frontiers | Frontiers in Epigenetics and Epigenomics

TYPE Original Research
PUBLISHED 28 February 2025
DOI 10.3389/freae.2025.1535838

OPEN ACCESS

EDITED BY
Zhenyu Xuan,
The University of Texas at Dallas, United States

REVIEWED BY
Mushan Li,
The Pennsylvania State University (PSU), United States
Jin Woo Oh,
Lunit, Republic of Korea

*CORRESPONDENCE
T. A. Hore,
tim.hore@otago.ac.nz
D. M. Bond,
donna.bond@otago.ac.nz

†These authors have contributed equally to this work and share last authorship

RECEIVED 27 November 2024
ACCEPTED 05 February 2025
PUBLISHED 28 February 2025

CITATION
Dunnet MJ, Morison IM, Bond DM and Hore TA (2025) High-throughput, pan-leukocyte biomarkers for the detection of inflammation in human breastmilk and stool.
Front. Epigenet. Epigenom. 3:1535838.
doi: 10.3389/freae.2025.1535838

High-throughput, pan-leukocyte biomarkers for the detection of inflammation in human breastmilk and stool

M. J. Dunnet[1,2], I. M. Morison[3], D. M. Bond[1*†] and T. A. Hore[1*†]

[1]Department of Anatomy, University of Otago, Dunedin, New Zealand, [2]Department of Biochemistry, University of Otago, Dunedin, New Zealand, [3]Department of Pathology, University of Otago, Dunedin, New Zealand

Background: DNA methylation can be used to track cellular identity. We have previously developed a high-throughput, cost-effective DNA methylation pipeline containing two loci, *HOXA3* and *MAP4K1*, that can quantify leukocyte proportion amongst a range of background tissues. Here, we apply this pipeline to two clinically relevant tissue samples: breastmilk and stool.

Results: We report that our leukocyte methylation assay can quantify the proportion of leukocytes in breastmilk, and find leukocyte levels fluctuate dramatically in concert with infection severity. We benchmarked our leukocyte methylation pipeline in stool samples against the commonly used faecal calprotectin assay. Our results show a high concordance between the two methods indicating the viability of our DNA methylation biomarkers in the context of intestinal inflammation.

Conclusion: The data presented here emphasise the clinical applicability of our high-throughput DNA methylation assay in the context of mastitis and intestinal inflammation.

KEYWORDS

DNA methylation, biomarker, leukocytes, inflammation, mastitis, inflammatory bowel disease

Introduction

DNA methylation is the covalent bonding of a methyl group to the 5′ carbon of a cytosine nucleotide. DNA methylation is dynamically added and removed throughput development; indeed, approximately 20% of all human autosomal CpG sites are differentially methylated or unmethylated based upon cell type and stage of development (Ziller et al., 2013). As such, cell-specific DNA methylation patterns can be used to identify the cell-of-origin for a particular DNA molecule. In cases where there is no genetic difference between healthy and diseased tissue or where information on cellular origin is required, DNA methylation assays will likely be useful for detection, diagnosis, and management decisions. For example, cell-free DNA (cfDNA), which can originate from

Abbreviations: CpG, CG dinucleotide; cfDNA, cell-free DNA; FCA, Faecal Calprotectin Assay; ROC, Receiver Operating Characteristic; AUC, Area Under the Curve; IBD, Inflammatory Bowel Disease; IBS, Irritable Bowel Syndrome; UC, Ulcerative Colitis; CD, Crohn's Disease; PEG, polyethylene glycol.

apoptotic or necrotic cells, can enter the circulatory system and be collected easily and non-invasively for diagnostic purposes. Analysis of the methylation pattern of cfDNA in healthy individuals shows its cellular origins: 55% from leukocytes, 30% from erythrocyte progenitors, 10% from vascular endothelial cells, and 1% from hepatocytes (Moss et al., 2018; Lam et al., 2020). In addition, in cases of trauma, autoimmunity, ischaemia, infection, or cancer, previously undetectable cfDNA molecules from affected organs can be measured (Lehmann-Werman et al., 2016; Cisneros-Villanueva et al., 2022; Zemmour et al., 2018; Lehmann-Werman et al., 2018; Cheng et al., 2021).

Local inflammation is difficult to detect with cfDNA because of the large amount of blood-cell-derived DNA already present in plasma. Several deconvolution algorithms based on Illumina 450K methylation array and EPIC array systems have been utilised for this analysis; for example, EPIDish (Teschendorff et al., 2017), can determine the proportion of leukocyte subpopulations from a given sample. However, these systems require thousands of CpG sites, are relatively expensive, and are low throughput. We have previously described a high throughput, cost-effective, pipeline that uses the locus-specific methylation to detect immune cells from a mixed sample (Dunnet et al., 2022). This involves bisulfite amplicon sequencing of *HOXA3* and *MAP4K1* (Dunnet et al., 2022). To expand on this work here, we examined the ability of this pipeline to infer the level of inflammation from clinically relevant samples, specifically, breastmilk and stool. These samples were chosen because of their relevance to common inflammatory conditions: mastitis and inflammatory bowel disease.

Mastitis is defined as the inflammation of the breast tissue, which typically, but not exclusively, occurs during lactation (Boakes et al., 2018). The fraction of leukocytes markedly increases from less than 5% in healthy mothers to up to 90% during mastitis (Hassiotou et al., 2013a). Flow cytometry approaches to measure the leukocyte fraction in breastmilk have already been proposed as a diagnostic tool to assess the health status of the mother/infant dyad (Hassiotou et al., 2013a). However, cellular composition in breastmilk is highly dynamic and sampling may be required more frequently than flow cytometry can conveniently provide.

Intestinal inflammation can be caused by a large number of factors, including cancer (von Roon et al., 2007), inflammatory bowel disease (Laserna-Mendieta and Lucendo, 2019), and coeliac disease (Ertekin et al., 2010). Detection of intestinal inflammation is most commonly performed with the faecal calprotectin assay (FCA), which measures the concentration of calprotectin, an antimicrobial complex highly abundant in the granules of neutrophils and, to a lesser extent in monocytes and macrophages (Odink et al., 1987; Røseth PNSMKF, 1999). The amount of calprotectin is proportional to the severity of an immune response, making it a useful non-invasive marker for intestinal inflammation (Røseth et al., 1992). While an extremely valuable tool, the FCA is limited in several ways. Firstly, the FCA kits are proprietary, and the inter-kit variability is large (Whitehead et al., 2013; Kittanakom et al., 2017; Labaere et al., 2014). Second, the calprotectin complex degrades in stool after 48–72 h at room temperature and requires refrigeration to remain stable (Labaera et al., 2014; Lasson et al., 2015; Oyaert et al., 2017). Finally, the age of the patient, medications they are taking, and some pathologies (in particular, pancreatic insufficiency) can result in variation in faecal calprotectin levels (Laserna-Mendieta and Lucendo, 2019; Degraeuwe et al., 2015; Henderson et al., 2014; Padoan et al., 2018; Lundgren et al., 2019; Ellemunter et al., 2017).

The high-throughput DNA methylation biomarker pipeline we have previously described (Dunnet et al., 2022) has the potential to address limitations in the current diagnostic assays. In particular, the ability to sample repeatedly, to avoid proprietary kits, and to capitalize on the stability of DNA methylation have the potential to improve the diagnostic capabilities for mastitis and intestinal inflammation.

In this study, we sampled numerous breastmilk samples from the same individual over the course of 1 month. We show the proportion of leukocytes varied markedly, even within a 24-hour period. The two biomarker loci, *HOXA3* and *MAP4K1*, were strongly correlated in milk samples, and the total level of estimated leukocytes was consistent with instances of mastitis. Furthermore, we examined the leukocyte proportion in stool samples with associated faecal calprotectin scores. We observed that the proportion of leukocyte-derived reads was extremely high (>70%) in all samples with even slightly elevated faecal calprotectin, suggesting that leukocyte DNA vastly exceeds epithelial cell DNA in stool with even minor inflammation. Our high-throughput, cost-effective DNA methylation assay has clinical relevance in the context of mastitis and intestinal inflammation.

Methods

Human breastmilk sample collection and cell isolation

Human breastmilk was obtained from a single donor with informed consent per the New Zealand Human Tissue Act 2008. Over 1 month, from the 22nd of November 2018 to the 21st of December 2018, 1 mL of excess pump-expressed milk was collected by the donor where possible. Pumping occurred either exclusively from one breast or was a mixture of the two breasts; while breast of origin was often recorded by the donor, in many cases it was not. The start of the collection date was approximately 5 weeks postpartum and seventy-five milk samples were collected in total. Milk was stored at −20°C, or colder, until use.

Cell isolation was performed by centrifugation. First, milk samples were centrifuged at 500 x g for 15 min to pellet the cells. Next, milk fats and liquid were removed, and the pellet was resuspended in 500 µL of 0.01 M PBS. Centrifugation and washing were repeated an additional two times. Finally, the cells were pelleted at 500x g for 15 min before resuspension in a lysis buffer consisting of 8 µL of 0.01 M PBS, 8 µL of the Zymo 2x M-Digestion Buffer, and 1 µL of 20 mg/mL Proteinase K (Zymo EZ-96 DNA Methylation-Direct™ MagPrep Kit) and incubated at 50°C for 20 min. Cell lysates were directly added to the bisulphite conversion protocol (see below) with no additional handling.

Saliva cell isolation and DNA extraction

Saliva samples were collected as previously described (Dunnet et al., 2022; Theda et al., 2018). Briefly, in the 30 min prior to

collection only water was consumed. Five mL of saliva was collected via passive drool and transferred to a 15 mL plastic centrifuge tube. The saliva was centrifuged at 400 g and supernatant removed to isolate the cells. The cell pellets were resuspended and washed with 0.01 M PBS three times. DNA was extracted using the BOMB.bio protocol 6.3: extraction of TNA from mammalian tissues (Oberacker et al., 2019) (bomb.bio/protocols/). DNA quality was assessed by gel electrophoresis and concentration by the HS dsDNA Qubit assay (Thermo Scientific).

K562 cell culture and DNA extraction

K562 cells were cultured as previously described (Pencovich et al., 2011). The cells were lysed with 1 mL of GITC lysis buffer (4 M GITC, 2% w/v SDS, 50 mM Tris-HCl pH 8.0, 0.1% v/v antifoam 204 (Sigma-Aldrich), and 20 mM EDTA). DNA extraction was performed with the BOMB.Bio protocol 6.1: TNA extraction of mammalian cells with GITC (Oberacker et al., 2019) (bomb.bio/protocols/). DNA quality was assessed by gel electrophoresis and concentration by the HSdsDNA Qubit assay.

Leukocyte isolation and DNA extraction

Leukocytes were isolated from saliva (see above) by sequential cellular filtration as previously described (Dunnet et al., 2022). Briefly, cellular isolates from saliva were first filtered through a 40-µm, then a 20-µm mesh filter to exclude buccal cells. The purity of the isolated leukocytes was assessed by microscopy. We counted a minimum of 100 cells across two fields per slide. We observed a purity of >99% for each sample.

Stool sample preparation and DNA extraction

The use of human stool samples was approved by the University of Otago Human Ethics Committee (Health) (approval number H21/138). Stool samples were obtained from Southern Community Laboratories (SCL), Dunedin. SCL had previously processed the stool samples for use in a FCA; specifically, they were homogenised and diluted in a proprietary extraction buffer. After the faecal calprotectin samples were received, they were heat-inactivated at 75 °C before DNA extraction with the Zymo Research™ Quick-DNA Fecal/Soil Microbe Miniprep Kit. DNA extractions were performed starting at step four of the protocol (steps 1-3 are for homogenisation). Furthermore, to improve overall DNA yield per sample, the volume of homogenised stool sample and genomic lysis buffer was increased four-fold from 400 µL to 1,200 µL to 1,600 µL and 4,800 µL, respectively. The entirety of each sample was run through the extraction column over multiple spin cycles. The remainder of the DNA extraction was performed as above. Other than these modifications, extraction was carried out as per the manufacturer's instructions. Total DNA concentration was measured with the HS dsDNA Qubit assay, and DNA quality was assessed by gel electrophoresis.

Bisulfite conversion of DNA

Bisulphite conversion for all samples was performed with the Zymo EZ-DNA Methylation Direct MagPrep kit per the manufacturer's instructions. This protocol suggests various initial conversion parameters specific to the amount and quality of input DNA. Lysates with breastmilk-derived DNA were converted under standard conditions for optimal cytosine conversion: 8 min at 98°C followed by 3 h and 30 min at 64°C. Stool-derived DNA samples were converted for 8 min at 98°C followed by 3 h and 30 min at 53°C to reduce DNA degradation at the cost of less efficient conversion. Bisulfite converted DNA was quantified using the Qubit ssDNA Qubit assay (Thermo Scientific).

Bisulfite amplicon sequencing

A dual-index, two-step PCR protocol was used to amplify bisulphite-converted DNA with the KAPA HiFi HotStart Uracil + kit (Roche) as described here (Dunnet et al., 2022). Briefly, in the first round of PCR, the target region is amplified with primers containing an overhanging linker sequence. The second round of PCR uses primers comprised of the complementary linker sequence attached to the Illumina P5 and P7 adapters and TruSeq indexes. Each PCR reaction contained the KAPA HiFi HotStart Uracil + ReadyMix, 0.3 µM of each primer, and 200 ng of bisulphite-converted DNA topped up to 25 µL with nuclease-free water. All primer sequences are shown in Table 1.

Breastmilk-derived DNA was first amplified with the *HOXA3_long* and *MAP4K1_long* primer pairs. The following cycle parameters were used: 95°C for 2 min, 23 cycles of 98°C for 20 s, 59°C for 10 s, and 72°C for 20 s. A final elongation step was performed for 5 min at 72°C. Reactions were centrifuged briefly to remove condensate on the tube walls. Next, the products were cleaned using solid-phase reverse immobilisation of carboxyl-coated magnetic beads suspended in standard PEG buffer (18% w/v polyethylene glycol 8000 (PEG), 1 M NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA, 0.05% v/v Tween-20), followed by two washes in 70% ethanol. The cleaned PCR products were eluted in 11.5 µL of filter sterile Milli-Q® water. The eluted DNA was combined with 12.5 µL of the KAPA HiFi HotStart Uracil + ReadyMix and 0.5 µL of each indexing primer (Illumina P5 and P7 adapters with TruSeq index combined with the complementary linker sequence). Amplification was repeated as above for an additional five cycles.

Saliva, K562, and human stool-derived DNA were amplified with the *MAP4K1_short* primer pair. The same approach as above was followed with modifications to thermocycling to improve PCR efficiency. The first amplification step was performed with the following parameters: 95°C for 2 min, 25 cycles of 98°C for 20 s, 59°C for 40 s, and 72°C for 40 s. A final elongation step was performed for 5 min at 72 °C. The PCR product solid-phase reverse immobilisation clean-up was performed similarly as above. The second amplification step used the following parameters: 95°C for 2 min, 5 cycles of 98°C for 20 s, 59°C for 40 s, and 72°C for 40 s, followed a final 5 min elongation step at 72°C.

Prior to sequencing, all amplicons were cleaned using solid-phase reverse immobilisation of carboxyl-coated magnetic beads suspended in standard PEG buffer (as above). Sequencing was

TABLE 1 The primer sequences used in this study. Bold text indicates linker sequences for bisulphite amplicon sequencing. Note that this linker sequence is the reverse complement of the P5 and P7 adapters.

Primer ID	Sequence	Target strand	Target location (hg38)
Bisulfite HOXA3_long Forward	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**GGTTTTGTTTGGGTTAGTGGTAT	+	chr7:27,113,959-27,114,117
Bisulfite HOXA3_long Reverse	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**CCAACAAAAAAACCCCTTTATAAA	+	chr7:27,113,959-27,114,117
Bisulfite MAP4K1_long Forward	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**TGTTTTTATATGGAAGTTATATTTATT	-	chr19: 38,596,411–38,596,696
Bisulfite MAP4K1_long Reverse	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**AACAACTCAAAACCTAACCC	-	chr19: 38,596,411–38,596,696
Bisulfite MAP4K1_short Forward	**ACACTCTTTCCCTACACGACGCTCTTCCGATCT**TTAGAAATGTTAGGGGATAAGGTTT	+	chr19:38,596,606-38,596,816
Bisulfite MAP4K1_short Reverse	**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**TTCCTCAAACTCAATACTACCACTC	+	chr19:38,596,606-38,596,816
Illumina P5 adapter with TruSeq index	AATGATACGGCGACCACCGAGATCTACACNNNNNN**ACACT CTTTCCCTACACGACGCTCTTCCGATCT**	N/A	N/A
Illumina P7 adapter with TruSeq index	CAAGCAGAAGACGGCATACGAGATNNNNNN**GTGACTGGAG TTCAGACGTGTGCTCTTCCGATCT**	N/A	N/A

performed on the Illumina iSeq100 as per the manufactures instructions.

Bioinformatic and statistical analyses

Raw read adapter and quality trimming were performed with Cutadapt and TrimGalore (v.0.6.7) (Martin, 2011). Reads were mapped to a custom "genome" consisting of only the amplicon sequence using Bismark (v0.14.3) (Krueger and Andrews, 2011). The sequences used for mapping were obtained from the UCSC genome browser (*HOXA3_long*: chr7:27,113,957–27,114,300 (hg38); *MAP4K1_long*: chr19:38,596,411–38,596,696 (hg38); *MAP4K1_short*: chr19:38,596,606-38,596,816 (hg38)). Heatmaps, linear regression, bionomical regression, cell-of-origin read classification, and ROC curve generation were performed with custom R scripts and the pROC package. The cut-off for a read to be classified as leukocyte derived was ≥6 methylated CpG sites for the *HOXA3_long* amplicon, ≥3 for the *MAP4K1_long* amplicon, and ≥4 for the *MAP4K1_short* amplicon. All statistical analyses were performed in R.

Results

DNA methylation at the *HOXA3* and *MAP4K1* loci can determine the proportion of blood-derived cells in human breastmilk.

Human breastmilk contains three principal cell populations (Ziller et al., 2013): blood-derived cells, comprised of both mature leukocytes and hematopoietic stem cells (Moss et al., 2018), breast-derived cells, comprised of lactocytes, myoepithelial cells, and progenitor cells; and (Lam et al., 2020) probiotic bacteria (Witkowska-Zimny and Kaminska-El-Hassan, 2017). In a healthy mother/child dyad, the proportion of leukocytes is high in the colostrum (between 13% and 70%), defined as the first milk until approximately 4 days postpartum, but rapidly decreases after that to less than 1% of the total cells in breastmilk (Hassiotou et al., 2013a; Witkowska-Zimny and Kaminska-El-Hassan, 2017). In contrast, when either the mother or child is ill, leukocyte proportions increase; during mastitis, leukocytes can constitute upwards of 90% of total breastmilk cells (Hassiotou et al., 2013a). Therefore we first aimed to determine if the previously described *HOXA3* and *MAP4K1* DNA methylation biomarkers (Dunnet et al., 2022) can accurately deconvolute blood-derived cells from breast-derived cells in human breastmilk. Seventy-five breastmilk samples from a single individual were gathered over the course of 1 month with various levels of meta-data (Figure 1A). Samples were recorded as originating from either the left or right breast during the first week of collection; samples after week one were either pooled together from both breasts or had no associated information. We examined DNA methylation at the *HOXA3* and *MAP4K1* loci and applied a previously described leukocyte estimation pipeline for each sample (Dunnet et al., 2022). Reads were either highly methylated or highly unmethylated with considerable concordance between both loci ($R^2 = 0.95$, Figures 1B, C), suggesting both biomarkers function similarly in breastmilk.

Next, we examined the leukocyte proportions from samples with available meta-data on the breast of origin (Figure 2). During the time these samples were taken, there was mastitis in the left breast and an injury to the right breast resulting from breastfeeding. We observed a high proportion of leukocytes in the left-breast-associated milk throughout the week and elevated leukocyte proportions in right-breast-associated milk occurring at the peak of mastitis in the left breast and during the injury to the right breast (Figure 2A). Outside of these time points, the number of leukocytes in the right breast decreased to approximately 60% of the total cell population. These results suggest that during mastitis, the proportion of leukocytes is elevated primarily in the infected

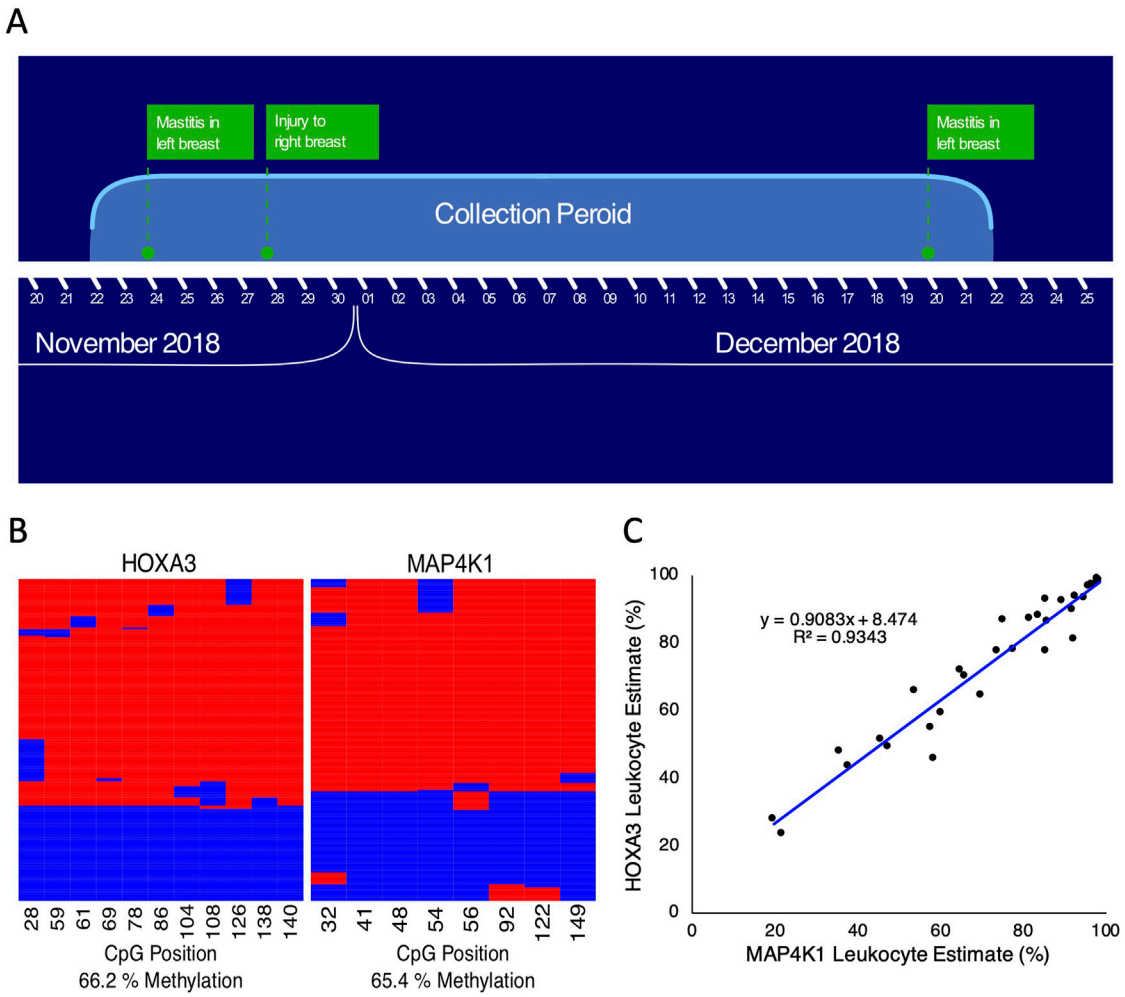

FIGURE 1
HOXA3 and *MAP4K1* leukocyte estimates correlate in human milk samples. **(A)** Overview of the milk collection period **(B)** Example of bisulphite amplicon heatmaps for *HOXA3* (left) and *MAP4K1* (right) from human milk samples. Sequencing reads run from left to right across the rows, and columns represent individual CpG positions. Red indicates a methylated CpG, and blue indicates an unmethylated CpG. **(C)** Scatterplot of leukocyte estimates for *MAP4K1* (X-axis) vs. *HOXA3* (Y-axis).

breast; however, a severe infection can produce elevated leukocyte proportions in both breasts regardless of where the infection is located.

Throughout the month-long period of sample collection, we observed a high proportion of leukocytes. Indeed, the mean leukocyte estimation across the collection period was 72.3% for *HOXA3* (median 76.6%) and 70.9% for *MAP4K1* (median = 73.8%); the sample with the lowest leukocyte proportion was estimated at 23.5% by *HOXA3* and 21.6% by *MAP4K1*. Nevertheless, this was not unexpected because of continued mastitis in the left breast and damage to the right breast. Interestingly, leukocyte estimates did not shift dramatically on a day-to-day basis when either the breast of origin was known or the samples were comprised of milk from both breasts (Figures 2A, C); however, we observed striking changes between milk samples collected on the same day when breast-of-origin meta-data was not available (maximum difference of 69%, Figure 2D, see methods). Therefore, we hypothesise that these rapid changes in cellular composition are partly the result of milk samples obtained from different breasts. Indeed, the cellular composition of breast milk is dynamic and changes based on the period of lactation, infection status, and infant feeding habits (Hassiotou et al., 2013a; Witkowska-Zimny and Kaminska-El-Hassan, 2017; Riskin et al., 2012). It is, therefore, unsurprising that each breast's cellular composition can differ depending on circumstance, especially in cases of infection.

Validation of a *MAP4K1* bisulfite amplicon for application with stool-derived human DNA

Total genomic DNA extracted from stool samples contains an amalgamation of bacterial-, fungal-, viral-, and host-derived DNA. In healthy individuals, the human component constitutes less than 1% of the total DNA (Vincent et al., 2015). In contrast, chronic intestinal inflammation damages the surrounding tissues and subsequently results in cell death (Anderton et al., 2020), leading to the hypothesis that more host-derived DNA will be present in the stool. Indeed, it has been shown that both mitochondrial DNA and genomic DNA amounts increase

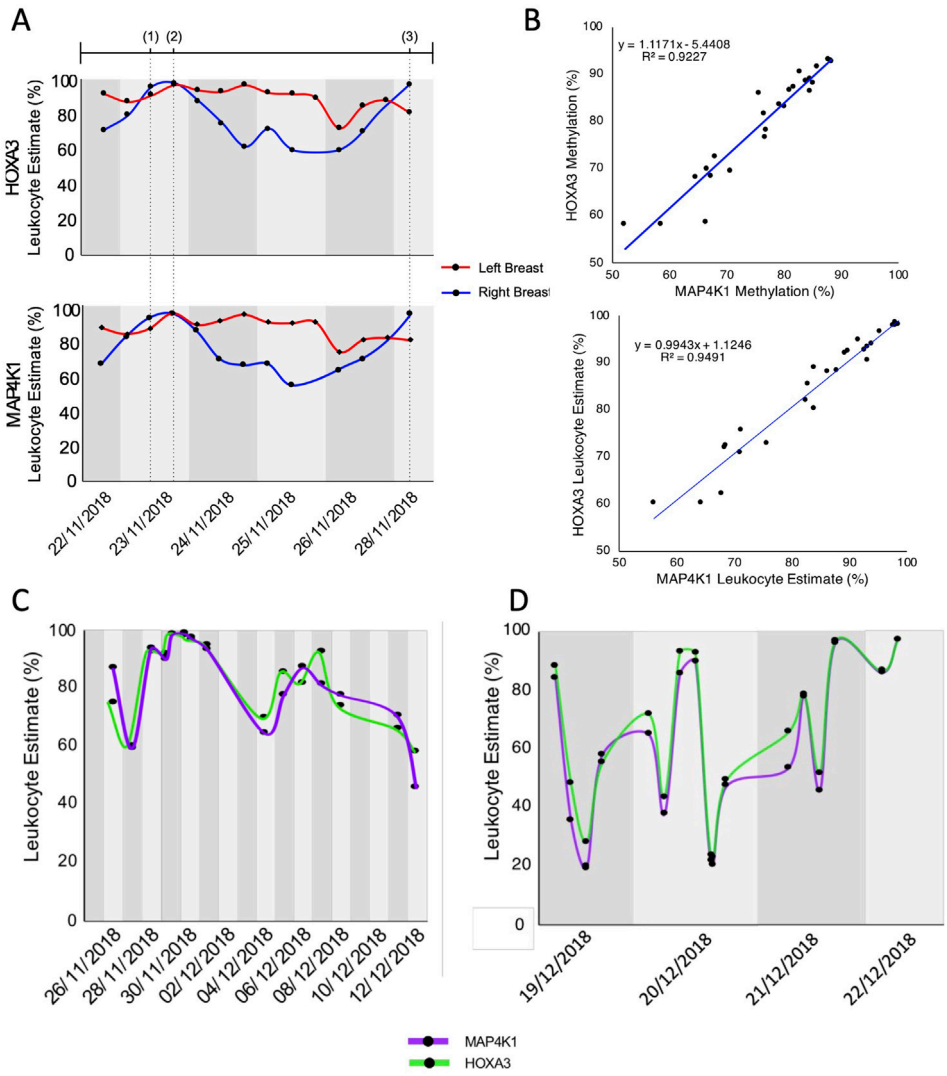

FIGURE 2
Tracking leukocytes in breastmilk over time **(A)** The proportion of leukocytes in milk from either the left (red line) or right (blue line) breast over 1 week as measured by *HOXA3* (top) and *MAP4K1* (bottom). Dotted vertical lines indicate the timepoint of events (Ziller et al., 2013): onset of flu-like symptoms as a result of left breast mastitis (Moss et al., 2018), peak of left-breast mastitis, including fever (Lam et al., 2020), injury to the right breast as a result of breastfeeding. **(B)** Scatterplot of *HOXA3* vs. *MAP4K1* methylation (top) and leukocyte estimate (bottom) for each sample with breast-of-origin meta-data. **(C)** Proportion of leukocytes measured by *HOXA3* and *MAP4K1* from milk known to be pooled from both breasts. **(D)** Proportion of leukocytes measured by *HOXA3* and *MAP4K1* from milk where breast of origin (or its mixing) was not recorded.

based on the severity of inflammatory bowel disease (IBD)-related intestinal inflammation (Casellas et al., 2004; Vrablicova et al., 2020). We hypothesized that under inflammatory conditions, DNA from neutrophils that have migrated into the intestinal lumen, could be detected with our DNA methylation assay. We obtained 48 faecal samples that had previously undergone FCA testing for our analysis. We initially examined a subset of these samples with the *HOXA3* and *MAP4K1* (Dunnet et al., 2022) amplicons described above; however, we were unable to obtain sufficient amplicon copies for sequencing within 30 PCR cycles. Over-amplification with more than 30 cycles can lead to clonal amplification of only a few DNA molecules, leading to a misinterpretation of DNA methylation patterns. As a result, shorter amplicons were designed to improve PCR yield (Table 1; Bisulfite *MAP4K1_short*). We sequenced leukocyte and K562 DNA (the latter is unmethylated in the *HOXA3* and *MAP4K1* loci of interest (Dunnet et al., 2022)) to ensure the DNA methylation patterns are consistent between the short and long amplicons. As expected, we observed that leukocytes were highly methylated (84.9%), while K562 DNA was virtually devoid of DNA methylation (4.6%) (Figure 3A). The near absence of methylation of K562 cells parallels the methylation of these loci in colonic epithelial cells (Dunnet et al., 2022). The two cell types cluster separately from one another based upon the number of methylated CpG sites per read (Figure 3B). To classify reads as either derived from mature leukocyte DNA or K562 DNA we employed a binomial logistic regression model and constructed a receiver operating characteristic (ROC) curve using the 'pROC' package in R. The optimal threshold for leukocyte classification was if ≥4 of seven CpG sites were methylated (sensitivity = 98.3%, specificity = 99.98%, Figure 3C). Furthermore, the area under the ROC curve was 0.998, indicating that the methylation

10.3389/freae.2025.1535838

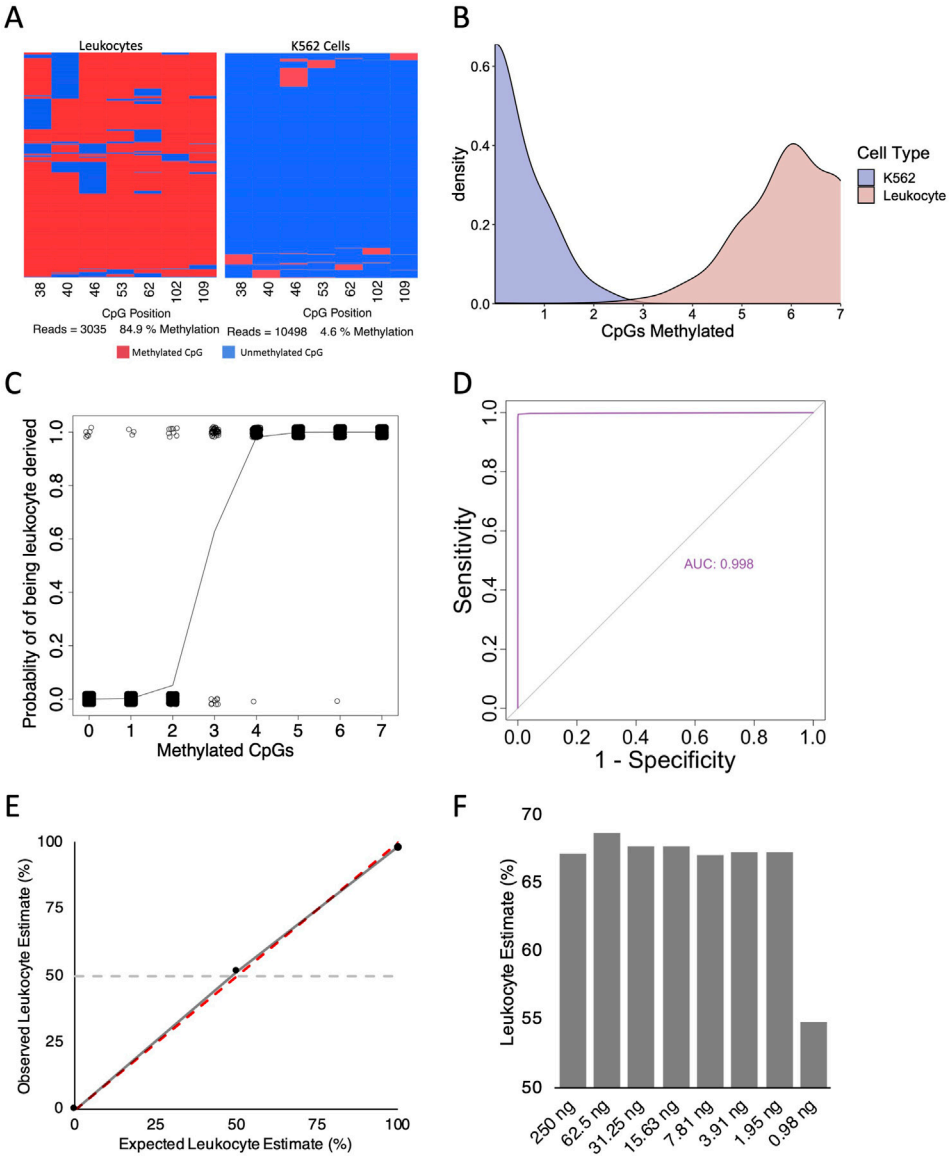

FIGURE 3
Validity of short *MAP4K1* amplicon primer sets. **(A)** Example bisulfite amplicon sequencing heatmaps from leukocyte derived DNA (left) and K562 DNA (right). **(B)** Density plot for the number of methylated CpG sites per read in K562- (blue) and Leukocyte-derived (red) DNA samples. The Y-axis represent the probability per unit on the X-axis such that the area under the curve for a specific interval is equal to the probability of the number of methylated CpG sites for that interval. Bandwidth 0.45 **(C)** Binomial regression analysis for the *MAP4K1* short amplicon using leukocyte and K562 DNA. **(D)** ROC curve for the *MAP4K1* short amplicon using leukocyte and K562 DNA. **(E)** Scatterplot of observed vs. expected read classification for leukocyte and K562 DNA mixes. The dotted red line indicates a perfect fit, the solid grey line indicates the observed trendline, and the horizontal dotted grey line indicates a 50% leukocyte estimate. **(F)** Bar chart of leukocyte estimates from two-fold, serially diluted saliva-derived DNA.

patterns obtained from this amplicon perform exceptionally well at distinguishing blood leukocyte-derived DNA from K562 DNA (Figure 3D).

We next asked if the short *MAP4K1* PCR showed preferential amplification of methylated or unmethylated alleles and if leukocyte estimates are consistent across differing amounts of input DNA. We combined blood leukocyte derived DNA and K562 DNA at a 50: 50 ratio and observed a leukocyte estimate of 51.5%. With leukocyte and K562 DNA alone we observed leukocyte estimates of 98.3% and 0.2% respectively (Figure 3E). Finally, we performed a two-fold serial dilution with saliva-derived DNA (a combination of salivary leukocytes and buccal epithelium (Theda et al., 2018)) and observed consistent levels of leukocyte estimates until 0.98 ng of DNA, where the result became discordant (Figure 3F). This suggests the assay can accurately determine the proportion of leukocyte-derived DNA with as little as 2 ng of input DNA.

MAP4K1 DNA methylation can identify intestinal inflammation as measured by faecal calprotectin

To determine if our bisulfite amplicon sequencing pipeline could be used to detect intestinal inflammation we applied the pipeline

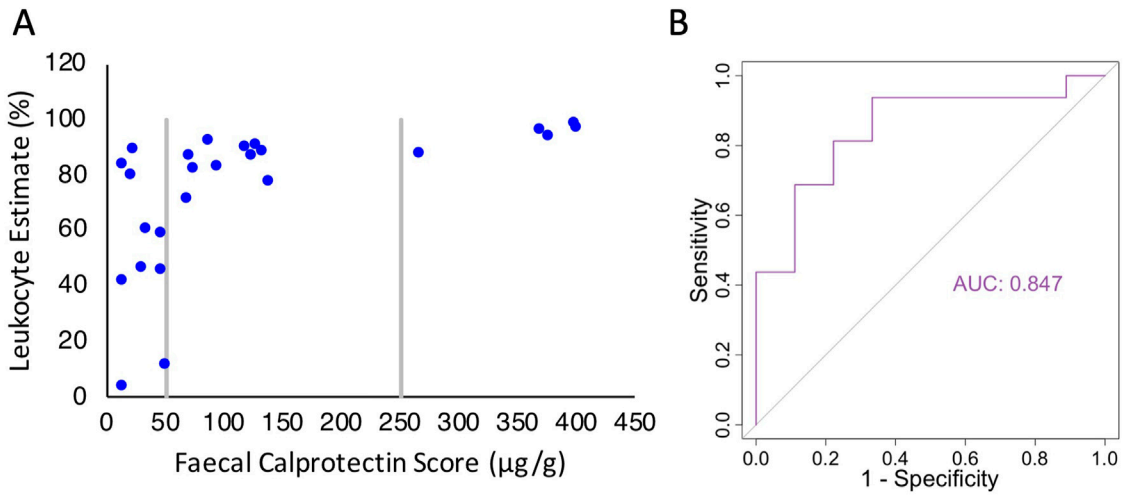

FIGURE 4
Comparison of the *MAP4K1* bisulfite amplicon sequencing assay to the FCA. **(A)** Scatterplot of leukocyte estimates based on *MAP4K1* methylation in stool samples vs. associated faecal calprotectin score. **(B)** ROC curve for the *MAP4K1* bisulfite amplicon sequencing assay using FCA as the reference standard.

(with the shorter *MAP4K1* amplicon) to stool samples that had previously undergone a FCA. The cutoff for a positive faecal calprotectin test result in the laboratory we obtained samples from is 50 µg/g. However, various cut-off thresholds have been suggested in the literature to best balance sensitivity and specificity (Jha et al., 2018; Rogler et al., 2013; Ricciuto and Griffiths, 2019). The reference interval of calprotectin is 10–100 µg/g for Caucasians, depending on the kit (Bjarnason, 2017), although the interval between 50 and 100 µg/g is thought to have some diagnostic value (Ricciuto and Griffiths, 2019). In contrast, a calprotectin score of up to 200 µg/g can be considered normal for people of African-Caribbean descent (Bjarnason, 2017). Furthermore, thresholds as high as 250 µg/g have been shown to markedly increase specificity at the cost of sensitivity (Jha et al., 2018; Rogler et al., 2013). For this reason, we examined the leukocyte estimates in three discrete bands: calprotectin scores less than 50 µg/g (low), between 50 and 250 µg/g (moderate), and greater than 250 µg/g (high). Low, moderate, and high calprotectin samples had mean leukocyte estimates of 51.8% (S.D. = 28.4%), 77.9% (S.D. = 22.9%), and 94.2% (S.D. = 4.2%), respectively (Figure 4A). The percentages of leukocyte-derived DNA in stool from low calprotectin samples were extremely variable, ranging from 3.4% to 88.7%. In contrast, leukocyte estimates in moderate and high samples were extremely similar and all above 70% (with the exception of one moderate sample with a calprotectin level of 52 µg/g and a leukocyte estimate of 11.4%). We observed a statistically significant difference between groups when either 50 µg/g (one tailed t-test, p-value = 4.7E-03) or 250 µg/g (one tailed t-test, p-value = 0.02) were used as the threshold for a positive test; however, there was no significant difference between moderate and high calprotectin samples (one tailed t-test, p-value = 0.07). Our results suggest that human DNA in stool from even a mildly inflamed intestine, as measured by faecal calprotectin, is almost completely derived from leukocytes. In a low inflammatory environment, the proportion of leukocytes in stool is extremely variable, although this might be the result of a relatively small amount of colonic epithelial cells being sloughed out of the lumen.

We next examined how well the methylation assay could function as a diagnostic marker. Since we did not have access to the clinical outcomes of each patient, we used the calprotectin level. With a 50 µg/g threshold, the FCA has a sensitivity and specificity of approximately 97% and 80%, respectively (Laserna-Mendieta and Lucendo, 2019; Degraeuwe et al., 2015; Henderson et al., 2014), so while not perfect, offers some indication of assay applicability. We constructed an ROC curve and observed an area under the curve of 0.847, indicating an high level of agreement between assays (Figure 4B). Using these data, the optimal threshold for the methylation assay to identify elevated FCA is a leukocyte estimate of ≥54.9%, with a sensitivity of 0.94 and specificity of 0.67. While promising, future studies will need to apply this in the context of clinical outcomes.

Discussion

In this proof-of-concept study we have applied two previously identified pan-leukocyte biomarkers (Dunnet et al., 2022), one each at the *HOXA3* and *MAP4K1* loci, to clinically relevant tissue samples: breastmilk and stool. With high-throughput sampling of breastmilk, we have shown that the proportion of leukocytes is extremely variable over short periods of time and between each breast. In the process, we have incorporated a new primer set into the pipeline, highlighting its ease of use and modifiability. Finally, we have demonstrated that even slight intestinal inflammation results in the majority of stool-derived DNA to be of leukocyte origin. Overall, our results show the assay pipeline can accurately and precisely determine the proportion of leukocytes from low levels of DNA.

TABLE 2 Total reagent cost of performing the *HOXA3* and *MAP4K1* bisulfite amplicon sequencing assay. Per-sample costs are calculated as the total cost of purchasing the reagent divided by the proportion of the reagent used for a single reaction. Kit prices are at retail price. All costs are in New Zealand Dollars (NZD).

Reagent	Use	Total cost	Number of reactions	Cost per sample
Zymo Research™ Quick-DNA Fecal/Soil Microbe Miniprep Kit	DNA extraction	$604.00	50	$12.08
Zymo EZ-DNA Methylation Direct MagPrep kit (4 × 96 reactions)	Bisulfite conversion	$1,349.00	384	$3.51
PCR reagents	PCR amplification	$540.00	100	$5.40
iSeq100 v2 cartridge and flow cell	DNA Sequencing	$990.00	1	varies

TABLE 3 The cost per sample of the *HOXA3* and *MAP4K1* bisulfite amplicon sequencing assay. Breastmilk cell lysis is performed in conjunction with the bisulfite conversion as part of the Zymo EZ-DNA Methylation Direct MagPrep kit. All costs are in New Zealand Dollars (NZD).

Sample type	Number of multiplex sequencing samples	Per sample DNA extraction cost	Per sample bisulfite conversion cost	Per sample PCR amplification cost	Per sample DNA sequencing cost	Total per sample cost
Stool	50	$12.08	$3.51	$5.40	$19.80	$40.79
	100	$12.08	$3.51	$5.40	$9.90	$30.89
	250	$12.08	$3.51	$5.40	$3.96	$24.95
	500	$12.08	$3.51	$5.40	$1.98	$22.97
Breastmilk	50	-	$3.51	$5.40	$19.80	$28.71
	100	-	$3.51	$5.40	$9.90	$18.81
	250	-	$3.51	$5.40	$3.96	$12.87
	500	-	$3.51	$5.40	$1.98	$10.89

Demonstration of cost-effectiveness

Cost-effectiveness is crucial in screening assay design; an assay that is too expensive to run has limited clinical value regardless of its performance metrics. The total reagent cost of performing either the *HOXA3* or *MAP4K1* dual-index PCR is the sum of costs for DNA extraction, bisulfite conversion, PCR amplification, and DNA sequencing per sample (Tables 2, 3). The per-sample cost is dramatically reduced as more individual samples are multiplexed together; thus, the high-throughput nature of the assay enables consumable cost reduction, down to $23 NZD per stool sample or $11 NZD per breastmilk sample.

High-throughput assaying of leukocyte fractions in human breastmilk

Leukocytes comprise a relatively small proportion of cells within human breastmilk. For example, in colostrum, leukocytes make up 8%–10% of total cells but are drastically reduced to approximately 1% of cells after 1-week post-partum (Hassiotou et al., 2013a; Trend et al., 2015). However, the fraction of leukocytes increases during illness to the mother or child, with the highest increase attributed to mastitis, where the percentage can exceed 90% (Hassiotou et al., 2013a). Thus, leukocyte proportions are an adequate measure of the overall health of the mother/infant dyad.

We applied the *HOXA3* and *MAP4K1* DNA methylation biomarkers (Dunnet et al., 2022) to milk samples collected from one individual during a lengthy bout of mastitis and were able to track increases and decreases in the leukocyte fraction with considerable consistency between the two biomarker loci. At times during the study period, leukocyte proportions approached normal baseline levels (Figure 2D); however, mostly we recorded high leukocyte production on account of the ongoing infection and injury.

Over the timepoints where breast-of-origin for milk samples could be tracked, we observed clear differences in leukocyte proportions. The primary site of mastitis (the left breast) maintained very high leukocyte levels throughout the peak of infection and in the days immediately following. Interestingly however, milk from the unaffected right breast produced greater than 90% leukocytes at the peak of the infection in the left breast (Figure 2A), but this dropped to 60% in the days immediately following infection.

Previous studies have shown that the levels of leukocytes in breastmilk increase when either the mother or infant is unwell and are dependent on the illness (Hassiotou et al., 2013a). Gastrointestinal infections and vaginal thrush in the mother induce small increases in leukocyte fraction, as do infant-only infections. However, illness in the mother can induce an increase in leukocyte fractions (Hassiotou et al., 2013a). When the mother has mastitis, leukocyte proportions in the unaffected breast have been reported to increase above baseline levels (Hassiotou et al., 2013a). In our time series, the unaffected right breast had a greatly

elevated proportion of leukocytes, perhaps indicative of a non-symptomatic infection or injury in that breast.

Leukocyte fractions in breastmilk have been suggested as a diagnostic marker to assess the health of mother/infant dyads (Hassiotou et al., 2013a). However, cellular composition studies employing flow cytometry often only examine milk from one or two timepoints (Riskin et al., 2012; Indumathi et al., 2013). One study on mastitis follows a case/control study design but does not explore the initiation and recovery from mastitis (Riskin et al., 2012). Furthermore, longitudinal studies frequently collect samples weeks apart (Hassiotou et al., 2013a; Trend et al., 2015; Hassiotou et al., 2013b; Nyquist et al., 2022). Given that the cellular components of human breastmilk are dynamic based on the stage of lactation (Trend et al., 2015), degree of breast fullness (Hassiotou et al., 2013b), infant feeding habits (Witkowska-Zimny and Kaminska-El-Hassan, 2017), the health of the mother infant dyad (Hassiotou et al., 2013a), and that the macronutrients within breastmilk are variable throughout the day (Hahn-Holbrook et al., 2019; Paulaviciene et al., 2020), it is not unexpected that the cell proportions of breastmilk change on a near daily basis. Indeed, the breastmilk samples employed in this study were collected almost daily over a 1-month period where continuous collection enabled the tracking of marked leukocyte dynamics within very short timeframes. High-throughput technology such as DNA methylation biomarkers can enable extensive sampling of breastmilk to characterise the cellular dynamics before, during, and after mastitis events with minimal cost and resources. While the process of bisulphite conversion and DNA sequencing is more time-consuming than flow cytometry for individual samples, batch testing of hundreds of samples simultaneously affords considerable efficiency of testing. As such, there is potential for this technology to be applied in a population-wide screen to track breast-related illness in lactating mothers or to further the understanding of the events leading up to mastitis provides more opportunities to prevent severe illness and guide treatment protocol.

Benchmarking a DNA methylation-based assay for the detection of leukocytes in stool against the faecal calprotectin assay

Stool is a tissue source for identifying intestinal tract pathologies and the clinical benefits of the FCA are well documented. For example, it may be used to aid diagnosis of infection (Sýkora et al., 2010; Shastri et al., 2008), colorectal cancer (von Roon et al., 2007), coeliac disease (Ertekin et al., 2010), or inflammatory bowel disease (IBD) (Laserna-Mendieta and Lucendo, 2019; Ayling and Kok, 2018). While FCAs often do not determine the cause of inflammation, they are extensively used because of their ability to distinguish between IBD and functional bowel disorders, such as irritable bowel syndrome (IBS) (Banerjee et al., 2015).

Inflammatory bowel disease describes a group of chronic inflammatory diseases affecting the gastrointestinal tract. There are two main types of IBD: ulcerative colitis (UC) and Crohn's disease (CD). Ulcerative colitis explicitly occurs in the colon and manifests as inflammation of the mucosal layer, while CD can present in any area of the gastrointestinal tract and is distinguished by granulomatous inflammation that can penetrate deep into the surrounding tissues (Ayling and Kok, 2018; Matsuoka and Kanai, 2015; Baumgart and Carding, 2007). In contrast, IBS has no identifiable pathophysiological features, and accordingly, no inflammation (Saha, 2014). Despite this, abdominal pain and diarrhoea are common symptoms in both IBD and IBS. Therefore, an endoscopy, occasionally followed by a histological examination of biopsied tissue, is required to diagnose IBD (Bharadwaj et al., 2018). Unfortunately, endoscopies are time-consuming, highly invasive for the patient (with some risk of complications), and require healthcare resources (Fisher et al., 2011). Using faecal calprotectin to distinguish between IBD and IBS minimises unnecessary endoscopies by screening prospective patients and monitoring patient remission to ensure that current treatment regimens are effective (Ayling and Kok, 2018; van Rheenen et al., 2010). Several meta-analyses have reported high sensitivities and excellent negative predictive values for the FCA (reviewed in Laserna-Mendieta and Lucendo (2019)) for diagnosing and monitoring IBD.

Rational for benchmarking a new intestinal inflammation detection tool

While the FCA has clinical utility when IBD is suspected, several limitations exist. Most commercial assays recommend ≥50 μg/g of calprotectin as the threshold for a positive test (Ayling and Kok, 2018). However, young children and adults over 65 have a higher baseline faecal calprotectin level. An initial meta-analysis suggested that the sensitivity of the FCA was significantly reduced in paediatric populations (van Rheenen et al., 2010). Contrary to this, more recent meta-analyses demonstrate that paediatric sensitivity is high (0.97–0.98) but specificity is reduced in comparison to adults for a 50 μg/g cut-off (Laserna-Mendieta and Lucendo, 2019; Degraeuwe et al., 2015; Henderson et al., 2014). The FCA has a lower sensitivity for adults over 65, resulting in lower diagnostic accuracy in this demographic (Padoan et al., 2018). Several studies have proposed various increases to cut-off thresholds in these groups, but these have not yet been implemented in a clinical setting (Padoan et al., 2018; Joshi et al., 2010; Mindemark and Larsson, 2012; Ezri and Nydegger, 2011). The exact reason elderly individuals have an increase in faecal calprotectin is currently unknown; however, one possible explanation is 'inflammaging' – an age-associated low-grade chronic inflammation in the absence of infection (Leonardi et al., 2018). In our study, methylation-based analysis detected a leukocyte signal in faecal samples that had a negative FCA result. Our previous work (Dunnet et al., 2022) showed specificity of the methylation pattern to leukocytes, whereas calprotectin is relatively specific to neutrophils raising the possibilities that we have detected non-neutrophlic leukocytes, such as monocyte/macrophages, eosinophils or lymphocytes. Alternatively the negative FCA result might reflect some of the factors that result in its low sensitivity.

The stability of faecal calprotectin over time has been extensively studied since the test was first proposed. An initial study reported that faecal calprotectin is stable at room temperature for up to 7 days (Røseth et al., 1992). Since then, several studies have shown that calprotectin is significantly degraded after 48–72 h at 20°C (Haisma

et al., 2019; Lasson et al., 2015; Oyaert et al., 2017). Storage at 4°C significantly reduces sample degradation for up to a week (Haisma et al., 2019). Calprotectin stability may prove problematic when samples must be transported long distances to testing facilities (for example, from rural communities) unless samples can be chilled during travel. In contrast, DNA methylation is a very stable epigenetic mark, although DNA itself is susceptible to hydrolysis when stored in an aqueous solution. However, storage of DNA in an appropriate conservation buffer can drastically decrease degradation. For example, in an experiment measuring DNA stability in stool samples, Olson et al. (2005) (Olson et al., 2005) show no DNA loss after 144 h of storage (the longest time point measured) at room temperature in three of four buffers tested. Their data suggest that DNA can be stable at least twice as long as calprotectin at room temperature.

Measuring faecal calprotectin is performed with an enzyme immunoassay, and each manufacturer uses unique proprietary antibodies. Different FCAs correlate well to patient outcomes, but absolute values cannot be compared between assays (Laserna-Mendieta and Lucendo, 2019; Ayling and Kok, 2018). For example, one study showed a 3.8-fold difference in the amount of calprotectin measured by different assays in a quality assurance scheme sample (Whitehead et al., 2013), while significant inter-assay differences have been observed in paediatric and adult IBD cases (Kittanakom et al., 2017; Labaere et al., 2014). Therefore, to achieve accurate follow-up testing during patient treatment, medical testing facilities must use a standardised FCA (Ayling and Kok, 2018). Alternatively, this may be alleviated with an open-source PCR-based DNA methylation assay where primer sequences and reaction conditions are known.

Calprotectin is also elevated in patients with pancreatic insufficiency (such as those with cystic fibrosis) (Ellemunter et al., 2017). This is likely due to a reduction of trypsin, which readily degrades calprotectin (Dumoulin et al., 2015). Here, a DNA methylation-based approach to identify leukocyte DNA might provide a more accurate representation than faecal calprotectin.

In healthy individuals, host DNA in stool is thought to predominately originate from sloughed colonic epithelial cells (He et al., 2019); however, cancers of the intestinal tract also shed DNA into the intestinal lumen (Whitney et al., 2004). Amplification of stool-derived DNA has been proposed as a non-invasive cancer detection tool (Mojtabanezhad Shariatpanahi et al., 2018). We hypothesized that the *MAP4K1* pan-leukocyte DNA biomarker coupled with the our high-throughput pipeline would detect an increase in leukocytes during inflammation which may be clinically relevant. Indeed, the high level of agreement between our pipeline and the FCA emphasises its clinical potential. We have shown that under even mild inflammatory states (as assessed by FCA), DNA from leukocytes is the dominant fraction of human-derived DNA. Interestingly, low calprotectin stool samples (under 50 µg/g) exhibited extremely variable proportions of leukocytes. This observation could be the result of a relatively low amount of sloughed epithelial cells such that even minor increases in leukocytes dramatically alters the overall proportion of cell types in each sample. Alternatively, this could be attributed to imprecisions in either the faecal calprotectin or DNA methylation assays. Patients with high levels of intestinal inflammation contain increased human DNA in their stool (Vincent et al., 2015), so it is likely the assay will perform better under higher levels of inflammation where more DNA is accessible for amplification. The FCA has high negative predictive value, but has been reported to generate a significant number of false positives (Freeman et al., 2021). As such, one possible application of our pipeline is for it to be used in tandem with the FCA to reduce the number of false positive results and subsequently the number of patients requiring an endoscopy.

Conclusion

In conclusion, we have applied a high-throughput, cost-effective DNA methylation biomarker pipeline to identify leukocytes from a mixed tissue. By employing a rapid sampling approach, we show that leukocyte proportions in breastmilk vary greatly with short periods of time. Leukocyte proportions in breastmilk are a useful health indictor of the mother/infant dyad, so a rapid sampling approach may provide the most accurate clinical information in this context. In addition, we have also shown that leukocyte-derived DNA can be detected in stool using our pipeline. We observed that even with low levels of intestinal inflammation as measured by faecal calprotectin, leukocyte-derived DNA dominates the human DNA fraction in stool samples.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://github.com/TimHore-Otago/Dunnet_NZ_Pan-leukocyte_Inflammation_Biomarkers.

Ethics statement

The use of human stool samples was approved by the University of Otago Human Ethics Committee (Health) (approval number H21/138). Human breastmilk and saliva were collected from the researchers with informed consent. The studies were conducted in accordance with the local legislation and institutional requirements.

Author contributions

MD: Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing–original draft, Writing–review and editing. IM: Conceptualization, Resources, Supervision, Writing–review and editing. DB: Conceptualization, Data curation, Investigation, Methodology, Project administration, Supervision, Writing–review and editing, Formal Analysis, Software. TH: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing–review and editing, Methodology, Software.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was funded by the University of Otago.

Acknowledgments

We thank Awanui Labs (formerly Southern Community Laboratories) for providing faecal calprotectin samples, and Jisha Antony for supplying K562 cells.

Conflict of interest

TH and DB are shareholders and directors of Totovision/Totogen Ltd, a small agricultural and biotechnology consultancy.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Anderton, H., Wicks, I. P., and Silke, J. (2020). Cell death in chronic inflammation: breaking the cycle to treat rheumatic disease. *Nat. Rev. Rheumatol.* 16 (9), 496–513. doi:10.1038/s41584-020-0455-8

Ayling, R. M., and Kok, K. (2018). Fecal calprotectin. *Adv. Clin. Chem.* 87, 161–190. doi:10.1016/bs.acc.2018.07.005

Banerjee, A., Srinivas, M., Eyre, R., Ellis, R., Waugh, N., Bardhan, K. D., et al. (2015). Faecal calprotectin for differentiating between irritable bowel syndrome and inflammatory bowel disease: a useful screen in daily gastroenterology practice. *Frontline gastroenterol.* 6 (1), 20–26. doi:10.1136/flgastro-2013-100429

Baumgart, D. C., and Carding, S. R. (2007). Inflammatory bowel disease: cause and immunobiology. *Lancet* 369 (9573), 1627–1640. doi:10.1016/s0140-6736(07)60750-8

Bharadwaj, S., Narula, N., Tandon, P., and Yaghoobi, M. (2018). Role of endoscopy in inflammatory bowel disease. *Gastroenterol. Rep.* 6 (2), 75–82. doi:10.1093/gastro/goy006

Bjarnason, I. (2017). The use of fecal calprotectin in inflammatory bowel disease. *Gastroenterol. Hepatol. (N Y).* 13 (1), 53–56.

Boakes, E., Woods, A., Johnson, N., and Kadoglou, N. (2018). Breast infection: a review of diagnosis and management practices. *Eur. J. breast health* 14 (3), 136–143. doi:10.5152/ejbh.2018.3871

Casellas, F., Antolı´n, M., Varela, E., Garcı´a-lafuente, A., Guarner, F., Borruel, N., et al. (2004). Fecal excretion of human deoxyribonucleic acid as an index of inflammatory activity in ulcerative colitis. *Clin. Gastroenterology Hepatology* 2 (8), 683–689. doi:10.1016/s1542-3565(04)00291-5

Cheng, A. P., Cheng, M. P., Gu, W., Sesing Lenz, J., Hsu, E., Schurr, E., et al. (2021). Cell-free DNA tissues of origin by methylation profiling reveals significant cell, tissue, and organ-specific injury related to COVID-19 severity. *Med* 2 (4), 411–422.e5. doi:10.1016/j.medj.2021.01.001

Cisneros-Villanueva, M., Hidalgo-Pérez, L., Rios-Romero, M., Cedro-Tanda, A., Ruiz-Villavicencio, C. A., Page, K., et al. (2022). Cell-free DNA analysis in current cancer clinical trials: a review. *Br. J. cancer* 126 (3), 391–400. doi:10.1038/s41416-021-01696-0

Degraeuwe, P. L. J., Beld, M. P. A., Ashorn, M., Canani, R. B., Day, A. S., Diamanti, A., et al. (2015). Faecal calprotectin in suspected paediatric inflammatory bowel disease. *J. Pediatr. Gastroenterology Nutr.* 60 (3), 339–346. doi:10.1097/mpg.0000000000000615

Dumoulin, E. N., Van Biervliet, S., Langlois, M. R., and Delanghe, J. R. (2015). Proteolysis is a confounding factor in the interpretation of faecal calprotectin. *Clin. Chem. Laboratory Med. (CCLM)* 53 (1), 65–71. doi:10.1515/cclm-2014-0568

Dunnet, M. J., Ortega-Recalde, O. J., Waters, S. A., Weeks, R. J., Morison, I. M., and Hore, T. A. (2022). Leukocyte-specific DNA methylation biomarkers and their implication for pathological epigenetic analysis. *Epigenetics Commun.* 2 (1), 5. doi:10.1186/s43682-022-00011-z

Ellemunter, H., Engelhardt, A., Schüller, K., and Steinkamp, G. (2017). Fecal calprotectin in cystic fibrosis and its relation to disease parameters: a longitudinal analysis for 12 years. *J. Pediatr. Gastroenterology Nutr.* 65 (4), 438–442. doi:10.1097/mpg.0000000000001544

Ertekin, V., Selimoğlu, M. A., Turgut, A., and Bakan, N. (2010). Fecal calprotectin concentration in celiac disease. *J. Clin. gastroenterology* 44 (8), 544–546. doi:10.1097/mcg.0b013e3181cadbc0

Ezri, J., and Nydegger, A. (2011). Pédiatrie. 4. La calprotectine fécale en pédiatrie: utilisation et interprétation. *Rev. Médicale Suisse* 7 (277), 69–70. doi:10.53738/revmed.2011.7.277.0069

Fisher, D. A., Maple, J. T., Ben-Menachem, T., Cash, B. D., Decker, G. A., Early, D. S., et al. (2011). Complications of colonoscopy. *Gastrointest. Endosc.* 74 (4), 745–752. doi:10.1016/j.gie.2011.07.025

Freeman, K., Taylor-Phillips, S., Willis, B. H., Ryan, R., and Clarke, A. (2021). Test accuracy of faecal calprotectin for inflammatory bowel disease in UK primary care: a retrospective cohort study of the THIN data. *BMJ open* 11 (2), e044177. doi:10.1136/bmjopen-2020-044177

Hahn-Holbrook, J., Saxbe, D., Bixby, C., Steele, C., and Glynn, L. (2019). Human milk as "chrononutrition": implications for child health and development. *Pediatr. Res.* 85 (7), 936–942. doi:10.1038/s41390-019-0368-x

Haisma, S. M., Verkade, H. J., Scheenstra, R., van der Doef, H. P. J., Bodewes, F., and van Rheenen, P. F. (2019). Time-to-reach target calprotectin level in newly diagnosed patients with inflammatory bowel disease. *J. Pediatr. Gastroenterol. Nutr.* 69 (4), 466–473. doi:10.1097/mpg.0000000000002458

Hassiotou, F., Hepworth, A. R., Metzger, P., Tat Lai, C., Trengove, N., Hartmann, P. E., et al. (2013a). Maternal and infant infections stimulate a rapid leukocyte response in breastmilk. *Clin. and Transl. Immunol.* 2 (4), e3–e. doi:10.1038/cti.2013.1

Hassiotou, F., Hepworth, A. R., Williams, T. M., Twigger, A.-J., Perrella, S., Lai, C. T., et al. (2013b). Breastmilk cell and fat contents respond similarly to removal of breastmilk by the infant. *PLOS ONE* 8 (11), e78232. doi:10.1371/journal.pone.0078232

He, K., Fujiwara, H., Zajac, C., Sandford, E., Reddy, P., Choi, S. W., et al. (2019). A pipeline for faecal host DNA analysis by absolute quantification of LINE-1 and mitochondrial genomic elements using ddPCR. *Sci. Rep.* 9 (1), 5599. doi:10.1038/s41598-019-41753-6

Henderson, P., Anderson, N. H., and Wilson, D. C. (2014). The diagnostic accuracy of fecal calprotectin during the investigation of suspected pediatric inflammatory bowel disease: a systematic review and meta-analysis. *Am. J. Gastroenterology* 109 (5), 637–645. doi:10.1038/ajg.2013.131

Indumathi, S., Dhanasekaran, M., Rajkumar, J. S., and Sudarsanam, D. (2013). Exploring the stem cell and non-stem cell constituents of human breast milk. *Cytotechnology* 65 (3), 385–393. doi:10.1007/s10616-012-9492-8

Jha, A. K., Chaudhary, M., Dayal, V. M., Kumar, A., Jha, S. K., Jha, P., et al. (2018). Optimal cut-off value of fecal calprotectin for the evaluation of ulcerative colitis: an unsolved issue? *JGH Open* 2 (5), 207–213. doi:10.1002/jgh3.12074

Joshi, S., Lewis, S. J., Creanor, S., and Ayling, R. M. (2010). Age-related faecal calprotectin, lactoferrin and tumour M2-PK concentrations in healthy volunteers. *Ann. Clin. Biochem. Int. J. Laboratory Med.* 47 (Pt 3), 259–263. doi:10.1258/acb.2009.009061

Kittanakom, S., Shajib, M. S., Garvie, K., Turner, J., Brooks, D., Odeh, S., et al. (2017). Comparison of fecal calprotectin methods for predicting relapse of pediatric inflammatory bowel disease. *Can. J. Gastroenterology Hepatology* 2017, 1–10. doi:10.1155/2017/1450970

Krueger, F., and Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27 (11), 1571–1572. doi:10.1093/bioinformatics/btr167

Labaere, D., Smismans, A., Van Olmen, A., Christiaens, P., D'Haens, G., Moons, V., et al. (2014). Comparison of six different calprotectin assays for the assessment of inflammatory bowel disease. *United Eur. gastroenterology J.* 2 (1), 30–37. doi:10.1177/2050640613518201

Lam, D., Luu, P. L., Song, J. Z., Qu, W., Risbridger, G. P., Lawrence, M. G., et al. (2020). Comprehensive evaluation of targeted multiplex bisulphite PCR sequencing for validation of DNA methylation biomarker panels. *Clin. Epigenetics* 12 (1), 90. doi:10.1186/s13148-020-00880-y

Laserna-Mendieta, E. J., and Lucendo, A. J. (2019). Faecal calprotectin in inflammatory bowel diseases: a review focused on meta-analyses and routine usage limitations. *Clin. Chem. laboratory Med.* 57 (9), 1295–1307. doi:10.1515/cclm-2018-1063

Lasson, A., Stotzer, P.-O., Öhman, L., Isaksson, S., Sapnara, M., and Strid, H. (2015). The intra-individual variability of faecal calprotectin: a prospective study in patients with active ulcerative colitis. *J. Crohn's and colitis* 9 (1), 26–32. doi:10.1016/j.crohns.2014.06.002

Lehmann-Werman, R., Magenheim, J., Moss, J., Neiman, D., Abraham, O., Piyanzin, S., et al. (2018). Monitoring liver damage using hepatocyte-specific methylation markers in cell-free circulating DNA. *JCI Insight* 3 (12), e120687. doi:10.1172/jci.insight.120687

Lehmann-Werman, R., Neiman, D., Zemmour, H., Moss, J., Magenheim, J., Vaknin-Dembinsky, A., et al. (2016). Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proc. Natl. Acad. Sci. U. S. A.* 113 (13), E1826–E1834. doi:10.1073/pnas.1519286113

Leonardi, G. C., Accardi, G., Monastero, R., Nicoletti, F., and Libra, M. (2018). Ageing: from inflammation to cancer. *Immun. and Ageing* 15 (1), 1. doi:10.1186/s12979-017-0112-5

Lundgren, D., Eklöf, V., Palmqvist, R., Hultdin, J., and Karling, P. (2019). Proton pump inhibitor use is associated with elevated faecal calprotectin levels. A cross-sectional study on subjects referred for colonoscopy. *Scand. J. Gastroenterology* 54 (2), 152–157. doi:10.1080/00365521.2019.1566493

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J.* 17 (1), 10. doi:10.14806/ej.17.1.200

Matsuoka, K., and Kanai, T. (2015). The gut microbiota and inflammatory bowel disease. *Seminars Immunopathol.* 37 (1), 47–55. doi:10.1007/s00281-014-0454-4

Mindemark, M., and Larsson, A. (2012). Ruling out IBD: estimation of the possible economic effects of pre-endoscopic screening with F-calprotectin. *Clin. Biochem.* 45 (7), 552–555. doi:10.1016/j.clinbiochem.2011.10.015

Mojtabanezhad Shariatpanahi, A., Yassi, M., Nouraie, M., Sahebkar, A., Varshoee Tabrizi, F., and Kerachian, M. A. (2018). The importance of stool DNA methylation in colorectal cancer diagnosis: a meta-analysis. *PLoS One* 13 (7), e0200735. doi:10.1371/journal.pone.0200735

Moss, J., Magenheim, J., Neiman, D., Zemmour, H., Loyfer, N., Korach, A., et al. (2018). Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat. Commun.* 9 (1), 5068. doi:10.1038/s41467-018-07466-6

Nyquist, S. K., Gao, P., Haining, T. K. J., Retchin, M. R., Golan, Y., Drake, R. S., et al. (2022). Cellular and transcriptional diversity over the course of human lactation. *Proc. Natl. Acad. Sci. U. S. A.* 119 (15), e2121720119. doi:10.1073/pnas.2121720119

Oberacker, P., Stepper, P., Bond, D. M., Höhn, S., Focken, J., Meyer, V., et al. (2019). Bio-On-Magnetic-Beads (BOMB): open platform for high-throughput nucleic acid extraction and manipulation. *PLOS Biol.* 17 (1), e3000107. doi:10.1371/journal.pbio.3000107

Odink, K., Cerletti, N., Brüggen, J., Clerc, R. G., Tarcsay, L., Zwadlo, G., et al. (1987). Two calcium-binding proteins in infiltrate macrophages of rheumatoid arthritis. *Nature* 330 (6143), 80–82. doi:10.1038/330080a0

Olson, J., Whitney, D. H., Durkee, K., and Shuber, A. P. (2005). DNA stabilization is critical for maximizing performance of fecal DNA-based colorectal cancer tests. *Diagn. Mol. Pathol.* 14 (3), 183–191. doi:10.1097/01.pas.0000176768.18423.7e

Oyaert, M., Van den Bremt, S., Boel, A., Bossuyt, X., and Van Hoovels, L. (2017). Do not forget about pre-analytics in faecal calprotectin measurement. *Clin. Chim. Acta* 473, 124–126. doi:10.1016/j.cca.2017.08.025

Padoan, A., D'Incà, R., Scapellato, M. L., De Bastiani, R., Caccaro, R., Mescoli, C., et al. (2018). Improving IBD diagnosis and monitoring by understanding preanalytical, analytical and biological fecal calprotectin variability. *Clin. Chem. laboratory Med.* 56 (11), 1926–1935. doi:10.1515/cclm-2018-0134

Paulaviciene, I. J., Liubsys, A., Molyte, A., Eidukaite, A., and Usonis, V. (2020). Circadian changes in the composition of human milk macronutrients depending on pregnancy duration: a cross-sectional study. *Int. Breastfeed. J.* 15 (1), 49. doi:10.1186/s13006-020-00291-y

Pencovich, N., Jaschek, R., Tanay, A., and Groner, Y. (2011). Dynamic combinatorial interactions of RUNX1 and cooperating partners regulates megakaryocytic differentiation in cell line models. *Blood* 117 (1), e1–e14. doi:10.1182/blood-2010-07-295113

Ricciuto, A., and Griffiths, A. M. (2019). Clinical value of fecal calprotectin. *Crit. Rev. Clin. laboratory Sci.* 56 (5), 307–320. doi:10.1080/10408363.2019.1619159

Riskin, A., Almog, M., Peri, R., Halasz, K., Srugo, I., and Kessel, A. (2012). Changes in immunomodulatory constituents of human milk in response to active infection in the nursing infant. *Pediatr. Res.* 71 (2), 220–225. doi:10.1038/pr.2011.34

Rogler, G., Aldeguer, X., Kruis, W., Lasson, A., Mittmann, U., Nally, K., et al. (2013). Concept for a rapid point-of-care calprotectin diagnostic test for diagnosis and disease activity monitoring in patients with inflammatory bowel disease: expert clinical opinion. *J. Crohn's Colitis* 7 (8), 670–677. doi:10.1016/j.crohns.2013.02.014

Røseth, A. G., Fagerhol, M. K., Aadland, E., and Schjønsby, H. (1992). Assessment of the neutrophil dominating protein calprotectin in feces: a methodologic study. *Scand. J. Gastroenterology* 27 (9), 793–798. doi:10.3109/00365529209011186

Røseth Pnsmkf, A. G. (1999). Correlation between faecal excretion of indium-111-labelled granulocytes and calprotectin, a granulocyte marker protein, in patients with inflammatory bowel disease. *Scand. J. Gastroenterology* 34 (1), 50–54. doi:10.1080/00365529950172835

Saha, L. (2014). Irritable bowel syndrome: pathogenesis, diagnosis, treatment, and evidence-based medicine. *World J. gastroenterology* 20 (22), 6759–6773. doi:10.3748/wjg.v20.i22.6759

Shastri, Y. M., Bergis, D., Povse, N., Schäfer, V., Shastri, S., Weindel, M., et al. (2008). Prospective multicenter study evaluating fecal calprotectin in adult acute bacterial diarrhea. *Am. J. Med.* 121 (12), 1099–1106. doi:10.1016/j.amjmed.2008.06.034

Sýkora, J., Siala, K., Huml, M., Varvařovská, J., Schwarz, J., and Pomahačová, R. (2010). Evaluation of faecal calprotectin as a valuable non-invasive marker in distinguishing gut pathogens in young children with acute gastroenteritis. *Acta Paediatr. (Oslo, Nor. 1992)* 99 (9), 1389–1395. doi:10.1111/j.1651-2227.2010.01843.x

Teschendorff, A. E., Breeze, C. E., Zheng, S. C., and Beck, S. (2017). A comparison of reference-based algorithms for correcting cell-type heterogeneity in Epigenome-Wide Association Studies. *BMC Bioinforma.* 18 (1), 105. doi:10.1186/s12859-017-1511-5

Theda, C., Hwang, S. H., Czajko, A., Loke, Y. J., Leong, P., and Craig, J. M. (2018). Quantitation of the cellular content of saliva and buccal swab samples. *Sci. Rep.* 8 (1), 6944. doi:10.1038/s41598-018-25311-0

Trend, S., de Jong, E., Lloyd, M. L., Kok, C. H., Richmond, P., Doherty, D. A., et al. (2015). Leukocyte populations in human preterm and term breast milk identified by multicolour flow cytometry. *PLoS One* 10 (8), e0135580. doi:10.1371/journal.pone.0135580

van Rheenen, P. F., Van de Vijver, E., and Fidler, V. (2010). Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. *BMJ Clin. Res.* 341, c3369. doi:10.1136/bmj.c3369

Vincent, C., Mehrotra, S., Loo, V. G., Dewar, K., and Manges, A. R. (2015). Excretion of host DNA in feces is associated with risk of *Clostridium difficile* infection. *J. Immunol. Res.* 2015, 1–7. doi:10.1155/2015/246203

von Roon, A. C., Karamountzos, L., Purkayastha, S., Reese, G. E., Darzi, A. W., Teare, J. P., et al. (2007). Diagnostic precision of fecal calprotectin for inflammatory bowel disease and colorectal malignancy. *Am. J. gastroenterology* 102 (4), 803–813. doi:10.1111/j.1572-0241.2007.01126.x

Vrablicova, Z., Tomova, K., Tothova, Ľ., Bábíčková, J., Gromova, B., Konečná, B., et al. (2020). Nuclear and mitochondrial circulating cell-free DNA is increased in patients with inflammatory bowel disease in clinical remission. *Front. Med.* 7, 593316. doi:10.3389/fmed.2020.593316

Whitehead, S. J., French, J., Brookes, M. J., Ford, C., and Gama, R. (2013). Between-assay variability of faecal calprotectin enzyme-linked immunosorbent assay kits. *Ann. Clin. Biochem. Int. J. Laboratory Med.* 50 (Pt 1), 53–61. doi:10.1258/acb.2012.011272

Whitney, D., Skoletsky, J., Moore, K., Boynton, K., Kann, L., Brand, R., et al. (2004). Enhanced retrieval of DNA from human fecal samples results in improved performance of colorectal cancer screening test. *J. Mol. Diagnostics* 6 (4), 386–395. doi:10.1016/s1525-1578(10)60536-3

Witkowska-Zimny, M., and Kaminska-El-Hassan, E. (2017). Cells of human breast milk. *Cell. and Mol. Biol. Lett.* 22, 11. doi:10.1186/s11658-017-0042-4

Zemmour, H., Planer, D., Magenheim, J., Moss, J., Neiman, D., Gilon, D., et al. (2018). Non-invasive detection of human cardiomyocyte death using methylation patterns of circulating DNA. *Nat. Commun.* 9 (1), 1443. doi:10.1038/s41467-018-03961-y

Ziller, M. J., Gu, H., Muller, F., Donaghey, J., Tsai, L. T., Kohlbacher, O., et al. (2013). Charting a dynamic DNA methylation landscape of the human genome. *Nature* 500 (7463), 477–481. doi:10.1038/nature12433