



## EDITORIAL

# Moving the needle on colorectal cancer genetics: it takes more than two to TANGO

British Journal of Cancer <https://doi.org/10.1038/s41416-018-0219-2>

### SUMMARY

Colorectal cancer (CRC) is a leading cause of cancer incidence and mortality. It is well known that an important fraction of CRC risk is accounted by genetic factors. Identifying and characterising such factors may aid in identifying high-risk individuals who will benefit from effective preventive interventions.

### PREVENTATIVE MEASURES

Colorectal cancer (CRC) is a leading global cause of cancer incidence and mortality. After lung and breast cancer, it is the third most commonly diagnosed malignancy, with 1.36 million new patients diagnosed annually.<sup>1</sup> Despite this large public health burden, most CRCs could be prevented through screening for pre-malignant lesions (polyps).<sup>2</sup> Several CRC prevention programmes are already in place in countries with socialised medical systems, and these have resulted in a shift towards higher incidence of early-stage CRCs and a concomitant improvement in survival outcomes.<sup>3–5</sup> Preventive interventions could benefit further from tools to identify high-risk groups. About a third of the CRC risk variance is attributable to inherited factors,<sup>6</sup> and hence, genetic-based strategies offer great opportunities to identify high-risk individuals who could benefit from prevention and early detection.

The past decade has seen several advances in our understanding of the genetic aetiology of CRC. Through approaches that include linkage analysis, which involves genetic studies of disease co-segregation among affected relatives, and genome-wide association studies (GWAS), which involve comparing gene frequency data between CRC patients and healthy controls, 14 highly penetrant genes and >40 low-penetrance variants have so far been identified for CRC.<sup>7,8</sup> These genes and variants only explain around 14% of the inherited risk of CRC<sup>9</sup> and with most of these remaining relatively poorly understood. There is thus a clear need for additional discovery and functional studies to allow translation of these genetic findings into improved prevention strategies. In this issue of the *British Journal of Cancer*, four independent studies provide new data that furthers our understanding of CRC genetics. Two of these studies focus on novel gene discovery and the mechanistic understanding of highly penetrant CRC forms, whereas the other two papers focus on GWAS and post-GWAS approaches to study CRC aetiology and survival.

### NEW GENES AND NEW MECHANISMS

Evidence for a new risk variant on chromosome 1q32 was provided by Schubert et al.,<sup>10</sup> building on previous studies that demonstrated the association of a nearby region (1q41) with the risk of CRC and multiple adenomas.<sup>11,12</sup> The authors used two

independent gene identification approaches (homozygosity mapping and linkage analysis) to identify a rare non-synonymous single-nucleotide variant (nsSNP, p.Asp1432Glu) in the *MIA3* gene (also known as *TANGO*, a gene not previously known to be associated with CRC). Co-segregation studies together with tumour gene and protein expression analyses provided supporting data for a role of *TANGO* p.Asp1432Glu in CRC tumourigenesis. Further independent studies are warranted to confirm the causal role of this *TANGO* variant in CRC.

The second study focussed on understanding the mechanisms by which epigenetic mutations lead to CRC in Lynch Syndrome (LS), the most common CRC-predisposing syndrome.<sup>13</sup> Estela Dámaso and collaborators<sup>14</sup> investigated the mechanism underlying constitutional primary *MLH1* epimutations. Individuals with these epimutations represent around 2% of all mutation-negative cases suspected of LS and typically have severe CRC manifestations.<sup>15</sup> In this condition, *MLH1* epimutations typically arise de novo and lead to the soma-wide inactivation of the *MLH1* allele by methylation, which in turns leads to CRC. *MLH1*, a DNA mismatch repair gene, is also methylated in sporadic CRC where it explains most CRCs with microsatellite instability. Yet the manifestation is different between sporadic CRC and constitutional primary *MLH1* epimutation, with the sporadic manifestation exhibiting a colorectal tissue-specific alteration, rather than soma-wide. In their study, 12 constitutional primary *MLH1* epimutation carriers along with 61 LS patients and 41 controls were analysed with genome-wide methylation arrays to identify differentially methylated regions (DMRs) between epimutation and non-epimutation carriers. This analysis found that the only DMR was a CpG island encompassing the *MLH1* and *EPM2AIP1* genes, suggesting that constitutional primary *MLH1* epimutation-driven CRCs are different from sporadic *MLH1* methylated CRCs, where they exhibit a focal, rather than genome-wide, methylation pattern.

Analysing the heritability of *MLH1* epimutations, the authors revealed that the epimutations were neither inherited nor passed to descendants, because they were not detected in affected relatives despite harbouring the epimutation-bearing haplotype. This latter result is important, because it shows that constitutional primary *MLH1* epimutations experience inter-generation erasure. Such information is important for genetic counselling of *MLH1* epimutation carriers and for our understanding of CRC tumourigenesis in patients with this condition.

### GENETIC AETIOLOGY, FUNCTIONAL STUDIES AND STUDYING THE GENETICS OF CRC SURVIVAL

Most GWAS-identified variants map to non-coding regions, and they likely increase risk by regulating transcript levels of nearby genes.<sup>8</sup> Identifying genes whose transcripts are under genetic

Received: 3 July 2018 Revised: 3 July 2018 Accepted: 11 July 2018  
Published online: 04 October 2018

control (called eGenes) is thus useful to discover new CRC variants and to understand the mechanisms by which GWAS alleles increase CRC risk. In order to link genetic variation with gene expression patterns, Moreno et al.<sup>16</sup> discovered expression quantitative trait loci (eQTLs) in colon tissue from paired colon tumour/adjacent normal samples and from normal colonic biopsies obtained from healthy donors. The study identified 363 colonic eGenes, many of which overlapped with those in GTEx (a publicly available eQTL database) or with those identified in a previous study.<sup>17</sup> Interestingly, of the 37,099 eQTLs discovered in this study, only 4,858 were identified both in normal and tumour tissue, raising the possibility that some of these might be good candidates for CRC tumorigenesis studies. These data are now publicly available through a dedicated website and they represent a great resource in post-GWAS studies of CRC. The final study was aimed at identifying nsSNPs affecting CRC survival.<sup>18</sup> Despite having sufficient statistical power in this GWAS study to detect nsSNPs with modest effects on survival, the study failed to find genome-wide significant associations, suggesting that survival may be influenced by rarer nsSNPs or by common non-coding variants. Future studies, therefore, should not only increase sample size to enable the detection of milder effects but should also include additional types of genetic variation that may have effects on prognosis.

The studies published in this issue help move the needle towards an improved understanding of the role of genetic factors in CRC aetiology and survival. They show that multiple approaches are needed to identify and characterise CRC genes and suggest that future studies, particularly those focussing on complex phenotypes such as survival, will require collaborative efforts by the international research community.

#### DISCLAIMER


The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

#### ACKNOWLEDGEMENTS

L.G.C.-C. receives funding from the University of California Davis School of Medicine (Dean's Fellowship in Precision Medicine) and from the National Cancer Institute (P30CA093373) and National Institute of Environmental Health Sciences (P30ES023513) of the National Institutes of Health.

#### ADDITIONAL INFORMATION

**Competing interests:** The author declares no competing interests.

Luis G. Carvajal-Carmona <sup>1,2,3</sup>

<sup>1</sup>Genome Center, University of California, Davis, CA, USA; <sup>2</sup>UC Davis Comprehensive Cancer Center, School of Medicine, University of California, Davis, CA, USA and <sup>3</sup>Department of Biochemistry and Molecular Medicine, School of Medicine, University of California, Davis, CA, USA

Correspondence: Luis G. Carvajal-Carmona (lgcarvajal@ucdavis.edu)

#### REFERENCES

1. Ferlay, J. et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136**, E359–386 (2015).
2. Zauber, A. G. et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N. Engl. J. Med.* **366**, 687–696 (2012).
3. Logan, R. F. et al. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. *Gut* **61**, 1439–1446 (2012).
4. Morris, E. J. et al. A retrospective observational study examining the characteristics and outcomes of tumours diagnosed within and without of the English NHS Bowel Cancer Screening Programme. *Br. J. Cancer* **107**, 757–764 (2012).
5. Cole, S. R. et al. Shift to earlier stage at diagnosis as a consequence of the National Bowel Cancer Screening Program. *Med. J. Aust.* **198**, 327–330 (2013).
6. Lichtenstein, P. et al. Environmental and heritable factors in the causation of cancer — analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* **343**, 78–85 (2000).
7. Carvajal Carmona, L. G. & Tomlinson, I. The hunting of the snark: whither genome-wide association studies for colorectal cancer? *Gastroenterology* **150**, 1528–1530 (2016).
8. Sud, A., Kinnersley, B. & Houlston, R. S. Genome-wide association studies of cancer: current insights and future perspectives. *Nat. Rev. Cancer* **17**, 692–704 (2017).
9. Jiao, S. et al. Estimating the heritability of colorectal cancer. *Hum. Mol. Genet.* **23**, 3898–3905 (2014).
10. Schubert, S. A. et al. Evidence for genetic association between chromosome 1q loci and predisposition to colorectal neoplasia. *Br. J. Cancer* **117**, 1215–1223 (2017).
11. Montazeri, Z. et al. Systematic meta-analyses and field synopsis of genetic association studies in colorectal adenomas. *Int. J. Epidemiol.* **45**, 186–205 (2016).
12. Spain, S. L. et al. Refinement of the associations between risk of colorectal cancer and polymorphisms on chromosomes 1q41 and 12q13.13. *Hum. Mol. Genet.* **21**, 934–946 (2012).
13. Lynch, H. T. et al. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin. Genet.* **76**, 1–18 (2009).
14. Dámaso, E. et al. Primary constitutional mlh1 epimutations: a focal epigenetic event. *Br. J. Cancer* (2018).
15. Hitchins, M. P. Finding the needle in a haystack: identification of cases of Lynch syndrome with MLH1 epimutation. *Fam. Cancer* **15**, 413–422 (2016).
16. Moreno, V., et al. Colon-specific eQTL analysis to inform on functional SNPs. *Br. J. Cancer* (2018).
17. Ongen, H. et al. Putative cis-regulatory drivers in colorectal cancer. *Nature* **512**, 87–90 (2014).
18. Theodoratou, E., et al. Genome-wide scan of the effect of common nsSNPs on colorectal cancer survival outcome. *Br. J. Cancer* (2018).