Characterization of candidate genes in unexplained polyposis and colorectal cancer

Abayzeed Elsayed Osman, F.

Chapter 7

Discussion and future perspectives

Discussion

The work described in this thesis aims to determine the underlying genetic causes of polyposis and colorectal cancer (CRC) in unexplained cases by screening known high-risk genes such as *POLE*, *POLD1*, *APC* and *NTHL1*.

POLE and *POLD1*

Palles et al. identified that germline variants affecting the proofreading domains of *POLE* and *POLD1* predispose to colorectal adenomas and carcinomas [1]. *POLE* p.(L424Val) and *POLD1* p.(Ser478Asn) were established as new high-penetrance causes of germline CRC predisposition with an autosomal dominant pattern of inheritance [1]. In **Chapter 2**, we show that germline variants in *POLE* are also associated with early-onset mismatch repair (MMR)-deficient colorectal cancer [2]. In a cohort of 1188 unexplained index patients enriched for inherited CRC and polyposis, we identified three *POLE* p.(Leu424Val) carriers at a frequency (0.25%), comparable to reported frequencies [1, 3]. Interestingly, *POLE* carriers from two families displayed a Lynch syndrome-like phenotype with MMR-deficient tumors. MMR deficiency in these tumors resulted from secondary somatic MMR variants due to the proofreading defect. In a study by Jansen et al. [4], a similar Lynch syndrome-like phenotype in *POLE* variant carriers was described.

DNA proofreading defects result in ultramutated tumor phenotypes with an increase in C:T>A:G mutations [5]. Recently, genomic sequencing of tumors with concurrent activity loss of one of the MMR genes and *POLE* or *POLD1* revealed the distinct mutational signatures SBS14 and SBS20, respectively, different from the signatures SBS10 or SBS6 for *POLE* or MMR deficiency, respectively [6-9]. Previously, MMR-deficient tumors with somatic MMR variants or *MLH1* promoter hypermethylation have been reported for patients with biallelic variants in the base excision repair (BER) gene *MUTYH* [10, 11]. The somatic MMR variants were MAP-specific G>T variants, indicating that impaired BER was the primary defect followed by MMR deficiency [10]. *POLE* DNA analysis now seems warranted in microsatellite-unstable CRC, especially in the absence of a causative DNA mismatch repair germline variant.

In **Chapter 3**, in search for additional *POLE*/*POLD1* pathogenic variants other than Leu424Val and Ser478Asn, we sequenced the exonuclease domains of *POLE* and *POLD1* in unexplained patients with multiple colorectal polyps. We describe two variants of unknown significance (VUS) in *POLD1* [12]. However, the available evidence is insufficient to evaluate the pathogenicity of these variants due to a

lack of cosegregation information and functional analysis. Sequencing of *POLE* and *POLD1* results in VUS variants rather than pathogenic variants, suggesting that pathogenic variants in *POLE* and *POLD1* probably occur at low frequencies. The assessment of the pathogenicity of variants of unknown significance remains a significant challenge in the investigation of hereditary CRC (and any other cancer syndrome). Interestingly, we found that one patient in addition to the *POLD1* VUS variant also carried a monoallelic *MUTYH* pathogenic variant, possibly suggesting that both genes could act cooperatively and together to confer an increased CRC risk. Hamzaoui et al. reported the cooccurrence of a *POLE* VUS variant and a pathogenic *MSH2* variant in CRC patients [13].

APC

In addition to classic *APC* germline variants, a few deep intronic variants contribute substantially to the *APC* mutation spectrum [14, 15]. In a study by Spier et al., the first systematic analysis of intronic variants that may affect RNA splicing in *APC* was performed. They investigated the frequency and type of deep intronic splice variants of *APC* in polyposis patients and highlighted the relevance of studying deep intronic *APC* splice variants in FAP, which cannot be identified by conventional routine screening methods [14]. In a study by Nieminen et al., pseudoexons in *APC* were successfully identified using next-generation sequencing, and this was the second study to reveal *APC*-related pseudoexons in FAP [15]. In **Chapter 4**, we attempt to investigate the roles of these deep intronic germline *APC* variants described by Spier et al. and Nieminen et al. [14, 15] as possible genetic causes of colorectal polyposis in a Dutch cohort of unexplained patients with more than 50 polyps. We did not detect any one of these variants in our cohort as a cause of colorectal polyposis. It is possible that either the frequency of intronic variants is lower in the Dutch population and the sample size of our cohort is not large enough or these intronic *APC* variants are local founder variants [16].

In 10-25% of the index patients with FAP, a de novo *APC* variant is identified [17-19]. Among those, there is a substantial but still underestimated proportion of mosaic carriers [20, 21]. Recent reports using methods that are able to detect germline variants with low allele frequencies, as well as variants only present in tumor material, indicate that many mosaic patients are undiagnosed [22, 23]. With the advantage of NGS technology, which allows for deep sequencing of selected regions, mosaic variants in *APC* are detected more frequently [22, 23]. In **Chapter 4**, we investigate the role of mosaic *APC* variants as possible genetic causes of colorectal polyposis in the same cohort that we screened

for deep intronic germline *APC* variants. We performed deep NGS of *APC* to identify possible undetected pathogenic mosaic variants in leukocyte DNA of unexplained index patients with colorectal polyposis. We did not detect mosaic *APC* variants. A limitation of this study is that we screened only the available leukocyte DNA for mosaicism due to the scarcity of tumor tissue for our study cohort [16]. The strategy of sequencing multiple adenomas of the same patients has been proven to be more sensitive and specific than sequencing leukocyte DNA for variants with low variant allele frequencies and can detect mosaicism confined to the colon [22-24].

Biallelic *NTHL1* LoF variants

In 2015, it was shown that germline biallelic loss-of-function (LoF) variants in *NTHL1* predispose to adenomatous polyposis and CRC, but the phenotypic spectrum remained to be elucidated, as patient numbers for this rare syndrome were low [25, 26]. Hence, large-scale studies are needed to further delineate this recently identified syndrome. In **Chapter 5**, using a large cohort of patients, we aimed to define the molecular and clinical characteristics of individuals with germline *NTHL1* LoF variants, and we found that NTHL1 deficiency predisposes them to multiple tumor types, including colon and breast cancer.

We screened our cohort for the most common LoF variant in *NTHL1* (p.Q90*) and studied the genotype-phenotype relationship in *NTHL1* biallelic LoF variant carriers. For a comprehensive analysis with sufficient cases, our data were combined with the data from an international consortium. In this chapter, we present a molecular and clinical characterization of the tumor spectrum of a total of 29 individuals with biallelic LoF variants in *NTHL1* from 17 unrelated families, including 11 previously unreported families, of which 26 developed one (n=10) or multiple (n=16) malignancies in 14 different tissues. We found that the majority of individuals developed one or more CRCs (59%). We show that 55% of the individuals with biallelic LoF variants in *NTHL1* developed multiple primary tumors at various sites, of which the majority were extracolonic (66%), while for MUTYH-associated polyposis, no more than 13% of the individuals developed an extracolonic malignancy [27]. An unexpectedly high breast cancer incidence was observed in female carriers (60%). In addition to breast cancer, we encountered endometrial (pre)malignancies, urothelial cell cancers, brain tumors, hematologic malignancies, basal cell carcinomas, head and neck squamous cell carcinomas, cervical cancers in multiple individuals and five other cancers in single individuals, including duodenal cancer.

We obtained additional evidence for causality of NTHL1 deficiency for specific malignancies by analyzing somatic mutational patterns using whole-exome sequencing from 14 tumors from seven different tissues (adenomatous/colorectal cancer, breast cancer, endometrial cancer, head and neck squamous cell carcinoma, meningioma, thyroid cancer, and urothelial cell cancer). We identified signature SBS30 in 13 out of the 14 tumors (93%). This signature is associated with NTHL1 deficiency and is characterized by C:G>T:A transitions at non-CpG sites. This suggests that deficiency of NTHL1 elicits the same mutational process in multiple tissues. The tumor without signature SBS30 was a urinary cell carcinoma in which signature 2 was the most prominent signature. This signature is commonly observed in sporadic urothelial cell cancers and suggests that this tumor developed sporadically [28]. A study in which *NTHL1* was knocked out in human intestinal organoids revealed that NTHL1 deficiency is the mutational process underlying signature SBS30 [29]. Signature SBS30 was previously identified in a single breast cancer case [30]. Retrospective analysis of that single breast cancer sample revealed an *NTHL1* germline LoF variant with loss of heterozygosity in tumors [29]. We show that in four breast cancer samples from four individuals with biallelic LoF variants in *NTHL1* that were sequenced, more than 80% of the mutations can be assigned to signature SBS30, suggesting that this base excision repair defect has driven breast cancer formation in these patients. We found a high incidence of breast cancer among women with biallelic *NTHL1* LoF variants (60%), and the median age at diagnosis for breast cancer in these women was found to be lower than in the general population (48.5 years [range: 38-63] compared with 62 years, respectively). This observation suggests a high penetrance for breast cancer for individuals with biallelic *NTHL1* LoF variants compared to, for example, the risks of breast cancer for *BRCA1* and *BRCA2* carriers of 57% and 49% by the age of 70 years, respectively [31]. We estimated the cumulative risk for extracolonic cancer to be between 35% and 78% by the age of 60 years, which highlights the importance of surveillance for extracolonic malignancies in patients with NTHL1 deficiency.

The tumor spectrum of individuals with biallelic *NTHL1* LoF variants was shown to be broader than polyposis and colorectal carcinomas, as has also been observed for other CRC syndromes associated with DNA repair defects. For example, MUTYH-associated polyposis patients have an increased lifetime risk of developing duodenal, ovarian, bladder, skin and possibly breast cancer [27]. Lynch syndrome patients have an increased lifetime risk of developing cancer of the endometrium, small bowel, urinary tract, stomach and ovaries [32, 33]. It has been postulated that polymerase proofreading-associated polyposis patients

may, next to endometrial cancer, be at an increased lifetime risk of developing brain tumors and cutaneous tumors [1, 34].

We conclude that biallelic germline *NTHL1* LoF variants predispose patients to multiple primary tumors, including colon cancer and breast cancer (**Chapter 5**)[28], and recent studies confirmed our findings [35-37]. Consequently, germline testing of *NTHL1* for individuals with multiple primary malignancies, either with or without adenomatous polyposis and/or a family history of cancer, might be considered.

Additionally, in **Chapter 5**, we demonstrate that mutational signatures in tumors can be used as a tool to corroborate a genetic predisposition. We found tumor mutational signature analysis to be suitable for obtaining additional support for a causative link between NTHL1 deficiency and tumor development. We showed that the presence of a unique mutational signature that is associated with a germline defect can distinguish these tumors from those that developed sporadically, as somatic inactivation of *NTHL1* is not a frequent event.

Monoallelic *NTHL1* LoF variants in polyposis and CRC

The list of genes associated with adenomatous polyposis and colorectal cancer now includes two recessive cancer-predisposing base-excision repair genes, i.e., *MUTYH* and *NTHL1*. For *MUTYH*, it is suggested that individuals with monoallelic LoF variants may have an increased, albeit small, risk of developing CRC compared to the general population [38-40]. Thus far, it is unknown whether monoallelic *NTHL1* LoF variants increase the risk of polyposis and/or CRC and whether carriers of monoallelic *NTHL1* LoF variants and their family members need additional counseling. While the prevalence of biallelic *NTHL1* LoF variants is low, the identification of monoallelic *NTHL1* LoF variant carriers from multigene panel testing is more common. The most common LoF variant in *NTHL1* is p.(Gln90*), which is heterozygous in approximately 0.28% of the general population [41]. The analysis of a breast cancer from an individual with a monoallelic *NTHL1* LoF variant suggests that these alleles may play a potential role in tumor development [29]. Therefore, it is of clinical importance to know whether carriers of monoallelic LoF variants in *NTHL1* are at increased risk of developing polyposis and/or CRC.

In **Chapter 6**, we investigated whether individuals with polyposis and/or CRC more frequently carry monoallelic LoF variants in *NTHL1* than the general population and whether monoallelic *NTHL1* LoF variants increase the risk of

polyposis and/or CRC in carriers. To address this question, an international collaboration between various research groups (the Netherlands, the United Kingdom, Poland, Germany, North Macedonia, North America, Canada and Australia) established a large cohort of 5,942 cases. The cohort consisted of individuals with unexplained polyposis, familial CRC, or sporadic CRC at a young age or those suspected of having Lynch syndrome with CRC or multiple adenomas. The cohort was investigated for monoallelic LoF variants in *NTHL1.* We did not find significant enrichment of monoallelic *NTHL1* LoF variant carriers in our cohort compared to control datasets. Furthermore, mutational signature analysis of 13 colorectal tumors from monoallelic *NTHL1* LoF variant carriers did not show a somatic second hit, and we did not find evidence of a main contribution of the mutational signature SBS30, the signature associated with NTHL1 deficiency, suggesting that monoallelic loss of *NTHL1* does not substantially contribute to colorectal tumor development [42]. Thus, we found no evidence that monoallelic *NTHL1* LoF variant carriers are at increased risk of developing polyposis and/or CRC; consequently, no additional surveillance is currently warranted. However, we cannot rule out that a small risk for CRC, similar to what is observed for *MUTYH* carriers, still exists. To date, screening cohorts of patients and tumors with a monoallelic pathogenic variant in *MUTYH* have been larger than those for *NTHL1*. Therefore, screening more patients for *NTHL1* is needed. From our data, we suggest that inactivation of the *NTHL1* wild-type allele (via LOH) is a rare event in colorectal tumors, which is in agreement with the observation that loss of 16p, the chromosome arm on which *NTHL1* is located, does not frequently occur in CRC [43]. Monoallelic LoF variants in *MUTYH* with LOH (on chromosome arm 1p) and high levels of signature SBS18 or combined SBS18/SBS36 have been reported in colorectal tumors [44, 45]. Loss of 1p is reported to occur in only approximately 10% of CRCs [46], which may explain the only slightly increased CRC risk reported for *MUTYH* [40]. In a recent study, molecular analysis of breast cancers from carriers indicated that *NTHL1* may be included in the growing list of low-penetrance breast cancer genes that appear to function via haploinsufficiency rather than the somatic biallelic inactivation mechanism almost universally observed for high-risk breast cancer predisposition genes [47]. The absence of a second hit in *NTHL1* may be a generic feature of low- to moderate-penetrance alleles, and these alleles are less prone to obtain second hits leading to a complete loss of function, always retaining some activity in the tumor [47]. To conclude, there is no evidence that monoallelic germline *NTHL1* LoF variant carriers are at increased risk of developing polyposis and/or CRC. To date, there is no evidence supporting specific surveillance for monoallelic carriers.

Monoallelic *NTHL1* LoF variants in the risk of extracolonic cancer

The biallelic *NTHL1* LoF variants predispose to a multitumor phenotype, but whether monoallelic carriers are at increased risk of developing other extracolonic malignancies remains to be elucidated. We investigated the role of the monoallelic *NTHL1 c.268C*>T, p.(Gln90*) variant in the risk of extracolonic cancers, but we found that the monoallelic *NTHL1* p.(Gln90*) variant does not seem to predispose patients to extracolonic cancer (unpublished data). In a cohort of cases with extracolonic cancer and suspected Lynch syndrome (N= 327), two monoallelic *NTHL1* p.(Gln90*) carriers were detected (2/327, 61%). One patient developed urothelial cell cancer (UCC), and the second patient developed adenosquamous carcinoma (ASC) of the mouth. We found no significant enrichment of monoallelic *NTHL1* p.(Gln90*) carriers in our cohort compared to a genome aggregation database (gnomAD) non-Finnish European control population (2/327; 0.61% versus 250/64,328; 0.39%; *P* = 0.36). Further exome sequencing for the available tumor (ASC) did not detect the *NTHL1* deficiency-related mutational signature SBS30 and LOH of the wild-type *NTHL1* allele, which indicates that monoallelic *NTHL1* did not play a role in tumor development in this patient. Following the initial discovery that biallelic LoF variants in *NTHL1* predispose to breast cancer, we genotyped *NTHL1* p.(Gln90*) in a cohort of 692 individuals with ductal carcinoma in situ (DCIS) and detected one biallelic (1/692; 0.14%) and three monoallelic carriers (3/692; 0.4%). The frequency of monoallelic *NTHL1* p.(Gln90*) was not significantly enriched in our DCIS cohort compared to gnomAD non-Finnish European controls (3/692; 0.4% versus 250/64,328; 0.39%; *P* = 0.75). We found no evidence that monoallelic *NTHL1* p.(Gln90*) carriers are at increased risk of developing DCIS. A recent study suggested that carriers of monoallelic *NTHL1* p.(Gln90*) do not have an increased risk for breast cancer [48]. An even more recent study suggested that monoallelic LoF variants in *NTHL1* may be associated with a low to moderate increased risk of breast cancer [47]. Salo-Mullen et al. identified a woman with high-grade serous ovarian carcinoma harboring monoallelic *NTHL1* p.(Gln90*) with corresponding LOH of the wild-type allele in the tumor resulting in signature 30 [49]. Based on data from cBioPortal, loss of 16p, the chromosome arm on which *NTHL1* is located, mainly occurs in ovarian serous cystadenocarcinoma and uterine carcinosarcoma, while in colorectal adenocarcinoma and breast invasive ductal carcinoma, this loss is only 6%. It is possible that monoallelic *NTHL1* carriers are at risk of developing ovarian cancer when loss of 16p occurs as an early event in tumorigenesis. Salo-Mullen et al. identified a prostate cancer patient with monoallelic *NTHL1* p.(Gln90*) and signature 30 but without LoF of the wild-type allele [49]. The contradictory results from these studies may be

explained by differences in tumorigenesis, including that different mechanisms can drive tumor development in monoallelic carriers, such as the timing of a potential second hit. In conclusion, our results indicate that monoallelic *NTHL1* p.(Gln90*) is unlikely to be a significant contributor to extracolonic cancer, which is in line with results obtained for CRC cancer in **Chapter 6**.

Future perspectives

In this thesis, we illustrate the power of mutational signature analysis in defining tumor phenotypes in rare cancer predisposition syndromes and provide proof of principle for recognizing new patients with cancer syndromes based on tumor sequencing data. In the future, mutational signature analysis will assist in the identification of novel cancer syndromes, including adenomatous polyposis and/or CRC syndromes caused by DNA repair deficiency.

Studying the mutation signatures in tumors could confirm the pathogenicity of VUS variants and mark them as causal variants in the predisposition for multiple colorectal polyps.

Recent reports using methods that are able to detect germline variants with low allele frequencies, as well as variants only present in tumor material, indicate that many mosaic patients are undiagnosed. Testing tumor DNA, rather than leukocyte DNA, will provide greater knowledge about the true incidence of mosaicism in *APC*. In-depth analysis of adenomas of patients could lead to the detection of more mosaic *APC* carriers. Recently, the recurrent *APC* splice variant c.835-8A>G in a patient with unexplained colorectal polyposis fulfilling the colibactin mutational signature was reported [50]. The presence of pks + E coli, causing a specific mutational signature, might be an additional explanation for unexplained polyposis patients.

The use of novel sequencing techniques will possibly enable the detection of rare variants and germline aberrations in noncoding regions in the near future. Well-defined patient cohorts and families with multiple affected members will help in the identification of novel polyposis- and CRC-predisposing germline aberrations. Joint efforts screening for variants in larger cohorts and data sharing are essential to find underlying genetic causes of colorectal polyposis and CRC. Hopefully, the results and knowledge gathered will ultimately contribute to the significant clinical management and prevention of CRC.

References

1. Palles C, Cazier JB, Howarth KM, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. Nat Genet 2013;45:136-44.
2. Elsayed FA, Kets CM, Ruano D, et al. Germline variants in POLE are associated with early onset mismatch repair deficient colorectal cancer. Eur J Hum Genet 2015;23:1080-4.
3. Chubb D, Broderick P, Frampton M, et al. Genetic diagnosis of high-penetrance susceptibility for colorectal cancer (CRC) is achievable for a high proportion of familial CRC by exome sequencing. J Clin Oncol 2015;33:426-32.
4. Jansen AM, van Wezel T, van den Akker BE, et al. Combined mismatch repair and POLE/POLD1 defects explain unresolved suspected Lynch syndrome cancers. Eur J Hum Genet 2016;24:1089-92.
5. Shinbrot E, Henninger EE, Weinhold N, et al. Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. Genome Res 2014;24:1740-50.
6. Campbell BB, Light N, Fabrizio D, et al. Comprehensive Analysis of Hypermutation in Human Cancer. Cell 2017;171:1042-1056.e10.
7. Haradhvala NJ, Kim J, Maruvka YE, et al. Distinct mutational signatures characterize concurrent loss of polymerase proofreading and mismatch repair. Nat Commun 2018;9:1746.
8. Castellsagué E, Li R, Aligue R, et al. Novel POLE pathogenic germline variant in a family with multiple primary tumors results in distinct mutational signatures. Hum Mutat 2019;40:36-41.
9. Alexandrov LB, Kim J, Haradhvala NJ, et al. The repertoire of mutational signatures in human cancer. Nature 2020;578:94-101.
10. Morak M, Heidenreich B, Keller G, et al. Biallelic MUTYH mutations can mimic Lynch syndrome. Eur J Hum Genet 2014;22:1334-7.
11. Colebatch A, Hitchins M, Williams R, et al. The role of MYH and microsatellite instability in the development of sporadic colorectal cancer. Br J Cancer 2006;95:1239-43.
12. Elsayed FA, Tops CMJ, Nielsen M, et al. Low frequency of POLD1 and POLE exonuclease domain variants in patients with multiple colorectal polyps. Mol Genet Genomic Med 2019:e603.
13. Hamzaoui N, Alarcon F, Leulliot N, et al. Genetic, structural, and functional characterization of POLE polymerase proofreading variants allows cancer risk prediction. Genet Med 2020;22:1533-1541.
14. 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.
15. Nieminen TT, Pavicic W, Porkka N, et al. Pseudoexons provide a mechanism for allele-specific expression of APC in familial adenomatous polyposis. Oncotarget 2016;7:70685-70698.
16. Elsayed FA, Tops CMJ, Nielsen M, et al. Use of sanger and next-generation sequencing to screen for mosaic and intronic APC variants in unexplained colorectal polyposis patients. Fam Cancer 2021.
17. Bisgaard ML, Fenger K, Bülow S, et al. Familial adenomatous polyposis (FAP): frequency, penetrance, and mutation rate. Hum Mutat 1994;3:121-5.
18. Ripa R, Bisgaard ML, Bülow S, et al. De novo mutations in familial adenomatous polyposis (FAP). Eur J Hum Genet 2002;10:631-7.

19. Aretz S, Uhlhaas S, Caspari R, et al. Frequency and parental origin of de novo APC mutations in familial adenomatous polyposis. Eur J Hum Genet 2004;12:52-8.
20. Hes FJ, Nielsen M, Bik EC, et al. Somatic APC mosaicism: an underestimated cause of polyposis coli. Gut 2008;57:71-6.
21. Jansen AML, Goel A. Mosaicism in Patients With Colorectal Cancer or Polyposis Syndromes: A Systematic Review. Clin Gastroenterol Hepatol 2020;18:1949-1960.
22. 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.
23. Jansen AM, Crobach S, Geurts-Giele WR, et al. Distinct Patterns of Somatic Mosaicism in the APC Gene in Neoplasms From Patients With Unexplained Adenomatous Polyposis. Gastroenterology 2017;152:546-549.e3.
24. Aretz S, Stienen D, Friedrichs N, et al. Somatic APC mosaicism: a frequent cause of familial adenomatous polyposis (FAP). Hum Mutat 2007;28:985-92.
25. 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.
26. Rivera B, Castellsagué E, Bah I, et al. Biallelic NTHL1 Mutations in a Woman with Multiple Primary Tumors. N Engl J Med 2015;373:1985-6.
27. Vogt S, Jones N, Christian D, et al. Expanded extracolonic tumor spectrum in MUTYH-associated polyposis. Gastroenterology 2009;137:1976-85.e1-10.
28. 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.
29. Drost J, van Boxtel R, Blokzijl F, et al. Use of CRISPR-modified human stem cell organoids to study the origin of mutational signatures in cancer. Science 2017;358:234-238.
30. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534:47-54.
31. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. J Clin Oncol 2007;25:1329-33.
32. Vasen HF, Wijnen JT, Menko FH, et al. Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. Gastroenterology 1996;110:1020-7.
33. Watson P, Vasen HF, Mecklin JP, et al. The risk of endometrial cancer in hereditary nonpolyposis colorectal cancer. Am J Med 1994;96:516-20.
34. Aoude LG, Heitzer E, Johansson P, et al. POLE mutations in families predisposed to cutaneous melanoma. Fam Cancer 2015;14:621-8.
35. Boulouard F, Kasper E, Buisine MP, et al. Further delineation of the NTHL1 associated syndrome: A report from the French Oncogenetic Consortium. Clin Genet 2021;99:662-672.
36. Beck SH, Jelsig AM, Yassin HM, et al. Intestinal and extraintestinal neoplasms in patients with NTHL1 tumor syndrome: a systematic review. Fam Cancer 2022.
37. Altaraihi M, Gerdes AM, Wadt K. A new family with a homozygous nonsense variant in NTHL1 further delineated the clinical phenotype of NTHL1-associated polyposis. Hum Genome Var 2019;6:46.
38. Win AK, Hopper JL, Jenkins MA. Association between monoallelic MUTYH mutation and colorectal cancer risk: a meta-regression analysis. Fam Cancer 2011;10:1-9.
39. Win AK, Cleary SP, Dowty JG, et al. Cancer risks for monoallelic MUTYH mutation carriers with a family history of colorectal cancer. Int J Cancer 2011;129:2256-62.
40. Win AK, Dowty JG, Cleary SP, et al. Risk of colorectal cancer for carriers of mutations in MUTYH, with and without a family history of cancer. Gastroenterology 2014;146:1208-11. e1-5.

41. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434-443.
42. 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.
43. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401-4.
44. Pilati C, Shinde J, Alexandrov LB, et al. Mutational signature analysis identifies MUTYH deficiency in colorectal cancers and adrenocortical carcinomas. J Pathol 2017;242:10-15.
45. Georgeson P, Pope BJ, Rosty C, et al. Evaluating the utility of tumour mutational signatures for identifying hereditary colorectal cancer and polyposis syndrome carriers. Gut 2021;70:2138-2149.
46. De Angelis PM, Clausen OP, Schjølberg A, et al. Chromosomal gains and losses in primary colorectal carcinomas detected by CGH and their associations with tumour DNA ploidy, genotypes and phenotypes. Br J Cancer 1999;80:526-35.
47. Li N, Zethoven M, McInerny S, et al. Evaluation of the association of heterozygous germline variants in NTHL1 with breast cancer predisposition: an international multi-center study of 47,180 subjects. NPJ Breast Cancer 2021;7:52.
48. Kumpula T, Tervasmäki A, Mantere T, et al. Evaluating the role of NTHL1 p.Q90* allele in inherited breast cancer predisposition. Mol Genet Genomic Med 2020:e1493.
49. Salo-Mullen EE, Maio A, Mukherjee S, et al. Prevalence and Characterization of Biallelic and Monoallelic NTHL1 and MSH3 Variant Carriers From a Pan-Cancer Patient Population. JCO Precis Oncol 2021;5.
50. Terlouw D, Suerink M, Boot A, et al. Recurrent APC Splice Variant c.835-8A>G in Patients With Unexplained Colorectal Polyposis Fulfilling the Colibactin Mutational Signature. Gastroenterology 2020;159:1612-1614.e5.