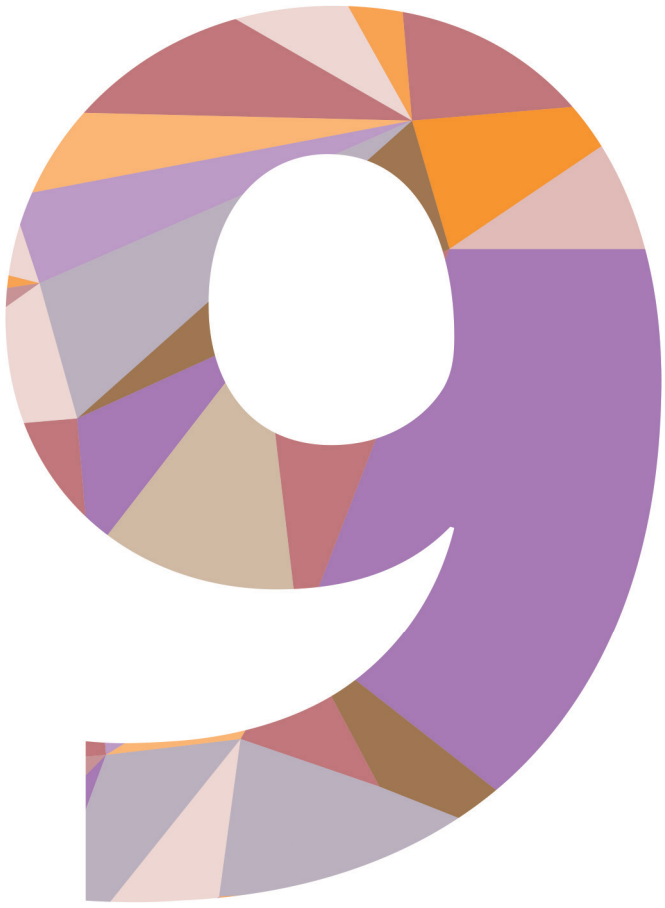
Exploring APC mosaicism: prevalence, clinical consequences and underlying causes

Terlouw, D.



General discussion

The aim of this thesis was, firstly, to evaluate the proportion of *APC* and *MUTYH* pathogenic variants in colorectal polyposis patients and subsequently identify the proportion of unexplained polyposis patients (Part I, **chapter 2**). Furthermore, three studies aimed to elucidate the significance of *APC* mosaicism and suggest testing and surveillance guidelines (Part II; **chapters 3-5**). Lastly, this thesis aimed to assess another explanation for the development of colorectal adenomatous polyps; the presence of *pks*$^{+}$ *E. coli* and colibactin-associated mutational signatures. (Part III, **chapters 6-8**).

Pathogenic germline variant detection rate in polyposis patients

To determine germline pathogenic *APC* and biallelic *MUTYH* variant detection rates in a Dutch cohort, we collected all patients tested in the Leiden University Medical Center between 1992 and 2017 in **chapter 2**. Comparable to most previous studies, a prevalence of 70% for FAP and 7% for MAP in patients with more than 20 adenomas was determined.[1-7] One previously performed study reported lower variant detection rates throughout the entire cohort.[6] This discrepancy could be explained by the clinical differences between the cohorts, such as age of first adenoma development. A unique aspect of our study is the large patient group with less than 20 adenomas, which could be used to evaluate testing guidelines.

Besides number of adenomas developed, the odds of finding a pathogenic germline variant in *APC* or *MUTYH* increased with a younger age of first adenoma diagnosis. A personal history of CRC only increases the odds of finding biallelic *MUTYH* variants. This can likely be explained by the (sub)total colectomy performed at an early age in FAP patients.[7] Lastly, the odds increased upon having a first-degree relative (FDR) with more than 10 adenomas only for *APC*, which is explained by the dominant and recessive inheritance pattern of FAP and MAP respectively.

Based on these findings, testing for germline pathogenic *APC* and *MUTYH* variants is indicated in patients with more than 10 adenomas before the age of 60 years and more than 20 adenomas before the age of 70 years. Other indications for testing are FAP-related extracolonic manifestations, CRC aged <40, a somatic *KRAS* c.34G>T transversion, or a FDR with >10 adenomas. These suggested guidelines are comparable to the Dutch and National Comprehensive Cancer Network (NCCN) guidelines for hereditary colorectal cancer and polyposis.[8, 9] Guidelines issued by the American College of Gastroenterology (ACG), on the other hand, might result in unnecessary testing.[10]

Our cohort also showed an increasing number of patients undergoing genetic testing for *APC* and *MUTYH* over time. This increase might, first of all, be due to the start of *MUTYH* testing in 2004, which led to more patients with milder phenotypes to be tested. Another reason for more genetic testing in polyposis patients is increased adenomas detection rates caused by more sensitive colonoscopy techniques, improved equipment and bowel preparation and introduction of population based screening in the Netherlands.[11-13] This suggests that prevalence of colorectal adenomas in the general population was possibly underestimated and we now gain relevant insight into the actual numbers. Also, modifiable risk factors like diet, alcohol and smoking, attribute to the development of about a third to half of all CRC.[14-16] This so-called Western lifestyle increases throughout both Western and non-Western countries contributing to CRC prevalence.[17] Therefore, a Western lifestyle may also contribute to the increase in colorectal adenomas in the general population.

Moreover, in **chapter 2**, a large proportion of colorectal polyposis patients remain unexplained, no germline pathogenic *APC* or biallelic *MUTYH* variants. The last decades lots of other colorectal cancer and polyposis associated genes were identified.[18-23] Due to increasing amount of genes included in Next Generation Sequencing (NGS) panels, the proportion of unexplained polyposis patients will eventually decrease. Moreover Whole Exome Sequencing (WES), analyzing the entire exosome, is used to find both newly discovered colorectal cancer or polyposis associated genes and to easily re-analyze patients in the future. Also, nowadays, the use of Whole Genome Sequencing (WGS) is more broadly introduced in the clinic, which compared to WES gives insight into possible pathogenic deep intronic variants, large genomic rearrangements or variants in the non-protein-coding sequences like regulatory sequences as promotors and enhancers, untranslated regions or Mitochondrial Iron-Regulated (MIR) genes.[24-27] Also, WES and WGS will provide data on (single nucleotide) polymorphisms which might add up to the risk of developing colorectal polyposis and cancer.[28] In the future, WGS on DNA from neoplastic tissue will provide knowledge about mutational signatures.[29] These signatures might hint towards an underlying (genetic) cause of the developed neoplasm. The broad use of these extensive sequencing techniques will eventually further decrease the prevalence of germline unexplained polyposis patients.

Prevalence of *APC* mosaicism in unexplained polyposis patients

Besides germline pathogenic *APC* and biallelic *MUTYH* variants and variants in other more rare or not yet discovered genes, a significant part of the unexplained polyposis patients are explained by *APC* mosaicism.[30-35] Especially, analysis of DNA isolated from multiple colorectal adenomatous polyps is efficient to detect *APC* mosaicism.[33] To assess the prevalence of *APC* mosaicism in patients with adenomas, we performed targeted NGS on DNA from colorectal adenomas or carcinomas of 458 patients in **chapter 3**. Moreover, this chapter

provides suggestions of *APC* mosaicism testing and surveillance guidelines. A detection rate of about 17% was found in patients falling inside the Dutch hereditary colorectal polyposis and cancer guidelines. This rate is much lower, about 3%, in patients falling outside these guidelines.

Based on the detection rates per phenotypic subgroup, we recommend *APC* mosaicism testing in all patients with (1) adenomas before the age of 50 years, (2) ≥20 adenomas before the age of 60 years or (3) ≥30 adenomas before the age of 70 years.

The broad spectrum of *APC* mosaicism phenotypes complicates an universal surveillance guideline suggestion. Still, in our opinion, *APC* mosaic patients should receive regular colonoscopies, for example every one or two years, comparable to FAP patients.[36] Re-evaluation of the follow-up could be considered in patients with effective polypectomies.

Furthermore, 28% of mosaic patients undergoing a esophagogastroduodenoscopy developed duodenal or gastric neoplasms. In **chapter 5**, we showed that the upper intestinal adenomas all harbored the mosaic variant. We therefore recommend offering at least one gastroduodenoscopy for all *APC* mosaicism patients. In chapter 5 we moreover present a case of duodenal *APC* mosaicism not affecting the colorectum. This shows the possibility of duodenal *APC* mosaicism despite colorectal adenomas and emphasizes the broad spectrum of *APC* mosaicism and its phenotype.

Moreover, children of 13 mosaic patients did not inherit the *APC* variant. Notable, of 10 patients leukocyte, urine and buccal swab was tested and nine showed a mosaicism restricted to the colorectum. Also, the mosaic variant was detected in 15% to 18% in semen DNA tested of a patient with child wish. Therefore, although chances of hereditability are small[33], we still recommend testing children especially in cases with mosaicism detected in other tissues next to the colorectum.

The family presented in **chapter 4** furthermore highlights **t**he significance of *APC* mosaicism in unexplained polyposis patients. Two first-degree relatives have different mosaic *APC* variants with distinct patterns throughout the body and distinct phenotypes. No underlying defect in DNA repair systems or mutational signatures could be identified using WES and WGS respectively.

A formula adapted from Le Caignec et al[37] determined the probability of finding two *APC* different mosaicism cases in one family to be small. Still, this family shows the value of testing for *APC* mosaicism in unexplained polyposis cases even if a FDR has a comparable phenotype.

Although important in genetic diagnostics, there are challenges in testing for (*APC*) mosaicism. In countries other than the Netherlands, in and outside Europe, *APC* mosaicism is underestimated and not regularly tested. One of the main issues are resources for sequencing multiple samples of one patient. Testing normal colorectal mucosa was a hypothesized solution. However, only in 50% of patients the mosaic variant was detected in a normal colorectal tissue sample.

Another challenge are the so-called hybrid mosaic cases were encountered. These cases have a shared variant in multiple but not all analyzed adenomas. Although this underlines necessity of analyzing more than two colorectal adenomas or carcinomas, the clinical impact remains unknown. Multiple possible explanations for hybrid mosaicism are hypothesized. We considered clonal relationship as an explanation whenever two lesions share the same precursor lesion; two adenomas or carcinomas located close to each other and share (multiple) variants.[38] Contamination, mixing two adenomas during polypectomy or mixing DNA samples during isolation or library preparation, was considered whenever multiple (*APC*) variants were shared between two adenomas or carcinomas and one of the samples also have additional (*APC*) variants. Another hypothesis was field cancerization, a mechanism in which normal tissue is replaced by tumor clones with identical *TP53* variants throughout the colon.[33, 39] This is typically described in inflammatory bowel disease in which chronic inflammation leads to crypt fission. A last explanation is that just by chance common *APC* variants occur in two adenomas of the same patient. In conclusion, no universal explanation could be found and case by case evaluation is required.

Interestingly, hybrid mosaic cases are phenotypically comparable to non-mosaic patients and significantly different from mosaic patients. Therefore, we suggest to treat hybrid mosaic cases as non-mosaic patients in surveillance and family testing guidelines for now. Although rare, an exception to this suggestion should be patients with the hybrid variant in normal colon mucosa or other tissues. In these cases, sporadic adenomas possibly developed in a background of *APC* mosaicism.

The prevalence of *APC* mosaicism might suggest a relevant role of mosaicism in other tumor syndromes. No mosaicism in any other gene included our targeted NGS panel was detected but this might be different in cohorts with other phenotypes than adenomatous polyposis. For example, mosaicism of *SMAD4* and *BMPR1A* might be present in unexplained juvenile polyposis patients.[40] Interestingly, *de novo* variant rates for genes like *BMPR1A*, *PTEN*, *SMAD4*, *STK11* and *TP53* are more than 10% of germline patients, suggestive for occurrence of mosaicism.[41-47]

Presence of colibactin as an additional explanation of colorectal adenomas

As described in **chapter 6**, a large proportion of hybrid mosaic patients shared the *APC* splice variant c.835-8A>G in multiple colorectal adenomas or carcinomas. Furthermore, this variant is the most common somatic *APC* variant detected in our cohort. The variant is predicted likely pathogenic as it leads to a premature stop codon and is detected in about 3% of sporadic colorectal carcinomas.[48-50] The c.835-8A>G variant has a (transcriptional) sequence context of AAAATT, where the underlined thymine is substituted by a cytosine, which perfectly fits the colibactin-associated mutational signature.

Colibactin is a genotoxin known to cause DNA crosslinks, double strand breaks and chromosomal aberrations.[51-53] Colibactin-associated mutational signatures are characterized as SBS88 and ID18.[54, 55]

Publicly available datasets showed that colibactin-associated mutational signatures are present in colorectal, head and neck and urinary tract cancer.[54, 55] Interestingly, the mutational signature is detected in normal colonic crypts with a variable mutational burden between individuals and even between crypts, not attributable to age. Using phylogenetics, the mutational signature was proposed to occur early in life.[56] This is supported by in vitro evidence showing genomic alterations after a short-term exposure to colibactin and the colonization of colibactin-encoding *E. coli* happening the first months after birth.[57, 58] The number of affected normal crypts was variable between patients, with some patients having a more affected left colon while in others the entire colon was affected.[56] This might explain why in our cohort, in **chapter 6 and 7**, the c.835-8A>G variant was detected both as a hybrid and 'real' mosaicism. A recent preprint supports our findings and shows that the c.835-8A>G might act as a biomarker for colibactin influence in the development of the adenoma or carcinoma.[59]

In our unexplained polyposis cohort, 110 patients had with at least one somatic *APC* variant fitting SBS88 or ID18. In **chapter 7,** fecal metagenomics and WGS of colorectal adenomas was performed to further assess the influence of colibactin. Fecal metagenomics detected *pks* genes in 25% of negative controls and 59% of patients with colibactin-associated *APC* variants. This is comparable to 19% to 29% of healthy individuals and approximately 60% of FAP and colorectal cancer patients in previous studies.[53, 60-63] Also, WGS showed an enrichment of colibactin-associated mutational signatures in 39% of cases compared to 11.1% of negative controls.

No clear correlation between presence of *pks* in feces and SBS88 and ID18 in colorectal lesions was detected. There are multiple hypotheses for this finding:

Due to the short-term effect of colibactin, affecting the colon early in life, eradication of the bacteria before feces sampling could be an explanation for patients with colibactin-associated signature without *pks* in feces.[56, 57]

Colonization of *pks*$^{+}$ bacteria after developing adenomas is unlikely in patients with *pks* in feces but no colibactin-associated signatures, as *pks*$^{+}$ *E. coli* is proposed to be transmitted during birth.[58, 64] These patients might, however, be able to inhibit colibactin from entering the host cell or protect cells against the DNA damage. For example, the autophagy-related protein ATG16L1 is associated with preventing colorectal tumorigenesis in presence of *pks*$^{+}$ *E. coli* and oxygen is associated with inhibition of colibactin production.[65, 66] On the other hand, oligosaccharides and co-colonization with enterotoxigenic *Bacteroides fragilis* are described to increase the genotoxic effect of colibactin.[60, 67] Further research should be performed to gain more knowledge about which patients are prone to the carcinogenic effect of colibactin.

Technically, especially WGS on Formalin-Fixed Paraffin Embedded tissue samples affects the performance and interpretation of mutational signature analyses due to fragmentation and deamination artefacts.[68-70] Also, complications detecting *pks* in feces due to possible low abundance of *E. coli* could be circumvented using more sensitive techniques like a specific quantitative PCR.

Chapter 8 emphasizes colibactin as a risk factor in hereditary colorectal cancer and polyposis syndromes. A biallelic *NTHL1* patient is described with *pks* in fecal metagenomics and colibactin-associated mutational signature in WGS data. A small cohort of patients with biallelic *MUTYH* variants showed one somatic *APC* variant in one lesion fitting the colibactin-associated mutational signature. Previous described WES of carcinomas of both biallelic as monoallelic *NTHL1* patients did however not show somatic variants suiting the colibactin-associated mutational signature.[71, 72] The *NTHL1* and MAP patient combined with previously described enrichment of *pks*$^{+}$ *E. coli* in FAP patients[60] and our polyposis cohort results described in **chapter 7,** suggest colibactin as an additional risk factor for development of colorectal malignancies in both sporadic colorectal neoplasms and patients with a known predisposition to CRC or polyposis. Future research should elaborate on this association but also on possible inhibition of colibactin or eradication of *pks+ E. coli*. Besides this, new research is set up to determine whether *pks*$^{+}$ *E. coli* could be used as a biomarker to neoadjuvant treatment response showing the increasing interest and implications of gut microbiome on colorectal cancer.[73]

Future perspectives

Future research into unexplained intestinal polyposis patients will be able to use faster, broader and hopefully better analysis methods. The increasing use of NGS might help in minimizing the number of unexplained polyposis patients. With the use of broad DNA sequencing, besides germline variants, somatic mutational signatures can be detected in tumor cells. These mutational signatures might hint towards the underlying known or unknown genetic cause. Additionally, whole genome sequencing can give insights into non-protein-coding sequences which can be used for finding (intronic splice site) variants in known colorectal polyposis-associated genes and possibly lead to the discovery of new candidate genes. Moreover, research into polygenic risk scores, combining pathogenic variants and single nucleotide polymorphisms in multiple genes, will help in delineating the risk of developing colorectal adenomas in individual patients or families.

Furthermore, future research in *APC* mosaicism should focus on explanations for the so-called hybrid mosaic cases. Also, more knowledge is needed about the association between variant allele frequency of the mosaic variants in different tissues or germ layers, phenotype, and risk of transmitting the variant to offspring. Based on the insights presented in this thesis, *APC* mosaicism will hopefully be recognized as an explanation for colorectal polyposis and be included as regular diagnostics in colorectal polyposis patients.

The association of current or past *pks*$^{+}$ *E. coli* (or other bacteria) derived colibactin exposure and having multiple colorectal adenomas should be more elaborately investigated in larger patient cohorts, even at a population level. Furthermore, the possible association with lifestyle factors should be studied. Moreover, future research should focus on the identification of patients carrying the colibactin-producing bacteria, for example via population-based screening programs. Possible inhibition of the DNA damaging effects of colibactin or eradication of the bacteria involved should be explored.

References

1. Kinzler KW, Nilbert MC, Su LK, et al. Identification of FAP locus genes from chromosome 5q21. Science 1991;253:661-5.
2. Powell SM, Petersen GM, Krush AJ, et al. Molecular diagnosis of familial adenomatous polyposis. N Engl J Med 1993;329:1982-7.
3. Miyoshi Y, Ando H, Nagase H, et al. Germ-line mutations of the APC gene in 53 familial adenomatous polyposis patients. Proc Natl Acad Sci U S A 1992;89:4452-6.
4. Sieber OM, Lipton L, Crabtree M, et al. Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. N Engl J Med 2003;348:791-9.
5. Wang L, Baudhuin LM, Boardman LA, et al. MYH mutations in patients with attenuated and classic polyposis and with young-onset colorectal cancer without polyps. Gastroenterology 2004;127:9-16.
6. Stanich PP, Pearlman R, Hinton A, et al. Prevalence of Germline Mutations in Polyposis and Colorectal Cancer-Associated Genes in Patients With Multiple Colorectal Polyps. Clin Gastroenterol Hepatol 2019;17:2008-2015.e3.
7. Grover S, Kastrinos F, Steyerberg EW, et al. Prevalence and phenotypes of APC and MUTYH mutations in patients with multiple colorectal adenomas. Jama 2012;308:485-492.
8. Gupta S, Provenzale D, Regenbogen SE, et al. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment: Colorectal, Version 3.2017. J Natl Compr Canc Netw 2017;15:1465-1475.
9. Tumoren SOE, Nederland VKG, Oncogenetica W. Erfelijke tumoren: Richtlijnen voor diagnostiek en preventie. https://www.stoet.nl/wp-content/uploads/2017/04/STOET-Richtlijnenboekje-april2017_DEF.pdf 2017.
10. Syngal S, Brand RE, Church JM, et al. ACG clinical guideline: Genetic testing and management of hereditary gastrointestinal cancer syndromes. Am J Gastroenterol 2015;110:223-62; quiz 263.
11. Brown SR, Baraza W, Din S, et al. Chromoscopy versus conventional endoscopy for the detection of polyps in the colon and rectum. Cochrane Database Syst Rev 2016;4:Cd006439.
12. Aranda-Hernández J, Hwang J, Kandel G. Seeing better--Evidence based recommendations on optimizing colonoscopy adenoma detection rate. World J Gastroenterol 2016;22:1767-78.
13. Brenner H, Altenhofen L, Kretschmann J, et al. Trends in Adenoma Detection Rates During the First 10 Years of the German Screening Colonoscopy Program. Gastroenterology 2015;149:356-66.e1.
14. Aleksandrova K, Pischon T, Jenab M, et al. Combined impact of healthy lifestyle factors on colorectal cancer: a large European cohort study. BMC Med 2014;12:168.
15. Erdrich J, Zhang X, Giovannucci E, et al. Proportion of colon cancer attributable to lifestyle in a cohort of US women. Cancer Causes Control 2015;26:1271-1279.
16. Vajdic CM, MacInnis RJ, Canfell K, et al. The Future Colorectal Cancer Burden Attributable to Modifiable Behaviors: A Pooled Cohort Study. JNCI Cancer Spectr 2018;2:pky033.
17. Azeem S, Gillani SW, Siddiqui A, et al. Diet and Colorectal Cancer Risk in Asia--a Systematic Review. Asian Pac J Cancer Prev 2015;16:5389-96.
18. Weren RD, Ligtenberg MJ, Kets CM, et al. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. Nat Genet 2015;47:668-71.
19. Valle L, Hernandez-Illan E, Bellido F, et al. New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. Hum Mol Genet 2014;23:3506-12.

20. Gala MK, Mizukami Y, Le LP, et al. Germline mutations in oncogene-induced senescence pathways are associated with multiple sessile serrated adenomas. Gastroenterology 2014;146:520-9.
21. Valle L, de Voer RM, Goldberg Y, et al. Update on genetic predisposition to colorectal cancer and polyposis. Mol Aspects Med 2019;69:10-26.
22. Schubert SA, Morreau H, de Miranda N, et al. The missing heritability of familial colorectal cancer. Mutagenesis 2020;35:221-231.
23. Golubicki M, Bonjoch L, Acuna-Ochoa JG, et al. Germline biallelic Mcm8 variants are associated with early-onset Lynch-like syndrome. JCI Insight 2020;5.
24. Spier I, Horpaopan S, Vogt S, et al. Deep intronic APC mutations explain a substantial proportion of patients with familial or early-onset adenomatous polyposis. Hum Mutat 2012;33:1045-50.
25. Te Paske I, Mensenkamp AR, Neveling K, et al. Noncoding Aberrations in Mismatch Repair Genes Underlie a Substantial Part of the Missing Heritability in Lynch Syndrome. Gastroenterology 2022;163:1691-1694 e7.
26. Makrythanasis P, Antonarakis SE. Pathogenic variants in non-protein-coding sequences. Clin Genet 2013;84:422-8.
27. Perez-Becerril C, Evans DG, Smith MJ. Pathogenic noncoding variants in the neurofibromatosis and schwannomatosis predisposition genes. Hum Mutat 2021;42:1187-1207.
28. Middeldorp A, Jagmohan-Changur S, van Eijk R, et al. Enrichment of low penetrance susceptibility loci in a Dutch familial colorectal cancer cohort. Cancer Epidemiol Biomarkers Prev 2009;18:3062-7.
29. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415-21.
30. Aretz S, Stienen D, Friedrichs N, et al. Somatic APC mosaicism: a frequent cause of familial adenomatous polyposis (FAP). Hum Mutat 2007;28:985-992.
31. Ciavarella M, Miccoli S, Prossomariti A, et al. Somatic APC mosaicism and oligogenic inheritance in genetically unsolved colorectal adenomatous polyposis patients. Eur J Hum Genet 2018;26:387-395.
32. Hes FJ, Nielsen M, Bik EC, et al. Somatic APC mosaicism: an underestimated cause of polyposis coli. Gut 2008;57:71-76.
33. Jansen AML, Crobach S, Geurts-Giele WRR, et al. Distinct Patterns of Somatic Mosaicism in the APC Gene in Neoplasms From Patients With Unexplained Adenomatous Polyposis. Gastroenterology 2017;152:546-+.
34. Out AA, Minderhout IJHMv, Stoep Nvd, et al. High-resolution melting (HRM) re-analysis of a polyposis patients cohort reveals previously undetected heterozygous and mosaic APC gene mutations. 14 2015.
35. Spier I, Drichel D, Kerick M, et al. Low-level APC mutational mosaicism is the underlying cause in a substantial fraction of unexplained colorectal adenomatous polyposis cases. J Med Genet 2016;53:172-9.
36. Yen T, Stanich PP, Axell L, et al. APC-Associated Polyposis Conditions. In: Adam MP, Everman DB, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, Gripp KW, Amemiya A, eds. GeneReviews((R)). Seattle (WA), 1993.
37. Le Caignec C, Kwiatkowski DJ, Kury S, et al. Three independent mutations in the TSC2 gene in a family with tuberous sclerosis. Eur J Hum Genet 2009;17:1165-70.
38. Greaves M, Maley CC. Clonal evolution in cancer. Nature 2012;481:306-13.
39. Galandiuk S, Rodriguez-Justo M, Jeffery R, et al. Field cancerization in the intestinal epitheli-

um of patients with Crohn's ileocolitis. Gastroenterology 2012;142:855-864 e8.

40. Calva-Cerqueira D, Chinnathambi S, Pechman B, et al. The rate of germline mutations and large deletions of SMAD4 and BMPR1A in juvenile polyposis. Clin Genet 2009;75:79-85.
41. Coburn MC, Pricolo VE, DeLuca FG, et al. Malignant potential in intestinal juvenile polyposis syndromes. Ann Surg Oncol 1995;2:386-91.
42. Mester J, Eng C. Estimate of de novo mutation frequency in probands with PTEN hamartoma tumor syndrome. Genet Med 2012;14:819-22.
43. Hernan I, Roig I, Martin B, et al. De novo germline mutation in the serine-threonine kinase STK11/LKB1 gene associated with Peutz-Jeghers syndrome. Clin Genet 2004;66:58-62.
44. Gonzalez KD, Buzin CH, Noltner KA, et al. High frequency of de novo mutations in Li-Fraumeni syndrome. J Med Genet 2009;46:689-93.
45. Gammon A, Jasperson K, Pilarski R, et al. PTEN mosaicism with features of Cowden syndrome. Clin Genet 2013;84:593-5.
46. Jansen AML, Goel A. Mosaicism in Patients With Colorectal Cancer or Polyposis Syndromes: A Systematic Review. Clin Gastroenterol Hepatol 2020;18:1949-1960.
47. McKay V, Cairns D, Gokhale D, et al. First report of somatic mosaicism for mutations in STK11 in four patients with Peutz-Jeghers syndrome. Fam Cancer 2016;15:57-61.
48. Fostira F, Thodi G, Sandaltzopoulos R, et al. Mutational spectrum of APC and genotype-phenotype correlations in Greek FAP patients. BMC Cancer 2010;10:389.
49. Jarry J, Brunet JS, Laframboise R, et al. A survey of APC mutations in Quebec. Fam Cancer 2011;10:659-65.
50. Yaeger R, Chatila WK, Lipsyc MD, et al. Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. Cancer Cell 2018;33:125-136 e3.
51. Nougayrede JP, Homburg S, Taieb F, et al. Escherichia coli induces DNA double-strand breaks in eukaryotic cells. Science 2006;313:848-51.
52. Arthur JC, Perez-Chanona E, Muhlbauer M, et al. Intestinal inflammation targets cancer-inducing activity of the microbiota. Science 2012;338:120-3.
53. Buc E, Dubois D, Sauvanet P, et al. High prevalence of mucosa-associated E. coli producing cyclomodulin and genotoxin in colon cancer. PLoS One 2013;8:e56964.
54. Pleguezuelos-Manzano C, Puschhof J, Huber AR, et al. Mutational signature in colorectal cancer caused by genotoxic pks(+)E. coli. Nature 2020;580:269-+.
55. Boot A, Ng AWT, Chong FT, et al. Characterization of colibactin-associated mutational signature in an Asian oral squamous cell carcinoma and in other mucosal tumor types. Genome Res 2020;30:803-813.
56. Lee-Six H, Olafsson S, Ellis P, et al. The landscape of somatic mutation in normal colorectal epithelial cells. Nature 2019;574:532-537.
57. Iftekhar A, Berger H, Bouznad N, et al. Genomic aberrations after short-term exposure to colibactin-producing E. coli transform primary colon epithelial cells. Nat Commun 2021;12:1003.
58. Tsunematsu Y, Hosomi K, Kunisawa J, et al. Mother-to-infant transmission of the carcinogenic colibactin-producing bacteria. BMC Microbiol 2021;21:235.
59. Georgeson P, Steinfelder RS, Harrison TA, et al. Genotoxic colibactin mutational signature in colorectal cancer is associated with clinicopathological features, specific genomic alterations and better survival. This article is a preprint and has not been certified by peer review. MedRxiv. DOI: https://doi.org/10.1101/2023.03.10.23287127.
60. Dejea CM, Fathi P, Craig JM, et al. Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. Science 2018;359:592-597.

61. Dubinsky V, Dotan I, Gophna U. Carriage of Colibactin-producing Bacteria and Colorectal Cancer Risk. Trends Microbiol 2020;28:874-876.

62. Putze J, Hennequin C, Nougayrede JP, et al. Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae. Infect Immun 2009;77:4696-703.

63. Nooij S, Ducarmon QR, Laros JFJ, et al. Faecal microbiota transplantation influences procarcinogenic Escherichia coli in recipient recurrent Clostridioides difficile patients. Gastroenterology 2021.

64. Payros D, Secher T, Boury M, et al. Maternally acquired genotoxic Escherichia coli alters offspring's intestinal homeostasis. Gut Microbes 2014;5:313-25.

65. Lucas C, Salesse L, Hoang MHT, et al. Autophagy of Intestinal Epithelial Cells Inhibits Colorectal Carcinogenesis Induced by Colibactin-Producing Escherichia coli in Apc(Min/+) Mice. Gastroenterology 2020;158:1373-1388.

66. Nadège Bossuet-Greif N, Guyonnet C, Chagneau CV, et al. Oxygen concentration modulates colibactin production. This article is a preprint and has not been certified by peer review. BioRxiv. DOI: https://doi.org/10.1101/2022.06.20.496773.

67. Oliero M, Calve A, Fragoso G, et al. Oligosaccharides increase the genotoxic effect of colibactin produced by pks+ Escherichia coli strains. BMC Cancer 2021;21:172.

68. de Schaetzen van Brienen L, Larmuseau M, Van der Eecken K, et al. Comparative analysis of somatic variant calling on matched FF and FFPE WGS samples. BMC Med Genomics 2020;13:94.

69. Robbe P, Popitsch N, Knight SJL, et al. Clinical whole-genome sequencing from routine formalin-fixed, paraffin-embedded specimens: pilot study for the 100,000 Genomes Project. Genet Med 2018;20:1196-1205.

70. Do H, Dobrovic A. Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. Clin Chem 2015;61:64-71.

71. Grolleman JE, de Voer RM, Elsayed FA, et al. Mutational Signature Analysis Reveals NTHL1 Deficiency to Cause a Multi-tumor Phenotype. Cancer Cell 2019;35:256-266 e5.

72. Elsayed FA, Grolleman JE, Ragunathan A, et al. Monoallelic NTHL1 Loss-of-Function Variants and Risk of Polyposis and Colorectal Cancer. Gastroenterology 2020;159:2241-2243 e6.

73. Taoum C, Carrier G, Jarlier M, et al. Determination of biomarkers associated with neoadjuvant treatment response focusing on colibactin-producing Escherichia coli in patients with mid or low rectal cancer: a prospective clinical study protocol (MICARE). BMJ Open 2022;12:e061527.