Characterization of candidate genes in unexplained polyposis and colorectal cancer

Abayzeed Elsayed Osman, F.

Chapter 8

Summary
Nederlandse samenvatting
List of publications
Curriculum vitae
Acknowledgements

Summary

Heritable factors account for approximately 35% of colorectal cancer (CRC) risk. Around 5 to 10% of CRC cases are associated with highly penetrant dominant or recessive inherited syndromes, caused by germline variants in known high-penetrance CRC genes. The etiology of the remaining 20%-30% of inherited CRC risk is not completely understood. In recent years, advances were made in discovering the genetic causes for CRC and polyposis. Germline variants in *POLE*, *POLD1* and biallelic variants in *NTHL1* were discovered underlying polymerase proofreading associated polyposis syndrome and NTHL1-assciated tumor syndrome respectively, and new genes are still being described. A precise understanding of the genetics of inherited CRCs is important for identifying at risk individuals, improving cancer surveillance and prevention strategies, and developing better diagnostic and therapeutic approaches. The studies described in this thesis focus on characterizing variants in known high risk genes such as *POLE*, *POLD1*, *APC* and *NTHL1 as* genetic causes of polyposis and CRC in unexplained cases.

Germline pathogenic variants in DNA polymerase ε (*POLE*) and ẟ (*POLD1*) have been identified in families with multiple colorectal adenomas and CRC, in **Chapter 2** we screened the pathogenic germline variants in *POLE* and *POLD1* that were identified by Palles et al. in our cohort of unexplained familial, early onset CRC and polyposis cases. The frequency of the variants we report is comparable to those previously reported, despite an enrichment in our cohort for inherited CRC and polyposis. Interestingly we showed that the tumors associated with *POLE* germline variants can show a Lynch syndrome-like phenotype with mismatch repair (MMR) deficiency due to somatic mutation in MMR genes which results from the proofreading deficiency caused by *POLE* inactivation.

In **Chapter 3**, with the aim to find additional pathogenic variants in *POLE* and *POLD1* using next-generation sequencing (NGS) we sequenced the exonuclease domains of *POLE* and *POLD1* on a cohort of unexplained index patients diagnosed with multiple colorectal polyps. Germline variants of uncertain significance were found in *POLD1*, but no further testing was possible to assess the functional relevance of these variants as tumors were not available for further studies. This study confirms the low frequency of causal variants in these genes in the predisposition for multiple colorectal polyps, and established that these genes are a rare cause of colorectal polyps or CRC.

In **Chapter 4** we screened for previously reported pathogenic deep intronic germline *APC* variants in a cohort of unexplained colorectal polyposis patients. Using deep NGS we furthermore screened for *APC* mosaic variants. We did not detect mosaic or intronic *APC* variants in the screened unexplained colorectal polyposis patients. The limitation of this study was that we screened only leukocyte DNA for mosaic variants. Consequently, *APC* mosaic variants solely confined to the colon could have been missed with this approach because we could not screen the DNA from adenomas of the patients.

In 2015, biallelic germline loss-of-function (LoF) variants in *NTHL1* were shown to predispose to adenomatous polyposis and CRC, but the exact clinical phenotype was unclear as the patient numbers for this syndrome were low. In **Chapter 5** we characterized *NTHL1* tumor syndrome with the use of mutational signature analysis. To define the molecular and clinical characterization of tumor spectrum of the individuals with biallelic germline LoF variants in *NTHL1*, a large collaborative study involving research groups from Netherlands, United Kingdom, Poland, Germany, Norway, Spain and Macedonia was established. We collected clinical and molecular data of 29 individuals with biallelic germline *NTHL1* LoF variants from 17 families. We found that 55% of the individuals developed multiple primary tumors at various sites, of which the majority was extracolonic (66%). In addition to colorectal tumors we found tumors in 13 tissue types. Most individuals developed one or more CRCs (59%) and high breast cancer incidence was observed in female carriers (60%). We identified a unique mutational signature (SBS30) that was associated with NTHL1-deficiency in 13 tumors from seven organs. Our study demonstrates that *NTHL1* is a multi-tumor predisposition gene with a high lifetime risk for extracolonic cancers.

While biallelic germline *NTHL1* LoF variants are causal to adenomatous polyposis and CRC, the adenomatous polyposis and CRC risk for carriers of monoallelic germline *NTHL1* LoF variants remained to be established. As carriers of monoallelic germline LoF variants in *MUTYH* were previously found to have a small increased risk for CRC as well, we investigated the role of monoallelic germline LoF variants in *NTHL1* on the risk of adenomatous polyposis and CRC in **Chapter 6.** To establish a large cohort to investigate the monoallelic *NTHL1* LoF variants role we established the collaborative *NTHL1* study group. In total 5,942 individuals with unexplained polyposis and/or CRC were screened. We demonstrated that monoallelic LoF variants in *NTHL1* are not enriched in individuals with polyposis and/or CRC compared to the general populations. Furthermore, mutational signature analysis on 13 colorectal tumors of

individuals with a monoallelic *NTHL1* LoF variant did not show a somatic second hit, nor did we find evidence of a main contribution of mutational signature SBS30, the signature associated with NTHL1 deficiency, indicating that monoallelic loss of *NTHL1* does not substantially contribute to colorectal tumor development. Thus, we found no evidence that monoallelic *NTHL1* LoF variant carriers are at increased risk to develop polyposis and/or CRC.