Single-cell RNA-sequencing reveals a distinct population of proglucagon-expressing cells specific to the mouse upper small intestine

Leslie L. Glass1, Fernando J. Calero-Nieto2, Wajid Jawaid1, Pierre Larrau1, Richard G. Kay1, Berthold Göttgens2, Frank Reimann1,*, Fiona M. Gribble1,*. 

ABSTRACT

Objectives: To identify sub-populations of intestinal preproglucagon-expressing (PPG) cells producing Glucagon-like Peptide-1, and their associated expression profiles of sensory receptors, thereby enabling the discovery of therapeutic strategies that target these cell populations for the treatment of diabetes and obesity.

Methods: We performed single cell RNA sequencing of PPG-cells purified by flow cytometry from the upper small intestine of 3 GLU-Venus mice. Cells from 2 mice were sequenced at low depth, and from the third mouse at high depth. High quality sequencing data from 234 PPG-cells were used to identify clusters by tSNE analysis. qPCR was performed to compare the longitudinal and crypt/villus locations of cluster-specific genes. Immunofluorescence and mass spectrometry were used to confirm protein expression.

Results: PPG-cells formed 3 major clusters: a group with typical characteristics of classical L-cells, including high expression of Gcg and Pyy (comprising 51% of all PPG-cells); a cell type overlapping with Gip-expressing K-cells (14%); and a unique cluster expressing Tph1 and Pzp that was predominantly located in proximal small intestine villi and co-produced 5-HT (35%). Expression of G-protein coupled receptors differed between clusters, suggesting the cell types are differentially regulated and would be differentially targetable.

Conclusions: Our findings support the emerging concept that many enteroendocrine cell populations are highly overlapping, with individual cells producing a range of peptides previously assigned to distinct cell types. Different receptor expression profiles across the clusters highlight potential drug targets to increase gut hormone secretion for the treatment of diabetes and obesity.

© 2017 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Keywords: Single cell RNA seq; GLP-1; PPG-cells; L-cells; Mass spectrometry; 5-HT

1. INTRODUCTION

Enteroendocrine preproglucagon-expressing PPG-cells (traditionally known as L-cells) secrete the gut hormones glucagon-like peptide 1 (GLP-1) and peptideYY (PYY) and are important regulators of glucose metabolism and appetite [1]. GLP-1 promotes glucose-dependent insulin secretion and satiety, and has been translated into therapies for type 2 diabetes and obesity through the development of long acting GLP-1 mimetics and inhibitors of dipeptidyl peptidase 4 (DPP4) [2]. PYY promotes satiety and is under investigation as a potential basis for new anti-obesity therapeutics [2].

These observations raised the prospect of developing treatments that promote the secretion of GLP-1 and other gut hormones. The strategy depends on understanding the different enteroendocrine cell populations and their relative expression of target receptors. Classically, a subpopulation of enteroendocrine cells defined by immunohistochemical staining for proglucagon-derived peptides (including GLP-1), often together with PYY, was designated as L-cells, referred to here as PPG-cells because of their characteristic expression of preproglucagon [3]. Recent studies investigating genetically tagged PPG-cells have shown that, as a population, they also express gut hormones previously thought to be expressed in distinct enteroendocrine cell types, including Cck (cholecystokinin, I-cells), Scg (secretin, S-cells), and Gip (glucose-dependent insulinotropic polypeptide, K-cells) [4]. However, it remained unclear whether cells expressing different hormone combinations represent fundamentally distinct cell populations.

Variability within the PPG-cell population is physiologically interesting because PPG-cell peptides show different post-prandial plasma profiles [5]. It has been proposed recently that within a single enteroendocrine cell, vesicle pools containing different hormones might be differentially responsive to stimuli [6], but it is also likely that expression of hormones, ion channels, transporters, and receptors varies between PPG-cell sub-populations. The aim of the present study was to use single cell RNA sequencing to determine whether PPG-cells

*Corresponding author. E-mail: fmg23@cam.ac.uk (F.M. Gribble).
**Corresponding author.

Received June 27, 2017 • Revision received July 21, 2017 • Accepted July 27, 2017 • Available online 1 August 2017

http://dx.doi.org/10.1016/j.molmet.2017.07.014
can be sub-divided into clusters with distinct expression of gut hormones, receptors, and other nutrient sensing proteins.

2. EXPERIMENTAL PROCEDURES

2.1. Animal welfare and ethical statements

This research has been regulated under the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge Animal Welfare and Ethical Review Board (AWERB). Mice were housed in individually ventilated cages with ad libitum access to water and chow. Mice were killed by cervical dislocation prior to tissue harvesting. Both male and female GLU-Venus mice [7] on a C57BL6 background were used.

2.2. Small intestine for FACS sorting

For single cell RNAseq, tissue was prepared from 3 male mice, aged 20—21 weeks. For FACS sorting, tissue pieces from the proximal 10 cm of the small intestine were stripped of the outer muscle layers. Tissue was chopped into 1—2 mm pieces and digested to single cells with 1 mg/ml collagenase in calcium-free Hanks Buffered Salt Solution (HBSS). Single cell suspensions were separated by FACS using an Influx Cell Sorter (BD Bioscience, USA). Side scatter, forward scatter, pulse width gates, and DAPI-staining were used to exclude debris and aggregates. Single fluorescent and non-fluorescent (control) cells were collected into individual wells of a 96-well plate containing lysis buffer 0.2% (v/v) Triton X-100 and 2 U/μl RNase inhibitor (Ambion) and stored at −80 °C.

2.3. Single-cell RNA sequencing (further details in supplementary material)

scRNA-seq analysis was performed using the Smart-seq2 protocol [8] as previously described [9]. Two mice were sequenced at low depth and one mouse at high depth. Cells with >20% reads mapping to mitochondrial genes were removed from downstream analyses. For the deeper sequenced sample, all cells with <750,000 reads mapping to endogenous RNA were excluded. Out of the 288 cells sorted across the 3 experiments, 94 and 95 passed quality control from the first 2 mice, and 75 cells passed from the deeper sequenced experiment with increased quality control stringency (78%). Data were normalized for sequencing depth and RNA quantity using size factors calculated on endogenous genes [10]. Clustering was performed on the dimensionality reduced ISNE coordinates using the R package, Mclust (v 5.1) using cells that passed QC from all 3 mice. This defined 6 populations of cells. Only clusters that contained cells from all 3 mice and only containing Venus positive cells were used for further analysis.

Differential expression analysis was limited to cells from the sample sequenced at higher depth. Differentially expressed genes were identified by performing pair-wise and unique comparisons between the 3 clusters using DESeq2 (v. 3.4). Hierarchical clustering was performed using the union of the top 15.

2.4. Cell collection for qPCR analysis

PPG-cells were isolated as above, with the variation that tissue pieces were incubated in 10 mM EDTA in Ca²⁺ free PBS for 5 min, then transferred to 10 ml Ca²⁺ free PBS and gently inverted to dissociate the villi. This was repeated 4 more times, with incubations 3—5 shaken more vigorously in PBS. The fractions were spun at 300 rcf, resuspended in HBSS, then re-centrifuged. For collecting mixed PPG-cell populations, these fractions were combined and digested in 1 mg/ml Collagenase in HBSS. For separate villus/crypt sorts, fractions 1—2 were retained separately to generate the villus-enriched fraction, and fractions 3—5 were filtered through 50 μm filters prior to centrifugation to generate crypt-enriched fractions. The two fractions were then separately incubated in collagenase. Digested tissue was centrifuged at 300 rcf, resuspended in HBSS with 0.1% BSA, spun again, and resuspended in ~1 ml of HBSS supplemented with 10% FBS. Cell suspensions were FACS sorted using a MoFlo Beckman Coulter Cytomation sorter (FL, USA) to obtain populations of Venus-positive or Venus-negative (control) cells that were collected directly into lysis buffer for mRNA extraction. The successful separation of crypts from villi was confirmed by qPCR for Apoα4 (villi) and Defa5 (crypts) using cDNA prepared from Venus negative cells.

2.5. RNA extraction and quantitative RT-PCR

RNA extraction from FACS-sorted cells was performed using an RNeasy Micro Plus Kit (Qiagen). Samples were reverse transcribed according to standard protocols (Superscript III, Life Technologies). Quantitative RT-PCR was performed with 7900 HT Fast Real-Time PCR system (Applied Biosystems, Life Technologies), using TaqMan primer/probe sets supplied by Applied Biosystems. Expression was compared with that of β-actin measured on the same sample in parallel on the same plate, giving a CT difference (ΔCT) for β-actin minus the test gene. Mean, standard error, and statistics were performed on the ΔCT data and only converted to relative expression levels (2ΔΔCT) for presentation in the figures. All experiments were performed on 4 independently isolated cDNA samples (n = 4 mice). A table of TaqMan primer-probes used is listed in the supplementary material.

2.6. Immunofluorescence microscopy

Duodenum from GLU-Venus mice was fixed in 4% paraformaldehyde, dehydrated in 15% and 30% sucrose, and frozen in O.C.T. media (VWR, UK). Cryostat-cut sections (7—10 μm) were mounted directly onto polylysine-covered glass slides (WWR, Belgium). Slides were incubated for 1 h in blocking solution containing 5% donkey serum/1% BSA/0.05% Tween-20 and overnight in blocking solution with primary antisera of interest. A table of antisera used is listed in the supplementary material. Sections were washed with blocking solution and incubated with appropriate secondary antisera (AlexaFluor, Life Technologies, UK) diluted 1:300. Control sections were stained with secondary antisera alone. Sections were washed with PBS and mounted with Prolong Gold (Life Technologies, UK) before confocal microscopy (Leica TCS SP8 X, Germany).

2.7. Peptidomics

16,000—20,000 PPG-cells from the proximal 10 cm of the small intestine were FACS-sorted directly into a lysis solution of acetonitrile (ACN) 80% (v/v) in a Protein LoBind tube (Eppendorf). Proteins were precipitated and after centrifugation (5 min, 10,000 g), supernatants containing small proteins and peptides were collected and dried in a centrifugal evaporator (Eppendorf). Peptides were reduced and alkylated to break and cap disulphide bonds. Samples were analyzed using a ThermoFisher Ultimate 3000 nano LC system coupled to a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Scientific, San Jose, USA). Peptides were separated on a nano easy column (Thermo Fisher Scientific) by a ramp of increasing concentration of ACN from 2 to 40% (v/v) in 0.1% formic acid solution over 145 min. A full scan range of 400—1600 m/z was performed, and the top 10 ions of each spectrum were selected for MS/MS analysis. The acquired data were analyzed using Peaks 8.0 software (Waterloo, ON, Canada) against the mouse Swissprot database (downloaded on 06/05/2016) with a no-digest setting, allowing the matching of endogenous peptides of up to 65 amino acids.
2.8. Data analysis

Statistical analysis was performed using GraphPad Prism 7 package (San Diego, CA, USA). Values were regarded as significant when p < 0.05.

3. RESULTS

Venus-positive PPG-cells from the top 10 cm of the small intestines of two GLU-Venus mice [7] were collected by flow cytometry together with negative (non-fluorescent) control cells. Venus-positive cells were collected across a range of fluorescence intensities, including cells that were only dimly fluorescent (Figure S1). The transcriptional profiles of a total of 184 cells (159 Venus positive and 25 Venus negative) were obtained as previously described [8]. 3381 genes exhibited expression variability exceeding technical noise [11]. Dimensionality reduction was performed using the t-Distributed Stochastic Neighbour Embedding (t-SNE) method using only these 3381 genes. Different clusters could be observed that included cells from both mice. As expected, Venus positive and negative cells appeared in separate clusters. Venus-positive cells appeared as a heterogeneous population that could be separated into distinct subpopulations.

Ninety six Venus positive cells from a third mouse were analyzed at greater depth, of which 75 passed strict quality control criteria (see methods). Combining all 259 cells from the 3 mice demonstrated 3381 genes. Different clusters could be observed that included cells from all three mice. As expected, Venus positive and negative cells appeared in separate clusters. Venus-positive cells appeared as a heterogeneous population that could be separated into distinct subpopulations.

3.1. Data analysis

To examine whether Cl1-3 might represent PPG-cell populations characteristic of different positions along the 10 cm of small intestine examined, we measured expression of genes that distinguished the clusters by qPCR in Venus-positive cells purified by FACs from five sequential 2 cm segments of the upper small intestine. We found no evidence for statistically significant expression gradients of Gcg, Gip, Tph1, Pzp, or Pyy in PPG-cells along the longitudinal axis of the upper small intestine, although Pzp showed a trend of decreasing and Pyy a trend of increasing expression distally (Figure 3A).

To examine whether Cl1-3 might represent PPG-cell populations characteristic of different positions along the 10 cm of small intestine examined, we measured expression of genes that distinguished the clusters by qPCR. Gcg, Pyy, Gpr119, Trpa1, and Necab2 were used as markers for Cl1, but did not exhibit consistent patterns between PPG-cells from crypts and villi: Gcg was higher in crypt PPG-cells, Trpa1 was higher in villi, and the other 3 markers did not differ between the two regions (Figure 3B). Gip, Fabp5, Th, Pikl, and Lpl were used to track Cl2, but similarly exhibited no consistent crypt/villus variation: Gip was more highly expressed in villus than crypt PPG-cells, but the other markers did not differ (Figure 3C). For Cl3, we examined Tph1, Pzp, Casr, Mc4r, and Cyp2j5. Three of these markers were significantly more highly expressed in PPG-cells from villi than crypts, and the other two showed a similar trend that was not significant (Figure 3D).

To examine whether hormonal products predicted by gene expression in Figure 1E were detectable at the peptide level, we performed mass spec analysis of PPG-cells purified from the upper small intestine. Peptides from the prohormones of proglucagon, GIP, PYY, CCK, neurtensin, ghrelin, and secretin, but not Urocortin3 were detected (Figure S4). PYY production in PPG-cells from the top 2 cm of the small intestine was also confirmed by immunostaining (Figure 4A). As Cl1 was characterized by expression of Gpr119, and Cl3 by Pzp, we performed immunostaining of mouse duodenum to confirm the existence of the predicted Gcg/Venus positive cell populations producing PZP and GPR119 (Figure 4B and C) and to assess whether the two subpopulations could be differentiated at the level of protein expression. Co-staining indicated that very few Venus positive cells (<2%) co-expressed both GPR119 and PZP (Figure 4C), consistent with the

Figure 1: PPG-cell clusters reflect hormone expression profiles. A, C, D. ISNE analysis showing: Cluster allocation of the 259 included cells from all 3 mice. B. Relative proportions of cells located in Cl1-3, 203 cells in total. C. Venus fluorescence intensity of individual cells (redness intensity denotes relative brightness, grey = negative, black = no data). D. Donor origin of cells used in the single-cell RNA-Seq analysis (deeper sequenced mouse is depicted in blue). E. Relative gene expression levels of Gcg, Gip, Tph1, Pyy, Cck, Nts, Ucn3, Ghr1, and Sct in Cl1-3. All values displayed are the log_{10} (normalized counts), bars indicate median values and interquartile range. Statistical analysis was performed using Kruskal–Wallis with Dunn’s multiple comparisons test. *p < 0.05, **p < 0.01, ***p < 0.001 between the indicated clusters.
cluster analysis. We confirmed production of 5-HT in mouse cells immunopositive for proglucagon (Figure 4D), and in human tissue slices, we demonstrated co-staining of the Cl3 marker CASR together with 5-HT and GLP-1 (Figure 4E).

4. DISCUSSION

Enteroendocrine cells are traditionally classified according to the major one or two hormones they produce, but recent evidence has suggested that each enteroendocrine cell can express a number of different hormones, and that there is strong overlap between the cells producing GLP-1, PYY, CCK, GIP, secretin, and NTS [4,14]. Single cell RNA sequencing has been used to study a variety of cell types, including pancreatic islet cells, neurons, and dissociated human intestinal organoids [15-17]. PPG-cells comprise less than 1% of the intestinal epithelium; thus, a previous study of intestinal organoids identified most of the cells as enterocytes and only a small number as enteroendocrine cells [17]. This analysis grouped enteroendocrine cells into a few small clusters, each including only a few cells, so did not produce an overall picture of PPG-cell diversity. Here we show that PPG-cells from the proximal small intestine can be divided into 3 major clusters with overlapping but distinct hormonal profiles. Fourteen percent of Venus positive cells were located in a cluster characterized by high Gip expression (Cl2), consistent with our previous quantification of cells co-producing GLP-1 and GIP by FACS analysis [4]. The remainder of the Venus positive cells could be subdivided into two similar-sized groups, one with higher Gcg (Cl1) and the other with lower Gcg but high Tph1 (Cl3). Only 4 cells from the deeply sequenced mouse were found in Cl2, so only the highly enriched genes in this group were identifiable with high probability. Among the characteristic Cl2 genes were several that we and others have previously identified as playing functional roles in K-cells, including Cnr1 and Fabp5 [12,13]. Also appearing in this cluster were the scavenger receptor Scarb1, and lipoprotein lipase (Lpl), the roles of which for GLP-1 and GIP secretion remain unknown.

Cl1 exhibited significantly higher expression of Gcg and Pyy than the other two clusters, and also had the highest levels of Chga and Chgb (Figure S6), vesicular markers of enteroendocrine cells [4].
expression of \textit{Gcg}, \textit{Pyy}, and other gut hormones suggests that this cluster represents classical “L-cells”. We were surprised that the tSNE analysis did not separate this group into more clusters that might reflect subtle differences between cells coexpressing different hormonal combinations. It is possible, however, that analysis of PPG-cells from additional regions of the gut might identify more distinct clusters.

\textbf{C1} was characterized by expression of a number of receptors and ion channels highlighted in previous studies, including \textit{Gpr119}, \textit{Ffar1}, \textit{Il6ra}, \textit{Gucy2c}, and \textit{Trpa1}. \textit{Gpr119} was detectable in \~{}50\% of PPG-cells, corresponding with our previous finding that about half of PPG-cells in primary culture exhibited cAMP responses to \textit{Gpr119} agonists and that GLP-1 secretion was stimulated by \textit{Gpr119} agonists \cite{18}. \textit{Ffar1}, by contrast, was found in almost all PPG-cells, although mean expression levels were highest in \textbf{C1}. This is a well-studied receptor for long chain free fatty acids, shown in a number of studies to contribute to fat-triggered GLP-1 secretion \cite{19}. \textbf{C1} was also enriched for \textit{Il6ra}, which has been implicated in the modulation of PPG-cell function by interleukin-6 in response to high fat diet and exercise \cite{20}.

We were initially concerned that \textbf{C3} might be an immature PPG-cell population, as it expressed lower levels of \textit{Gcg} and chromogranins. Separate qPCR analysis of PPG-cells purified from crypts and villi, however, revealed that most genes that characterized \textbf{C3} were more highly expressed in PPG-cells from villi than crypts. The predominant villus location of \textbf{C3} markers suggests that this is a mature cell population, as immature cells should be located closer to their site of generation in the crypts. Alternatively, these cells might reflect a PPG-population gaining \textit{Tph1} and \textit{Pzp} expression with aging. The expression in this cluster of \textit{Tph1} is consistent with a recent report that the majority of enterochromaffin cells in the mouse small intestine coexpress other gut hormones including \textit{Pyy}, GLP-1, and \textit{Cck} \cite{21}. The strongest marker of \textbf{C3} was \textit{Pzp}, an alpha2 macroglobulin of uncertain function \cite{22}. \textbf{C3} appeared distinct to the upper small intestine, as mRNA for \textit{Pzp} was low in PPG-cells purified from the lower small intestine; a trend for higher expression in PPG-cells from the proximal compared to the distal intestine was also observed for another \textbf{C3} enriched gene \textit{Tph1} (Figure S7), although this does not exclude that a relatively smaller sub-population of \textit{Pzp} and/or \textit{Tph1} coexpressing cells is present in the more distal PPG population.

Notable among the receptors found in \textbf{C3} were the calcium sensing receptor (\textit{Casr}) and melanocortin 4 receptor (\textit{Mc4r}). We were surprised that \textit{Casr} was more highly expressed in \textbf{C3} than \textbf{C1}, as a number of studies have linked \textit{CASR} activation to the secretion of GLP-1 \cite{23,24}. However, we demonstrated colocalization of \textit{CASR} with 5-HT and GLP-1 in human intestinal slices as also reported previously \cite{25}, and exposure of human colonic tissue to a phenylalanine/tryptophan mixture has been reported to increase pCAMKII labeling in 5-HT and GLP-1 expressing cells \cite{25}. \textit{Mc4r} expression has been reported previously in colonic PPG-cells, where its activation was linked to PPG-cell secretion \cite{26}.

\section*{5. CONCLUSIONS}

In summary, we conclude that upper small intestinal PPG-cells can be separated into at least 3 major clusters that exhibit differential expression of \textit{Gcg}, \textit{Cck}, \textit{Pyy}, \textit{Gip}, and \textit{Tph1}. Receptor and ion channel expression profiles differed across the clusters suggesting that these PPG-cell sub-populations likely contribute to the differential responsiveness of gut hormones to nutritional and local signals.

\section*{AUTHOR CONTRIBUTIONS}

LLG and FJC-N performed experiments, analyzed data and helped to write the manuscript. PL and RGK performed population RNAseq and
mass spec analyses. BG took responsibility for single cell RNAsseq protocols and bioinformatics analysis. FR and FMG jointly designed and supervised the project and wrote the manuscript. All authors approved the final manuscript.

ACKNOWLEDGEMENTS

Research in the FR/FMG lab was funded by the Wellcome Trust (106262/Z/14/Z and 106263/2/14/2) and MRC (MRC_MC_U1_12012/3, MRC_MC_U1_12012/5 and MRC/M008941/1). Work in the Göttgens laboratory was supported by grants from Bloodwise (12029), Cancer Research UK (C1163/A12765 and C1163/A21762), the Wellcome Trust (105031/D/14/Z), the MRC (MR/M008975/1), NIH-NIDDK (R24 DK106766-01), and core funding from the Wellcome Trust to the Cambridge Stem Cell Institute (097922/Z/11/Z).

CONFLICT OF INTEREST

None declared.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data related to this article can be found athttp://dx.doi.org/10.1016/j.molmet.2017.07.014.

REFERENCES

[1] Gribble, F.M., Reimann, F., 2016. Enteroendocrine cells: chemosensors in the intestinal epithelium. Annual Review of Physiology 78:277–299.
[2] Tan, T., Bloom, S., 2013. Gut hormones as therapeutic agents in treatment of diabetes and obesity. Current Opinion in Pharmacology 13:996–1001.
[3] Sjölund, K., Sandén, G., Håkanson, R., Sundler, F., 1983. Endocrine cells in human intestine: an immunocytochemical study. Gastroenterology 85:1120–1130.
[4] Habib, A.M., Richards, P., Cairns, L.S., Rogers, G.J., Bannon, C.A., Parker, H.E., et al., 2012. Overlap of enteroendocrine hormone expression in the mouse intestine revealed by transcriptional profiling and flow cytometry. Endocrinology 153:3054–3065.
[5] Little, T.J., Doran, S., Meyer, J.H., Smout, A.J., O’Donovan, D.G., Wu, K.L., et al., 2006. The release of GLP-1 and ghrelin, but not GIP and CCK, by glucose is dependent upon the length of small intestine exposed. American Journal of Physiology — Endocrinology And Metabolism 291:E647–E655.
[6] Fothergill, L.J., Callaghan, B., Hunne, B., Bravo, D.M., Furness, J.B., 2017. Co-storage of enteroendocrine hormones evaluated at the cell and subcellular levels in male mice. Endocrinology. http://dx.doi.org/10.1210/en.2017-00243.
[7] Reimann, F., Habib, A.M., Tolhurst, G., Parker, H.E., Rogers, G.J., Gribble, F.M., 2008. Glucose Sensing in L cells: a primary cell study. Cell Metabolism 6:532–539.
[8] Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., Sandberg, R., 2014. Full-length RNA-seq from single cells using Smart-seq2. Nature Protocols 9:171–181.
[9] Scialdone, A., Tanaka, Y., Jawaid, W., Moignard, V., Wilson, N.K., Macaulay, I.C., et al., 2016. Resolving early mesoderm diversification through single-cell expression profiling. Nature 535:289–293.
[10] Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biology 11:R106.