



An internal pilot study of a novel rectal mucocellular sampling device to allow next-generation sequencing for colorectal disease

H. N. Humphrey¹ · A. Diodato² · J.-C. Isner² · E. Walker² · J. Lacy-Colson³ · B. Nedjai^{2,4} · I. R. Daniels^{1,2} on behalf of ORI-EGI-02 Study Group²

Received: 9 August 2022 / Accepted: 3 September 2022
© Springer Nature Switzerland AG 2022

Abstract

Background The ORI-EGI-02 study was designed to test the hypothesis that rectal mucus collected using a novel rectal sampling device (OriCol™), contains sufficient human deoxyribonucleic acid (DNA) of the required quality for Next Generation Sequencing (NGS), for colorectal disease genetic signature discovery.

Methods Using National Institute for Health and Care Research methodology, an internal pilot study was performed in January 2020–May 2021, at four sites in the United Kingdom, to assess the process of recruitment, consent, specimen acquisition and viability for analysis. Following an OriCol™ test, the sample was stabilized with a buffer solution to preserve the material, which was posted to the laboratory. Samples were processed using QIAamp® DNA Blood Midi kit to extract DNA and Quant-iT™ PicoGreen® dsDNA Reagent to quantify the retrieved DNA. DNA integrity was measured by Agilent TapeStation system. 25 ng of human amplifiable DNA was prepared for Next Generation Sequencing (NGS), which was performed on an Illumina NextSeq550 sequencer using the 300-cycle high output kit v2.5.

Results This study assessed the first 300 patients enrolled to the ORI-EGI-02 Study ($n = 800$). 290/300 (96.67%) were eligible to undergo OriCol™ sampling procedure and 285/290 (98.27%) had a successful OriCol™ sample taken. After transportation, extraction and quantification of DNA, 96.20% (279/290) of the samples had NGS successfully performed for bioinformatic analysis.

Conclusions Our internal pilot study demonstrated that the OriCol™ sampling device can capture rectal mucus from unprepared bowel in subjects who could undergo a digital rectal examination. The technique could be applied irrespective of age, frailty, or co-morbidity. Completion of the study to 800 patients and analysis of NGS data for colorectal cancer mutations will now proceed.

Keywords Rectal mucus · Colorectal cancer · Two-week referral · Next-generation sequencing

Introduction

Currently, patients in the United Kingdom (UK) with suspected colorectal cancer are referred via the 2-Week Wait (2WW) Pathway based upon National Institute for Health and Care Excellence (NICE) criteria [1]. In England, 373,204 patients were referred by the 2WW pathway for suspected lower gastrointestinal (GI) cancer [2] with a cancer diagnosis being confirmed in approximately 3% [3, 4]. Overall, 54% of colorectal cancer diagnoses result from the 2WW referral pathway, 20% present as emergency admissions and ~10% through the UK National Bowel Cancer Screening Programme (BCSP), the remainder via other routes [5]. In the last 10 years, the numbers of referrals have increased by 45%, impacting the delivery of the

The members of the “ORI-EGI-02 Study Group” was processed under Acknowledgements section.

✉ I. R. Daniels
ian.daniels@originsciences.com

¹ Royal Devon University Healthcare NHS Foundation Trust, Barrack Road, Exeter EX53BJ, Devon, UK

² Origin Sciences, Granta Park, Cambridge CB21 6AD, UK

³ Shrewsbury and Telford NHS Trust, Shrewsbury SY3 8XQ, Shropshire, UK

⁴ Centre for Prevention, Diagnosis and Detection, Queen Mary University of London, WIPH, Cancer Prevention Unit, London, UK

“Gold-standard” diagnostic test (colonoscopy), with 747 per 100,000 population colonoscopy procedures completed in 2017–18 [6–8]. Performance of a successful colonoscopy is reliant on the ability to undergo oral bowel preparation and achieve a “clean bowel”, which is impacted by both patient frailty and co-morbidity as well as technical ability to assess the mucosa of the entire colon and rectum. Recent work assessed the value of adding quantitative faecal immunochemical test (qFiT) to the established pathway [9] as a triage tool to select patients for urgent investigation [10]. The continued impact of the coronavirus disease (COVID-19) pandemic has added additional pressures on service delivery and lengthened patient waits, leading to the search for new solutions for improved diagnostic tests. These include the “Colon Capsule” and detection of volatile organic compounds in urine, which are both undergoing clinical trials [11] as well as studies assessing the detection of circulating cell free deoxyribonucleic acid (cfDNA) in the blood and initiatives to expand research in this area have been launched [12, 13].

However, currently, the only stool-based alternative to a qFiT is the Cologuard® test (Exact Sciences Corporation, Madison, WI, USA) assessing multi-target stool DNA detection in combination with qFiT [14]. With several studies demonstrating that significant numbers of exfoliated cells and their products are retained in the mucus layer overlying the rectal mucosa, we hypothesized that analysis of rectal mucus could be used to detect significant colorectal disease [15–17]. Having developed a novel sampling device (OriCol™) to collect rectal mucus from the distal rectum, without the need for prior bowel preparation, we have performed an internal pilot study using the first 300 of the 800 participants recruited to the Ori-EGI-02 Study. Using National Institute for Health and Care Research NIHR pilot study methodology [18], we aimed to assess patient recruitment, consent, specimen acquisition and viability of specimens for Next Generation Sequencing (NGS) from human DNA (huDNA) retrieved from the mucus.

Materials and methods

The study was conducted in January 2020—May 2021 at four sites in the UK: The Shrewsbury & Telford NHS Trust, The Royal Devon & Exeter NHS Foundation Trust, The Royal Cornwall NHS Foundation Trust, and Oxford Hospitals NHS Foundation Trust. The study is registered at clinicaltrials.gov (NCT04659590) as ORI-EGI-02.

Patient and public engagement statement

Patients who had been triaged to be assessed in an Outpatient Clinic prior to investigation were invited to take part in

the study. No patients were involved in this stage of the study design as we were assessing the ability to collect material for analysis as well as assessing the number of patients who could potentially undergo an Oricol™ test. Patients provided feedback about the acceptability of both the passage of the proctoscope and the inflation of the balloon. Future studies will assess patient acceptability of the test compared to qFiT, colonoscopy and computed tomography (CT) scanning, addressing patient preference and confidence in the test. Patient involvement will be necessary for future phases of development.

Patient selection

Patients referred through the 2WW process to the colorectal outpatient clinic and those with a confirmed colorectal cancer managed through the multidisciplinary team (MDT) process with suspected or confirmed colorectal cancer were approached. Symptoms for referral to the colorectal clinic triaged via the NICE guidelines (NG-12) are; aged 40 or over with unexplained weight loss or abdominal pain, aged under 50 years with rectal bleeding in addition to abdominal pain/change in bowel habit/weight loss/iron deficiency anemia, aged 50 or over with unexplained rectal bleeding, aged 60 or over with iron-deficiency anemia or changes in their bowel habit or tests show occult blood in their faeces, all ages with rectal or abdominal mass and a positive qFiT. General inclusion criteria required participants to be over 18 years of age and to be able to give voluntary, written, informed consent. Exclusion criteria included previous history of cancer, previously received pelvic radiotherapy, induction/neo-adjuvant chemotherapy or are receiving concurrent immunotherapy, history of allergic reactions to polypropylene and/or nitrile, positive pregnancy test, any form of bowel preparation/oral contrast medium within the last 14 days, any painful perianal conditions that would make proctoscopy inappropriate as determined by rectal examination, participation in an interventional study if treatment/intervention has already started and/or known Hepatitis B or C, human immunodeficiency virus (HIV), or any other similarly classified human pathogen including prion diseases (Creutzfeldt-Jakob disease), or COVID-19.

Ethics approvals

Informed written consent was obtained from all participants. Ethical approval for this study was granted by the Health Research Authority, East Midlands, Nottingham 1 Research Ethics Committee (REC Reference 19/EM/0266). Research contact occurred in one or two visits within the patient pathway and had minimal participant burden.

Oricol™ preparation

The Oricol™ Sampling device (Origin Sciences, Cambridge, UK) comes prepared in a clam shell pack containing the sampling device and a standard proctoscope. As per manufacturing instructions, the Oricol™ Sampling device is checked by test inflation of 80mls and retraction of 80mls of air.

Oricol™ sample collection

A digital rectal examination was performed on all participants prior to the use of the Oricol™ Sampling device. The sampling device incorporates a nitrile membrane that, upon insertion into the unprepared rectum via a standard proctoscope, is inflated with 80 ml of air using a syringe. The inflated membrane is of a diameter that makes circumferential contact with the rectal mucosa and remains in contact for 10 s before deflation, it is then fully retracted prior to removal from the patient and the membrane is then inverted into the device (<https://www.originsciences.com/oricol/device>). Following the test, a buffer solution (0.5 M Tris, 0.1 M EDTA, 10 mM NaCl, 0.1% SDS, pH 8.0) is added directly to the inverted membrane to preserve the material until received and processed by the laboratory. The device is then sealed for storage and transportation. Samples are initially stored and shipped at ambient temperature. Upon arrival in the laboratory, the samples are immediately processed. The samples are aliquoted and stored at -80°C .

DNA extraction and quality control

A QIAamp® DNA Blood Midi kit (Qiagen, 51,185) was used to extract the DNA according to the manufacturers' protocol, except that an extra centrifugation step was added (5000 g, 10 min) before the lysis step, which was performed at 65°C to decrease the levels of prokaryotic material. Total DNA was quantified by using Quant-iT™ PicoGreen® dsDNA Reagent (Molecular Probes, P7581). The proportion of human DNA (huDNA) present in the sample was determined with a quantitative real-time polymerase chain reaction (PCR) with a 94 bp probe of the beta Globin gene (PrimerBG_F: AGC AACCTCAAACAGACACCAT, PrimerBG_R: CCAACTTCATCCACGTTACCTT). The DNA integrity was measured on a TapeStation (Agilent 5067–5366). This method measures the level of fragmentation of the genomic DNA (gDNA). The TapeStation Analysis Software determines a DNA Integrity Number (DIN) as a measure of gDNA integrity, from 1 to 10, where 1 indicates strongly degraded gDNA, and 10 indicates high integrity of gDNA. A cut-off of $>1\%$ huDNA was set for each sample to proceed to NGS.

Next-generation sequencing

A targeted NGS-based multigene panel focused on colorectal cancer (Cell3™ Target Cancer Panel from Nonacus, Birmingham, UK). The library was prepared according to the manufacturer recommendation starting with 25 ng of huDNA, adding unique molecular indices (UMI) adapter to the DNA fragments before the PCR amplification (eight cycles). Ten libraries were pooled together, and a subsequent 12 cycle PCR was performed. The sequencing was performed on a NextSeq550 sequencer using the 300-cycle high output kit v2.5 (Illumina, San Diego, USA). The reads were preprocessed using fastp, the sequences were then aligned against the human genome hg38 with bwa-mem2 [19]. The optical duplicates were removed. The percentage of mapped reads was determined with SAMtools [20].

Statistical analysis

Statistical analysis was performed on the generated data, Pearson correlation coefficient and Mann–Whitney was used to assess significance. Scatter and boxplots were created using Python 3.8 (Open Source) [21]. The study has been reported by the Strobe checklist [22].

Results

The Ori-EGI-02 study aims to recruit 800 patients referred with suspected bowel cancer through the 2WW pathway. This internal pilot study of the initial 300 recruited patients assesses the sample collection, transfer and the quality of human DNA extraction, amplification, and analysis and patient acceptance (Fig. 1).

Patient population

Of the 300 patients, 151 (50.03%) were male. Full demographic detail is given (Table 1). The patients approached for the study were those referred for face-to-face assessment from 2WW Triage and patients diagnosed with colorectal cancer identified at the Colorectal Cancer MDT meeting.

Patient frailty

The population group encompassed patients with degrees of frailty assessed against the nine-point Rockwood Clinical Frailty Score [23]. The median frailty score was 2 (range 1–7).

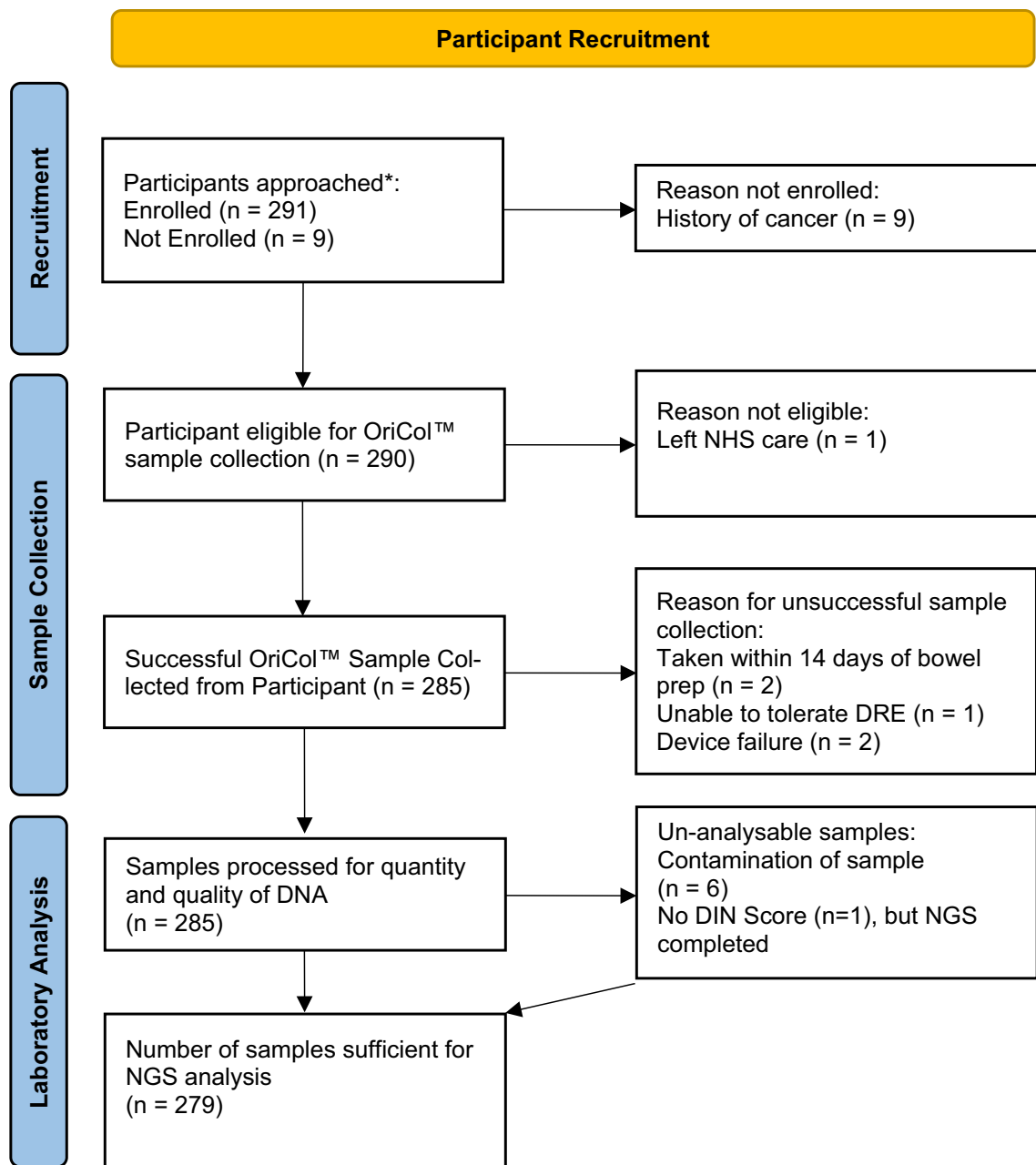


Fig. 1 PRISMA flow diagram. *PRISMA* Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Sampling

Samples were collected by both clinicians (consultant surgeons and surgical research fellows, $n = 5$) and clinical nurse specialists ($n = 3$). Of the 300 patients approached, OriCol™ tests were successfully performed on 290/300 (96.67%). Nine patients were excluded as protocol deviators with a previous history of cancer and one patient did not complete their care within the NHS and was withdrawn from the study. Samples were successfully taken from 285/290 (98.27%) of patients, 2/290 patients were removed as they had received

bowel preparation within 14 days of the test (0.69%), 1/290 patient (0.34%) was unable to tolerate a rectal examination and device failure occurred in 2/290 patients (0.69%).

Patient acceptability

Pain scores measured on a visual analogue scale (scores 1–10) were assessed for both the passage of the proctoscope and the inflation of the balloon. The mean proctoscope pain score was 1.95 (median 1, range 1–8) and for inflation of the balloon the mean pain score was 1.64

Table 1 Patient demographics

Characteristic	N= 300 (%)
Sex	
Male	151 (50.3%)
Female	149 (49.7%)
BMI (kg/m ²)	
< 18.5	3 (1.0%)
≥ 18.5–< 25	79 (26.3%)
≥ 25–< 30	99 (33.0%)
≥ 30–< 40	68 (22.7%)
≥ 40	8 (2.7%)
Unknown	43 (14.3%)
Smoker	
Never	166 (55.3%)
Ex-smoker	112 (37.3%)
Smoker	22 (7.3%)
Age (years)	
18–49	16 (5.3%)
50–59	33 (11.0%)
60–69	50 (16.7%)
70–79	106 (35.3%)
≥ 80	95 (31.7%)
Rockwood scale	
1 = Very fit	91
2 = Well	273
3 = Managing well	141
4 = Vulnerable	57
5 = Mildly Frail	30
6 = Moderately frail	15
7 = Severely frail	3
8 = Very severely frail	0
9 = Terminally ill	0

BMI body mass index

(median 1, range 1–8). Overall, 99.60% of participants would be willing to have an OriCol™ test again.

Previous bowel movement

Data were collected on the relationship between time from last bowel movement ($n = 284$) and sampling to assess the quantity of the huDNA included in the total DNA retrieved. Based upon time from last bowel movement, the groups were < 30 min ($n = 6$), 30–60 min ($n = 7$), 60–240 min ($n = 112$) and > 240 min ($n = 159$), there was no effect on the fraction of huDNA when assessed by Mann–Whitney U test (Fig. 2a).

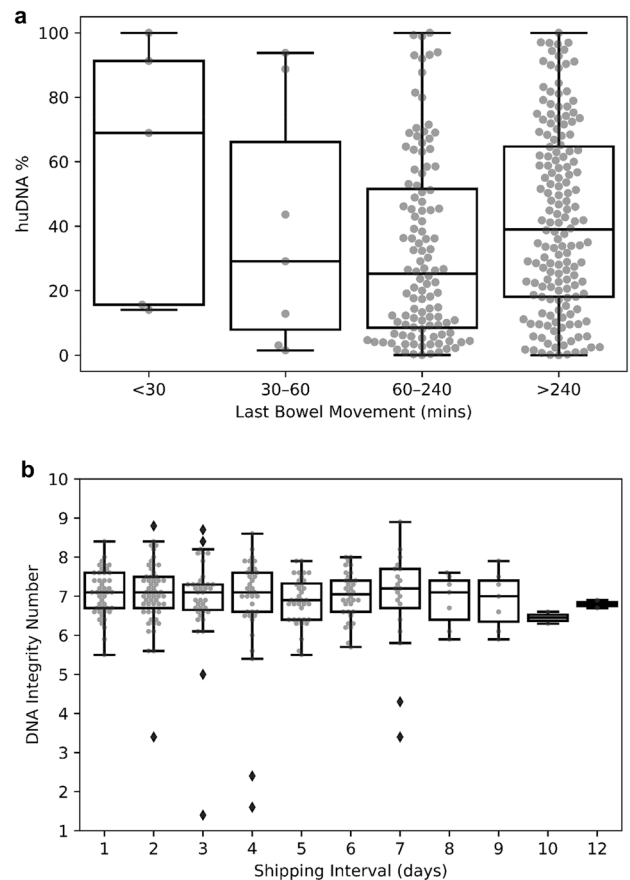


Fig. 2 Quantity and quality of human DNA. **a** Relationship between time from last bowel movement and quantity of human DNA. **b** The relationship between time from collection to arrival in laboratory and the integrity of the DNA. DNA deoxyribonucleic acid

Transfer of material

All samples arrived successfully in the laboratory through the Royal Mail Postage System. The longest delay between sample collection and arrival for processing in the laboratory was 12 days. The median time to arrival was 4 days (range 1–12 days) from OriCol™ sampling to arrival in the laboratory. No samples were lost in transit, one sample had a loose screw top, although material was still retrieved. No statistically significant difference was seen in the integrity of DNA (DIN) and time from sampling to laboratory receipt (Pearson R value – 0.07) (Fig. 2b).

Quantity and quality of extracted DNA

Upon arrival in the laboratory (Origin Sciences, Cambridge, UK) samples were stored at – 80 °C. Total DNA extracted ranged from 1 to 50 ug/ml rectal mucus. We identified that a high DNA amount correlated with high levels of non-human DNA (from residual food stuffs). Six samples (6/285, 2.11%)

had a high degree of contamination and were not suitable for analysis. Overall, 279/285 (97.89%) of samples were available for downstream analyses. The quantity and quality of DNA suitable for downstream applications were assessed using the PicoGreen® assay. This demonstrated that 232 out of 281 samples (82.56%) had more than 1 µg of total DNA per ml of rectal mucus (range 1–26.7 µg/ml). The quality of DNA was assessed on TapeStation (Agilent 2200 TapeStation system) and was high in 97.86% of the samples with a DIN between 5 and 8.9. Only 6 samples (2.14%) had a DIN lower than 5 (Fig. 3). These six samples arrived with a high content of stool as described in the recorded comments from visual assessment. Of these six samples, only two samples could not be used on NGS as they had a very low content of huDNA. There was negligible correlation between the fraction of huDNA and the integrity of the DNA (DIN) (Pearson R value 0.29). However, a relevant quality parameter for using a sample for NGS in our setting is to have a minimum of 1% of huDNA. The DIN does not impact on the success of NGS, for example, one sample had a DIN 3.4 (2.43% of huDNA).

Human amplifiable DNA

huDNA has been calculated by analyzing the beta globin gene by real-time quantitative PCR. The ratio between the amount of beta globin gene detected and the total amount of DNA measured by PicoGreen®, gives an estimation of the content of human DNA in Oricol™ samples. Overall, 96.09% of samples have huDNA percentage in a range

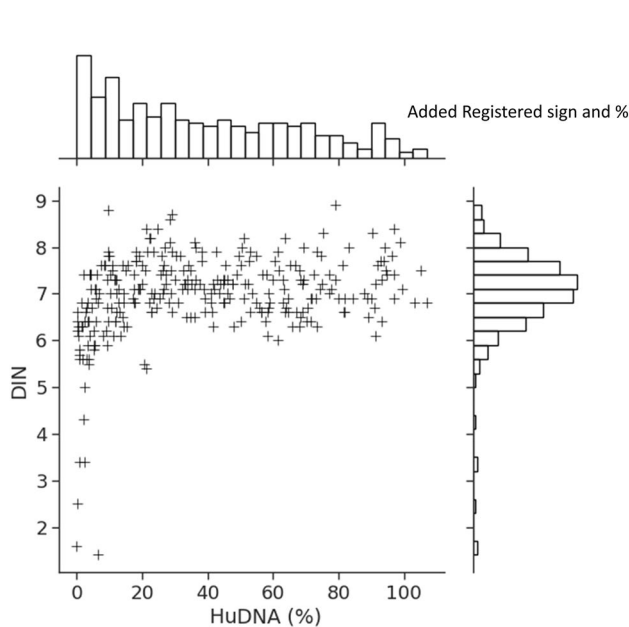


Fig. 3 Correlation between DIN and the huDNA percentage. *DNA* deoxyribonucleic acid, *huDNA* human DNA, *DIN* DNA integrity number

highly suitable for the NGS experiment. Only 3.91% of samples have huDNA below or equal to 1%. All the samples were processed through NGS to investigate whether huDNA had an impact on the sequencing quality.

NGS

The libraries were prepared using exome capture on a panel of 50 genes using Unique Molecular Indices (UMI) to improve the specificity. Ten libraries were pooled together and sequenced on a high-throughput cell on a Nextseq550. Reads were aligned to the human genome and 248/285 (87%) samples had a high percentage (> 80%) of reads mapped to the human genome. When huDNA was below 20%, we identified an increasing number of unmapped reads (Fig. 4). Four samples had very low huDNA % (0.16, 0.43, 0.67 and 0.86%). When we used them, we were able to achieve a minimum average coverage of 8000× on the sample with 0.86% huDNA. However, the percentage of unmapped reads for the three other samples was very high (> 75%). We concluded that 1% huDNA is the minimum requirement to process a sample via NGS sequencing using hybrid capture. Overall, 279/285 (97.89%) samples were suitable for NGS and subsequent bioinformatic analysis.

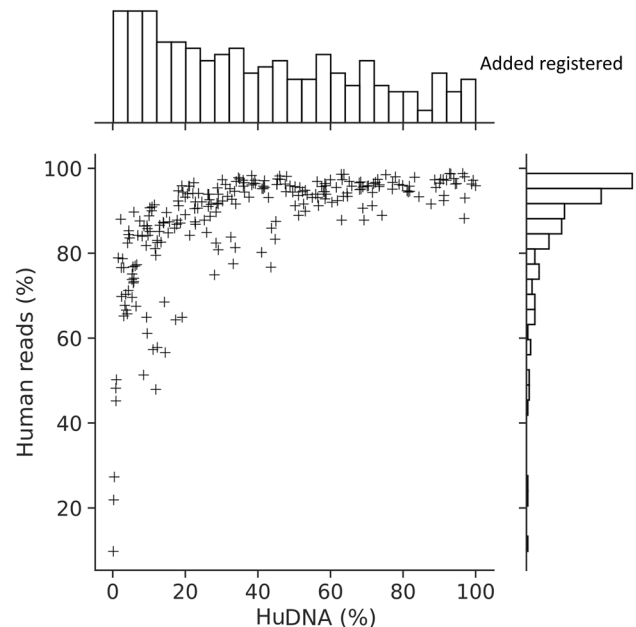


Fig. 4 Correlation between the human reads from the NGS analysis and huDNA percentage. *huDNA* human deoxyribonucleic acid, *NGS* next-generation sequencing

Discussion

In this internal pilot study of the first 300 patients recruited to the ORI-EGI-02 Study, 279/300 (93%) patients successfully underwent a test, had their sample transferred, the human DNA amplified and NGS sequencing successfully performed. The reasons for failure were protocol deviation in 12 (4%), device failure in 2 (0.67%), and faecal contamination in 6 (2%) cases. During this internal study, a manufacturing fault was identified in the double seal, which holds the balloon to the base of the device, leading to loss of the balloon in the rectum of two patients. Neither patient came to harm with the balloon being passed at the next bowel movement. This led to a review and alteration of the balloon seal pressure testing procedure by the manufacturer. The manufacturer's recommendations state the requirement of test inflation by the sampler to check the OriCol™ Sampling Device. In the two cases of device failure, this was not apparent during test inflation at the time of sampling. All episodes of device failure were recorded as adverse events and were reported to the Site Monitor and Origin Sciences Limited within 24 h. All devices from this batch were recalled, and with the modification of the double seal pressure ring, no further device failure occurred in the remainder of the pilot study.

Overall, the performance of the test resulted in a low level of discomfort to the patient with only one patient unable to tolerate a digital rectal examination, which is a contraindication to performing the OriCol™ test. Digital rectal examination is a requirement of the test, and like endoscopic examination, a painful or incomplete rectal examination in a patient should lead to alternative form of assessment, e.g., examination under anaesthesia to assess the anal canal for pathology.

The Rockwood Frailty Score [23] data also support the use of the test in patients who are unfit, frail or multiply co-morbid and whom would not be suitable for colonoscopic assessment. The potential to retrieve genetic material for analysis, to identify significant pathology such as a colorectal cancer allows very focused investigation such as CT or flexible sigmoidoscopy in patients whom curative options are excluded due to their poor health.

The focus of the complete study is to identify human DNA in rectal mucus. The majority of colorectal disease has a luminal origin with material shed into the faecal stream, huDNA in rectal mucus could offer an alternative to cfDNA in blood. The DNA extracted in this 300-participant cohort was not affected by recent bowel movement, however, the amount of non-human DNA in the sample does impact on the level of huDNA and separate bacterial analysis may be performed from the sample taken—a future potential area of microbiome research and disease.

However, even with low levels of huDNA, NGS sequencing was successful thus demonstrating the robust nature of the retrieved mucus sample for analysis. A level of > 1% huDNA allowed successful sequencing for bioinformatic analysis. Overall, seven samples had huDNA below 1% and yet had a DNA concentration of more than 5000 ng/ml of mucus demonstrating that those samples are highly contaminated with non-human DNA. The sequencing of these samples using hybrid capture produced good-quality human sequences, yet most of the reads (> 50%) did not align to the human genome. We, therefore, set the limit of 1% as the minimum percentage of huDNA to provide a sufficient number of reads for tertiary analyses of gene mutations for the remainder of the complete study.

Conclusions

The potential use of the material extracted from the rectal mucus is to demonstrate DNA transferred through the colorectum from proximal disease, allowing recognition of known cancer mutations and thus triage for interventional colonoscopy. Successful use of the OriCol™ sampling device by different users and transfer of the material demonstrated that retrieval of the material allowed high rates of NGS analysis paving the way to further interrogation of key genes involved in colorectal cancer pathway and other bowel diseases. Overall, the high rate of recruitment, sample retrieval and safe transfer and low rate of sample loss from contamination will allow the study to proceed to full recruitment and complete sequencing of samples for genomic bioinformatic analysis is on-going.

Acknowledgements **Members of ORI-EGI-02 Study Group:** F. D. McDermott: Royal Devon University Healthcare NHS Foundation Trust, Barrack Road, Exeter, Devon, EX2 5DW; E. T. Walker: Royal Devon University Healthcare NHS Foundation Trust, Barrack Road, Exeter, Devon, EX2 5DW; N. J. Battersby: Royal Cornwall Hospital NHS Trust, Treliske, Truro, TR1 3LJ; H. Sisodia: Royal Cornwall Hospital NHS Trust, Treliske, Truro, TR1 3LJ; H. Rottenburg: Royal Cornwall Hospital NHS Trust, Treliske, Truro, TR1 3LJ; C. Cunningham: Oxford University Hospitals NHS Trust, Headley Way, Headington, Oxford OX3 9DU; S. Bird: Oxford University Hospitals NHS Trust, Headley Way, Headington, Oxford OX3 9DU; G. A. R. Jones: Shrewsbury & Telford NHS Trust, Shrewsbury, Shropshire, SY3 8XQ; D. Wise: Origin Sciences, Granta Park, Cambridge, CB21 6AD; S. J. Spencer: Origin Sciences, Granta Park, Cambridge, CB21 6AD.

Author contribution All authors contributed to the study conception and design. Clinical data collection was undertaken by HN Humphrey, J Lacey-Coulson, FD McDermott, ET Walker, NJ Battersby, H Sisodia, H Rottenburg, C Cunningham, S Bird, GAR Jones. Laboratory and data analysis was performed by A Diodato, JC Isner, B Nedjai, D Wise, E Walker, SJ Spencer and IR Daniels. The first draft of the manuscript was written by HN Humphrey and A Diodato, these authors contributed equally. All authors and contributors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding The ORI-EGI-02 study was funded by Origin Sciences Ltd.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest Financial Interests ORI-EGI-02 Study Group: Author HN Humphrey is a surgical research fellow funded by an unrestricted research grant from Origin Sciences. Author A Diodato is employed by Origin Sciences. Author J-C Isner is employed by Origin Sciences. Author E Walker is employed by Origin Sciences. Author J Lacey-Coulson holds Origin Sciences share options. Author IR Daniels holds Origin Sciences share options and is a paid consultant to Origin Sciences. Author B Nedjai holds Origin Sciences share options and is a paid consultant to Origin Sciences. ET Walker is a surgical research fellow funded by an unrestricted research grant from Origin Sciences. H Sisodia is a surgical research fellow funded by an unrestricted research grant from Origin Sciences. H Rottenberg is a surgical research fellow funded by an unrestricted research grant from Origin Sciences. S Bird is a research nurse funded by an unrestricted research grant from Origin Sciences. GAR Jones is a surgical research fellow funded by an unrestricted research grant from Origin Sciences. D Wise is employed by Origin Sciences. SJ Spencer holds Origin Sciences share options and is contracted to Origin Sciences as Clinical Study Manager.

FD McDermott has no financial disclosures relevant to this study. NJ Battersby has no financial disclosures relevant to this study. C Cunningham has no financial disclosures relevant to this study.

Ethical approval for research involving human participants All procedures performed the ORI-EGI-02 study involving human participants were in accordance with the ethical standards of the national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the Health Research Authority/National Research Ethics Service/ East Midlands-Nottingham 1 Research Ethics Committee (REC Reference: 19/EM/0266). The study was completed in accordance with STROBE guidelines.

Ethics Application Health Research Authority/National Research Ethics Service/East Midlands-Nottingham 1 Research Ethics Committee. *REC reference:* 19/EM/0266; *Protocol number:* Ori-EGI-02; *IRAS project ID:* 263745; *ClinicalTrials.gov:* NCT04659590.

Consent to participate Written informed consent was obtained from all individual participants included in the study.

Consent to publish Written informed consent obtained from all individual participants included a signed statement regarding permission to publish.

References

- National Institute for Health and Care Excellence, *Suspected cancer recognition and referral: site or type of cancer* (2021 <https://pathways.nice.org.uk/pathways/suspected-cancer-recognition-and-referral/suspected-cancer-recognition-and-referral-site-or-type-of-cancer#content=view-node%3Anodes-gastrointestinal-tract-lower-cancers>).
- Cancer Services* (2022 <https://fingertips.phe.org.uk/profile/cancerservices/data#page/1/gid/1938133085/pat/159/par/K0200001/ati/15/are/E92000001/yr/1/cid/4/tbm/1>).
- Kidney E et al (2015) Preliminary results of a feasibility study of the use of information technology for identification of suspected colorectal cancer in primary care: the CREDIBLE study. *Br J Cancer* 112:S70–S76
- Thompson M et al (2020) Have large increases in fast track referrals improved bowel cancer outcomes in UK? *BMJ* 371:m3273
- Flashman BEK, OLeary D, Senapati A, Thompson MR (2015) OC-099 The two-week clinic for bowel cancer isn't working; a re-audit of referrals twelve years on. *Gut*. <https://doi.org/10.1136/gutjnl-2015-309861.99>
- PH England, in Information on the programme, including commissioning, quality assurance, education and training. (gov.uk, 2015), vol. NHS bowel cancer screening (BCSP) programme.
- Public Health England, *Cancer services profiles; short statistical commentary, December 2018* (2018 <https://www.gov.uk/government/statistics/cancer-services-profiles-2018-annual-update/cancer-services-profiles-short-statistical-commentary-december-2018#further-information>).
- Round T, Gildea C, Ashworth M, Møller H (2020) Association between use of urgent suspected cancer referral and mortality and stage at diagnosis: a 5 year national cohort study. *Br J Gen Pract* 70:e389–e398
- Turvill JL et al (2021) The fast track FIT study: diagnostic accuracy of faecal immunochemical test for haemoglobin in patients with suspected colorectal cancer. *Br J Gen Pract* 71:e643–e651
- NICE, *Quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care* (2017 <https://www.nice.org.uk/guidance/dg30>).
- MacLeod C et al (2021) ScotCap—A large observational cohort study. *Colorectal Dis* 24:411
- Chan RH et al (2021) Clinical utility of a cell-free DNA assay in patients with colorectal cancer. *Front Oncol* 11:589673
- Zhang Y et al (2021) Pan-cancer circulating tumor DNA detection in over 10,000 Chinese patients. *Nat Commun*. <https://doi.org/10.1038/s41467-020-20162-8>
- Imperiale TF et al (2014) Multitarget stool DNA testing for colorectal-cancer screening. *N Engl J Med* 370:1287–1297
- Anderson N et al (2011) Protein biomarkers in exfoliated cells collected from the human rectal mucosa: implications for colorectal disease detection and monitoring. *Int J Colorectal Dis* 26:1287–1297
- Loktionov A et al (2009) Colorectal cancer detection by measuring DNA from exfoliated colonocytes obtained by direct contact with rectal mucosa. *Int J Oncol* 34:301–311
- Pelaseyed T et al (2014) The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunol Rev* 260:8–20
- NIHR, *Pilot and Feasibility Studies* (2022 <https://www.journalslibrary.nihr.ac.uk/information-for-authors/report-contents/report-types/pilot-and-feasibility-studies.htm>).
- V. Md, S. Misra, H. Li, S. Aluru. 2019 Efficient Architecture-Aware Acceleration of BWA-MEM for Multicores Systems. arXiv preprint arXiv. <https://doi.org/10.48550/arXiv.1907.12931>
- Danecek P et al (2021) Twelve years of SAMtools and BCFtools. *GigaScience*. <https://doi.org/10.1093/gigascience/giab008>
- Python. (2022).
- Von Elm E et al (2007) The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 370:1453–1457

23. Rockwood K (2005) A global clinical measure of fitness and frailty in elderly people. *Can Med Assoc J* 173:489–495

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.