



OPEN

## Gut microbiota alternation under the intestinal epithelium-specific knockout of mouse *Piga* gene

Aditi Jangid<sup>1</sup>, Shinji Fukuda<sup>2,3,4,5</sup>, Masahide Seki<sup>6</sup>, Yutaka Suzuki<sup>6</sup>, Todd D. Taylor<sup>7</sup>, Hiroshi Ohno<sup>3,4</sup> & Tulika Prakash<sup>1,7</sup>✉

Crosstalk between the gut microbiota and intestinal epithelium shapes the gut environment and profoundly influences the intestinal immune homeostasis. Glycosylphosphatidylinositol anchored proteins (GPI – APs) contribute to a variety of gut-associated immune functions, including microbial surveillance and defense, and epithelial cell polarity. Properly polarised epithelial cells are essential for the establishment of the barrier function of gut epithelia. The *Piga* gene is one among seven genes that encode for an enzyme which is involved in the first step of GPI-anchor biosynthesis. This is the first study reporting a knockout of the intestinal epithelial cell-specific *Piga* gene (*Piga*<sup>-/-</sup>) and its association with the gut microbiota in mice using a whole metagenome shotgun-based sequencing approach. An overall reduced microbiota diversity has been observed in the *Piga*<sup>-/-</sup> group as compared to the control group (ANOVA  $p = 0.34$ ). The taxonomic biomarkers, namely: Gammaproteobacteria (class), Enterobacteriales (order), *Enterobacteriaceae* (family), *Escherichia* (genus), *Proteus* (genus) and *Escherichia coli* (species), increased more in the *Piga*<sup>-/-</sup> mice as compared to in the control group. Further, the pathogenic *E. coli* strains, namely *E. coli* O157:H7 str. EDL 933 (EHEC), *E. coli* CFT073 (UPEC) and *E. coli* 536 (UPEC), were found in the *Piga*<sup>-/-</sup> mice which also harbored virulence factor transporters. In addition, the taxa responsible for short chain fatty acid production were decreased in the *Piga*<sup>-/-</sup> group. The *Piga*<sup>-/-</sup> mice gut harbored an increased number of microbial functions responsible for the survival of pathogens in the inflamed gut environment. Our observations clearly indicate that the *Piga*<sup>-/-</sup> mice gut might have an overall enhancement in pathogenic behaviour and reduced capabilities beneficial to health.

The microbes and their cumulative genomes, called the microbiome, modulate metabolic phenotype and influence the host immune system<sup>1</sup>. The interactions between gut microbiota and the immune system commence at birth. The gut microbiota influence the development of the immune system; and the immune system in turn shapes the gut microbiota composition<sup>2</sup>. Commensal bacteria act on the host's immune system to induce protective responses that prevent colonization and invasion by pathogens<sup>3</sup>. In genetically susceptible individuals atypical interactions between the microbiome and the host's immune system may contribute to the development of complex immune-mediated diseases, including inflammatory bowel disease (IBD), systemic autoimmune diseases, and cancer, etc.<sup>4</sup>. For example, the dysregulated microbiome-immunity interaction associated with *NOD2* mutation is assumed to play an important role in Crohn's disease pathogenesis<sup>4</sup>.

Glycosylphosphatidylinositol (GPI) is a glycolipid, which is produced in the endoplasmic reticulum (ER) and to the proteins<sup>5</sup>. The GPI-anchors are involved in important processes, including protein trafficking, sorting, and

<sup>1</sup>BioX Centre and School of Basic Sciences, Indian Institute of Technology Mandi, Kamand, Mandi, Himachal Pradesh 175005, India. <sup>2</sup>Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata 997-0052, Japan. <sup>3</sup>Laboratory for Intestinal Ecosystem, RIKEN Center for Integrative Medical Sciences, Yokohama, Kanagawa 230-0045, Japan. <sup>4</sup>Intestinal Microbiota Project, Kanagawa Institute of Industrial Science and Technology, Kawasaki, Kanagawa 210-0821, Japan. <sup>5</sup>Transborder Medical Research Center, University of Tsukuba, Tsukuba, Ibaraki 305-8575, Japan. <sup>6</sup>Department of Computational Biology and Medical Sciences, The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa Chiba 277-8562, Japan. <sup>7</sup>Laboratory for Microbiome Sciences, RIKEN Center for Integrative Medical Sciences, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. ✉email: tulika@iitmandi.ac.in

dynamics. A large number of GPI-anchored proteins (GPI-APs) have been identified in higher eukaryotes including human, *Arabidopsis thaliana*, mouse, and rat. The GPI-APs function as enzymes and complement regulators, adhesion molecules, and co-receptors in the signal transduction pathways<sup>6</sup>. These proteins play pivotal roles in several physiological processes such as embryogenesis, neurogenesis, fertilization, and immune responses<sup>7–11</sup>.

With respect to the intestinal immune response, some GPI-APs are known to activate important immune molecules and play pivotal roles in host gut health. On the other hand, there are some GPI-APs which are known to promote the growth of commensal microbiota<sup>12</sup>. In addition to these, the GPI anchor has also been found as one of the dominant sorting signals in intestinal epithelial cells (IECs)<sup>13</sup> and GPI-APs may partially contribute to IECs polarity<sup>14</sup>. The polar nature of IECs is essential for the establishment of the epithelial barrier function, which is required to segregate the commensal microbiota from the host's internal milieu<sup>14</sup>. Some of the GPI-APs, including glycoprotein 2 (GP2) and Ly6/PLAUR domain containing protein 8 (Lypd 8) are reported to be involved in microbial defense and play important roles in the intestinal immune system and homeostasis<sup>15,16</sup>.

For the biosynthesis of GPI-APs, eukaryotes have conserved post-translational mechanism, which is critical for attaching these proteins to the cell membrane<sup>5</sup>. Over 20 phosphatidylinositol glycan biosynthesis protein (PIG) subclasses involved in the biosynthesis and remodeling of GPI-anchors have been discovered so far<sup>5</sup>. The first step of GPI-anchor biogenesis is the formation of N-acetylglucosamine-phosphatidylinositol with the help of an enzyme which is encoded by a cluster of seven genes including the *PIGA* (MIM 311,770) gene<sup>5</sup>. Mutations in the GPI biosynthesis genes have been linked to a growing number of human diseases. For example, somatic and germline mutations in the *PIGA* gene have led to GPI deficiencies<sup>6,17</sup> in PNH (paroxysmal nocturnal hemoglobinuria)<sup>17</sup> and XLIDD (X-linked intellectual developmental disorder)<sup>6</sup>, respectively.

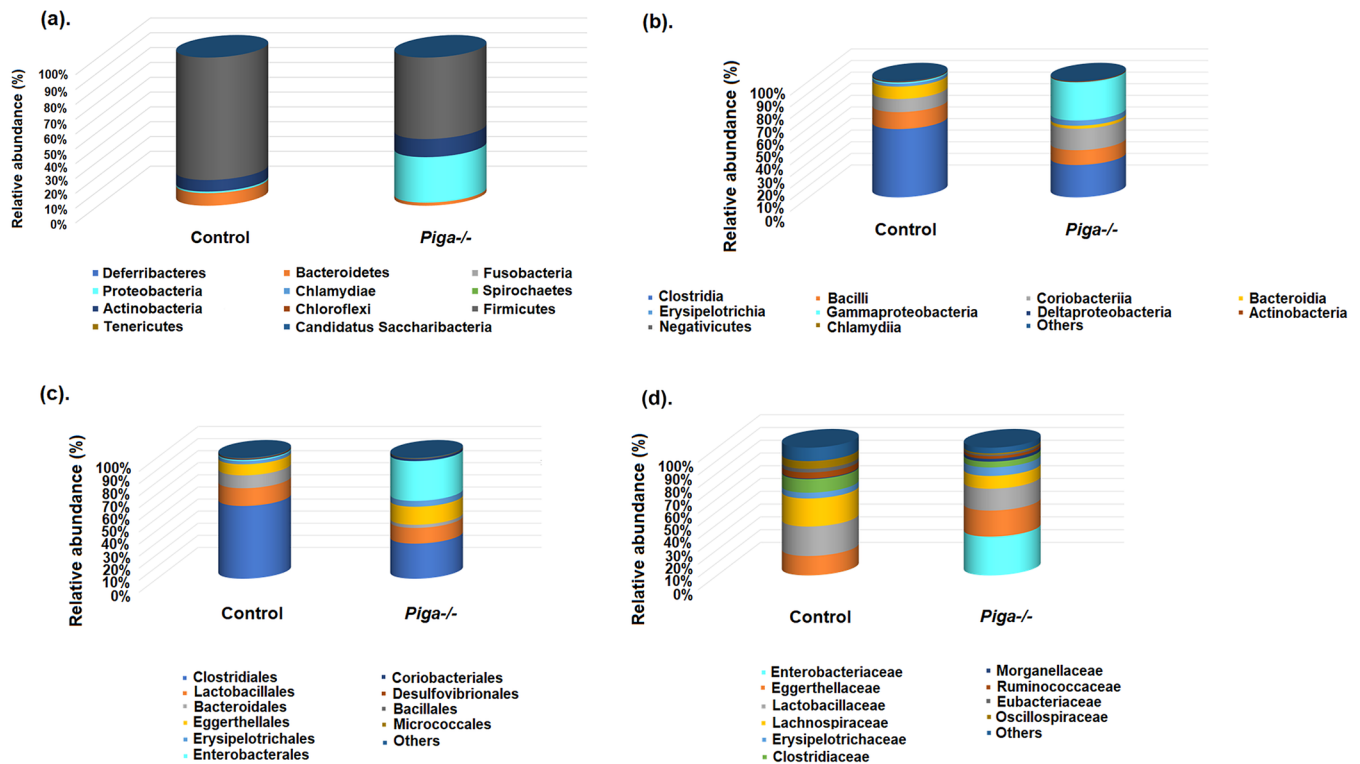
Given such important roles of the GPI-APs and gut microbiome in intestinal immune homeostasis, an exploration of the effect of one on the other will help in deciphering the possible cross-talks between these two important contributors to host health. To this end, for the first time, we developed an intestinal epithelium-specific *Piga* gene knockout in mice since this gene is known to play an important role in GPI-anchor biogenesis process. We hypothesize that by knocking out the *Piga* gene, the expression and localisation of the intestine-specific GPI-APs may be hampered, leading to a compromise in the intestinal immune system which may be associated with gut microbiota dysbiosis. The main aim of the present study is to perform a comparative analysis of the gut microbiota of the *Piga* gene knockout mice with the control mice to explore the microbial community dynamics and the respective microbial functional alternations and to gain important insights on the possible cross-talk between the GPI-APs and mice gut microbiota. To address this, we have used the whole metagenome shotgun sequencing (WMGS) based approach to perform the metagenomic sequencing of the gut microbiomes of *Piga*<sup>-/-</sup> and control mice. The scope of this analysis is to perform a comprehensive comparative metagenomic analysis of the gut microbiota of the two groups to elucidate the important microbial changes and their functional dynamics which may be associated with the impact of the knockout of the *Piga* gene on intestinal immune homeostasis. Such analysis will help us in the identification of the possible cross-talk between the two important contributors of intestinal immunity namely, the gut microbiota and GPI-APs.

## Results

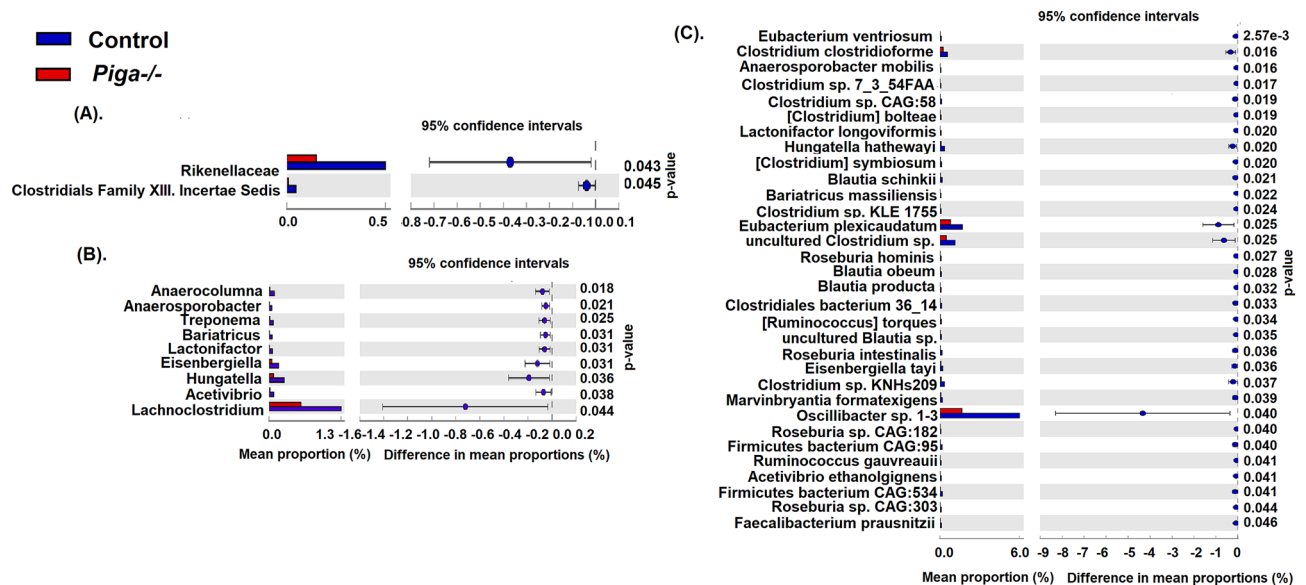
**Taxonomic identification of metagenomic samples in control and *Piga*<sup>-/-</sup> mice.** We investigated the effect of *Piga* gene knockout on the composition of fecal microbiota using whole metagenome shotgun sequencing (WMGS) analysis at all taxonomical levels (S1 Fig. and S1 Table). One of the major advantages of the WMGS approach lies in the availability of functional information and taxonomic data for the microbiome. The sequenced reads were processed to remove low quality reads, duplicated reads, and contaminations of host DNA. The remaining good quality metagenomic reads were mapped on bacterial genomic data to obtain taxonomic and functional feature counts. The obtained feature counts (taxonomic and functional) were normalized and statistical analyses were performed. A total of 2.29 billion raw reads were obtained as a result of metagenomic sequencing which resulted in 1.54 billion good quality reads. Based on the alpha-diversity analysis, the diversity of the *Piga*<sup>-/-</sup> samples is found to be lower than the controls, albeit with a p-value greater than the significance cut-off (p = 0.34 ANOVA S2 Fig.).

Metagenomic analysis of the samples showed that the relative abundance of phyla Proteobacteria (*Piga*<sup>-/-</sup> → 30.67%; control → 1.01%) and Actinobacteria (*Piga*<sup>-/-</sup> → 12.22%; control → 7.78%) increased whereas that of Firmicutes (*Piga*<sup>-/-</sup> → 54.93%; control → 82.59%) and Bacteroides (*Piga*<sup>-/-</sup> → 2.07%; control → 8.45%) decreased in the *Piga*<sup>-/-</sup> samples as compared to the control group (Fig. 1). We also explored which taxa were predominantly altered at the class, order, and family levels using the taxa relative abundances (Fig. 1 and S2 Table). An F-test analysis indicated significant alterations in the relative abundance of the order Enterobacteriales (p-value 0.001561) and family Enterobacteriaceae (p-value 3.87995 E-06) (S2 Table). At the other levels alterations in the taxa abundances are clearly observed although no statistical significance was obtained using the F-test.

As shown in Fig. 1, the relative abundances of family Lactobacillaceae, Clostridiaceae, and Lachnospiraceae belonging to phylum Firmicutes and order Bacteroidales belonging to phylum Bacteroidetes are higher in the control group than in the *Piga*<sup>-/-</sup> samples. This corroborates earlier observations that a healthy gut is dominated by Firmicutes and Bacteroides taxa in mice<sup>18</sup>. On the other hand, Enterobacteriaceae, Eggerthellaceae, and Erysipelotrichaceae of phyla Proteobacteria, Actinobacteria, and Firmicutes, respectively, are increased in the *Piga*<sup>-/-</sup> group more than in the control metagenomic samples. An increased abundance of taxa Actinobacteria and Proteobacteria correlates with a diseased gut as in the case of IBD<sup>19</sup>. However, an increased abundance of Erysipelotrichaceae of phylum Firmicutes is noteworthy. Erysipelotrichaceae have been found to be enriched in colorectal cancer and changes in their abundance have also been identified in patients with IBD or animal models of IBD<sup>20</sup>. These studies emphasize the importance of Erysipelotrichaceae in inflammation-related gastrointestinal disorders.

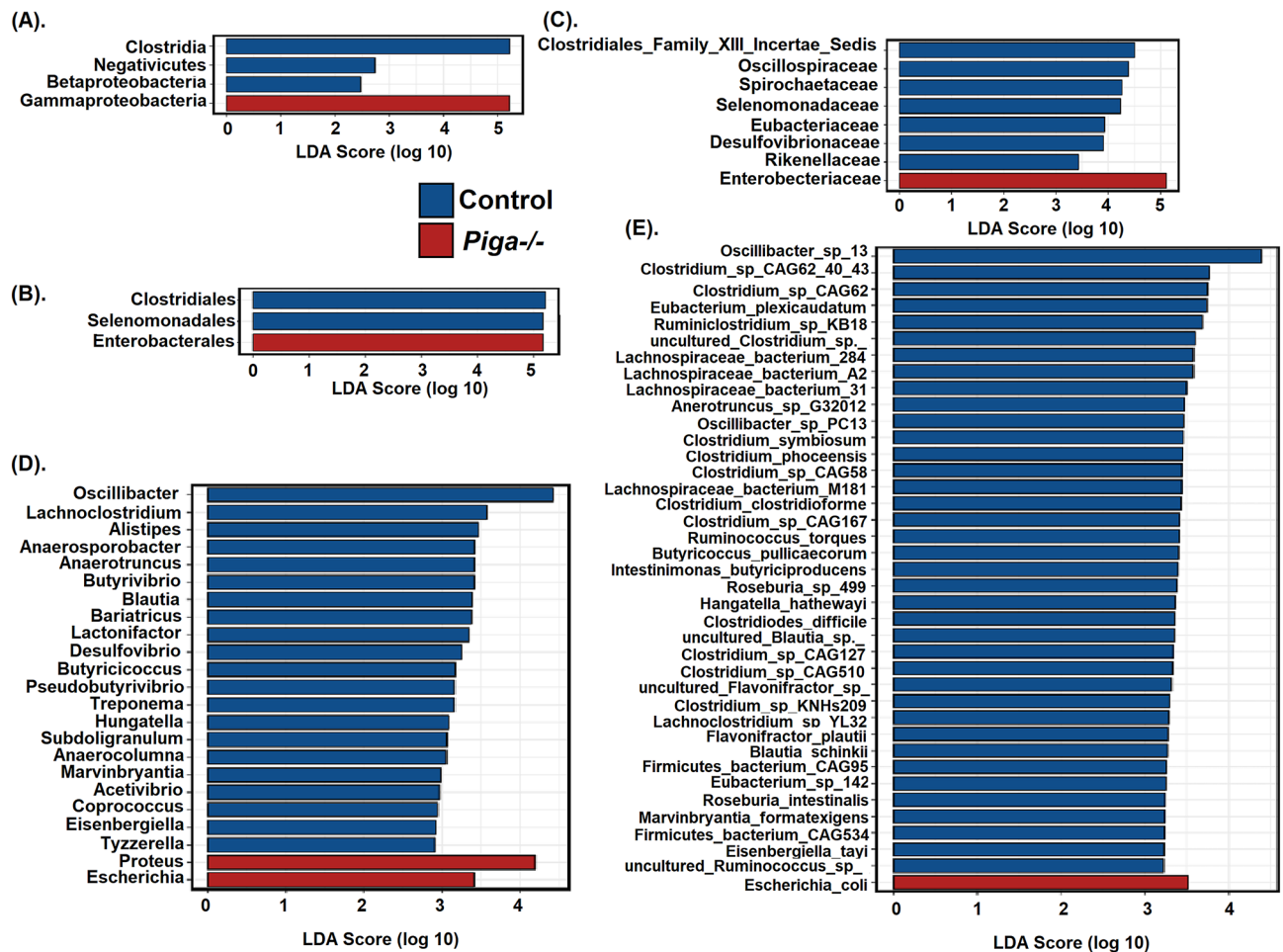


**Figure 1.** Relative abundances of taxa at the (A) phylum, (B) class, (C) order, and (D) family levels in the *Piga*<sup>-/-</sup> and control metagenomic samples.



**Figure 2.** Extended bar chart represents statistically differing taxa at (A) family, (B) genus, and (C) species levels. Welch's t-test was performed with  $p < 0.05$  and a confidence interval of 95%.

We used another method, Welch's t-test on the differences in mean-proportions of taxa abundance to examine significant differences in the microbiota composition at various taxonomic levels between the *Piga*<sup>-/-</sup> and control metagenomic samples (Fig. 2). Using this method, at the phylum, class, and order levels no taxa were found to be significantly altered statistically. At the family, genus, and species level all the significantly differing taxa were found to be decreasing in the *Piga*<sup>-/-</sup> as compared to the control metagenomic samples (Fig. 2A–C). Most of the altered taxa belong to the *Lachnospiraceae*, *Ruminococcaceae*, *Clostridiaceae*, and *Eubacteriaceae* families (S3 Table). Some members of these families have been noted for their active participation in the production of short chain fatty acids (SCFAs), specifically butyrate<sup>21</sup>. The remaining non-butyrogenic taxa may influence butyrate production by making other butyrate precursors such as acetate, lactate, and succinate (S3 Table) while



**Figure 3.** Differentiating biomarkers identified using the LEfSe analysis on different levels of taxonomy, namely, (A) class, (B) order, (C) family, (D) genus, and (E) species.

some others may be involved in propionate production<sup>22</sup>. This indicates that in the *Piga*<sup>-/-</sup> mice the population of beneficial gut commensal microbes is reduced and subsequently their potential for butyrate production in addition to the other SCFAs also is diminished.

To determine the differentiating biomarker taxa specific to the *Piga*<sup>-/-</sup> and control groups a LEfSe analysis was performed on the metagenomic reads (Fig. 3). This analysis identified taxonomic biomarkers at all levels, except at the phylum level. At the species level only one taxa, namely *Escherichia coli*, was found to be a biomarker specific to the *Piga*<sup>-/-</sup> group. Consistent with this observation, *Escherichia* was identified as one of the biomarkers at the genus level in the *Piga*<sup>-/-</sup> group as well. Further, *Enterobacteriaceae*, Enterobacteriales, and Gammaproteobacteria were identified as the biomarkers at the family, order, and class levels, respectively, in the *Piga*<sup>-/-</sup> group. In addition to *Escherichia*, *Proteus* were also identified as the biomarker for the *Piga*<sup>-/-</sup> group at the genus level. These observations clearly demonstrate that the gram-negative bacterial population is profoundly increased in the *Piga*<sup>-/-</sup> mice as compared to the control. Interestingly, some of the members of genus *Escherichia*, and *Proteus* are known to be potential pathogens in gut associated diseases. For example, *E. coli* and *Proteus spp.* have been suggested to play a role in the pathogenesis of IBD<sup>23</sup>. This indicates that in the absence of the *Piga* gene the epithelial barrier breach by pathogenic microbes might be increased, leading to a rise in their population in the *Piga*<sup>-/-</sup> mice.

**Functional identification in Control and *Piga*<sup>-/-</sup> metagenomic samples.** The functional prediction for the metagenomic reads was performed by mapping the reads on the EggNOG functional classes. In order to explore the microbial functional variations in the *Piga*<sup>-/-</sup> as compared to the control samples, Welch's t-test was performed on the predicted functions based on the EggNOG classification of the metagenomic reads. The statistically significantly different functions at the levels -1 and -2 were not identified (S3 Fig.). Using Welch's t-test analysis, at the level -3, we found 226 significantly altered COGs/NOGs belonging to cellular processes and signaling, information storage and processing, and metabolism functional classes (S4 Table, S4 Fig.).

In this analysis, at level -3 a large majority of the COGs/NOGs were found to be decreasing in the *Piga*<sup>-/-</sup> as compared to the control samples. Among these are the genes involved in cobalamin (vitamin B12) biosynthesis (COG2243, COG1492, COG2875, COG1270, COG1797, COG2242, COG2082, COG2038), biotin metabolism



| COGs   | Roles  |
|--|--|
| Fumarate hydratase class II (COG0114)  | Involved in the TCA cycle and its expression was observed in conditions of iron limitation and oxidative stress <sup>25</sup>  |
| YhcH YjgK YiaL family protein (COG2731)  | YjgK (TabA) influences biofilm formation by toxin-antitoxin systems <sup>26</sup> . YhcH is a possible sugar isomerase of sialic acid catabolism <sup>27</sup>                               |
| Part of the ABC transporter complex PotABCD involved in spermidine putrescine import (COG3842) | Spermidine accumulation increases resistance to oxidative stress in gram-negative bacteria and their survival in macrophages <sup>28</sup>   |
| Polypeptide-transport-associated domain protein, ShlB-type (COG2831)                           | ShlB is important in the secretion and activation of the haemolysin ShlA (cell bound hemolysin, releases heme-iron from erythrocytes by interaction with erythrocyte membrane) <sup>29</sup> |

**Table 1.** Significantly increased microbial functions in the *Piga*<sup>-/-</sup> mice gut microbiota that might help these bacteria in their survival in an inflamed gut.

(COG1654, ENOG410XQH3), riboflavin metabolism (COG0807), nicotinate and nicotinamide metabolism (COG1712) and propanoate metabolism (COG1803, COG0511). This suggests that the gut microbiota associated with the *Piga*<sup>-/-</sup> mice have lost their capability to produce cofactor and SCFAs as compared to the control mice. In addition, COGs involved in other processes such as defense mechanisms (ENOG410ZX9X, ENOG410XTHI, ENOG410ZYVS, ENOG410ZVEA, ENOG4111FSE, ENOG410XQ2T, ENOG410Y2F5, ENOG410Y12H, ENOG41120ZG, ENOG410ZWUI, ENOG410YATZ, ENOG410XTDY, ENOG410ZWZY), intracellular trafficking, secretion, and vesicular transport (ENOG41122HV, ENOG410XR7W, ENOG41124QY, COG3505) and signal transduction mechanisms (ENOG4111VXS, ENOG4111HM8, COG0643, COG4252, ENOG4111GNT, COG2172, ENOG4111Y53, ENOG410XPTB, ENOG410YSNH, ENOG410ZWIB, COG5001, ENOG41120GI) are also found to be significantly reduced in the *Piga*<sup>-/-</sup> as compared to the control samples.

Only 11 COGs/NOGs were found to be significantly increased in the *Piga*<sup>-/-</sup> as compared to the control metagenomic samples (S4 Table). Ribose-5-phosphate isomerase A (RpiA: COG0120) (S4 Table) is found to be significantly increased in the *Piga*<sup>-/-</sup> mice gut as compared to the control. This enzyme interconverts ribose-5-phosphate and ribulose-5-phosphate and plays critical roles in the anabolism and catabolism of carbohydrates (especially in the pentose-phosphate pathway). Pentose-phosphates are the crucial intermediates of the cellular-metabolism of *E. coli*<sup>24</sup>. The production of the heptose components of lipopolysaccharide (LPS), as well as other essential components, requires RpiA activity<sup>24</sup>. This indicates an enhanced growth of LPS covered bacteria in the knockout mice gut. The other functions that may aid bacterial survival in an inflamed gut are also found to be significantly increased in the *Piga*<sup>-/-</sup> mice gut as compared to the control (Table 1). It is interesting to note that the taxonomic biomarkers in the *Piga*<sup>-/-</sup> mice were gram-negative bacterial populations. The LPS covering of these gram-negative bacteria might be responsible for the inflammation in the gut and subsequently these bacteria might produce the observed genes that will allow them to reside in the inflamed gut. These findings indicate that in the gut of knockout mice, where the immune system was already weakened, the microbiota lowered the expression of genes implicated in processes like defense, signaling and transport and trafficking and might utilize their energy in enhancing those genes that aid in their overall survival in the inflamed gut.

**Backtracing of significantly altered functions on significantly altered taxa.** Aiming to understand which taxa may be primarily responsible for the alternations in microbial functions observed in the *Piga*<sup>-/-</sup> mice gut microbiota as compared to the control, we performed a backtracing analysis. In addition, the backtracing analysis helps in identifying the microbial strain level differences. In this analysis, we explored which significantly altered bacterial taxa harbored the significantly altered functions observed in the *Piga*<sup>-/-</sup> mice gut microbiota. Out of the 215 COGs/NOGs, which are found to be significantly increased in the control metagenomic samples, 142 functions backtraced on 12 taxa belonging to eight distinct bacterial species (S5 Table). These species include *Eubacterium ventriosum* ATCC 27,560, *Clostridium bolteae* ATCC BAA-613, *Clostridium symbiosum* WAL-14163, *Marvinbryantia formatexigens* DSM 14469, *Roseburia hominis* A2-183, *Roseburia intestinalis* XB6B4, *Ruminococcus torques* ATCC 27756, *Ruminococcus torques* L2-14, *Faecalibacterium prausnitzii* L2-6, *Faecalibacterium cf. prausnitzii* KLE1255, *Faecalibacterium prausnitzii* SL3/3, and *Faecalibacterium prausnitzii* A2-165. As it is evident from S3 Table that several of these bacterial species are actively involved in SCFA production.

It is interesting to note that all the 11 significantly altered COGs/NOGs, which are found to be increased in the *Piga*<sup>-/-</sup> metagenomic samples, backtraced on different strains of the *E. coli* (S6 Table). It is noteworthy that *E. coli* is found as the taxonomic biomarker for the *Piga*<sup>-/-</sup> group as compared to the control in our analysis. This clearly indicates that *E. coli* is one of the prime taxa which is increased in the *Piga*<sup>-/-</sup> mice and might start to express those functions which will help it to survive in the inflamed gut. The strains of *E. coli* which are identified in this analysis include *E. coli* O157:H7 str. EDL 933 (EHEC), *E. coli* CFT073 (UPEC), *E. coli* 536 (UPEC), *E. coli* str. K-12 substr. DH10B, *E. coli* str. K-12 substr. W3110, *E. coli* str. K-12 substr. MG1655, *E. coli* ATCC 8739, and *E. coli* BL21 (DE3). Among the above *E. coli* strains, *E. coli* O157:H7 str. EDL 933 (EHEC), *E. coli* CFT073 (UPEC), and *E. coli* 536 (UPEC) are the pathogenic strains<sup>30</sup>. In addition, these are the only strains which mapped on the polypeptide-transport-associated domain protein, ShlB-type (COG2831) function, which is a virulence factor transporter. This clearly indicates that in the absence of the *Piga* gene these three *E. coli* strains might be the prime pathogenic strains which might breach the epithelial barrier and further contribute to gut inflammation.

## Discussion

**Role of GPI-APs in gut-associated immune processes.** In order to understand the cross-talk between the gut microbiota and intestine specific GPI-APs, we have created intestinal epithelium-specific *Piga*<sup>-/-</sup> mice and explored the microbiota dynamics with respect to the control in the present analysis. Several members of the GPI-APs family are expressed on the apical or basolateral surfaces of the IECs. GPI-anchors play an important role as a signal for the transport of proteins to the apical surface of the fully polarized epithelial cells<sup>14</sup>. A loss of the *Piga* gene due to knockout is expected to result in an impaired GPI-anchoring process. This in turn might lead to diseases by affecting the expression and localisation of various GPI-APs in the intestine and the important functions imparted by them. For example, the mislocalisation of the apical proteins in IECs leads to malnutrition and diarrhoeal disorders<sup>31</sup>. On the other hand, the mislocalisation of basolateral proteins correlates with loss of epithelial architecture, cancer development, and IBD. Although not confirmed, the mislocalisation of  $\alpha 2$ -integrin, which is a basolateral GPI-AP, has been reported in some microvillus inclusion disease (MVID) patients and mouse models.

The GPI-APs are known to play several important roles in overall cell biology, including intestinal immunity (S7 Table). IECs are constantly challenged by a bombardment of foreign antigens and environmental microorganisms. The epithelial barrier not only carries out pathogenic microbial surveillance and prevents their breach from the lumen across the barrier, but also maintains immune tolerance to commensal microbes. A precise regulation of the intestinal epithelial barrier allows the maintenance of mucosal immune homeostasis and prevents the onset of uncontrolled inflammation. The functions of epithelial cells are executed by their polar nature, which are generated in part by GPI-APs<sup>14</sup>. A loss of epithelial cell polarity is associated with impaired epithelial homeostasis and has been linked with several diseases<sup>32,33</sup>. Similarly, a defective microbial surveillance mechanism in the intestinal lumen is known to result in a loss of mucosal immune homeostasis leading to an inflamed gut<sup>33</sup>. Those GPI-APs which are present on the apical surface come in contact with the gut microbiota or the other microbes that enter into our bodies through our digestive tracts. In this context, glycoprotein 2 (GP2), Ly6/PLAUR domain-containing protein 8 (Lypd8), uromodulin (UMOD), cellular prion protein (PrP<sup>C</sup>), and  $\beta 1$  Integrin are particularly interesting GPI-APs since they are all present on the apical surface of the M-cells of the epithelial layer (S7 Table).

Lypd8 is a GPI-AP which is a secretory protein specifically needed to prevent the invasion of gram-negative flagellated bacteria in the inner-mucus layer. It binds to the flagellum of bacteria, such as *Helicobacter*, *Escherichia*, and *Proteus* (*P. mirabilis*), thereby inhibiting bacterial motility in the intestinal lumen. Segregation of epithelial cells and intestinal bacteria is required to keep intestinal homeostasis in the colon<sup>16</sup>. *Lypd8*<sup>-/-</sup> mice were found to be extremely sensitive to dextran sulfate sodium (DSS) induced intestinal inflammation. Antibiotic administration eliminated the gram-negative flagellated bacteria, re-established the bacterial free state of the inner-mucus layer, and ameliorated intestinal inflammation induced by DSS in these mice. Recently, Lypd8 deficiency has been shown to promote the rapid colonization of *Citrobacter rodentium* (family *Enterobacteriaceae*) in the colon resulting in severe colitis with neutrophil and Th17-cell expansion in the lamina propria<sup>34</sup>. Lypd8 prevents *C. rodentium* from attaching to the intestinal colonic epithelial cells, thereby preventing intimin and translocated intimin receptor interaction<sup>34</sup>. Hence, Lypd8 plays an important role in early-phase defense against *C. rodentium*<sup>34</sup>.

The GP2 protein is a GPI-AP which is expressed on the M cells and acts as a bacterial uptake receptor. A previous study has demonstrated that it mediates antigen transcytosis by M cells in the mucosal-associated lymphoid tissue<sup>15</sup>. It specifically recognizes FimH, on the outer membrane of a subset of gram-negative enterobacilli such as *E. coli* and *Salmonella enterica*<sup>15,35</sup>. However, recently it has been found that the Pancreatic acinar cells and not M cells, are the source of intestinal GP2<sup>36</sup>. Pancreas-specific GP2-deficient colitis mice have a larger mucosal *E. coli* population than the intact mice, indicating that the digestive-tract GP2 binds the commensal *E. coli*, preventing the epithelial attachment and penetration by this pathogen. However, both the pancreas-specific and intestinal epithelium-specific GP2-deficient mice do not exhibit any significant alterations in the luminal microbiota content as compared to the wild type mice. Thus, although the intestinal epithelium M-cell specific GP2 have been shown to be involved in FimH + bacterial recognition and their subsequent translocation, their roles in the mechanism of microbial surveillance and clearance remains poorly understood and requires further exploration.

UMOD is another GPI-AP, which is closely related to GP2 and shows binding towards the FimH + Type 1 pili. This gene was first characterized as the Tamm-Horsfall protein expressed on the apical surface of renal tubular epithelial cells and capable of being shed into urine where it binds to uropathogenic *E. coli*, interfering with bacterial adhesion and pathogenicity<sup>37,38</sup>. However, the M-cells specific UMOD serve as an uptake receptor for *Lactobacillus acidophilus* (L-92 strain), a probiotic strain known to exert anti-allergic immunomodulatory effects upon oral administration in mice<sup>39</sup>. The interaction of the surface layer protein A (SlpA) on L-92 with UMOD on the M-cell surface is responsible for M-cell mediated uptake of L-92 and the bacteria is subsequently delivered to the Peyer's patches (PP) dendritic cells for immunomodulation. These observations suggest that bacterial access to the gut immune system is a crucial process to promote host immune responses and some GPI-APs play crucial roles in such processes. Corroborating with this role of UMOD, in our study, we have noted an overall reduced relative abundance of *Lactobacillus* genera in the *Piga*<sup>-/-</sup> mice group.

Although the proteins mentioned above serve a crucial function in microbial surveillance or immunomodulation processes, some GPI-APs have been found to be exploited by pathogenic bacteria. *Brucella abortus* is a Gram-negative bacterium causing a form of brucellosis in cattle and other domestic animals which results in their transmission to humans. PrP<sup>C</sup> is a GPI-AP which is expressed on M-cells and acts as an uptake receptor for *Brucella abortus* (via binding with the conserved Hsp60 proteins expressed on *B. abortus*)<sup>40</sup>. Following *B. abortus* uptake, the bacteria penetrate into M cells and are capable of surviving inside Dendritic Cells (accumulated beneath the M cells) by producing replicative vacuoles which then spread throughout the body. Similarly,

the GPI-linked protein integrin  $\beta_1$ , which is normally distributed in the basolateral membrane of enterocytes, localizes in the apical surface of M cells, and acts as a receptor for the pathogenic *Yersinia enterocolitica*<sup>38</sup>.

There are a number of intestine specific GPI-APs whose membrane location is not known. These include T-cell ecto-ADP-ribosyltransferase 1, T-cell ecto-ADP-ribosyltransferase 2, and low affinity immunoglobulin gamma Fc region receptor IV which are known to activate the important immune molecules such as macrophages and T-cells (S7 Table). GPI-APs including, hydrolases and proteases and their receptors, play important roles in host gut health (S7 Table). It is interesting to note that some of the basolateral proteins such as TLR5 and TLR9 can act as pattern recognition receptors for microbial pathogens<sup>41</sup>. Thus, the roles of the above-mentioned GPI-APs in microbial recognition and other similar processes cannot be ruled out and requires further exploration.

One of the GPI-APs, namely intestinal-type alkaline phosphatase (IAP), is found to promote the growth of commensal microbiota and maintain overall gut immunity. IAP is a small intestinal brush border enzyme which has been recognised as a gut mucosal defense factor. IAP has the ability to detoxify lipopolysaccharides (LPS) from Gram-negative bacteria and exogenous IAP has been shown to attenuate LPS-mediated toxicity<sup>42,43</sup>. Further, an IAP knockout (IAP-KO) mice study reported increased bacterial translocation to mesenteric lymph nodes when the intestine was subjected to a local or distant ischaemic injury<sup>44</sup>. The luminal environment of IAP-KO mice is unfavorable particularly for common enteric commensal *E. coli*<sup>44</sup>. Also the gut of IAP-KO mice have more Clostridia belonging to the Firmicutes phylum than the wild type<sup>44</sup>.

GPI-AP Intelectin have been implicated in innate immunity and are mainly produced by intestinal goblet cells, intestinal Paneth cells, and lungs<sup>45–47</sup>. In mice, expression of intelectin rises upon intestinal parasitic nematode infection<sup>48</sup>. Intelectins are also found abundantly in the mucus produced by allergic responses in humans<sup>49</sup>. Recently, human intelectins (hIntL-1) are found to differentiate between the glycans found on mammalian cells from those of microorganisms resulting in the recognition of the later<sup>50,51</sup>. The ligands of hIntL-1 include glycans from Gram-positive and Gram-negative bacteria, including *Streptococcus pneumoniae*, *Proteus vulgaris*, *Proteus mirabilis*, *K. pneumoniae*, and *Yersinia pestis*<sup>50,51</sup>. This selectivity of only microbial ligands suggests that hIntL-1 functions in the microbial surveillance process. Further, mouse IntL-1 is reported to interact with the  $\beta$ -Gal gene<sup>50,51</sup>. These findings support the hypothesis that IntLs from diverse organisms have evolved to bind specific microbial epitopes and may play defensive roles against microbes.

The above discussion suggests that several intestine specific GPI-APs have important roles to play in intestinal immune homeostasis. Some GPI-APs which are present on the basolateral membrane or plasma membranes of the intestine may be involved in gut immunity either by directly or indirectly associating with the microbial members. Also, several GPI-APs which are present on the apical surface play a variety of important roles in maintaining gut-associated immunity including microbial surveillance by acting as pathogen uptake receptors or immunomodulation processes. On the other hand, some apical GPI-APs are utilized by pathogens for their uptake into the host system. Therefore, the dysregulation of the expression and localisation of these GPI-APs due to the *Piga* knockout is expected to severely affect all these immune related process and affect the overall host health.

**Family *Enterobacteriaceae* and species *E. coli* as the differentiating biomarkers of the *Pig-/-* mice from the control.** The above discussions suggest that an impaired function of GPI-APs including, *Lypd8*, *UMOD*, and perhaps *GP2* and others is expected to lead to a faulty pathogenic microbial surveillance. Since these proteins are involved in defense against pathogenic bacteria such as *E. coli*, *Proteus*, *H. pylori*, and *S. enterica* by various mechanisms, a loss in the functioning of these proteins will lead to an increased abundance of these pathogens. This is mainly because these pathogens might escape from the recognition by the immune molecules which would have otherwise eliminated them. We have found the family *Enterobacteriaceae* and species *E. coli* as the taxonomic biomarkers in the *Piga-/-* mice group. Another pathogenic genera, namely, *Proteus* is also identified as a biomarker with the relative abundance of the species *Proteus mirabilis* increased in the *Piga-/-* mice. Further, in the *Piga-/-* mice microbial functions that help in the survival of pathogenic microbes in an inflamed gut are found to be significantly increased. In addition, we have observed an increased relative abundance of the fimbrial proteins (COG3539, COG4969) in the *Piga-/-* mice.

Interestingly, the functions increased in the *Piga-/-* mice in our backtracing analysis are found to be mapping on only *E. coli* species including the three pathogens namely, *E. coli* O157:H7 str. EDL 933 (EHEC), *E. coli* CFT073 (UPEC), and *E. coli* 536 (UPEC). All these pathogenic *E. coli* strains are found to harbor a virulence factor transporter. A recent study showed that Enteropathogenic *E. coli* (EPEC) can disrupt cell polarity, causing basolateral membrane proteins, in particular  $\beta_1$ -integrins, to migrate to the apical cell surface. These  $\beta_1$ -integrins on the apical cell surface can bind to the intimin of EPEC, which mediates the intimate attachment of EPEC to epithelial cells<sup>52</sup>. Although Uropathogenic *E. coli* (UPEC) is responsible for causing urinary tract infection, it is likely that the infection begins with the colonization of the bowel with a uropathogenic strain in addition to the commensal flora.

From these discussions, we conclude that due to the *Piga* gene knockout the pathogenic bacteria will escape the host immune system by preventing their recognition and transportation through the immune molecules (for example, due to the lack of *Lypd8* protein) and subsequently this will lead to a gut microbiota dysbiosis. Since the GPI-APs such as *GP2*, which are actively involved in the surveillance of pathogenic *E. coli* in gut are of pancreatic origin, a knockout of intestinal *Piga* is not expected to impact the function of the *GP2* of pancreatic origin. However, we observe a significantly higher abundance of pathogenic *E. coli* in *Piga-/-* mice group. This clearly indicates that the process of microbial surveillance is performed by other factors (*Lypd8* and *UMOD* and others) than the pancreatic *GP2* proteins. As the abundance of pathogenic *E. coli* species increases in the gut, strains such as EPEC can further disrupt the epithelial cell polarity, thereby aggravating the condition further.

This also indicates that gut inflammation may be one of the phenotypes associated with the intestine specific *Piga*<sup>-/-</sup> in mice, mainly due to the increase in the abundance of pathogenic microbes.

***Piga*<sup>-/-</sup> mice gut microbiome exhibits a reduced ability for SCFA production.** Another interesting observation is a significantly reduced abundance of the butyrate and propionate producing commensal bacteria in the *Piga*<sup>-/-</sup> mice group. These include *F. prausnitzii*, *R. intestinalis*, *R. hominis*, *E. ventriosum*, *E. plexicaudatum*, *R. torques*, *B. massiliensis*, *E. tayi*, *C. symbiosum*, *C. clostridioforme*, *Clostridium* sp. KNHs209 and *Hungatella hathewayi*. Further, the significantly increased microbial functions in the control mice backtraced predominantly on those commensal species which were responsible for SCFA production. These observations suggest that in the *Piga*<sup>-/-</sup> mice the population of beneficial gut commensal microbes may get reduced and subsequently their potential for butyrate and other SCFAs production may also get diminished. This suggests that the intestinal mucosa of the *Piga*<sup>-/-</sup> mice have lost their energy source that could have been provided by these commensal bacteria. Butyrate is a critical metabolite for the overall gut physiology and host wellbeing. It is the main source of energy for the colonic epithelial cells, and it has defensive properties against IBD and colorectal cancer (CRC). Butyrate can reduce mucosal inflammation through upregulating peroxisome proliferator-activated receptor (PPAR)-gamma, inhibiting NF-κB transcription factor activation, and interferon gamma<sup>53</sup>. In addition, propionate has been demonstrated to be converted into glucose through intestinal gluconeogenesis (IGN) thus further improving the energy homeostasis<sup>54</sup>.

*F. prausnitzii* is a representative commensal member of the phylum Firmicutes and is well-known for playing a number of important functions in gut homeostasis. In a previous study the administration of *F. prausnitzii* (strain A2-165) or its supernatant significantly reduced colitis severity induced by trinitrobenzenesulfonic acid (TNBS) and restored dysbiotic gut microbiota community<sup>55</sup>. In addition, *F. prausnitzii* (strain A2-165) exerted anti-inflammatory effects by inhibiting NF-κB activation and IL-8 production (on both cellular and colitis animal models)<sup>55</sup>, facilitating the IL-10 induction in murine and human dendritic cells and altering T-cell responses<sup>56</sup>. Furthermore, in colitis mice models, *F. prausnitzii* (strain A2-165) or its supernatant has been shown to reduce colitis severity by downregulating pro-inflammatory cytokines, myeloperoxidases, and T-cell levels, as well as maintaining the intestinal epithelial barrier<sup>57,58</sup>. *F. prausnitzii* may be a promising candidate for next-generation probiotics because of these characteristics<sup>59</sup>. In corroboration with these observations, we found significant reduction in *F. prausnitzii* in our *Piga*<sup>-/-</sup> group, leading us to suspect that *Piga*<sup>-/-</sup> mice gut might have lost the probiotic community of microbes and their subsequent beneficial anti-inflammatory effects.

*Roseburia* spp., along with *F. prausnitzii* and *Eubacterium* spp., form a prominent butyrate-producing Firmicutes group that accounts for 7–24% of all bacteria in a healthy human colon<sup>60</sup>. A previous study reported that the intestinal-abundance of *R. intestinalis* was much lower in untreated CD patients than in healthy controls<sup>61</sup>. To reduce intestinal inflammation, *R. intestinalis* raised the number of T-regulatory (Treg)-cells and enhanced the gene expression of the cytokines, namely, transforming growth factor-β (TGF-β) and thymic stromal lymphopoietin (TSLP)<sup>61</sup>. In addition, *R. intestinalis* supernatant was observed to ameliorate colitis in TNBS and DSS-induced mice models by decreasing the count of Th17 cells and inflammatory macrophages in the colon, and by suppressing the gene expression of signal transducer and activator of transcription 3 (STAT3) and IL-6<sup>62</sup>. In addition, *R. hominis* (strain A2-183) promotes and regulates innate immunity via upregulating genes related to gut barrier function, antimicrobial peptides, T-cell biology, and toll-like receptors (TLR) signaling<sup>63</sup>. Furthermore, treatment with *R. hominis* provided protection against DSS-induced colitis<sup>63</sup>. Additionally, patent rights have been granted for the use of several strains of *E. ventriosum* in the treatment of colitis and/or colorectal cancer<sup>64</sup>. In our study we found a significant reduction in *Roseburia* spp., and *Eubacterium* spp. in the *Piga*<sup>-/-</sup> group as compared to controls leading us to hypothesize that the commensal community of microbes and their subsequent beneficial anti-inflammatory and immunomodulatory effects are lost in *Piga*<sup>-/-</sup> mice gut.

On the other hand, we found *E. coli* O157:H7 str. EDL 933 (EHEC), *E. coli* CFT073 (UPEC) and *E. coli* 536 (UPEC) pathogenic strains as a taxonomic biomarker in *Piga*<sup>-/-</sup> group. These observations corroborate the fact that colonic SCFA regulate the growth and virulence of enteric pathogens, such as enterohemorrhagic *E. coli* (EHEC)<sup>65</sup>. A reduced ability to produce SCFAs has been shown to be linked with gut microbiota dysbiosis and an increased abundance of microbes with deleterious effects on host health<sup>66</sup>.

This is the first study of the intestinal epithelium specific *Piga* gene knockout creation in mice and its impact on the fecal microbiota by using the whole metagenome shotgun sequencing method. In *Piga*<sup>-/-</sup> mice, we were able to identify the main features of microbiota dysbiosis in terms of taxa and gene functions. The findings in this paper provide new directions for designing more focused experiments to confirm the microbial and functional changes associated with the *Piga*<sup>-/-</sup> and gut health.

## Materials and methods

**Animal experiments, knockout preparation, and sample collection.** All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of the RIKEN Yokohama Branch. Mice were maintained under specific pathogen-free (SPF) conditions in the animal facility at the RIKEN Yokohama Branch. 10-week-old male SPF (C57BL/6) mice were taken for experiment. Conditional *Piga* floxed mice were crossed with *villin*-Cre transgenic mice<sup>67</sup> to obtain mice with intestinal epithelium-specific deletion of *Piga* gene. A total of six fresh fecal samples were collected from three *Piga* floxed mice (Control) and three intestinal epithelium-specific *Piga* gene knockout (*Piga*<sup>-/-</sup>) mice. The fecal samples were stored at -80 °C before DNA extraction.

**DNA extraction.** Fecal DNA extraction was performed as described previously<sup>68</sup>. Briefly, 10 mg of freeze-dried fecal samples were disrupted with 3 and 0.1 mm zirconia/silica beads by vigorous shaking (1,500 r.p.m.



for 5 min) using a Shake Master (Biomedical Science) suspended in DNA extraction buffer containing 200  $\mu$ L of 10% (w/v) SDS/TE (10 mM Tris-HCl, 1 mM EDTA, pH8.0) solution, 400  $\mu$ L of phenol/chloroform/isoamyl alcohol (25:24:1), and 200  $\mu$ L of 3 M sodium acetate. After centrifugation, bacterial genomic DNA was purified by the standard phenol/chloroform/isoamyl alcohol protocol. RNAs were removed from the sample by RNase A treatment.

**Whole metagenome shotgun sequencing (WMGS) and read quality improvement.** The complete workflow of the metagenomic analysis is provided as S1 Fig. WMGS sequence libraries were developed using the Illumina TruSeq DNA Sample Preparation kit with catalog number PE-940–2001. Sequencing was carried out using the Illumina HiSeq2000 platform to produce paired-end reads of 125 bp. In a step of end repair, the fragments were purified using AMPureXP beads with gel-free method. The accuracy of raw reads was analysed using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). The removal of the adapter sequences and the filtering of reads based on the Q-score (Q = 30) was performed using FaQCs (v1.34). Finally, the duplicated reads and the reads mapped on the host DNA were removed using PRINSEQ (PRINSEQ-lite 0.20.4) (<http://prinseq.sourceforge.net/>) and Bowtie2 (v2.2.5)<sup>69</sup>, respectively.

**Metagenomic data annotations and statistical analysis.** The metagenomic data processing and analysis was done using the pipeline described previously<sup>33</sup>. Briefly, the high quality reads were aligned against the nr-db (as of 2017) with default parameters using the BLASTX option of DIAMOND (v0.9.9.110)<sup>70</sup>. The alignment files were then introduced into MEGAN6 (v6.8.18)<sup>71</sup>, where taxonomic and functional binning of the reads were carried out by selecting the Lowest Common Ancestor (LCA) algorithm with minimum bit score (50) and minimum support (50) parameters. The statistically significant differences between the control and *Piga*<sup>-/-</sup> mice metagenomic samples were identified using STAMP (v2.1.3)<sup>72</sup> and Calypso<sup>73</sup>. The differences between these two groups were analyzed using Welch's t-test with confidence interval 95% and p-value < 0.05. The taxonomic biomarkers between the groups were identified using Linear discriminant analysis Effect Size (LEfSe) analysis<sup>74</sup>. Alpha-diversity measures (on genus level) carried out using Shannon index and ANOVA (< 0.05) were used to determine the statistical significance among the two groups. In addition, on taxa belonging to dominant phylum F-test and T-test also applied to see the significant changes among the two groups. The statistically significant functions were backtraced on statistically significant taxa using EggNOG 4.5.1<sup>75</sup>.

**Animal study approval.** All mice experiment procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of the RIKEN Yokohama Branch and abide by all regulatory standards of IACUC of the RIKEN Yokohama Branch. We hereby confirm the study was carried out in compliance with the ARRIVE guidelines.

## Data availability

Metagenomic samples are available in the NCBI database with Bioproject ID PRJNA712972.

Received: 13 September 2021; Accepted: 22 April 2022

Published online: 25 June 2022

## References

- Gordon, J. I. Honor thy gut symbionts redux. *Science* **336**, 1251. <https://doi.org/10.1126/science.1224686> (2012).
- Nicholson, J. K. & Wilson, I. D. Understanding “Global” systems biology: Metabonomics and the continuum of metabolism. *Nat. Rev. Drug Discovery* **2**, 668–676. <https://doi.org/10.1038/nrd1157> (2003).
- Khan, R., Petersen, F. C. & Shekhar, S. Commensal bacteria: An emerging player in defense against respiratory pathogens. *Front. Immunol.* **10**, 1203 (2019).
- Zheng, D., Liwinski, T. & Elinav, E. Interaction between microbiota and immunity in health and disease. *Cell Res.* **30**, 492–506. <https://doi.org/10.1038/s41422-020-0332-7> (2020).
- Fujita, M. & Kinoshita, T. GPI-anchor remodeling: Potential functions of GPI-anchors in intracellular trafficking and membrane dynamics. *Biochem. Biophys. Acta.* **1050–1058**, 2012. <https://doi.org/10.1016/j.bbaliip.2012.01.004> (1821).
- Tarailo-Graovac, M. *et al.* The genotypic and phenotypic spectrum of PIGA deficiency. *Orphanet J. Rare Dis.* **10**, 23. <https://doi.org/10.1186/s13023-015-0243-8> (2015).
- Kawagoe, K. *et al.* Glycosylphosphatidylinositol-anchor-deficient mice: implications for clonal dominance of mutant cells in paroxysmal nocturnal hemoglobinuria. *Blood* **87**, 3600–3606 (1996).
- Nozaki, M. *et al.* Developmental abnormalities of glycosylphosphatidylinositol-anchor-deficient embryos revealed by Cre/loxP system. *Laboratory Invest. J. Tech. Methods Pathol.* **79**, 293–299 (1999).
- Hazenbos, W. L., Clausen, B. E., Takeda, J. & Kinoshita, T. GPI-anchor deficiency in myeloid cells causes impaired Fc $\gamma$ 3R effector functions. *Blood* **104**, 2825–2831. <https://doi.org/10.1182/blood-2004-02-0671> (2004).
- Almeida, A. M. *et al.* Hypomorphic promoter mutation in PIGM causes inherited glycosylphosphatidylinositol deficiency. *Nat. Med.* **12**, 846–851. <https://doi.org/10.1038/nm1410> (2006).
- Ueda, Y. *et al.* PGAP1 knock-out mice show otocephaly and male infertility. *J. Biol. Chem.* **282**, 30373–30380. <https://doi.org/10.1074/jbc.M705601200> (2007).
- Malo, M. S. *et al.* Intestinal alkaline phosphatase promotes gut bacterial growth by reducing the concentration of luminal nucleotide triphosphates. *Am. J. Physiol. Gastrointest. Liver Physiol.* **306**, G826–838. <https://doi.org/10.1152/ajpgi.00357.2013> (2014).
- Soole, K. L., Jepson, M. A., Hazlewood, G. P., Gilbert, H. J. & Hirst, B. H. Epithelial sorting of a glycosylphosphatidylinositol-anchored bacterial protein expressed in polarized renal MDCK and intestinal Caco-2 cells. *J. Cell Sci.* **108**(Pt 1), 369–377 (1995).
- Zurzolo, C. & Simons, K. Glycosylphosphatidylinositol-anchored proteins: Membrane organization and transport. *Biochimica et Biophysica Acta (BBA)—Biomembr.*, **1858**, 632–639. <https://doi.org/10.1016/j.bbame.2015.12.018> (2016).
- Hase, K. *et al.* Uptake through glycoprotein 2 of FimH(+) bacteria by M cells initiates mucosal immune response. *Nature* **462**, 226–230. <https://doi.org/10.1038/nature08529> (2009).

16. Okumura, R. *et al.* Lypd8 promotes the segregation of flagellated microbiota and colonic epithelia. *Nature* **532**, 117–121. <https://doi.org/10.1038/nature17406> (2016).
17. Takeda, J. *et al.* Deficiency of the GPI anchor caused by a somatic mutation of the PIG-A gene in paroxysmal nocturnal hemoglobinuria. *Cell* **73**, 703–711 (1993).
18. Rinninella, E. *et al.* What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms* **7**, 14. <https://doi.org/10.3390/microorganisms7010014> (2019).
19. Soltys, K. *et al.* Seasonal changes of circulating 25-hydroxyvitamin D correlate with the lower gut microbiome composition in inflammatory bowel disease patients. *Sci. Rep.* **10**, 6024. <https://doi.org/10.1038/s41598-020-62811-4> (2020).
20. Kaakoush, N. O. Insights into the role of erysipelotrichaceae in the human host. *Front Cell Infect Microbiol* **5**, 84. <https://doi.org/10.3389/fcimb.2015.00084> (2015).
21. Vital, M., Karch, A. & Pieper, D. H. Colonic butyrate-producing communities in humans: an overview using omics data. *mSystems* <https://doi.org/10.1128/mSystems.00130-17> (2017).
22. Ohara, T. Identification of the microbial diversity after fecal microbiota transplantation therapy for chronic intractable constipation using 16S rRNA amplicon sequencing. *PLoS ONE* **14**, e0214085–e0214085. <https://doi.org/10.1371/journal.pone.0214085> (2019).
23. Mirsepasi-Lauridsen, H. C., Vallance, B. A., Krogfelt, K. A. & Petersen, A. M. *Escherichia coli* Pathobionts Associated with Inflammatory Bowel Disease. *Clin. Microbiol. Rev.* **32**, e00060–e118. <https://doi.org/10.1128/CMR.00060-18> (2019).
24. Hove-Jensen, B. & Maigaard, M. *Escherichia coli* rpiA gene encoding ribose phosphate isomerase A. *J. Bacteriol.* **175**, 5628–5635. <https://doi.org/10.1128/jb.175.17.5628-5635.1993> (1993).
25. Park, S. J. & Gunsalus, R. P. Oxygen, iron, carbon, and superoxide control of the fumarase fumA and fumC genes of *Escherichia coli*: role of the arcA, fnr, and soxR gene products. *J. Bacteriol.* **177**, 6255–6262. <https://doi.org/10.1128/jb.177.21.6255-6262.1995> (1995).
26. Kim, Y., Wang, X., Ma, Q., Zhang, X. S. & Wood, T. K. Toxin-antitoxin systems in *Escherichia coli* influence biofilm formation through YjgK (TabA) and fimbriae. *J. Bacteriol.* **191**, 1258–1267. <https://doi.org/10.1128/jb.01465-08> (2009).
27. Teplyakov, A., Obmolova, G., Toedt, J., Galperin, M. Y. & Gilliland, G. L. Crystal structure of the bacterial YhCh protein indicates a role in sialic acid catabolism. *J. Bacteriol.* **187**, 5520–5527 (2005).
28. Di Martino, M. L. *et al.* Polyamines: emerging players in bacteria-host interactions. *Int. J. Med. Microbiol. IJMM* **303**, 484–491. <https://doi.org/10.1016/j.ijmm.2013.06.008> (2013).
29. Desvaux, M. *et al.* Pathogenicity factors of genomic islands in intestinal and extraintestinal *Escherichia coli*. *Front. Microbiol.* **11**, 2065 (2020).
30. Brzuszkiewicz, E. *et al.* How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. *Proc Natl Acad Sci USA* **103**, 12879–12884. <https://doi.org/10.1073/pnas.0603038103> (2006).
31. Schneeberger, K., Roth, S., Nieuwenhuis, E. E. S. & Middendorp, S. Intestinal epithelial cell polarity defects in disease: Lessons from microvillus inclusion disease. *Dis. Model Mech.* <https://doi.org/10.1242/dmm.031088> (2018).
32. Takahashi, D. *et al.* The epithelia-specific membrane trafficking factor AP-1B controls gut immune homeostasis in mice. *Gastroenterology* **141**, 621–632 (2011).
33. Jangid, A. *et al.* Association of colitis with gut-microbiota dysbiosis in clathrin adapter AP-1B knockout mice. *PLoS ONE* **15**, e0228358 (2020).
34. Okumura, R. *et al.* Lypd8 inhibits attachment of pathogenic bacteria to colonic epithelia. *Mucosal Immunol.* **13**, 75–85. <https://doi.org/10.1038/s41385-019-0219-4> (2020).
35. Ohno, H. & Hase, K. Glycoprotein 2 (GP2): Grabbing the FimH bacteria into M cells for mucosal immunity. *Gut Microbes* **1**, 407–410. <https://doi.org/10.4161/gmic.1.6.14078> (2010).
36. Kurashima, Y. *et al.* Pancreatic glycoprotein 2 is a first line of defense for mucosal protection in intestinal inflammation. *Nat. Commun.* **12**, 1067. <https://doi.org/10.1038/s41467-021-21277-2> (2021).
37. Iorember, F. M. & Vehaskari, V. M. Uromodulin: Old friend with new roles in health and disease. *Pediatr. Nephrol.* **29**, 1151–1158. <https://doi.org/10.1007/s00467-013-2563-z> (2014).
38. Kanaya, T., Williams, I. R. & Ohno, H. Intestinal M cells: Tireless samplers of enteric microbiota. *Traffic* **21**, 34–44. <https://doi.org/10.1111/tra.12707> (2020).
39. Yanagihara, S. *et al.* Uromodulin-SlpA binding dictates *Lactobacillus acidophilus* uptake by intestinal epithelial M cells. *Int. Immunol.* **29**, 357–363. <https://doi.org/10.1093/intimm/dxx043> (2017).
40. Nakato, G. *et al.* Cutting Edge: *Brucella abortus* exploits a cellular prion protein on intestinal M cells as an invasive receptor. *J. Immunol.* **189**, 1540–1544. <https://doi.org/10.4049/jimmunol.1103332> (2012).
41. Fukata, M. & Arditi, M. The role of pattern recognition receptors in intestinal inflammation. *Mucosal Immunol.* **6**, 451–463. <https://doi.org/10.1038/mi.2013.13> (2013).
42. Koyama, I., Matsunaga, T., Harada, T., Hokari, S. & Komoda, T. Alkaline phosphatases reduce toxicity of lipopolysaccharides in vivo and in vitro through dephosphorylation. *Clin. Biochem.* **35**, 455–461. [https://doi.org/10.1016/s0009-9120\(02\)00330-2](https://doi.org/10.1016/s0009-9120(02)00330-2) (2002).
43. Beumer, C. *et al.* Calf intestinal alkaline phosphatase, a novel therapeutic drug for lipopolysaccharide (LPS)-mediated diseases, attenuates LPS toxicity in mice and piglets. *J. Pharmacol. Exp. Ther.* **307**, 737–744. <https://doi.org/10.1124/jpet.103.056606> (2003).
44. Malo, M. S. *et al.* Intestinal alkaline phosphatase preserves the normal homeostasis of gut microbiota. *Gut* **59**, 1476–1484. <https://doi.org/10.1136/gut.2010.211706> (2010).
45. Tsuji, S. *et al.* Human intelectin is a novel soluble lectin that recognizes galactofuranose in carbohydrate chains of bacterial cell wall. *J. Biol. Chem.* **276**, 23456–23463. <https://doi.org/10.1074/jbc.M103162200> (2001).
46. French, A. T. *et al.* The expression of intelectin in sheep goblet cells and upregulation by interleukin-4. *Vet. Immunol. Immunopathol.* **120**, 41–46. <https://doi.org/10.1016/j.vetimm.2007.07.014> (2007).
47. Voehringer, D. *et al.* *Nippostrongylus brasiliensis*: identification of intelectin-1 and -2 as Stat6-dependent genes expressed in lung and intestine during infection. *Exp. Parasitol.* **116**, 458–466. <https://doi.org/10.1016/j.exppara.2007.02.015> (2007).
48. Pemberton, A. D., Knight, P. A., Wright, S. H. & Miller, H. R. Proteomic analysis of mouse jejunal epithelium and its response to infection with the intestinal nematode, *Trichinella spiralis*. *Proteomics* **4**, 1101–1108. <https://doi.org/10.1002/pmic.200300658> (2004).
49. Kerr, S. C. *et al.* Intelectin-1 is a prominent protein constituent of pathologic mucus associated with eosinophilic airway inflammation in asthma. *Am. J. Respir. Crit. Care Med.* **189**, 1005–1007. <https://doi.org/10.1164/rccm.201312-2220LE> (2014).
50. Wesener, D. A. *et al.* Recognition of microbial glycans by human intelectin-1. *Nat. Struct. Mol. Biol.* **22**, 603–610. <https://doi.org/10.1038/nsmb.3053> (2015).
51. Wesener, D. A., Dugan, A. & Kiessling, L. L. Recognition of microbial glycans by soluble human lectins. *Curr. Opin. Struct. Biol.* **44**, 168–178. <https://doi.org/10.1016/j.sbi.2017.04.002> (2017).
52. Kaper, J. B., Nataro, J. P. & Mobley, H. L. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* **2**, 123–140. <https://doi.org/10.1038/nrmicro818> (2004).
53. Lopez-Siles, M., Duncan, S. H., Garcia-Gil, L. J. & Martinez-Medina, M. *Faecalibacterium prausnitzii*: From microbiology to diagnostics and prognostics. *ISME J.* **11**, 841–852. <https://doi.org/10.1038/ismej.2016.176> (2017).
54. Abdul Rahim, M. B. H. *et al.* Diet-induced metabolic changes of the human gut microbiome: importance of short-chain fatty acids, methylamines and indoles. *Acta Diabetol.* **56**, 493–500. <https://doi.org/10.1007/s00592-019-01312-x> (2019).
55. Sokol, H. *et al.* Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflamm. Bowel Dis.* **15**, 1183–1189 (2009).

56. Rossi, O. *et al.* Faecalibacterium prausnitzii A2–165 has a high capacity to induce IL-10 in human and murine dendritic cells and modulates T cell responses. *Sci. Rep.* **6**, 1–12 (2016).
57. Carlsson, A. H. *et al.* Faecalibacterium prausnitzii supernatant improves intestinal barrier function in mice DSS colitis. *Scand. J. Gastroenterol.* **48**, 1136–1144 (2013).
58. Martín, R. *et al.* Faecalibacterium prausnitzii prevents physiological damages in a chronic low-grade inflammation murine model. *BMC Microbiol.* **15**, 1–12 (2015).
59. Martín, R. *et al.* Functional characterization of novel Faecalibacterium prausnitzii strains isolated from healthy volunteers: a step forward in the use of F. prausnitzii as a next-generation probiotic. *Front. Microbiol.* **8**, 1226 (2017).
60. La Rosa, S. L. *et al.* The human gut Firmicute Roseburia intestinalis is a primary degrader of dietary  $\beta$ -mannans. *Nat. Commun.* **10**, 905. <https://doi.org/10.1038/s41467-019-08812-y> (2019).
61. Shen, Z. *et al.* Insights into Roseburia intestinalis which alleviates experimental colitis pathology by inducing anti-inflammatory responses. *J. Gastroenterol. Hepatol.* **33**, 1751–1760 (2018).
62. Luo, W. *et al.* Roseburia intestinalis supernatant ameliorates colitis induced in mice by regulating the immune response. *Mol. Med. Rep.* **20**, 1007–1016. <https://doi.org/10.3892/mmr.2019.10327> (2019).
63. Patterson, A. M. *et al.* Human gut symbiont Roseburia hominis promotes and regulates innate immunity. *Front. Immunol.* **8**, 1166. <https://doi.org/10.3389/fimmu.2017.01166> (2017).
64. Mukherjee, A., Lordan, C., Ross, R. P. & Cotter, P. D. Gut microbes from the phylogenetically diverse genus Eubacterium and their various contributions to gut health. *Gut. Microbes* **12**, 1802866. <https://doi.org/10.1080/19490976.2020.1802866> (2020).
65. Zhang, S. *et al.* Short Chain fatty acids modulate the growth and virulence of pathosymbiont *Escherichia coli* and host response. *Antibiotics (Basel)* **9**, 462. <https://doi.org/10.3390/antibiotics9080462> (2020).
66. Parada Venegas, D. *et al.* Short Chain Fatty Acids (SCFAs)-mediated gut epithelial and immune regulation and its relevance for inflammatory bowel diseases. *Front. Immunol.* **10**, 277 (2019).
67. Madison, B. B. *et al.* Cis elements of the villin gene control expression in restricted domains of the vertical (crypt) and horizontal (duodenum, cecum) axes of the intestine. *J. Biol. Chem.* **277**, 33275–33283. <https://doi.org/10.1074/jbc.M204935200> (2002).
68. Murakami, S. *et al.* The consumption of bicarbonate-rich mineral water improves glycemic control. *Evid Based Complement Alternat. Med.* **2015**, 824395. <https://doi.org/10.1155/2015/824395> (2015).
69. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359. <https://doi.org/10.1038/nmeth.1923> (2012).
70. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60. <https://doi.org/10.1038/nmeth.3176> (2015).
71. Huson, D. H. *et al.* MEGAN community edition—interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* **12**, e1004957. <https://doi.org/10.1371/journal.pcbi.1004957> (2016).
72. Parks, D. H., Tyson, G. W., Hugenholtz, P. & Beiko, R. G. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**, 3123–3124. <https://doi.org/10.1093/bioinformatics/btu494> (2014).
73. Zakrzewski, M. *et al.* Calypso: a user-friendly web-server for mining and visualizing microbiome-environment interactions. *Bioinformatics* **33**, 782–783. <https://doi.org/10.1093/bioinformatics/btw725> (2017).
74. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60. <https://doi.org/10.1186/gb-2011-12-6-r60> (2011).
75. Huerta-Cepas, J. *et al.* a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**, D286–293. <https://doi.org/10.1093/nar/gkv1248> (2016).

## Acknowledgements

This work was supported in part by JST ERATO grant no. JPMJER1902 to S.F.; AMED-CREST grant no. JP20gm1010009 to S.F.; the Takeda Science Foundation to S.F.; the Food Science Institute Foundation to S.F. The Council of Scientific and Industrial Research, India provided support in the form of a fellowship to A.J. T.P. acknowledges the HPC facility at IIT Mandi which has been extensively used for the bioinformatics analysis.

## Author contributions

H.O., and T.P. conceptualized the problem, S.F. performed the animal experiments, M.S. and Y.S. performed sequencing experiments, S.F., M.S., and Y.S. did the data curation, A.J. did the formal analysis, validation of data, and writing of the manuscript under the guidance of T.P., T.D.T., S.F. and T.P. reviewed and edited the manuscript. S.F. and T.P. acquired funding. T.P. arranged the resources and supervised the overall research.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-15150-5>.

**Correspondence** and requests for materials should be addressed to T.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022