasis to develop a prognostic scoring system. Using a list of 425 genes known to be involved in cancer, they were able to show that only a small portion of the genes studied correlated with malignancy. Genes such as MMP7, TGFβ3, and IGFBP3 were associated with the most aggressive tumors and were subsequently proposed as good diagnostic markers (27). Furthermore, studies by Hippo and Hasegawa and colleagues (21, 25) showed that similar genes involved in matrix remodeling (COLIA2, DDOST, FNI), cell adhesion (CDH17), and motility (IGF2) were increased in patients with lymph node metastasis. Similar methods have been used to compare gene expression profiles between the two general histological subtypes of gastric adenocarcinoma: diffuse and intestinal type (45). Not surprisingly, genes associated with the usually more infiltrative diffuse-type cancer were involved in the TGF-β signaling pathway (TGFβ1), cell migration (ITGB1), and ECM remodeling (MMP7, COLIA2, FNI), while the genes associated with the generally less invasive, intestinal-type cancer were skewed more toward enhancement of cell growth (21, 32). These studies highlight a certain tendency toward increased expression of a handful of ubiquitous genes (MMP7, FN1, TGFβ) in gastric cancer that are also implicated in progression of many other cancers and, moreover, had been identified before the advent of functional genomics.

One promising gene with more specific molecular function that was identified on a functional genomics screen and that may eventually acquire momentum as a molecular handle for following or classifying gastric cancer is PLA2G2A, a gene associated with inflammation, prostaglandin biosynthesis, and antibacterial activity (48). PLA2G2A was identified on a microarray screen of gastric adenocarcinomas and shown to correlate with less invasive cancer and better survival (10). In a subsequent meta-analysis of multiple gastric cancers, it was shown to be a target of the Wnt pathway (1). And, most recently, it was shown to be expressed at higher levels in early stage tumors and to correlate with decreased invasiveness in an in vitro assay (17). However, it should be noted that several other microarray-based surveys of gastric adenocarcinoma genes from other groups have not reported PLA2G2A as a cancer-related gene in their screens (6, 32).

Interestingly, even though functional genomic studies have not yet identified consensus gastric adenocarcinoma-regulated genes, multiple studies are in agreement on several genes downregulated in neoplasia. For example TFF1, a surface mucous cell gene (38, 59), was reported as downregulated by Jinawath et al. (32), Bousiouciats et al. (6), and Chen et al. (10), although decreased TFF1 in gastric cancer has also been well characterized independently of its identification on functional genomic screens (8, 49). TFF2, a neck cell gene (see below also), is another consensus downregulated in cancer gene, identified by Chen et al. (10) and Jinawath et al. (32). PGC, a ZC gene, was also reported to be decreased in cancer by both Jinawath et al. and Chen et al. (10, 32).

It is unclear why there has not been striking consensus among the extant functional genomic studies attempting to find specific genes that positively correlate with gastric cancer. Several technical problems may be confounding. For example, perhaps cancer cohorts have been too small to date, or perhaps it will be necessary to find better ways to purify carcinoma cells from surrounding, contaminating tissue (i.e., mesenchyme, inflammatory cells, necrotic debris). But the field may also be hampered by more conceptual or paradigmatic problems. For example, maybe assaying even pure populations of tumor cells will not help us understand how cancers are formed or promoted, because, as has been proposed, the vast majority of tumor cells are derivatives of much rarer cancer stem cells (78). Thus, the stem cells fueling gastric cancer may have to be profiled to identify a consensus cohort of cancer-associated genes, and the genes associated with the bulk tumor cells profiled thus far are red herrings. Finally, it’s also possible that the molecular variability among gastric tumors reflects the fact that there is no particular consensus pattern of gastric cancer...
origin or growth, other than that gastric tumor cells need to express general growth-related genes and extracellular matrix remodeling genes to grow and spread. Certainly, gastric tumors are histologically diverse well beyond the simple shorthand of “diffuse” vs. “intestinal” subtypes (45). Maybe they are equally molecularly diverse.

Clearly, more molecular dissection of gastric tumors (or their stem cells) is needed. In the meantime, one avenue of investigation that has not been adequately explored but might be more fruitful is profiling the gene expression changes underlying the preneoplastic changes gastric mucosa undergoes. Toward that end, Boussioutas et al. (6) and Chen et al. (10) have profiled gastric tissue with intestinal metaplasia. Perhaps not surprisingly, intestinal metaplasia seems to be distinguished from normal gastric mucosa at the molecular level by genes that are markers of normal intestinal development. For example, the prominent transcription factors increased in intestinal metaplasia are CDX1 and CDX2, which are critical for normal intestinal development (6, 10, 56). This, of course, represents a kind of molecular tautology: Intestinal metaplasia is different from normal gastric tissue by virtue of its expression of normal intestinal genes. It remains unclear if there are any consistent differences in gene expression between gastric intestinal metaplasia glands and normal intestine. In the end, consensus patterns in gene expression in intestinal metaplasia may be difficult to determine, because intestinal metaplasia appears in a variety of differentiation patterns, designated, for example, as complete or incomplete and tending to resemble small intestine (with Paneth cells) or colon or mixtures of the two.

Metaplasia

An almost completely unexplored avenue of investigation to date is the molecular dissection of the other common form of metaplasia, variably known as pseudopyloric or spasmolytic polypeptide expressing metaplasia (SPEM). SPEM correlates highly with progression to neoplasia (75), yet its molecular origin, at least in humans, is unclear. Histologically, SPEM is characterized by cells expressing considerable amounts of spasmolytic polypeptide (Tff2), whose expression is normally confined to the gastric gland neck, at the base of fundic mucosa (75). Trefoil family factor 2 (Tff2, a.k.a. spasmolytic polypeptide) is a small protease-resistant protein that normally functions to promote gastric mucosal healing and regulate inflammatory responses and may also play a role in maintaining normal growth of the gastric mucosa (4, 15, 40). Loss of PCs in response to Helicobacter infection and chemical or genetic ablation is associated with an increased expression of Tff2 and other markers of normal mucous neck cells in the metastatic cells (19, 60, 75, 83).

Recent studies show mice also develop SPEM and that the metastatic cells derive from zymogenic lineage cells, showing coexpression of Tff2 and normal markers of zymogenic differentiation, such as Pgc and Mist1, although Mist1 expression is eventually lost in SPEM cells (7, 61). Genomic profile studies of mouse SPEM cells indicate that these cells have an increased expression of minichromosome maintenance proteins, which are involved in cell cycle regulation, and HE4, which is a secreted mucosal factor that characterizes both intestinal metaplasia and ovarian neoplasia (61). The results suggest a model whereby normally postmitotic ZCs at the bases of units reacquire expression of precursor markers and, perhaps more importantly, begin to proliferate again. It would be interesting to use functional genomics to assess global gene expression in purified (e.g., by laser capture) human SPEM cells. Interestingly, HE4, though identified in mouse at the transcript level, also marks mouse and human SPEM cells at the protein level, showing how functional genomics in a mouse model might lead to new biomarkers for human cancer development.

Helicobacter Infection

One of the most common predisposing factors for gastric cancer is the loss of acid-secreting PCs or oxyntic atrophy. As discussed above, PCs play a critical role in maintaining proper growth and differentiation of the gastric unit. Hence, the development of oxyntic atrophy correlates with the emergence of both SPEM and intestinal metaplasia within the neck and base zones (64). In humans, infection with Helicobacter pylori can cause loss of PCs, with associated atrophy, metaplasia, and, in some cases, progression to gastric cancer (63, 64, 75). Work from Mills and coworkers (55) used Affymetrix Mu11K GeneChips to generate expression profiles from PCs isolated by lectin panning and magnetic bead sorting to show that, in response to H. pylori infection in germ-free mice, while the molecular profile of parietal cells remained relatively unchanged, the nonacid-secreting cells of the gastric unit showed increased expression of genes involved in cell motility, extracellular matrix remodeling, and growth factor signaling. A subsequent study by Falkow and colleagues (58) extended these findings by looking at cell type-specific responses and ZCs from mice infected with H. pylori; transcriptomes from each cell were determined by spotted mouse cDNA microarray. However, of the cell types, the mucus-secreting surface (pit) cell was the only one to show a significant change in gene expression upon exposure to H. pylori, consistent with an earlier study showing that PCs themselves do not respond directly to H. pylori (55). Upregulated genes were involved in the initiation of proinflammatory responses (Ccl5), iron processing (lactoferrin), and angiogenesis (Vegfa) (58). Recently, tissue (i.e., not gene-expression-based) microarray analysis from a large cohort of gastric adenocarcinomas has indicated a decreased expression of other surface mucous cell genes: gastrokine 1 and 2 (Gkn1, Gkn2) (57). While the function of these proteins is unknown, Gkn2 has been shown to interact with various members of the trefoil factor family, which have a known role in regulating gastric epithelial growth and differentiation (59, 69, 79, 82). This study implicates Gkn1 and Gkn2 as early molecular markers of gastric cancer transformation.

HIGHER-ORDER PATTERNS OF GENE EXPRESSION

As functional genomic analyses accumulate, investigators have begun to develop the tools to ask higher-order questions about patterns of gene expression. In other words, rather than simply identifying specific genes preferentially expressed by PCs, investigators can now begin to address what organizing principles are at work underlying the PC-specific pattern of gene expression. We have begun such analyses, using the gastric epithelium, where there are functional genomic analy-
ses of all the principle cell lineages as well as the multipotent progenitor. We have developed an approach wherein gene expression profiles are re-expressed in terms of the GO terms (http://www.geneontology.org) they define (13). This re-expression allows us to examine the underlying functions, pathways, and cellular locations of transcripts or proteins in gene expression profiles and compare different cell types on the basis of these patterns, a forest-to-forest rather than tree-to-tree analysis. Recently, we have shown how this approach enabled us to identify patterns of gene expression that characterize all tissue stem cells. Namely, adult tissue stem cells preferentially express genes that are associated with regulation of transcription and translation/biosynthesis that can be defined by GO terms such as “RNA binding,” “nucleic acid binding,” and “protein biosynthesis” (14). Interestingly, there was similar commonality in higher-order patterns of expression in differentiated cells, regardless of their tissue location. Differentiated cells exhibit patterns of gene expression that function in cell-cell communication and extracellular processes, with GO terms such as “membrane,” “extracellular space,” and “signal transduction” (14). These studies emphasize the practical application of integrated gene expression data in classifying unknown cell populations based on their state of differentiation and suggest a possible role for this technology in defining tissue states in various stages of metaplastic growth.

Functional genomic expression profiles of stomach cells closely matched the same stem and differentiated cell patterns with one exception. Thus, the microarray-defined GEP classifies with stem cells of other tissues, and surface mucous cells with one exception. Thus, the microarray-defined GEP classifies with stem cells of other tissues, and surface mucous cells exhibit patterns of gene expression that function in cell-cell communication and extracellular processes, with GO terms such as “membrane,” “extracellular space,” and “signal transduction” (14). These studies emphasize the practical application of integrated gene expression data in classifying unknown cell populations based on their state of differentiation and suggest a possible role for this technology in defining tissue states in various stages of metaplastic growth.

CONCLUDING REMARKS

We have shown how functional genomics has been used to dissect the molecular anatomy of diseased and normal stomach and has provided new insights and clues into cell function and regulation of differentiation. Upcoming analyses will likely help us to better understand cell differentiation in the antrum relative to the corpus and may include more refined analyses of differentiation patterns by analysis of precursor cells (e.g., surface mucous cell precursors and mucous neck cells, the ZC precursors). More detailed analyses of higher-order gene expression patterns might not only elucidate further genetic cassettes that specify critical functions but also the genes that may regulate those cassettes. For example, we have identified increased metabolic gene expression in PCs. Functional genomics has already identified Mist1 as a gene that regulates a gene cassette that induces serous secretory granules in ZCs. Are there a limited number of transcription factors analogous to Mist1 that regulate metabolic genes to help induce PC differentiation? Finally, there are still numerous genes already identified in published expression profiles whose function still needs to be elucidated by conventional gain-of-function and loss-of-function experiments. Stay tuned, because the best is yet to come.

ACKNOWLEDGMENTS

For GeneChip hybridization, we thank the Multiplexed Gene Analysis Core of the Siteman Cancer Center (supported in part by National Cancer Institute Grant P30-CA-91842). Further grant support: J. C. Mills (R01 DK-079798-01) and B. J. Capoccia (T32 CA-009547-23).

REFERENCES


Physiol Genomics • VOL 37 • www.physiolgenomics.org


82. Westley BR, Griffin SM, May FE. Interaction between TFF1, a gastric tumor suppressor trefoil protein, and TF2I1, a brixos domain-containing protein with homology to SP-C. *Biochemistry* 44: 7967–7975, 2005.


