Use of multigene-panel identifies pathogenic variants in several CRC-predisposing genes in patients previously tested for Lynch Syndrome

Maren F. Hansen1,2 | Jostein Johansen3 | Anna E. Sylvander2 | Inga Bjørnevoll2 | Bente A. Talseth-Palmer1,4,5 | Liss A. S. Lavik2 | Alexandre Xavier4 | Lars F. Engebretsen2 | Rodney J. Scott4,6 | Finn Drabløs3 | Wenche Sjursen1,2

Background: Many families with a high burden of colorectal cancer fulfill the clinical criteria for Lynch Syndrome. However, in about half of these families, no germline mutation in the mismatch repair genes known to be associated with this disease can be identified. The aim of this study was to find the genetic cause for the increased colorectal cancer risk in these unsolved cases.

Materials and methods: To reach the aim, we designed a gene panel targeting 112 previously known or candidate colorectal cancer susceptibility genes to screen 274 patient samples for mutations. Mutations were validated by Sanger sequencing and, where possible, segregation analysis was performed.

Results: We identified 73 interesting variants, of whom 17 were pathogenic and 19 were variants of unknown clinical significance in well-established cancer susceptibility genes. In addition, 37 potentially pathogenic variants in candidate colorectal cancer susceptibility genes were detected.

Conclusion: In conclusion, we found a promising DNA variant in more than 25% of the patients, which shows that gene panel testing is a more effective method to identify germline variants in CRC patients compared to a single gene approach.

KEYWORDS colorectal cancer, diagnostics, gene panel testing, inherited cancer, Lynch syndrome, next generation sequencing (NGS)

1 | INTRODUCTION

Colorectal cancer (CRC) is one of the most common cancers in the world with approximately 1.3 million new cases diagnosed each year, and is a significant cause of cancer mortality.1 Inherited factors are estimated to be involved in the development of one third of CRC cases. However, Mendelian CRC syndromes only explain about 5% of these cases.2 These syndromes are caused by mutations or epimutations in well-known cancer susceptibility genes that include MLH1, PMS2, MSH2, MSH6, EPCAM, APC, SMAD4, BMPR1A, STK11, MUTYH, PTEN, KLLN, PIK3CA, AKT1, POLE, POLD1, AXIN2, BUB1 and BUB3. Mutations in high penetrance genes such as TP53 and CDH1 resulting in other cancer aggregations reveal ambiguous results in terms of their association with colorectal cancer risk.3,4 Four other genes, ATM, CHEK2, MLH3, and EXO1 (all associated with some aspect of DNA repair), have been implicated in CRC susceptibility.5–8
ATM and CHEK2 are increasingly being recognized as moderate pene-
trance genes primarily associated with an increased risk of breast
cancer, but they have also been associated with CRC.5,7 The involve-
ment of MLH3 and EXO1 in CRC is still disputed and if any effect at
all, they are more likely to modify the risk of other high penetrant
genes.8,9 Previous low-throughput sequencing studies aimed at inves-
tigating genes potentially involved in CRC susceptibility have identi-

cified candidates like GALNT12 and PTPRJ.9,10 However, these studies
have not been replicated in additional independent cohorts and these
genes require further validation before being included in the clinical
management of CRC patients.

CRC is also considered as a complex disease, and low penetrant
variants together with environmental factors are likely to be asso-
ciated with the missing heritability apparent for the disease. Genome-
wide association studies (GWASs) have identified at least 31 common
low-penetrant genetic variants associated with CRC susceptibility
(reviewed in11). One GWAS has revealed that common variants in
BMP4 influence CRC risk12 which has been supported by a study that
has potentially identified pathogenic germline mutations in BMP4 in
early onset CRC patients with a family history of cancer.13 It is there-
fore possible that rare coding variants in genes identified by GWAS
can cause hereditary CRC.

Recent advances in sequencing technology have aided a high-
throughput approach in the search for new genes involved in hereditary
CRC. Four recent exome sequencing studies have identified several potential predisposition alleles.14–17 However, these studies only
implicate potential candidates and require verification before these genes
can be considered bona fide hereditary colorectal cancer genes.

In some families there is a clustering of CRC, which is suggestive
of a hereditary predisposition. These families typically fulfill the Amster-
dam I/II criteria (AM I/II) and/or the revised Bethesda guidelines (RBG),
which were devised to help identify patients with Lynch Syndrome
(LS) (MIM #609310, #120435, #614350, #614337)18,19 In this study,
we included 274 patients who fulfilled the AM I/II criteria and/or the
RBG. The patients had previously been referred for clinical genetic
testing of 1 or more of the MMR genes (MLH1, PMS2, MSH2, MSH6),
but no germline mutations were identified. The aim of this study was
to find the genetic cause for the increased CRC risk in these unsolved
cases, by using a gene-panel targeting 112 previously known or can-
didate CRC susceptibility genes.

2 | MATERIALS AND METHODS

2.1 | Samples

This study included DNA samples from 274 (82 Norwegian and 192 -
Australian) familial CRC patients. Some of the individuals were related
and altogether there were 8 families with 2 to 3 family members each
(19 individuals). All patients fulfilled AM I/II and/or RBG and had pre-
viously been screened for mutations in 1 or more of the MMR genes
(MLH1, PMS2, MSH2 and MSH6) without any pathogenic findings
(80 of the Norwegian samples were also screened by MLPA). Some
patients were also tested for other CRC-susceptibility genes, again
without any pathogenic germline mutations being identified. Table 1
shows the clinical characteristics of the patients included in the study.
The Norwegian samples were screened for mutations as part of their
standard patient healthcare, and all genetic testing was performed
only after written informed consent from the participants. The Aus-

stralian patients included in the study had previously given
informed consent for their de-identified DNA and clinical records to
be used in research related to their condition. Ethics approval was
obtained from the Hunter New England Human Research Ethics
Committee and the University of Newcastle’s Human Research Ethics
Committee. DNA was isolated from EDTA-preserved whole blood
using iPrep PureLink gDNA Blood kit (Thermo Fisher Scientific, Wal-
tham, Massachusetts) (Norwegian samples) or the salt precipitation
method20 (Australian samples).

2.2 | Gene panel sequencing

We designed a custom HaloPlex (Agilent Technologies, Santa Clara,
California) gene panel targeting 112 genes (Table S1, Supporting
information) including both well-known CRC genes and candidate
CRC susceptibility genes. The design was generated using the webt-
ool SureDesign (Agilent Technologies). Target enrichment was per-
formed according to manufacturer’s protocol. Briefly, the samples
were quantified on Qubit 2.0 Fluorometer (Life Technologies, Carls-
bad, California) using dsDNA BR Assay Kit (Life Technologies). DNA
was fragmented by restriction digestion, hybridized to HaloPlex
probes containing indexes and purified using magnetic beads. Frag-
ments were then ligated and amplified through 18 PCR cycles. Each
library was quantified on Agilent 2100 Bioanalyzer (Agilent Technolo-
gies) using the High Sensitivity DNA kit (Agilent Technologies) and
finally equimolarly pooled into sequencing ready libraries. The Nor-
wegian samples were sequenced using an Illumina HiSeq 2500 using
HiSeq Rapid SBS kit v2 (200 cycles) (Illumina, San Diego, CA). The
Australian samples were sequenced on a NextSeq (Illumina) using
NextSeq 500 High Output Kit (300 cycles).

2.3 | Data analysis

Analysis of sequencing data was performed as previously described
21, with only minor variation. PCR duplicates were not removed from
datasets due to the use of restriction enzymes in the HaloPlex
library preparation, leading to non-random fragmentation. Removing
PCR duplicates at this step can lead to removal of ~90% of reads.22
The variant caller used was HaploTypeCaller. For filtering variants we
used the filtering tool FILTUS version 0.99-91.23

2.4 | Filtering of variants

Our aim was to detect potentially pathogenic variants and therefore
our filtering strategy aimed at removing neutral variants and sequen-
cing errors. First, we selected variants tagged as ‘PASS’ after quality
control, present in 1000 Genomes Project with MAF <0.01 and with
a sequencing depth >10. To remove systematic sequencing errors
and variants common in the patients included in this study, we
excluded all variants detected in ≥10 individuals in these datasets
(if over 10 individuals carry a specific variant it can be regarded as
TABLE 1  Clinical characteristics of the patients included in this study

| Nationality | Total cohort (N = 274) |
|-------------|------------------------|
| Norwegian   | 82                     |
| Australian  | 192                    |
| Female      | 183                    |
| Male        | 91                     |
| Median age at first cancer\(^a\) | 51.5 (21-86) |
| Cancer history\(^b\) |                     |
| CRC         | 229                    |
| Other cancers\(^c\) | 28                  |
| Only adenomas | 14                   |
| Multiple primary cancers\(^d\) | 64                  |
| Amsterdam criteria |                 |
| Positive    | 262                    |
| Negative\(^e\) | 12                   |
| Microsatellite instability status\(^f\) |                  |
| MSS         | 38                     |
| MSI-L       | 6                      |
| MSI-H       | 27                     |
| IHC\(^g\)   |                        |
| Loss of MMR protein staining | 83                |
| Normal staining | 56                 |

Abbreviations: CRC, colorectal cancer; RBG, revised Bethesda guidelines; MSS, Microsatellite stable; MSI-L, Microsatellite instability low; MSI-H, Microsatellite instability high; MMR, mismatch repair.

\(^a\) Data missing for 6 patients.
\(^b\) Data missing for 3 patients.
\(^c\) Cancer in locations other than colon and rectum.
\(^d\) Patients with more than 1 case of cancer, regardless of location.
\(^e\) AM negative patients were RBG positive.
\(^f\) Only available for the Norwegian patients. Data missing for 203 patients.
\(^g\) Data available for 68 Norwegian and 71 Australian samples. Data missing for 135 patients.

common and therefore not likely to be pathogenic). Further, we included non-synonymous, splice-site and frameshift variants. The selected non-synonymous variants were located in conserved regions based on phastCons score, predicted to be at conserved sites by PhyloP and to be deleterious by SIFT, Polyphen2, LRT and MutationTaster. We also included all frameshift and splice-site variants. Following is a brief explanation of the thresholds used to define what is conserved: Annovar uses UCSC phastCons 46 species alignment to annotate variants that fall within conserved regions. It assigns a score ranging from 0 to 1000. The higher score, the more conserved. We selected all variants with any score. In addition, we used PhyloP for base level conservation scores where a score >0.95 is conserved.

The next steps in the filtering process was to review bam files to discover and remove artifacts and variant interpretation to only select variants most likely to be pathogenic. Variant interpretation was performed utilizing Alamut software (Interactive Biosoftware, Rouen, France) and evaluating the available literature. Detected variants were classified into 5 classes according to the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) guidelines.\(^24\)

2.5  Validation and segregation analysis by Sanger sequencing

Sanger sequencing was used to confirm detected variants remaining after applying filtering steps described above and to test for detected variants in additional family members. Sanger sequencing was done as previously described.\(^25\) The variants confirmed were submitted to Leiden Open Variation Database 3.0 (http://databases.lovd.nl/shared/genes).

3  RESULTS

3.1  Filtering results

The 95 Norwegian samples had a mean coverage of: 256.03. The 192 Australian samples had a mean coverage of: 320.26. This is per base coverage in the targeted sequenced regions. Prior to filtering we identified 13783 unique variants in the 274 samples, and after in silico filtering 148 unique variants remained. Manual filtering and interpretation to remove artefacts and to select variants most likely to be causal left 92 unique variants. Validation by Sanger sequencing confirmed 73 variants. Of these, 37 were found in known CRC susceptibility genes (Tables 2 and 3). The other 36 variants were found in candidate genes, where the association to CRC is yet to be clarified (Table 4). The 19 variants not confirmed by Sanger sequencing were mostly false positive frameshift variants, due to the remaining adapter sequences. All but 1 of the patients with Sanger validated variants fulfilled the Amsterdam criteria.

3.2  Pathogenic variants in known CRC susceptibility genes

We found 17 pathogenic variants in 21 samples (Table 2). Of these, there were 4 mono-allelic MUTYH mutation carriers and 1 mono-allelic BLM mutation carrier. The mono-allelic BLM mutation carrier did not fulfill the Amsterdam criteria. One patient (no. 203) was bi-allelic for MUTYH mutation (NM_001128425; c.1187G>A and c.1227_1228dup). When excluding the mono-allelic MUTYH and BLM mutation carriers, we found a most probable genetic explanation for the increased cancer risk in 16 (6%) of the patients’ families using this multigene panel.

We identified 3 pathogenic (class 5) variants in the MMR genes MLH1 and MSH6 in 3 patient samples. The MSH6 (NM_000179.2) variant, c.3261dup (p.Phe1088Leufs*5) had previously been identified in a diagnostic setting and was included as a positive control. The 2 other samples were originally classed as mutation negative for the MMR genes.

Two patients had pathogenic mutations in ATM, which is known to be a moderate penetrance gene that confers an increased risk of breast cancer. Both patients had a personal and family history of CRC, and 1 of the patients (no. 154) had breast cancer in the family. The ATM variant c.8584+2T>C (NM_000051.3) was also tested, but
| Sample ID | Gene       | Ref. seq.     | DNA             | Protein          | dbSNP138      | ExAC (ALL) | 1000 genomes | Class | Affected<sup>a</sup> | Unaffected<sup>a</sup> | ClinVar ID and interpretation |
|-----------|------------|---------------|------------------|------------------|---------------|------------|--------------|-------|----------------------|------------------------|--------------------------------|
| 154       | ATM        | NM_000051.3   | c.8494C>T       | p.Arg2832Cys     | rs587779872   | 8.24e-06   | 5            |       |                      |                        | 127459: P/LP                   |
| 34        | ATM        | NM_000051.3   | c.8584+2T>C     | p.?              | rs730881326   | NR         | 4            |       |                      |                        | 181899: P/LP                   |
| 112       | AXIN2      | NM_004655.3   | c.1987dup       | p.Trp663Leufs<sup>*</sup>44 | NR            | NR         | 5            |       |                      |                        |                               |
| 82        | BLM        | NM_000057.3   | c.2824-2A>T     | p.?              | rs745538883   | 1.65e-05   | 4            |       |                      |                        | 371621: LP                    |
| 7         | BRCA1      | NM_007294.3   | c.4096+3A>G     | p.?              | rs80358015    | NR         | 4            |       |                      |                        |                               |
| 157       | BRCA2      | NM_000059.3   | c.4415_4418del  | p.Lys1472Thrfs<sup>*</sup>6 | rs748716604   | NR         | 5            |       |                      |                        | 37902: P                      |
| 164       | BRCA2      | NM_000059.3   | c.2808_2811del  | p.Ala938Profs+21 | rs80359351    | 1.65e-05   | 5            |       |                      |                        | 9322: P                       |
| 291       | CHEK2      | NM_007194.3   | c.1100del       | p.Thr367Metfs<sup>*</sup>15 | 0.0018        | 0.0009984  | 5            |       |                      |                        | RCV000123265: P               |
| 116       | MLH1       | NM_000249.3   | c.2103+1G>T     | p.?              | rs267607888   | NR         | 5            |       |                      |                        | RCV000075531: LP             |
| 183       | MSH6       | NM_000179.2   | c.2079dup       | p.Cys694Metfs<sup>*</sup>4 | rs267608083   | NR         | 5            |       |                      |                        | RCV000210176: P              |
| 41        | MSH6       | NM_000179.2   | c.3261dup       | p.Phe1088Leufs<sup>*</sup>5 | rs748452299   | 0.0018     | 5            | 2 of 3 | 0 of 6               |                        | 89364: P                      |
| 135, 203<sup>b</sup>, 230, 245 | MUTYH    | NM_001128425.1 | c.1187G>A      | p.Gly396Asp     | rs36053993    | 0.0028     | 0.00239617  | 5     |                      |                        | 5294: P                       |
| 203<sup>b</sup> | MUTYH    | NM_001128425.1 | c.1227_1228dup  | p.Glu410Glyfs<sup>*</sup>43 | rs587780078   | 0.0001     | 5            |       |                      |                        | 127831: P                     |
| 186       | MUTYH      | NM_001128425.1 | c.934+2A>G     | p.?              | rs77542170    | 0.0010     | 0.0029951   | 5     |                      |                        | 41766: LP                     |
| 4, 27, 28<sup>c</sup> | POLE    | NM_006231.3   | c.1373A>T       | p.Tyr458Phe     | NR           | NR         | 5            |       |                      |                        | ref. 21                       |
| 42        | POLE       | NM_006231.3   | c.824A>T        | p.Asp275Val     | NR           | NR         | 5            | 1 of 1 | 0 of 1               |                        | NR                            |
| 33        | PTEN       | NM_000314.4   | c.377C>T        | p.Ala126Val     | NR           | NR         | 4            |       |                      |                        | ref 26, 27                    |

Abbreviations: P, pathogenic; LP, likely pathogenic; NR, not reported.

<sup>a</sup> Variant found in additional affected and unaffected individuals from the same family.

<sup>b</sup> Patient 203 has 2 pathogenic mutations in the MUTYH gene.

<sup>c</sup> Patient 4, 27 and 28 belong to the same family.
| Sample ID | Gene | Ref.seq. | DNA         | Protein       | dbSNP138   | ExAC       | ESP         | 1000 genomes | Affected\(^a\) | Unaffected\(^a\) | ClinVar ID and interpretation |
|-----------|------|----------|-------------|---------------|------------|------------|-------------|--------------|----------------|----------------|--------------------------------|
| 213       | APC  | NM_000038.5 | c.6136G>A    | p.Ala2046Thr  | rs770406711 | 1.65e-05   |             |              |                |                | 185089: US                     |
| 256       | APC  | NM_000038.5 | c.1139G>A    | p.Arg380Gln   | rs587782886  | 5.79e-05   |             |              |                |                | 143004: LB/US                   |
| 45        | BLM  | NM_000057.3 | c.2983T>C    | p.Tyr995His   | rs142723411  | NR         | 0.000077    |              |                |                |                                |
| 175       | BRCA2| NM_000059.3 | c.714_716dup | p.Glu238_Ser239insArg | rs80359640 | NR         |              |              |                |                | 126202: US                     |
| 249       | BUB1 | NM_001278616.1 | c.2458A>G    | p.Arg820Gly   | NR         | NR         |              |              |                |                |                                |
| 127       | FANCD2| NM_001018115.1 | c.3269C>T    | p.Ala1090Val  | NR         | NR         |              |              |                |                |                                |
| 73        | FLCN | NM_144997.5  | c.1508G>C    | p.Cys503Ser   | rs778904029  | 1.65e-05   |             |              |                |                |                                |
| 83        | FLCN | NM_144997.5  | c.1523A>G    | p.Lys508Arg   | rs199643834  | 0.0002     | 0.000030    |              |                |                | 41856: LB/US                   |
| 250       | MLH1 | NM_000249.3  | c.514G>A     | p.Glu172Lys   | NR         | NR         |              |              |                |                | RCV000075700: US               |
| 9         | MSH2 | NM_000251.2  | c.138C>G     | p.His46Gln    | rs33946261  | 0.0003     |              | 0 out of 3   | 3 out of 9     |                | 90654: US                      |
| 281       | MSH2 | NM_000251.2  | c.1045C>G    | p.Pro349Ala   | rs267607939  | 9.06e-05   | 0.000077    |              |                |                | 90512: US                      |
| 169       | MSH6 | NM_000179.2  | c.1282A>G    | p.Lys428Glu   | rs761822293  | 8.24e-06   |             |              |                |                |                                |
| 242       | PIK3CA| NM_006218.2 | c.1729A>G    | p.Arg577Gly   | NR         | NR         |              |              |                |                |                                |
| 24        | PMS2 | NM_0000535.5 | c.1004A>G    | p.Asn335Ser   | rs200513014  | 0.0003     |              | 0 out of 1    |                |                | 127751: US                     |
| 3, 21, 37\(^b\) | POLE | NM_006231.3  | c.229C>T     | p.Arg77Cys    | NR         | NR         |              | 1 out of 1    |                |                |                                |
| 147       | POLE | NM_006231.3  | c.844C>T     | p.Pro282Ser   | rs138207610  | 0.0001     | 0.000231    | 0.000399361   |                |                | RCV000229770: US               |
| 172       | POLE | NM_006231.3  | c.4168C>T    | p.Arg1390Cys  | rs768504121  | 1.65e-05   |             |              |                |                | 246319: US                     |
| 29        | PTEN | NM_000314.4  | c.-491_-486del | p.?         | NR         | NR         |              |              |                |                |                                |
| 45, 74    | PTEN | NM_000314.4  | c.-488_-486del | p.?         | NR         | NR         |              |              |                |                |                                |

Abbreviations: US, uncertain significance; NR, not reported; LB, likely benign.

\(^a\) Variant found in additional affected and unaffected individuals from the same family.

\(^b\) Patient 4, 27 and 28 belong to the same family.
| Sample ID | Gene     | Ref. seq. | DNA       | Protein               | ExAC      | dbSNP138 | ESP       | 1000 genomes | Affected* | Unaffected* | ClinVar ID and interpretation |
|-----------|----------|-----------|-----------|-----------------------|-----------|----------|-----------|--------------|-----------|-------------|--------------------------------|
| 204  | AXIN1    | NM_003502.3 | c.497G>T  | p.Ser166Ile            | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 190  | BMP4     | NM_001202.3 | c.250C>T  | p.Arg84Trp             | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 174  | CCDC18   | NM_206886.4 | c.366_3663del | p.Leu121Glnfs*23       | NR        | rs761268563 | NR        | NR           | NR        | NR          |                                |
| 21   | DCC      | NM_005215.3 | c.1817C>G  | p.Pro606Arg            | 1.647e-05 | rs773588703 | NR        | NR           | NR        | NR          |                                |
| 164  | DCC      | NM_005215.3 | c.3370C>T  | p.Arg1124Cys            | 0.00016   | rs547920182 | 0.00019968 | NR           | NR        | NR          |                                |
| 194  | DCC      | NM_005215.3 | c.4028G>A  | p.Arg1343His            | 0.00012   | rs149118168 | 0.000308  | NR           | NR        | NR          |                                |
| 60, 131 | DCLRE1A  | NM_014881.3 | c.412C>T  | p.Arg138               | 0.0028    | rs41292634 | 0.002384  | 0.00199681  | 0 out of 1 | 1 out of 2 |                                |
| 113  | DUSP4    | NM_001394.6 | c.824G>A  | p.Arg275His             | 6.88e-05  | rs372203752 | 0.000077  | NR           | NR        | NR          |                                |
| 66   | FAM166A  | NM_001001710.1 | c.41C>T  | p.Pro14Leu             | 5.06e-05  | rs140737708 | 0.000077  | NR           | NR        | NR          |                                |
| 146  | HELQ     | NM_133636.2 | c.2225G>T  | p.Cys742Phe             | 8.29e-06  | rs374570294 | 0.000077  | NR           | NR        | NR          |                                |
| 79   | LAMA3    | NM_198129.2 | c.8693A>G  | p.Gln298Ser             | 7.413e-05 | rs77988893 | NR        | NR           | NR        | NR          |                                |
| 213  | LAMA3    | NM_198129.2 | c.3712dup | p.Tyr1238Leufs*3        | 0.0001    | rs758832093 | NR        | NR           | NR        | NR          |                                |
| 276  | LAMA3    | NM_198129.2 | c.1273+2_1273+41del    | p.Arg142Cys | 0.0003    | rs751342772 | 0.0008    | NR           | NR        | NR          |                                |
| 223  | LAMA5    | NM_005560.3 | c.3964G>A  | p.Glu132Ser             | 0.00035   | rs150741810 | 0.000389  | NR           | NR        | NR          |                                |
| 136  | LAMB4    | NM_007356.2 | c.2468G>A  | p.Gly823Glu             | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 249  | LAMB4    | NM_007356.2 | c.1525G>C  | p.Asp509His             | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 76   | LAMC1    | NM_00293.3  | c.2426A>G  | p.Asp809Gly             | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 259  | LAMC1    | NM_00293.3  | c.1088A>G  | p.His363Arg             | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 9    | MAML3    | ENST0000059479.3 | c.1139C>T | p.Ser380Phe             | 0.0003    | rs200202141 | 0.000724  | 0.00019968  | 0 out of 1 | 0 out of 1 |                                |
| 14   | MLH3     | NM_001040108.1 | c.885del | p.His296Thrfs*12       | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 195  | MRPL3    | NM_007208.3 | c.506G>T  | p.Gly691Val             | 0.000124  | rs369657581 | 0.000384  | 0.00019968  | NR        | NR          |                                |
| 97   | MYH11    | NM_002474.2 | c.4603C>T  | p.Arg1535Trp            | 0.00012   | rs143402648 | 0.000077  | 0.00019968  | NR        | NR          | 372423: US |
| 149, 262 | NUDT7    | NM_001105663.1 | c.178C>T | p.Arg60Trp             | 0.00021   | rs199760367 | 0.000336  | 0.00019968  | NR        | NR          |                                |
| 185  | NUDT7    | NM_001105663.1 | c.272G>A  | p.Arg91Gln             | 0.00012   | rs768311455 | NR        | NR           | NR        | NR          |                                |
| 276  | PICALM   | NM_001008660.2 | c.130T>A  | p.Try44Asn              | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 97, 167 | PSHH    | NM_004577.3 | c.115G>A  | p.Gly39Ser             | 0.00089   | rs147075540 | 0.000769  | NR           | NR        | NR          |                                |
| 123  | PTPRJ    | NM_002843.3 | c.3678_3679del | p.Gln1293Leufs*28    | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 141  | PTPRJ    | NM_002843.3 | c.3793G>A  | p.Val1265Met            | 2.47e-05  | rs550632588 | 0.00019968 | NR           | NR        | NR          |                                |
| 295  | PTPRJ    | NM_002843.3 | c.3208C>A  | p.Arg1070Ser            | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 275  | PTPRJ    | NM_002843.3 | c.1085del | p.Phe362Serfs*14       | NR        | NR       | NR        | NR           | NR        | NR          |                                |
| 110  | SLC5A9   | NM_001011547.2 | c.1475del | p.Gly492Alafs*13      | 0.00037   | rs777247762 | NR        | NR           | NR        | NR          |                                |

(Continues)
not found, in a maternal cousin with 3 synchronous cancers and multiple polyps. The unaffected mother of the index patient has now been tested, and did not harbour the ATM variant. Therefore, the cousin might have another predisposing genetic variant leading to his high cancer burden.

One patient diagnosed with CRC at age 65 had a frameshift mutation in **AXIN2**. This patient is deceased, but abnormal dentition was reported, consistent with Oligodontia-colorectal cancer syndrome (MIM #608615).

One patient had a mutation in **BRCA1** (no. 7) and 2 individuals in **BRCA2** (no. 157 and 164). These 3 female patients were affected with early onset CRC. Two of them (nos 7 and 164) had a family history of CRC, breast and ovarian cancer, whereas the third (no. 157) had no family history of breast or ovarian cancer.

Two unique pathogenic variants were detected in 4 patients in **POLE** (NM_006231.3). In 3 of these patients a pathogenic **POLE** mutation c.1373A>T (p.Tyr458Phe) previously reported by Hansen et al21 was observed. These individuals are all related and belong to the previously reported family.21 Variant c.824A>T (p.Asp275Val) was identified in individual no. 42 affected with bilateral ovarian cancer at age 37. She was included in this study because of lack of blood sample from her deceased mother. The mother was affected with endometrial cancer at age 49 and CRC at age 88, and the **POLE** variant (c.824A>T) was detected in paraffin-embedded tissue sample from her surgery. This variant is previously found as a somatic change in endometrial cancer25, but not as a germline variant. Asp275 forms the exonuclease catalytic site of **POLE** and is involved in binding of metal ions important for exonuclease activity.

We found 1 **PTEN** (NM_000314.4) variant c.377C>T (p.Ala126Val) in a patient diagnosed with 4 metachronous tumours (CRC, clear cell renal carcinoma, thymoma and parathyroid adenoma), some of which overlap with the tumour spectrum of Cowden Syndrome (MIM #158350). CRC was the first cancer, diagnosed at 46 years of age. The **PTEN** missense variant is within a highly conserved catalytic domain, and it is reported to give rise to completely inactive protein.26,27

The **CHEK2** (NM_007194.3) variant (c.1100del, p.Thr367Metfs*15) was found in a patient who was diagnosed with CRC at age 37. This **CHEK2** variant is a well described, lower penetrant mutation, mainly associated with breast cancer, but also CRC and prostate cancer.28,29

### 3.3 Variants of unknown significance (VUS) in known CRC susceptibility genes

A total of 19 variants of unknown clinical significance were detected in 21 samples in known cancer susceptibility genes, and some of these may also prove to be pathogenic (Table 3).

**MLH1** variant c.514G>A (p.Glu172Lys) was found in a patient diagnosed with CRC at age 51 who has several family members affected with CRC. Residue Glu172 is highly conserved and located in the ATPase domain of **MLH1**, although not at the ATP binding site. This variant has previously been observed 3 times in the COSMIC database. Two times as a somatic change in breast and endometrial cancer and once in a cell culture from the large intestine. A **MSH6**
variant c.1282A>G (p.Lys428Glu) was found in a patient diagnosed with cancer at age 41 with a family history of CRC and uterine cancer. Lys428 is highly conserved and located in the MutS I domain. The variant has not been previously reported.

The POLE variant, c.229C>T (p.Arg77Cys), was identified in 3 affected individuals from the same family and in 1 obligate carrier. All 4 family members had early onset CRC and 1 had polyposis. Most of the previously identified pathogenic mutations in POLE are found in the DNA binding sites within the exonuclease domain.\textsuperscript{21,30,31} POLE p.Arg77fs conserved (up to \textit{S. cerevisiae}), and there is a large physicochemical difference between Arg and Cys (Grantham distance 180). However, it is not located in any exonuclease domain or at an active site, thus further investigation is needed in order to decide whether it is a causative variant.

A BUB1 (NM_001278616.1) variant c.2458A>G (p.Arg820Gly) was found in a patient affected with CRC at age 42. Residue Arg820 is highly conserved and located in the protein kinase catalytic domain of BUB1. The mutant residue potentially disturbs the domain and is predicted to abolish its function. Although, the physicochemical difference between Arg (positively charged) and Gly (no charge) is moderate (Grantham distance 125), the difference in size, hydrophobicity and charge between the wild-type and mutant residue is predicted to disturb hydrogen bonds (Cys891 and Asp932) and ionic interactions (salt bridges) (Glu819, Glu892 and Asp932) between residue 820 and these other internal residues. The loss of charge can also cause loss of interaction with other molecules.\textsuperscript{32} The mutation is therefore likely to affect the function of the protein.

PIK3CA (NM_006218.2) VUS c.1729A>G (p.Arg577Gly) was found in a patient diagnosed with CRC at age 58 and 3 metachronous melanomas. Arg577 is highly conserved, it is predicted to be pathogenic by 6 prediction programs (PolyPhen, SIFT, MutationTaster, Align GVD, SNPs3D and UMD Predictor), and it located in the PIK domain which has been suggested to be involved in substrate presentation. As described above for the BUB1 mutation, the physicochemical difference between Arg and Gly is moderate (Grantham distance 125). However, this change is predicted to disturb ionic interactions (salt bridges) between PIK3CA residue 577 and Aspartic acid at position 395 and 578, indicating an effect on the protein’s function.\textsuperscript{32}

Two PTEN variants c.-491_-486del (p.Arg163fs*) and c.-488_-486del are located in 5’ UTR (or exon 1 in transcript NM_001304717) at a binding site for RNA Polymerase II. Detecting mutations in this region in 3 unrelated Norwegian individuals suggests that these variants are common in the Norwegian population. However, because these patients are highly selected the 2 PTEN variants may be pathogenic if they disrupt RNA Polymerase II binding, but this needs further investigation.

The variants in Table 3 with reported minor allele frequencies are less likely to be pathogenic, except for that identified in BLM, which is associated with recessive disease. In addition, segregation analysis of the MSH2 variant c.138C>G (p.His46Gln) and PMS2 c.1004A>G (p.Asn335Ser) does not support pathogenicity. However, PMS2 is found to have much lower penetrance for CRC than the other MMR genes, and therefore mutations may not always be associated with disease.\textsuperscript{33} For the remaining variants listed in Table 3, there is no further information indicating whether they are pathogenic or benign.

### 3.4 Variants in candidate CRC genes

We identified 37 unique variants in 36 different patients in candidate genes that have a potential role in CRC susceptibility (Table 4). There was no evidence of autosomal recessive disease identified in this dataset. Variants with a reported allele frequency are less likely to cause a highly penetrant disorder, although moderately penetrant disorders are possible but more difficult to identify. Laminins are essential components of connective tissue basement membranes and influence cell differentiation, migration, and adhesion. Laminin is vital for the maintenance and survival of tissues and defective laminins can lead to the autosomal recessive disorders such as congenital muscular dystrophy (MIM #607855), junctional epidermolysis bullosa (MIM #226700 and #226650) and Pierson Syndrome (MIM #609049).\textsuperscript{34} We identified 8 variants in laminin genes; LAMA3, LAMA5, LAMB4 and LAMC1. Based on Laminins function, these variants are not the most probable candidates to play a role in CRC susceptibility.

Segregation analysis was only possible for the variants DCLRE1A (NM_014881.3) c.412C>T (p.Arg138*), MAML3 (ENST00000509479.3) c.1139C>T (p.Ser380Phe) and MLH3 ( NM_0014041081.3) c.885del (p.His296Thrfs*12) due to the availability of samples from additional family members. However, none of these variants seemed to segregate with disease. The MLH3 variant has previously been found in 2 CRC patients, 1 endometrial cancer patient and 1 unaffected below the age of 75 in a family,\textsuperscript{35} suggesting the variant has reduced penetrance. They further suggested MLH3 to be a low risk gene for CRC. DCC variant c.1817C>G (p.Pro606Arg) identified in patient no. 21 was not found in 2 affected family members (nos 3 and 37) who also were included in this study. Instead, these 3 family members all had the POLE VUS c.229C>T described above. Another DCC variant, c.3370C>T (p.Arg1124Cys), was identified in patient no. 164 who also has a pathogenic BRCA1 mutation. Consequently, these 2 DCC variants are not likely to be associated with a predisposition to CRC.

The remaining 14 variants in the genes AXIN1, BMP4, CCDC18, NUDT7, PICALM, PTEN, SLC5A9, TLR2, TWSG1, UBA2, USP6NL and ZFP14 have a potential role in CRC susceptibility (marked bold in the table). Of these, the missense variants in AXIN1, BMP4, NUDT7, PICALM, PTEN, TLR2, TWSG1, USP6NL and ZFP14 are located in protein functional domains and the residue (Arg91) affected in NUDT7 is a putative active site. Four variants in CCDC18, PTEN and SLC5A9 are frameshift variants. The most interesting candidates are the 2 frameshift and the missense variant (marked bold) in the PTEN gene. Epigenetic silencing of this gene due to an inherited duplication in a CRC family has previously been reported\textsuperscript{36} suggesting that this may be a new CRC susceptibility gene. The 2 frameshift mutations are predicted to disrupt the function of this gene and the missense variant alters a highly conserved amino acid involved in 2 functional domains (PTP type protein phosphatase and protein-tyrosine phosphatase-like). All the patients with PTEN alterations in this study were diagnosed with CRC above the age of 50 years and have several family members affected with CRC. Unfortunately no samples from additional family members were available at this stage.
In this study, we found several pathogenic or likely pathogenic (class 4-5) variants in known cancer susceptibility genes, which validates our approach for identifying disease causing variants. Some of the VUS’s revealed in this study may also prove to be pathogenic, as more becomes known about the functional impact of these variants.

Three variants in MLH1 and MSH6 as well as a number of variants of unknown significance (VUS) were identified in our sample set. The most likely explanation for this finding is the accuracy of some of the screening protocols that were used to identify variants in known MMR genes. Using high-throughput screening approaches that are significantly more accurate than previous methodologies it is to be expected some additional mutations in these genes will come to light. We recommend that samples screened by methodologies that do not employ direct DNA sequencing be re-evaluated by better more cost-effective and accurate assays.

The phenotype of hereditary cancer syndromes often overlap, because of the pleiotropy of cancer genes. For example in LS a wide spectrum of cancer types are associated with mutations in MMR genes, like ovary cancer. Increased risk of ovary cancer is also associated with mutations in BRCA1 and BRCA2. The spectra of cancer types associated with each cancer syndrome are not always totally determined either. Whether breast cancer is a part of the LS spectrum is not always totally determined either. Whether breast cancer is a part of the LS spectrum is not meaning that all patients with breast cancer have LS. However, breast cancer is a common feature of LS patients.

Another advantage by using a broader gene panel testing approach is that it may reveal whether there is more than 1 pathogenic candidate CRC susceptibility genes which should be the subject of further research to determine their involvement in CRC risk. Overall, the results show that gene panel sequencing is a more effective method by which to identify pathogenic germline variants in CRC patients compared with a single gene approach.

**ACKNOWLEDGMENTS**

The sequencing service was provided by the Genomics Core Facility, Norwegian University of Science and Technology (NTNU). The bioinformatics analyses were performed at the Bioinformatics core facility, Norwegian University of Science and Technology (NTNU). This work was supported by grants from the Liaison Committee between the Central Norway Regional Health Authority (RHA) and the Norwegian University of Science and Technology (NTNU); and travel support for scientific meetings and courses from COST Action BM1206.

**Conflict of interest**

The authors declare no conflict of interest.

**ORCID**

Maren F. Hansen http://orcid.org/0000-0001-5231-4034

Wenche Sjursen http://orcid.org/0000-0003-3880-6440
REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136:E359-E386.

2. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. Gastroenterology. 2010;138:2044-2058.

3. Wong P, Versells J, Garber JE, et al. Prevalence of early onset colorectal cancer in 397 patients with classic Li-Fraumeni syndrome. Gastroenterology. 2006;130:73-79.

4. Richards FM, McKee SA, Rajap M, et al. Germline E-cadherin gene (CDH1) mutations predispose to familial gastric cancer and colorectal cancer. Hum Mol Genet. 1999;8:607-610.

5. Thompson D, Duedal S, Kinner J, et al. Cancer risk and mortality in heterozygous ATM mutation carriers. J Natl Cancer Inst. 2005;97:813-822.

6. Liberti SE, Rasmussen LJ. Is heXO1 a cancer predisposing gene? Mol Cancer Res. 2004;2:427-432.

7. Xiang HP, Geng XP, Ge WW, Li H. Meta-analysis of CHEK2 1100delC variant and colorectal cancer susceptibility. Eur J Cancer. 2011;47:2546-2551.

8. Kim JC, Roh SA, Yoon YS, Kim HC, Park J. MLH3 and EXO1 alterations in familial colorectal cancer patients not fulfilling Amsterdam criteria. Cancer Genet Cytogenet. 2007;176:172-174.

9. Guda K, Moinova H, He J, et al. Inactivating germ-line and somatic mutations in polyadenylate N-acetylgalactosaminyltransferase 12 in human colon cancers. Proc Natl Acad Sci U S A. 2009;106:12921-12925.

10. Venkataraman R, Lichtenberg MJ, Hoogerbrugge N, et al. Germline epigenetic silencing of the tumor suppressor gene PTENL1 in early-onset familial colorectal cancer. Gastroenterology. 2010;139:2221-2224.

11. Esteban-Jurado C, Garre P, Vila M, et al. New genes emerging for colorectal cancer predisposition. World J Gastroenterol. 2014;20:1961-1971.

12. Tomlinson IP, Carvajal-Carmona LG, Dobbins SE, et al. Multiple common susceptibility variants near BMP pathway loci GREM1, BMP4, and BMP2 explain part of the missing heritability of colorectal cancer. PLoS Genet. 2011;7:e1002105.

13. Lubbe SJ, Pittman AM, Matjussen C, et al. Evaluation of germline BMP4 mutation as a cause of colorectal cancer. Hum Mutat. 2011;32:E1928-E1938.

14. Gyfle AE, Katainen R, Kondelin J, et al. Eleven candidate susceptibility genes for common familial colorectal cancer. PLoS Genet. 2013;9:e1003876.

15. Smith CG, Naven M, Harris R, et al. Exome resequencing identifies potential tumor-suppressor genes that predispose to colorectal cancer. Hum Mutat. 2013;34:1026-1034.

16. DeRycke MS, Gunawardena SR, Middha S, et al. Identification of novel variants in colorectal cancer families by high-throughput exome sequencing. Cancer Epidemiol Biomarkers Prev. 2013;22:1239-1251.

17. Esteban-Jurado C, Vila-Casadesus M, Garre P, et al. Whole-exome sequencing identifies rare pathogenic variants in new predisposition genes for familial colorectal cancer. Genet Med. 2015;17:131-142.

18. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst. 2004;96:261-268.

19. Vansen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPPC, Lynch syndrome) proposed by the International Collaborative group on HNPPC. Gastroenterology. 1999;116:1453-1456.

20. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16:1215.

21. Hansen MF, Johansen J, Bjornvoll I, et al. A novel POL2 mutation associated with cancers of colon, pancreas, ovaries and small intestine. Fam Cancer. 2015;14:437-448.

22. Coonrod EM, Durschi JD, VanSant Webb C, Voelkerding KV, Kumanovics A. Next-generation sequencing of custom amplicons to improve coverage of HaloPlex multigene panels. Biotechniques. 2014;57:204-207.

23. Vigeland MD, Gjotterud KS, Selmer KF. FTLUS: a desktop GUI for fast and efficient detection of disease-causing variants, including a novel autogyosity detector. Bioinformatics (Oxford, England). 2016;32:1592-1594.

24. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405-424.

25. Briggs S, Tomlinson I. Germline and somatic polymerase epsilon and delta mutations define a new class of hypermutated colorectal and endometrial cancers. J Pathol. 2013;230:148-153.

26. Rodriguez-Escudero I, Oliver MD, Andres-Pons A, Molina M, Cid VJ, Pulido R. A comprehensive functional analysis of PTEN mutations: implications in tumor- and autism-related syndromes. Hum Mol Genet. 2011;20:4132-4142.

27. Costa HA, Leitner MG, Sol ML, et al. Discovery and functional characterization of a neomorphic PTEN mutation. Proc Natl Acad Sci USA. 2015;112:13976-13981.

28. Gronwald J, Cybulski C, Piesiak W, et al. Cancer risks in first-degree relatives of CHEK2 mutation carriers: effects of mutation type and cancer site in proband. Br J Cancer. 2009;100:1508-1512.

29. Meijers-Heijboer H, Wijnjens J, Vansen H, et al. The CHEK2 1100delC mutation identifies families with a hereditary breast and colorectal cancer phenotype. Am J Hum Genet. 2003;72:1308-1314.

30. Palles C, Caizer JB, Howarth KM, et al. Germline mutations affecting the proofreading domains of POLE and POL1 predispose to colorectal adenomas and carcinomas. Nat Genet. 2013;45:136-144.

31. Rohlin A, Zagoras T, Nilsson S, et al. A mutation in POLE predisposing to colorectal cancers is associated with a tumor speed phenotype. J Oncol. 2014;45:77-81.

32. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inherited diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinf. 2010;11:548.

33. Moller P, Seppala T, Bernstein I, et al. Cancer incidence and survival in Lynch syndrome patients receiving colonoscopy and gynaecological surveillance: first report from the prospective Lynch syndrome database. Gut. 2015;66:446-472.

34. Yurchenco PD, Patton BL. Developmental and pathogenic mechanisms of basement membrane assembly. Curr Pharm Des. 2009;15:1277-1294.

35. Liu HX, Zhou XL, Liu T, et al. The role of hMLH3 in familial colorectal cancer. Cancer Res. 2003;63:1894-1899.

36. Sopik V, Phelan C, Cybulski C, Narod SA, BRCA1 and BRCA2 mutations and the risk for colorectal cancer. Clin Genet. 2015;87:411-418.

37. Mensenkamp AR, Vogelaar IP, van Zelst-Stams WA, et al. Somatic mutations in MLH1 and MSH2 are a frequent cause of mismatch-repair deficiency in Lynch syndrome-like tumors. Gastroenterology. 2014;146:643-646.e648.

38. Grindeel EM, Aarset H, Bjornvoll I, et al. The Norwegian PM52 founder mutation c.989-1G>T shows high penetrance of microsatellite instable cancers with normal immunohistochemistry. Hered Cancer Clin Pract. 2014;12:12.

39. Sjursen W, Hauknes BI, Grindeel EM, et al. Current clinical criteria for Lynch syndrome are not sensitive enough to identify MSH6 mutation carriers. J Med Genet. 2010;47:579-585.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Hansen MF, Johansen J, Sylvangler AE, et al. Use of multigene-panel identifies pathogenic variants in several CRC-predisposing genes in patients previously tested for Lynch Syndrome. Clin Genet. 2017;92:405-414. https://doi.org/10.1111/cge.12994