



Universiteit  
Leiden  
The Netherlands

**Next generation sequencing of ovarian metastases of colorectal cancer**  
Crobach, A.S.L.P.

**Citation**

Crobach, A. S. L. P. (2018, March 29). *Next generation sequencing of ovarian metastases of colorectal cancer*. Retrieved from <https://hdl.handle.net/1887/60909>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/60909>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The following handle holds various files of this Leiden University dissertation:  
<http://hdl.handle.net/1887/60909>

**Author:** Crobach, A.S.L.P.

**Title:** Next generation sequencing of ovarian metastases of colorectal cancer

**Issue Date:** 2018-03-29

NEXT GENERATION SEQUENCING  
OF OVARIAN METASTASES OF COLORECTAL CANCER

Colophon

© ASLP Crobach, Leiden, The Netherlands

Next generation sequencing of ovarian metastases of colorectal cancer

The studies described in this thesis were performed at the Department of Pathology (Head: Prof. V.T.H.B.M. Smit) of the Leiden University Medical Center, the Netherlands.

No part of this thesis may be reproduced, stored, or transmitted in any form or by any means without prior permission of the authors.

Produced by: F&N Eigen Beheer

ISBN: 978949254412 4

# **Next generation sequencing of ovarian metastases of colorectal cancer**

PROEFSCHRIFT

ter verkrijging van  
de graad van Doctor aan de Universiteit Leiden,  
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,  
volgens besluit van het College voor Promoties  
te verdedigen op donderdag 29 maart 2018 klokke 16.15 uur

door

Augustinus Servatius Lodewijk Pieter Crobach

geboren te Maastricht  
in 1983

Promotor: Prof. dr. H. Morreau  
Co-promotor: Dr. T. van Wezel

Leden promotiecommissie:

Prof. dr. H.I. Grabsch,  
afdeling Pathologie, Maastricht University  
Emeritus Prof. dr. F.T. Bosman,  
Université de Lausanne, Suisse.  
Prof. dr. V.T.H.B.M. Smit

# Contents

<b>Chapter 1</b>	General Introduction Parts of this chapter have been published previously and were adapted in a modified form. <i>Clinical laboratory international</i> . June 2015, page 36-38.	7
<b>Chapter 2</b>	Ovarian metastases of colorectal and duodenal cancer in familial adenomatous polyposis. <i>Familial Cancer</i> . 2012 Dec;11(4):671-3.	29
<b>Chapter 3</b>	Target-enriched next-generation sequencing reveals differences between primary and secondary ovarian tumors in formalin-fixed, paraffin-embedded tissue. <i>Journal of Molecular Diagnostics</i> . 2015 Mar;17(2):193-200	39
<b>Chapter 4</b>	Somatic mutation profiles in primary colorectal cancers and matching ovarian metastases: Identification of driver and passenger mutations. <i>The Journal of Pathology: Clinical Research</i> . 2016 Apr 15;2(3):166-74	79
<b>Chapter 5</b>	Next generation sequencing using the HaloPlex targeting method in formalin-fixed paraffin-embedded (FFPE) material. <i>Manuscript in preparation</i>	137
<b>Chapter 6</b>	Excluding Lynch syndrome in a female patient with metachronous DNA mismatch repair deficient colon - and ovarian cancers. <i>Familial Cancer</i> . 2017 Nov 9	167
<b>Chapter 7</b>	Concluding remarks and future perspectives.	181
<b>Chapter 8</b>	English Summary / Nederlandse samenvatting List of publications Curriculum vitae Dankwoord / Acknowledgements	193





# Chapter 1

## **General Introduction**

**Parts of this chapter have been published previously and were adapted in a modified form**

## **I. Diagnostically challenging areas: distinguishing primary from secondary ovarian malignancies**

### **Description of the problem**

The ovaries are a preferential location for metastases from, among others, colon, stomach, appendiceal, breast, and endometrium carcinomas.[1] The reported percentages of secondary ovarian tumors (metastases) vary from 8-30%.[2] Several reasons can be given why such percentages show a broad range. First, studies are different by design. Some studies are based on autopsy findings, while others are based on prophylactic oophorectomies. Second, differences in the incidence of primary tumors can cause a variance in the pattern of metastases. For example, stomach cancer has a higher incidence in Japan than in many other countries.[3] Therefore, metastases of stomach cancer to the ovaries are expected to be more common in Japan. Mostly the gastro-intestinal tract (GIT) seems to be the main source for ovarian metastases. The contribution of tumors from other organs is less clear. Breast and endometrial cancers are the second and third major sources, respectively, of ovarian metastases. Less frequent are metastases from cervical tumors. Correctly distinguishing between primary and secondary ovarian tumors using hematoxylin-eosin staining in combination with immunohistochemistry can be problematic but is crucial for correct treatment choice.[4, 5]

### **Macroscopic and histologic approach**

A gross distinction between primary and secondary ovarian tumors can be made by taking tumor size and unilaterality versus bilaterality into account.[6] Following the decision tree depicted in Figure 1, it is possible to estimate whether an ovarian tumor is a primary tumor or a metastasis. A unilateral ovarian tumor with a diameter larger than 10 cm is probably a primary tumor. All bilateral and unilateral tumors smaller than 10 cm are much more likely to be metastases.

Visceral organs are mostly affected by conventional adenocarcinomas, originating from the glandular epithelium. Some of the primary ovarian malignancies such as endometrioid and mucinous adenocarcinomas can show extensive histological and immunohistochemical similarities to these adenocarcinomas. Otherwise, the histologic characteristics of metastatic GIT ovarian tumors do not resemble serous papillary or clear cell tumors of the ovary. Consequently, based on histology, a subset of primary ovarian tumors has a clear origin and diagnosis is straightforward. In addition, other histologic findings can assist in defining the malignancy. For example, surface involvement by malignant epithelial cells is much more commonly seen in metastases than in primary ovarian tumors.[7] On the other hand, an expansile growth pattern is

more often seen in primary ovarian tumors. Therefore, with the help of histopathological findings, the difference between a primary origin or a metastatic process becomes clearer.

### **Immunohistochemical approaches**

The logical next step in differentiating primary ovarian tumors from metastases is applying the use of immunohistochemistry. For example, primary ovarian tumors are classically positive for keratin 7 and negative for keratin 20, while colorectal tumors show the opposite staining pattern (keratin 7 negative, keratin 20 positive).[8, 9] Other markers can also be used, not only to rule out an ovarian origin of the tumor but also to gain insight into the location of the primary tumor. Positivity for intestinal markers (such as *carcinoembryonic antigen* (CEA) and *caudal type homeobox 2* (CDX-2)) can be an argument for an intestinal origin of the tumor cells.[9, 10]

Furthermore, the staining profile of a possible metastasis can be compared with the primary tumor when the supposed primary location has already been discovered. However, it is reported that only in up to 38% of cases the detection of ovarian metastases precedes the detection of the primary tumors.[11] Finally, although infrequently occurring, unrelated primary ovarian tumors can arise in patients who anamnestically suffered from another malignancy, complicating the diagnostic procedures.

In routine diagnostics, the use of immunohistochemistry is frequently not fully discriminating. For example, primary ovarian tumors generally tend to have a Ker7+/Ker20- immunoprofile, while colonic metastases have a Ker7-/Ker20+ immunoprofile. Nevertheless, keratin 7 positivity can be seen in proximally located GIT tumors, and keratin 20 positivity can be seen in primary ovarian malignancies. A guided immunohistochemical decision scheme is shown for complex cases in Figure 2.

---

## II. Molecular subtyping of malignancies for diagnostic and therapeutic objectives

When the clinical information, histologic features and immunohistochemical staining patterns are combined, it is possible to differentiate between primary tumors and metastases in a substantial subset of cases. For example, when a patient with a history of a colorectal tumor subsequently presents with a large ovarian mass a few years later that displays a similar immunoprofile, it is not difficult to decide that the ovarian tumor is likely to be a metastasis from the CRC. Nevertheless, some cases are not as clear. In those cases, tumor size, unilaterality vs. bilaterality and the histologic findings are not enough to discriminate between primary tumors and secondary metastases.

In pathology, histology has always been the basis for the subtyping of malignancies. With the development of novel technologies (e.g., immunohistochemistry, expression array analysis, DNA and RNA sequencing), additional subtypes have been defined. Currently, the use of molecular characterization is advocated for all cancer types, leading to molecular subtyping that is based on the underlying biology. Molecular subtyping can help establish the correct primary diagnosis, give prognostic information and help stratify (neo-) adjuvant treatment decisions. Tumors can be typed on several levels (e.g., protein, DNA and RNA) and for multiple molecular features (e.g., protein expression, copy number alterations, mutations and methylation patterns). These characteristics of a tumor can be described by the multiple “omes” (also called “omics”).[12, 13] We describe these “omes” below.

The *proteome* is the complete set of proteins that partly reflects the transcriptome. The proteome shows both differences over time and differences per tissue type.[14] The proteome can be seen as the most functional profile of a cell, as all the other “omes” eventually influence the generation of proteins. A subclassification per cell compartment can be made (membrane, cytoplasm, and nucleus). A secretome, composed of proteins that are secreted, can be established using cell cultures. Proteins that are specifically produced by tumor cells can be useful as biomarkers if they are detectable in serum. However, a relatively unexplored level of complexity is the analysis of all post-translational modifications of proteins such as the addition of all kinds of glycan and lipid molecules.[15]

The term *genome* applies to the complete DNA sequence, including coding sequences, which is the blueprint for the formation of proteins. The introduction of next-generation sequencing (NGS) changed this field dramatically. In 2000, the

International Human Genome Sequencing Consortium revealed of a rough draft of the human genome sequence. [16] In 2003, a more detailed version of the human genome sequence was presented. The cost of the first version of the genome was 3 billion US dollars. Currently, with NGS techniques, a genome can be sequenced for a fraction of the cost (about \$1000). Another development is 'targeted sequencing' in which only genes of interest are selected from the genome.[17] In this way, those selected regions can be sequenced with high coverage, i.e., sequencing the same locus multiple times. This method improves the analysis and reduces false-positive and false-negative calls. In the past, Sanger DNA sequencing was used to detect mutations in clinically relevant genes. However, to screen complete genes and multiple genes in a sequential row is time-consuming. Currently, with the introduction of the revolutionary NGS technology, it is possible to sequence multiple or even all genes at the same time. NGS has become a standard technique in diagnostics for identifying gene variations.

The Catalogue Of Somatic Mutations In Cancer (COSMIC; <http://cancer.sanger.ac.uk/cosmic>) was the earliest database in which the mutational profiles of most cancer types were compiled.[18] These mutational profiles were constructed by sequencing data generated by Sanger sequencing. The enormous amount of data coming from NGS devices resulted in an immense increase in gene variants. These variants were compiled in hundreds of databases displaying overviews of pathogenic and non-pathogenic variants (e.g., SNPs).[19] Well known are the dbSNP database that aims to show non-pathogenic variants and the ClinVar database that lists genetic variations and their clinical relevance.[20, 21] However, databases polluted by false positive (suggested to be disease-causing) variants are problematic when analyzing sequence data and determining the clinical significance of variants.[22]

The *transcriptome* is the complete set of all RNA components (mRNA, rRNA, tRNA, microRNA and other non-coding RNAs).[23] A key feature of the transcriptome, in contrast with the stable genome, is its dynamics. Over time and per tissue type, the expression levels of all RNA subtypes can differ. The transcriptome can be examined by oligonucleotide arrays that use chip technology with complementary sequences to bind cDNA. When co-hybridizing a reference pool of cDNA labelled with a fluorescent signal, the amount of cDNA of the test sample labelled with a second fluorochrome influences the intensity of the read out signals and is thus informative about the expression levels.[24] NGS technology has created an alternative approach for analyzing the transcriptome. The number of transcripts obtained in an NGS analysis can be used as a read-out for the expression levels of genes, being a modern version of the classic serial analysis of gene expression (SAGE).[25] Sequencing of all RNA

(cDNA) molecules can also be informative about expressed pathogenic variants, alternative gene spliced transcripts and fusion transcripts that are sufficiently expressed and do not undergo nonsense-mediated decay.[26, 27]

The *epigenome* (also called the methylome) reflects all the epigenetic modifications, which are mainly alterations of DNA methylation patterns and histone modification of the genome.[28] Differential DNA hypermethylation regulates gene expression by the binding of methyl groups to specific regions in the DNA, the so called CpG islands. In cancer, a frequently observed phenomenon is hypermethylation of tumor suppressor genes. Tumor suppressor genes that are active in normal tissues have many regulatory roles and, once inactivated, can induce tumor formation. However, in addition to CpG island hypermethylation, global hypomethylation of widely dispersed DNA elements in the genome (for instance the LINE-1 elements) can also be seen.[29] Changes in global methylation patterns can affect three dimensional DNA structures through altered CCCTC-binding factor (encoded by *CTCF*) expression.[30] This in turn leads to altered mRNA expression patterns as a consequence of differential accessibility for all transcription factors.

The above described “omes” are only a selection of all the possibilities that can be recognized at present. To completely understand the underlying biology of cancer cells, a comprehensive analysis of all omics fields is theoretically needed in the *interactome* or *multiome*. Furthermore, complete and in-depth analysis of tumors at all these levels might lead to a better understanding of why tumors react or do not react to classic and targeted therapies. Additionally, new approaches might be revealed by absolute comprehensive analysis. In the context of this PhD thesis, extensive analysis might reveal the stratifying molecular profiles that undoubtedly indicate the true origin of metastasized tumors.

### **Molecular subtyping of sporadic colorectal cancer**

To some extent, comprehensive molecular profiling information of several tumor types is currently available. Large cohorts of, among others, colon, breast, endometrial and ovarian carcinomas have been studied in The Cancer Genome Atlas project (<https://cancergenome.nih.gov/>).[31] Not only somatic DNA variations are investigated but also methylation patterns, gene fusions and expression patterns.

For colon carcinoma, three classic molecular pathways implicated in colorectal tumorigenesis have been identified.[32] The chromosomal instability pathway (CIN) is the most prevalent of these three pathways, accounting for approximately 70-85% of colorectal cancer. The microsatellite instability (MSI) and the CpG island methylator

phenotype (CIMP) pathways are the other two.

### **Chromosomal instability pathway (CIN)**

The CIN pathway is associated with pathogenic variants of *APC* and *KRAS* and the loss of chromosome 5q, chromosome 17p and chromosome 18q.[33] *APC* variants are found in approximately 60-80% of colon carcinomas. *KRAS* mutations are found in approximately 35-42% of the cases. Other genes that are involved are *DCC*, *SMAD2*, *PIK3CA* and *SMAD4*, which are all located on chromosome 18q. Allelic loss of the 18q region is found in 60% of colorectal carcinomas. Functional loss of *TP53* (by combined mutation and loss of heterozygosity of the wild-type allele) is seen in approximately 50-75% of colorectal carcinomas. A comprehensive overview of the genetic profiles of CRC by next-generation sequencing (NGS) was recently published, and the results mostly confirmed the above prior knowledge.[34]

### **Microsatellite instability (MSI)**

During DNA replication, DNA polymerase makes copying errors in repetitive DNA elements, the micro-satellites. The DNA mismatch repair system (MMR) is meant to recognize and repair these mistakes. Inactivation of one of the genes responsible for MMR repair leads to a high incidence of sequence variation in these repetitive microsatellites, often 2-6 base pairs long, termed microsatellite instability. Tumors with a high incidence of somatic variation in microsatellites are typed as microsatellite instability-high tumors (MSI-H).[35] In Lynch syndrome, a colorectal and endometrium cancer susceptibility syndrome, *MSH2* and *MLH1* are the most frequently germ line altered genes. However, in recent years, *MSH6*, *PMS2* and *EPCAM* have also turned out to be important target genes in Lynch syndrome. Altered immunohistochemical staining patterns of the MMR proteins can be used as a screening tool to guide germ line testing of the MMR genes.

In addition to those resulting from Lynch syndrome, MSI-H tumors can also develop as a consequence of somatic hypermethylation of the *MLH1* gene promoter region or due to inactivation of MMR as a result of somatic pathogenic variants in the MMR genes with or without concomitant loss of heterozygosity of the wild-type alleles.

Tumors with low microsatellite instability (MSI-low or MSI-L tumors) often show instability at dinucleotide or tetranucleotide DNA repeats. These are not typically associated with inactivation of the 4 major MMR genes, although an association with *MSH3* inactivation was recently suggested.[36, 37] MSI-L tumors are furthermore associated with *KRAS* mutations and methylation of *MGMT*.

Recently, another molecular subtype of colorectal carcinoma was described that is characterized by an ultramutated phenotype. Mutations in DNA proofreading enzymes *polymerase*  $\epsilon$  and  $\delta$  (*POLE* / *POLD1*) cause colon cancers with high muta-

tional burdens, mostly comprising C>T base alterations. The pathogenic *POLE/D1* variation is mostly somatic in origin, with a small proportion being germ line-based. Comparable mutational phenotypes are also observed in endometrial adenocarcinomas.[38] Functional *POLE / POLD1* alteration can secondarily lead to MMR defects, thereby further contributing to ultramutated phenotypes.

### **CpG Island Methylator Phenotype (CIMP)**

The *CpG Island Methylator Phenotype (CIMP)* is associated with widespread promoter hypermethylation of numerous genes. CpG islands are DNA regions located in the promoter regions of housekeeping genes carrying high G:C contents.[32] CRCs with such characteristics are annotated as tumors with high frequency CpG island methylation (CIMP-high).[39] Hypermethylation of promoter regions can result in decreased transcription of target genes, resulting in inactivation of tumor suppressor genes, among others, and thereby contributing to tumorigenesis. A CIMP-high status is also associated with the presence of somatic *BRAF* activation due to gene variation, which is in itself associated with a poor clinical outcome.

Subclassifications of CRC have been proposed that take *BRAF* (and *KRAS*) gene variation into account.[40]

Revised subclassifications might better predict therapeutic response and prognosis.[41] To that end, the Colorectal Cancer Subtyping Consortium has classified colorectal cancer into four subtypes based on an integration of various levels.[42] Gene expression-based subclassification was integrated with genome and methylome information. Four molecular subtypes of colorectal cancer were identified: CMS1, microsatellite instability immune (14%); CMS2, canonical (37%); CMS3, metabolic (13%); and CMS4, mesenchymal (23%).[42, 43] However, the relevance of these revised classifications in a clinical setting has yet to be explored.

### **Molecular typing applied to ovarian cancer**

The classic categorization of subtypes of ovarian tumors is based on histological features. When taking molecular data into account, a different classification scheme emerges. Based on mutational profiles, ovarian tumors can be classified in type 1 and type 2 tumors.[44-46] Type 1 tumors consist of low-grade serous carcinomas, low/intermediate-grade endometrioid carcinomas and most clear cell and mucinous carcinomas. Type 2 tumors consist of high-grade serous carcinomas, high-grade endometrioid carcinomas and undifferentiated carcinomas. Type 1 tumors are slow growing and mostly found to be restricted to the ovaries. In addition, in type 1 tumors, precancerous stages (borderline lesions) are identified. Type 2 tumors are fast growing and have often already metastasized at the time of diagnosis. Precancerous lesions of type 2 tumors can be the intra-epithelial neoplasms of the fallopian tube.[47]



The classifications of type 1 and type 2 for endometrioid tumors can also be applied at the molecular level.[48] Type 1 endometrioid tumors often show *PTEN*, *PIK3CA*, *CTNNB1/β-catenin* and *ARID1a* pathogenic variants, while in type 2 endometrioid tumors, *TP53* pathogenic variants are often observed. Thus, low-grade ovarian endometrioid tumors are characterized by mutations that deregulate the *PI3K/PTEN* and the canonical *Wnt/β-cat* pathways and typically lack *TP53* mutations. High-grade ovarian tumors often show *TP53* mutations and lack *Wnt/β-cat* or *PI3K/PTEN* pathway defects. Additional analysis through the Cancer Genome Network revealed several subtypes within the group of high-grade serous ovarian carcinomas.[31]

A similar pattern is seen in serous carcinomas, where pathogenic variants in *KRAS*, *BRAF* and *ERBB2* oncogenes are observed. Inactivating variants in *TP53* are rare in type 1 serous tumors, in contrast with type 2 serous tumors. Ovarian cancer in the context of germline *BRCA1/2* gene variants also shows high grade serous histology. Interestingly, the mutations found in type 1 tumors show similarity with the mutations observed in their precursor lesions (such as borderline tumors and endometriosis). This finding would be an argument in favor of the stepwise development of type 1 tumors.

In mucinous ovarian tumors, *KRAS* pathogenic variants are often present.[49] Clear cell tumors frequently harbor *PIK3CA* and *ARID1a* mutations.[50] Furthermore, deletions in *PTEN* are regularly seen in the clear cell tumors.

Mismatch repair deficiency has been reported in all histological subtypes of ovarian cancer, although it seems most prevalent in endometrioid and mucinous adenocarcinomas.[51, 52] These mismatch repair-deficient tumors may show an improved survival and specific chemosensitivity. *POLE / POLD1* pathogenic variants are reported in a small subset of endometrioid tumors. Additionally, these tumors may be characterized by specific features.[53]

Serous ovarian carcinomas, the histological subtype that is most frequently diagnosed, have been extensively molecularly characterized.[54] However, those studies are still lacking for endometrioid and mucinous tumors.

### **Comparing molecular profiles of carcinomas**

Comparing in-depth mutational profiles of tumors derived from different organs or tissues has made it possible to test whether specific mutational patterns and/or mutation types in different tumor types could be revealed. Although distinctive mutational signatures were discovered, recent studies have shown that the mutational profiles do

---

not differ greatly between tumor types.[54, 55] The few well-known “cancer driver” genes seem to be important in many malignancies. Subsequent or subclonal gene variants that are seen during tumorigenesis are also seen in many tumor types. Looking at the gene variants described in COSMIC (<http://cancer.sanger.ac.uk/cosmic>) or The Cancer Genome Atlas (TCGA, <https://cancergenome.nih.gov/>), similar variants can be seen in both primary ovarian tumors and metastases from CRC, although with different frequencies.[18] *TP53*, *KRAS*, and *PIK3CA* are frequently mutated in both primary ovarian tumors and in metastases from CRC. Only the *APC* gene shows the potential to discriminate between types based on the described mutational profiles in the databases.

Looking at coding and non-coding DNA, but particularly the latter, at least 20 mutational signatures can be distinguished based on the type of DNA variations identified. These variations reflect the lifelong interaction with mutagenic influences.[56, 57] For example, skin exposure to UV radiation and the exposure of cells to mutagens present in tobacco smoke or certain food components is clearly reflected in characteristic DNA variants in tumor DNA. Signs of aging can be seen in typical DNA deamination patterns. In every tumor type, combinations of such mutational signatures are apparent with individual signatures dominating depending on the exposure to a certain mutagen. Unfortunately, in individual tumors DNA signature typing will not unequivocally reflect the origin of the lesion. Combining many “ome” patterns might eventually solve the issue of true typing of tumors. Previously, for the challenging characterization of cases from unknown primary tumors (UPT), alternatively named carcinomas of unknown primary site (CUPs), expression array-based assays were developed in order to identify the primary tumors.[58] So far, expression-based models seem to be the most suitable in determining the site of origin. Mutational analysis of the tumors might primarily play a role in choosing the optimal (targeted) treatment.[59, 60] For synchronously or metachronously presenting tumors at different sites, DNA variant analysis can reveal a clonal relationship. Once the in-depth comparison of all molecular features becomes available, the primary origin may be more easily identified. The use of comprehensive testing in clinical care is in the beginning phase. Technical innovation and novel bioinformatic pipelines should make the enormous amount of data (“big data”) accessible for clinical decision making.

### III. Intra-tumor and inter-tumor heterogeneity: differences between primary tumors and metastases

Although tumors arise as clonal outgrowths from one cell, they consist of a heterogeneous population of cells years later. Differences between cells within one tumor are described with the term intra-tumor heterogeneity (ITH).[61] General oncologic treatments such as radio- or chemotherapy that target tumors as one single entity will have different effects on heterogeneous cell populations within a tumor.[62] This finding also explains the often observed differential responses of tumors to such therapies, with certain tumor cells being less sensitive, resulting in residual tumor fractions and/or tumor recurrence.

In-depth NGS analysis has taught us that a tumor consists of multiple subclones, each with its own mutational profile.[63] To visualize the composition of a tumor, the comparison to a tree with all its branches is often made. The trunk of the tree represents the early “tumor-driving” gene variants, whereas the branches represent the different tumor subclones all originating from tumor cells with different additional gene variants.[64]

The effectiveness of targeted therapies is dependent on the presence of targetable gene variants in the tumor cells. Hypothetically, once all tumor cells carry a targetable variant, a complete response by the tumor can be expected, although the small molecules used to target the identified molecules will inhibit signal transduction and not be lethal per se. When targeting therapies were first introduced, there was hope that specifically targeting genetic variants would result in spectacular reduction of tumor load. In some cases, good initial results were achieved.[65, 66] However, the lack of long lasting responses to targeted therapy could be explained by the complex and hybrid mutational profiles of tumors.[67] Swanton et al. showed that sequencing different regions within renal cell carcinomas resulted in specific mutational profiles that differed per region.[68] Such “spatial tumor heterogeneity” can be seen within one tumor but also exists when comparing different metastatic sites. This finding would explain the differential responsiveness /resistance at different metastatic sites. Many studies have investigated the concordance of genomic variants between primary tumors and their metastases.[70] These studies did increase the understanding of the biological behavior of tumors. Primary tumors consist of large numbers of subclones, of which only a limited number of clones will show metastatic potential.

Previously, one single biopsy, often of the primary tumor, was believed to be representative of the entire tumor process.[69] The targeted treatment strategy was chosen based on the molecular profile of single biopsies. Currently, studies have been conducted that determine the mutation profiles of tumors at different regions or metastatic sites. Most early driver genes, such as *KRAS* and *P53*, show considerable overlap.[63, 70, 71] However, in lung carcinoma, additional mutated driver genes are de-

tected after subclonal diversification.[72] In addition, a large cohort of passenger genes differs per tumor location. An option to address subclonality patterns could also be through the analysis of circulating free tumor DNA or circulating single cells.[73] Mutational profiles of tumors change over time, which is called “temporal tumor heterogeneity”.[67] Repeated analysis is then needed to address this phenomenon. Recent research has taught us that in the first stage(s) of tumor development, subclones are already present that show resistance to targeted therapies.

## IV. Technical considerations using NGS

Next-generation sequencing has dramatically changed the field of molecular diagnostics. High-throughput approaches deliver sequencing results in a fast and cost-effective way. Generating large-scale DNA sequences by next generation sequencing takes distinct steps.[74] First, DNA has to be processed in what is termed the “library preparation”. For clinical purposes, targeted strategies are most often used.[75] In this way, only the genomic regions of interest, where mutations are known to frequently occur, are sequenced. By doing so, the bioinformatics analysis and interpretation of the data are less complex, as the number of variants that is detected will be limited in comparison to whole-genome sequencing. Furthermore, targeted sequencing requires less sequencing capacity and allows a higher throughput. For target enrichment, several approaches are available, based on hybridization, circularization or PCR.[75] Each approach has its pros and cons. Hybridization is most suitable for targeting large regions, while PCR shows better results for smaller targets. However, a larger amount of DNA is necessary with hybridization approaches than with PCR-based technologies. DNA of average quality can be used in PCR-based approaches. Very specific amplification of DNA regions can be obtained by circularization approaches.

During the process of targeting DNA, patient-specific barcodes can be added so that multiple samples can be analyzed in parallel, reducing costs. Additionally, molecular barcodes or single molecule tags can be added that mark each original template molecule. In this way, PCR duplicates, produced during the library preparation, can be distinguished from independent reads originating from original template molecules.[76]

A specific single molecule tag in each probe is informative to identify independent biological template molecules.[77] In targeted approaches, all targeted regions should be amplified in the same order. However, these are complex interactive chemical processes, which can lead to over- or under amplification of certain targets. Next, as the number of genes that are clinically relevant will increase, new genes of interest will have to be added in the enrichment step. As multiple primers can interact with each other, updated gene panels will have to be validated to control their performance.

After enrichment of the regions of interest, DNA can be sequenced. Multiple sequencing platforms are available, based on different techniques.[74, 78] Optical read-outs as the result of the incorporation of fluorescent nucleotides are most commonly used. With pyrosequencing, pyrophosphates are released and measured after the incorporation of a base. A non-optical method is semiconductor sequencing, which meas-

ures hydrogen ions that are released during the polymerization of DNA.

Each sequencing device has its advantages. In a clinical setting, fast turn-around times are desired. The Ion-PGM machine, based on semiconductor sequencing, delivers results in a couple of hours.[79, 80] However, this technique is more prone to making mistakes in small repetitive sequences. On the other hand, optical sequencing methods do not show these failures but have much higher turnaround times.

The last step in NGS, the analysis of the generated data, is the most complex.[81, 82] Next-generation sequencing produces an enormous amount of sequencing data. However, this huge amount of data causes problems with correct interpretation of the data. For example, single nucleotide polymorphisms (SNPs) are non-pathogenic variants that are present throughout the population. Many SNPs are detected with NGS. However, as no database exists in which biologically proven “true” SNPs are archived, these single nucleotide variants are difficult to evaluate. Therefore, parallel sequencing of normal tissue is useful in evaluating these variants.

Another problematic issue is thresholds. In analyzing NGS data, it is necessary to have thresholds to filter out false-positive data. A certain amount of reads with mutant alleles is desired for reliable mutation calling. However, the exact number of mutant reads that is necessary to call a mutation “true” is not known. Rules for thresholds are difficult to establish, as such numbers are also dependent on coverage that can differ not only per gene but also per experiment.

Finally, the interpretation of data creates challenges. With NGS, variants are detected in many genes for which no functional data are available. One example is the gene *FAT4*. This gene is frequently reported to be mutated in glioblastomas, colorectal carcinomas and head and neck carcinomas.[83] However, a clear mechanism by which *FAT4* is involved in colorectal carcinogenesis is not known. For these recently discovered genes, functional tests are necessary. Currently, mutational profiles of tumor types are formed by “census genes”, mutations in which have been causally implicated in cancer.[84] Genes that are known to be frequently mutated but for which functional data are lacking are not mentioned in the mutation profiles. It might be that future experiments reveal that previously unrecognized genes play an important role in tumorigenesis. These functional experiments are crucial to determine the role of mutated genes because the presence of genetic variants within a gene does not imply a carcinogenic effect. For example, the gene *TTN* (Titin) is a very large gene consisting of 363 exons that encodes the Titin protein. This protein is important in the contraction of striated muscle tissues, and due to the size of the gene, it shows very frequent genomic variations. However, these variants are probably sequencing artifacts or SNPs, as variants in *TTN* are not linked to carcinogenic processes.[85]

## Thesis Outline

Distinguishing between primary and secondary ovarian tumors (metastases) based on histological and immunohistochemical features is a known diagnostic problem.

**Chapter 2** describes a cohort of CRC and duodenal cancer cases that presented with metastases to the ovaries. The characteristics of this cohort, including the germline *APC* status, were investigated.

In **chapter 3** the comparison between the mutational profiles of primary ovarian tumors versus secondary ovarian tumor (metastases) were explored. Mucinous and endometrioid primary ovarian tumors were selected as these subtypes pose diagnostic difficulties in the differentiation from metastases of the gastrointestinal tract. A gene panel consisting of 115 genes was used for next generation sequencing (NGS). Besides, loss of heterozygosity (LOH) and methylation of the *APC* gene were investigated.

**Chapter 4** describes the comparison between the mutation profiles of primary colorectal tumor and the matching metastases to the ovaries. The same gene panel as described above was used to generate mutation profiles of the primary CRC and the matching metastases to the ovaries. After extensive bioinformatic analysis overlap and differences in mutations, in correlation with the time between detection of the primary tumor and metastasis, was studied.

In **chapter 5** two different targeting techniques were examined. The HaloPlex target enrichment (based on circularization) and the Ampliseq technique (based on PCR) were compared for efficiency, number of reads, and detection of variants.

**Chapter 6** gives a description of a patient that shows the complexity of the diagnostic difficulties of ovarian tumors and how molecular analysis can be helpful in achieving the right diagnosis.

---

## References

1. Prat, J., *Ovarian carcinomas, including secondary tumors: diagnostically challenging areas*. *Mod. Pathol*, 2005. **18 Suppl 2**: p. S99-111.
2. Bruin, S.C., et al., *Molecular alterations associated with liver metastases development in colorectal cancer patients*. *Br. J. Cancer*, 2011. **105(2)**: p. 281-287.
3. Naylor, G.M., et al., *Why does Japan have a high incidence of gastric cancer? Comparison of gastritis between UK and Japanese patients*. *Gut*, 2006. **55(11)**: p. 1545-52.
4. Young, R.H., *From Krukenberg to today: the ever present problems posed by metastatic tumors in the ovary. Part II*. *Adv. Anat. Pathol*, 2007. **14(3)**: p. 149-177.
5. Young, R.H. and R.E. Scully, *Metastatic tumors in the ovary: a problem-oriented approach and review of the recent literature*. *Semin. Diagn. Pathol*, 1991. **8(4)**: p. 250-276.
6. Yemelyanova, A.V., et al., *Distinction of primary and metastatic mucinous tumors involving the ovary: analysis of size and laterality data by primary site with reevaluation of an algorithm for tumor classification*. *Am. J. Surg. Pathol*, 2008. **32(1)**: p. 128-138.
7. Lee, K.R. and R.H. Young, *The distinction between primary and metastatic mucinous carcinomas of the ovary: gross and histologic findings in 50 cases*. *Am. J. Surg. Pathol*, 2003. **27(3)**: p. 281-292.
8. Ji, H., et al., *Cytokeratins 7 and 20, Dpc4, and MUC5AC in the distinction of metastatic mucinous carcinomas in the ovary from primary ovarian mucinous tumors: Dpc4 assists in identifying metastatic pancreatic carcinomas*. *Int. J. Gynecol. Pathol*, 2002. **21(4)**: p. 391-400.
9. Groisman, G.M., A. Meir, and E. Sabo, *The value of Cdx2 immunostaining in differentiating primary ovarian carcinomas from colonic carcinomas metastatic to the ovaries*. *Int. J. Gynecol. Pathol*, 2004. **23(1)**: p. 52-57.
10. Kir, G., et al., *Clinicopathologic and immunohistochemical profile of ovarian metastases from colorectal carcinoma*. *World J Gastrointest Surg*, 2010. **2(4)**: p. 109-16.
11. Petru, E., et al., *Nongenital cancers metastatic to the ovary*. *Gynecol. Oncol*, 1992. **44(1)**: p. 83-86.
12. Bhati, A., et al., *Omics of cancer*. *Asian Pac J Cancer Prev*, 2012. **13(9)**: p. 4229-33.
13. Garay, J.P. and J.W. Gray, *Omics and therapy - a basis for precision medicine*. *Mol Oncol*, 2012. **6(2)**: p. 128-39.
14. Kim, M.S., et al., *A draft map of the human proteome*. *Nature*, 2014. **509(7502)**: p. 575-81.
15. Munkley, J., I.G. Mills, and D.J. Elliott, *The role of glycans in the development and progression of prostate cancer*. *Nat Rev Urol*, 2016. **13(6)**: p. 324-33.
16. Bentley, D.R., *The Human Genome Project—an overview*. *Med Res Rev*, 2000. **20(3)**: p. 189-96.
17. Hadd, A.G., et al., *Targeted, high-depth, next-generation sequencing of cancer genes in*



- formalin-fixed, paraffin-embedded and fine-needle aspiration tumor specimens*. J. Mol. Diagn, 2013. **15**(2): p. 234-247.
18. Bamford, S., et al., *The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website*. Br. J. Cancer, 2004. **91**(2): p. 355-358.
  19. *Improving databases for human variation*. Nat Methods, 2016. **13**(2): p. 103.
  20. Sherry, S.T., et al., *dbSNP: the NCBI database of genetic variation*. Nucleic Acids Res, 2001. **29**(1): p. 308-11.
  21. Landrum, M.J., et al., *ClinVar: public archive of interpretations of clinically relevant variants*. Nucleic Acids Res, 2016. **44**(D1): p. D862-8.
  22. Brookes, A.J. and P.N. Robinson, *Human genotype-phenotype databases: aims, challenges and opportunities*. Nat Rev Genet, 2015. **16**(12): p. 702-15.
  23. Muers, M., *Gene expression: Transcriptome to proteome and back to genome*. Nat Rev Genet, 2011. **12**(8): p. 518.
  24. Bejjani, B.A. and L.G. Shaffer, *Application of array-based comparative genomic hybridization to clinical diagnostics*. J Mol Diagn, 2006. **8**(5): p. 528-33.
  25. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics*. Nat Rev Genet, 2009. **10**(1): p. 57-63.
  26. Ozsolak, F. and P.M. Milos, *RNA sequencing: advances, challenges and opportunities*. Nat Rev Genet, 2011. **12**(2): p. 87-98.
  27. Tattini, L., R. D'Aurizio, and A. Magi, *Detection of Genomic Structural Variants from Next-Generation Sequencing Data*. Front Bioeng Biotechnol, 2015. **3**: p. 92.
  28. Virani, S., et al., *Cancer epigenetics: a brief review*. Ilar j, 2012. **53**(3-4): p. 359-69.
  29. Torano, E.G., et al., *Global DNA hypomethylation in cancer: review of validated methods and clinical significance*. Clin Chem Lab Med, 2012. **50**(10): p. 1733-42.
  30. Ghirlando, R. and G. Felsenfeld, *CTCF: making the right connections*. Genes Dev, 2016. **30**(8): p. 881-91.
  31. *Integrated genomic analyses of ovarian carcinoma*. Nature, 2011. **474**(7353): p. 609-15.
  32. Kang, G.H., *Four molecular subtypes of colorectal cancer and their precursor lesions*. Arch Pathol Lab Med, 2011. **135**(6): p. 698-703.
  33. Pino, M.S. and D.C. Chung, *The chromosomal instability pathway in colon cancer*. Gastroenterology, 2010. **138**(6): p. 2059-2072.
  34. *Comprehensive molecular characterization of human colon and rectal cancer*. Nature, 2012. **487**(7407): p. 330-337.
  35. Boland, C.R. and A. Goel, *Microsatellite instability in colorectal cancer*. Gastroenterology, 2010. **138**(6): p. 2073-2087 e3.
  36. Haugen, A.C., et al., *Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer*. Cancer Res, 2008. **68**(20): p. 8465-72.
  37. Adam, R., et al., *Exome Sequencing Identifies Biallelic MSH3 Germline Mutations as a Recessive Subtype of Colorectal Adenomatous Polyposis*. Am J Hum Genet, 2016. **99**(2):

- p. 337-51.
38. Briggs, S. and I. Tomlinson, *Germline and somatic polymerase epsilon and delta mutations define a new class of hypermutated colorectal and endometrial cancers*. J Pathol, 2013. **230**(2): p. 148-53.
  39. Nazemalhosseini Mojarad, E., et al., *The CpG island methylator phenotype (CIMP) in colorectal cancer*. Gastroenterol Hepatol Bed Bench, 2013. **6**(3): p. 120-8.
  40. Sinicrope, F.A., et al., *Molecular markers identify subtypes of stage III colon cancer associated with patient outcomes*. Gastroenterology, 2015. **148**(1): p. 88-99.
  41. Muller, M.F., A.E. Ibrahim, and M.J. Arends, *Molecular pathological classification of colorectal cancer*. Virchows Arch, 2016. **469**(2): p. 125-34.
  42. Dienstmann, R., et al., *Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer*. Nat Rev Cancer, 2017.
  43. Guinney, J., et al., *The consensus molecular subtypes of colorectal cancer*. Nat Med, 2015. **21**(11): p. 1350-6.
  44. Bell, D.A., *Origins and molecular pathology of ovarian cancer*. Mod. Pathol, 2005. **18 Suppl 2**: p. S19-S32.
  45. Gorringer, K.L. and I.G. Campbell, *Large-scale genomic analysis of ovarian carcinomas*. Mol. Oncol, 2009. **3**(2): p. 157-164.
  46. McCluggage, W.G., *Morphological subtypes of ovarian carcinoma: a review with emphasis on new developments and pathogenesis*. Pathology, 2011. **43**(5): p. 420-32.
  47. Koshiyama, M., N. Matsumura, and I. Konishi, *Recent concepts of ovarian carcinogenesis: type I and type II*. Biomed Res Int, 2014. **2014**: p. 934261.
  48. Leong, H.S., et al., *Efficient molecular subtype classification of high-grade serous ovarian cancer*. J Pathol, 2015. **236**(3): p. 272-7.
  49. Brown, J. and M. Frumovitz, *Mucinous tumors of the ovary: current thoughts on diagnosis and management*. Curr Oncol Rep, 2014. **16**(6): p. 389.
  50. Cobb, L.P., et al., *Adenocarcinoma of Mullerian origin: review of pathogenesis, molecular biology, and emerging treatment paradigms*. Gynecol Oncol Res Pract, 2015. **2**: p. 1.
  51. Xiao, X., D.W. Melton, and C. Gourley, *Mismatch repair deficiency in ovarian cancer — molecular characteristics and clinical implications*. Gynecol Oncol, 2014. **132**(2): p. 506-12.
  52. Murphy, M.A. and N. Wentzensen, *Frequency of mismatch repair deficiency in ovarian cancer: a systematic review This article is a US Government work and, as such, is in the public domain of the United States of America*. Int J Cancer, 2011. **129**(8): p. 1914-22.
  53. Hoang, L.N., et al., *Polymerase Epsilon Exonuclease Domain Mutations in Ovarian Endometrioid Carcinoma*. Int J Gynecol Cancer, 2015. **25**(7): p. 1187-93.
  54. Zhang, Z., et al., *Molecular Subtyping of Serous Ovarian Cancer Based on Multi-omics Data*. Sci Rep, 2016. **6**: p. 26001.
  55. Vogelstein, B., et al., *Cancer genome landscapes*. Science, 2013. **339**(6127): p. 1546-

- 58.
56. Alexandrov, L.B., et al., *Signatures of mutational processes in human cancer*. Nature, 2013. **500**(7463): p. 415-421.
57. Alexandrov, L.B., et al., *Deciphering signatures of mutational processes operative in human cancer*. Cell Rep, 2013. **3**(1): p. 246-59.
58. Monzon, F.A. and T.J. Koen, *Diagnosis of metastatic neoplasms: molecular approaches for identification of tissue of origin*. Arch. Pathol. Lab Med, 2010. **134**(2): p. 216-224.
59. Economopoulou, P., et al., *Cancer of Unknown Primary origin in the genomic era: Elucidating the dark box of cancer*. Cancer Treat Rev, 2015. **41**(7): p. 598-604.
60. Ross, J.S., et al., *Comprehensive Genomic Profiling of Carcinoma of Unknown Primary Site: New Routes to Targeted Therapies*. JAMA Oncol, 2015. **1**(1): p. 40-9.
61. Swanton, C., *Intratumor heterogeneity: evolution through space and time*. Cancer Res, 2012. **72**(19): p. 4875-4882.
62. Brannon, A.R., et al., *Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic colorectal cancer lesions*. Genome Biol, 2014. **15**(8): p. 454-0454.
63. Murugaesu, N., S.K. Chew, and C. Swanton, *Adapting clinical paradigms to the challenges of cancer clonal evolution*. Am J Pathol, 2013. **182**(6): p. 1962-71.
64. Swanton, C., *Intratumor heterogeneity: evolution through space and time*. Cancer Res, 2012. **72**(19): p. 4875-4882.
65. Holderfield, M., et al., *Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond*. Nat Rev Cancer, 2014. **14**(7): p. 455-67.
66. Hiley, C.T. and C. Swanton, *Spatial and temporal cancer evolution: causes and consequences of tumour diversity*. Clin Med (Lond), 2014. **14** Suppl 6: p. s33-7.
67. Schmitt, M.W., L.A. Loeb, and J.J. Salk, *The influence of subclonal resistance mutations on targeted cancer therapy*. Nat Rev Clin Oncol, 2016. **13**(6): p. 335-47.
68. Gerlinger, M., et al., *Intratumor heterogeneity and branched evolution revealed by multi-region sequencing*. N. Engl. J. Med, 2012. **366**(10): p. 883-892.
69. Baas, J.M., et al., *Concordance of predictive markers for EGFR inhibitors in primary tumors and metastases in colorectal cancer: a review*. Oncologist, 2011. **16**(9): p. 1239-1249.
70. Vignot, S., et al., *Next-generation sequencing reveals high concordance of recurrent somatic alterations between primary tumor and metastases from patients with non-small-cell lung cancer*. J. Clin. Oncol, 2013. **31**(17): p. 2167-2172.
71. Vakiani, E., et al., *Comparative genomic analysis of primary versus metastatic colorectal carcinomas*. J Clin Oncol, 2012. **30**(24): p. 2956-62.
72. de Bruin, E.C., et al., *Spatial and temporal diversity in genomic instability processes defines lung cancer evolution*. Science, 2014. **346**(6206): p. 251-6.
73. Cabel, L., et al., *Circulating tumor cells: clinical validity and utility*. Int J Clin Oncol, 2017.

74. Morey, M., et al., *A glimpse into past, present, and future DNA sequencing*. Mol Genet Metab, 2013. **110**(1-2): p. 3-24.
75. Mertes, F., et al., *Targeted enrichment of genomic DNA regions for next-generation sequencing*. Brief Funct Genomics, 2011. **10**(6): p. 374-86.
76. Eijkelenboom, A., et al., *Reliable Next-Generation Sequencing of Formalin-Fixed, Paraffin-Embedded Tissue Using Single Molecule Tags*. J Mol Diagn, 2016. **18**(6): p. 851-863.
77. Jabara, C.B., et al., *Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID*. Proc Natl Acad Sci U S A, 2011. **108**(50): p. 20166-71.
78. Mardis, E.R., *Next-generation sequencing platforms*. Annu Rev Anal Chem (Palo Alto Calif), 2013. **6**: p. 287-303.
79. Singh, R.R., et al., *Clinical validation of a next-generation sequencing screen for mutational hotspots in 46 cancer-related genes*. J Mol Diagn, 2013. **15**(5): p. 607-22.
80. Beadling, C., et al., *Combining highly multiplexed PCR with semiconductor-based sequencing for rapid cancer genotyping*. J. Mol. Diagn, 2013. **15**(2): p. 171-176.
81. Shen, T., et al., *Clinical applications of next generation sequencing in cancer: from panels, to exomes, to genomes*. Front Genet, 2015. **6**: p. 215.
82. Xuan, J., et al., *Next-generation sequencing in the clinic: promises and challenges*. Cancer Lett, 2013. **340**(2): p. 284-95.
83. Morris, L.G., D. Ramaswami, and T.A. Chan, *The FAT epidemic: a gene family frequently mutated across multiple human cancer types*. Cell Cycle, 2013. **12**(7): p. 1011-2.
84. Futreal, P.A., et al., *A census of human cancer genes*. Nat Rev Cancer, 2004. **4**(3): p. 177-83.
85. Tamborero, D., et al., *Comprehensive identification of mutational cancer driver genes across 12 tumor types*. Sci Rep, 2013. **3**: p. 2650.

*Decision tree based on immunohistochemical staining patterns. A first classification ► in four subgroups can be made on keratin 7 (CK7) and keratin 20 (CK20) positivity or negativity. Colonic metastasis mostly shown a CK7-/CK20+ pattern (A). The assumption of a colonic origin can be strengthened by other markers. Primary ovarian tumors mostly shown a CK7+/CK20- pattern (D). Besides, metastases to the ovaries from other locations also can show a CK7+/CK20- pattern, but can be identified by additional markers. A CK7+/CK20+ pattern (B) is not discriminating between a primary origin or a metastatic process. A CK7-/CK20- pattern (C) is very uncommon.*

Figure 1

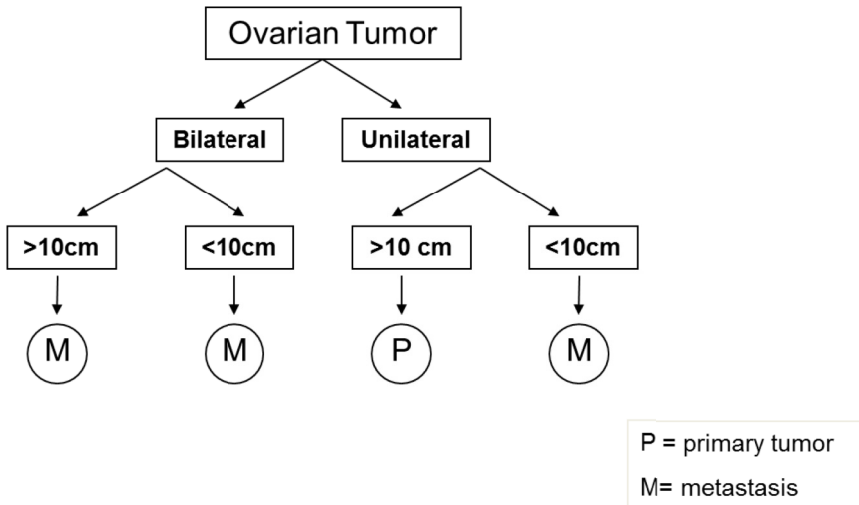
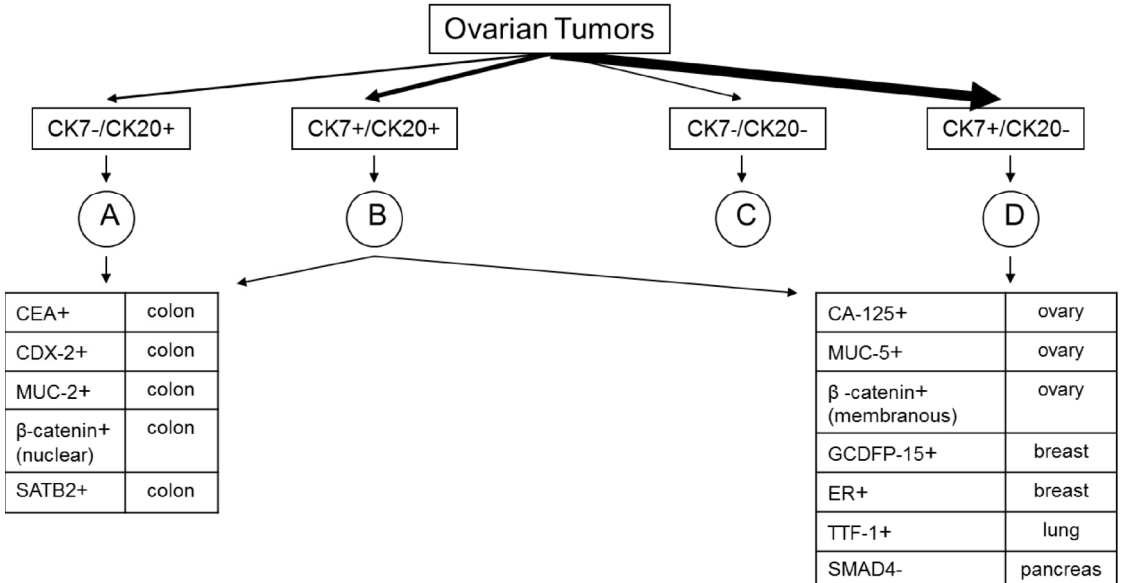


Figure 2





## Chapter 2

### **Ovarian metastases of colorectal and duodenal cancer in Familial Adenomatous Polyposis**

Stijn Crobach<sup>1</sup>, Tom van Wezel<sup>1</sup>, Hans F Vasen<sup>2</sup>, and Hans Morreau<sup>1</sup>

<sup>1</sup>Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands

<sup>2</sup>Department of Gastroenterology, University Medical Center, Leiden, the Netherlands

## **Abstract**

Metastases to the ovary occur in 0,8% - 9.7 % of colorectal cancer (CRC) cases.[1] The need to combine surgical resection of the primary tumor and bilateral oophorectomy is a matter of debate.[2] In a consecutive multi-hospital cohort of 30 CRC metastases to the ovary we came across four female patients (13%; 95% CI 3,6 – 34,1) with Familial Adenomatous Polyposis (FAP). This number is high since the estimated incidence of FAP CRC is far below 1% of all CRC and the expected incidence of FAP CRC that metastasized to the ovaries would thus be almost zero. In a second screen in nationwide databases we found that ovarian metastases occurred in at least 15% of female FAP CRC cases. We provide now first evidence that especially in female FAP CRC patients bilateral oophorectomy during surgery should be discussed.



## Main text

Reason for resection of the ovaries in case of a CRC is the chance to have synchronous or develop metachronous metastases. The highest percentages are found in autopsy series.[3, 4] In a multi-hospital based cohort of 30 CRC metastases to the ovary we came across four female patients (13%) with FAP at ages 34, 48, 56 and 56, respectively (Table 1). Two of four FAP patients (patients 1 and 2) with ovarian metastases presented with stage IV colon cancer (T3/N2/M1 and T4/N1/M1) at first surgery with synchronous metastases in lymph nodes, omental fat (one of two patients), liver (one of two patients) and ovaries. The remaining two patients (patients 3 and 4) presented with stage III and II cancer (T4/N2/M0 and T3/N0/M0) with ovarian metastases only metachronously occurring at two and eight years after first surgery, respectively.

As stated before, the estimated incidence of FAP CRC is low (far <1 % of all CRC). This number probably even decreased with stringent endoscopic surveillance and prophylactic colectomies. We hypothesized that ovarian metastases might be more common in female FAP-patients than in sporadic patients with CRC. The metastasis incidence and distribution in FAP CRC was not described in the last version of the WHO Classification of Tumours of the Digestive System (2010).[5] To address this matter we searched the literature for female FAP CRC and small bowel cancer cases (Table 2).

Three female FAP-patients with ovarian metastases were described.[6-8] The ovarian metastases originated from a rectum carcinoma, a colon ascendens tumor and from an unspecified location in the colon, respectively. Other site(s) of distant metastasis was lung in one of these cases. Furthermore we searched the files of the Netherlands Foundation for the Detection of Hereditary Tumors (NFDHT) for female FAP patients and crossed these data with data from PALGA; the nationwide Dutch network and registry of histopathology and cytopathology.[9] Of 575 FAP-patients, registered during the period 1971 till now, 63 had a history of a malignancy in the gastro-intestinal tract. Thirty female FAP patients were identified either with a colorectal carcinoma (27 patients) or a duodenal carcinoma (3 patients). Intramucosal lesions were excluded. Of the 27 female CRC patients (apart from the 4 cases already known to us and described above) no additional patient was documented with pathologically verified ovarian metastases, making the incidence at least 15% (4/27; 95% CI 4,0 – 37,9). The remaining 23 FAP CRC patients in the PALGA cohort did present with lymph node and liver metastases in 11 and 5 cases, respectively. In one patient metastasis to the femur occurred. No lung or metastases at other sites were de-

scribed. None of the three duodenal cancer cases showed distant metastases.

A decisive explanation for the relatively high frequency of ovarian metastases in female FAP-patients with CRC cannot be given. The route of dissemination of gastrointestinal tumors to the ovaries is a topic of discussion in the literature.[10] Different options are considered: dissemination through the lymph-angiogenic system or through direct peritoneal spreading. Tumor extension through the serosal surface (T-stage T4) would increase the chance of ovarian metastases. Looking at the seven cases, now described by us and in previous literature (supplemental Table) three out of seven patients were diagnosed with a T-stage T4 carcinoma. An option that also would explain the relatively high frequency of ovarian metastases in female FAP-patients is the overall cancer stage at presentation. Such details were not always evident for the patients included in our series. At least three out of seven patients presented with synchronous ovarian metastasis (stage IV). Whether the average (overall) tumor stage of female FAP patients with CRC at initial presentation is higher than in matched consecutive CRC series, is however unclear.

A possible link between previous colorectal surgery and the occurrence of ovarian metastases has not been reported in the literature. Also looking at the detailed description of the cases compiled in Table 2 the existence of such association could not be found (supplemental Table).

Two of four FAP patients described by us were possibly premenopausal. In the literature at least 11 studies have looked into the difference between pre- and postmenopausal status in relation to the incidence of ovarian metastases.[11-21] No significant difference between pre- and postmenopausal women seems evident in patient groups with ovarian metastases of CRC.

Next, CRC cancers in FAP-patients might have a slightly different behaviour in comparison to sporadic colorectal tumors. In FAP and sporadic CRC the principle of '*just-right signaling*' of the Wnt pathway as described by Fodde et al., plays an important role.[22] In this theory the altered signaling of APC through betacatenin binding must fulfill the rule that at least one betacatenin binding site is preserved in the cancer cell. Wnt signaling might slightly differ between FAP CRC and sporadic CRC, since the genetic hits on APC are different.

In conclusion: we found a relatively high percentage of female FAP-CRC that metastasized to the ovary. The overall estimated 15 percent of ovarian metastasis of female FAP CRC is above 0,8% - 9.7% that is reported in sporadic CRC. However, the cohort on which the estimated 15 percent is based is relatively small, leading to large confidence intervals. Furthermore an independent confirmatory series is needed. We provide now first evidence that especially in female FAP CRC patients bilateral oophorectomy during surgery should be discussed.

## References

- (1) Hanna NN, Cohen AM. Ovarian neoplasms in patients with colorectal cancer: understanding the role of prophylactic oophorectomy. *Clin Colorectal Cancer* 2004 Feb;3(4):215-22.
- (2) Erroi F, Scarpa M, Angriman I, Cecchetto A, Pasetto L, Mollica E, Bettiol M, Ruffolo C, Polese L, Cillo U, D'Amico DF. Ovarian metastasis from colorectal cancer: prognostic value of radical oophorectomy. *J Surg Oncol* 2007 Aug 1;96(2):113-7.
- (3) Fujiwara K, Ohishi Y, Koike H, Sawada S, Moriya T, Kohno I. Clinical implications of metastases to the ovary. *Gynecol Oncol* 1995 Oct;59(1):124-8.
- (4) Abrams HL, Sprio R, Goldstein N. Metastases in carcinoma; analysis of 1000 autopsied cases. *Cancer* 1950 Jan;3(1):74-85.
- (5) Bosman FT, Carneiro F, Hruban RH, Theise ND. WHO Classification of Tumours of the Digestive System, Fourth Edition. 2010.
- (6) Miyaki M, Iijima T, Konishi M, Sakai K, Ishii A, Yasuno M, Hishima T, Koike M, Shitara N, Iwama T, Utsunomiya J, Kuroki T, Mori T. Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. *Oncogene* 1999 May 20;18(20):3098-103.
- (7) Hosogi H, Nagayama S, Kanamoto N, Yoshizawa A, Suzuki T, Nakao K, Sakai Y. Biallelic APC inactivation was responsible for functional adrenocortical adenoma in familial adenomatous polyposis with novel germline mutation of the APC gene: report of a case. *Jpn J Clin Oncol* 2009 Dec;39(12):837-46.
- (8) Donnellan KA, Bigler SA, Wein RO. Papillary thyroid carcinoma and familial adenomatous polyposis of the colon. *Am J Otolaryngol* 2009 Jan;30(1):58-60.
- (9) Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de Pol A, van Krieken JH, Meijer GA. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol* 2007;29(1):19-24.
- (10) Sakakura C, Hagiwara A, Kato D, Hamada T, Yamagishi H. Manifestation of bilateral huge ovarian metastases from colon cancer immediately after the initial operation: report of a case. *Surg Today* 2002;32(4):371-5.
- (11) Kesic V, Radmila S. Surgery of ovarian metastases: special considerations. *J Gynecol Oncol* 2006;11:93-100.
- (12) Talebpour M, Zargar M. Synchronous surgical removal of suspicious ovarian metastases from colorectal cancer. *Med J Iran* 2005 Feb 18;4:285-288.
- (13) Moore RG, Chung M, Granai CO, Gajewski W, Steinhoff MM. Incidence of metastasis to the ovaries from nongenital tract primary tumors. *Gynecol Oncol* 2004 Apr;93(1):87-91.
- (14) McGill F, Ritter DB, Rickard C, Kaleya RN, Wadler S, Greston WM. Management of Krukenberg tumors: an 11-year experience and review of the literature. *Prim Care Update Ob Gyns* 1998 Jul 1;5(4):157-8.

- (15) Sielezneff I, Salle E, Antoine K, Thirion X, Brunet C, Sastre B. Simultaneous bilateral oophorectomy does not improve prognosis of postmenopausal women undergoing colorectal resection for cancer. *Dis Colon Rectum* 1997 Nov;40(11):1299-302.
- (16) Cutait R, Lesser ML, Enker WE. Prophylactic oophorectomy in surgery for large-bowel cancer. *Dis Colon Rectum* 1983 Jan;26(1):6-11.
- (17) Birnkrant A, Sampson J, Sugarbaker PH. Ovarian metastasis from colorectal cancer. *Dis Colon Rectum* 1986 Nov;29(11):767-71.
- (18) Omranipour R, Abasahl A. Ovarian metastases in colorectal cancer. *Int J Gynecol Cancer* 2009 Dec;19(9):1524-8.
- (19) Kim NK, Kim HK, Park BJ, Kim MS, Kim YI, Heo DS, Bang YJ. Risk factors for ovarian metastasis following curative resection of gastric adenocarcinoma. *Cancer* 1999 Apr 1;85(7):1490-9.
- (20) Renaud MC, Plante M, Roy M. Metastatic gastrointestinal tract cancer presenting as ovarian carcinoma. *J Obstet Gynaecol Can* 2003 Oct;25(10):819-24.
- (21) MacKeigan JM, Ferguson JA. Prophylactic oophorectomy and colorectal cancer in premenopausal patients. *Dis Colon Rectum* 1979 Sep;22(6):401-5.
- (22) Fodde R, Smits R, Clevers H. APC, signal transduction and genetic instability in colorectal cancer. *Nat Rev Cancer* 2001 Oct;1(1):55-67.

**Table 1**

Patient	Age	APC germline mutation	Stage colontumor	Previous surgery before oophorectomy
1	34	c.1192_1193delAA,	Stage IV	No
2	48	c.1548G>C, p.Lys516As	Stage IV	No
3	56	c.646-1G>A	Stage III	Yes
4	56	c.471G>A, p.Trp157X	Stage II	Yes

Overview of four female FAP-patients with metastases of CRC to the ovaries.

**Table 2**

		Number of Carcinomas	Metastases Described	Gender Patients With Metastases	Location Primary Tumor
Parc	2004	11	1 lung 1 liver	unknown	10 colon 1 rectum
He	2004	5	1 liver	unknown	colorectum
Vitelaro	2011	4	1 liver	unknown	colorectum
Campos	2009	53	6 location unknown	unknown	colorectum
Eigenbrod	2006	1	1 liver	female	small bowel
Panis	1996	2	1 liver	female	rectum
Jang	1997	23	11 location unknown	unknown	20 colon 3 rectum
Iizuka	2002	1	1 lymph node	female	ileostoma
Miyaki [8]	1999	1	1 ovary	female	colorectum
Hosogi [6]	2009	1	1 ovary 1 lymph node	female	colon
Donellan [9]	2009	1	1 ovary 1 lung	female	rectum

Literature search of female FAP patients with CRC / small bowel cancer showing metastases.

Supplemental Table

Article	Number of carcinomas	Metastases described	Initial stage of colorectal tumors	Previous surgery	Age	Gender
Parc	11	1 lung 1 liver	Stage I Stage II Stage III	No	Median age 26,5 years (range 10 – 67,5).	Not specified
He	5	1 liver	All at an advanced stage.	No	Average age 38 years.	Female 1 not specified
Vitelaro	4	1 liver	Stage 0 Stage I Stage IIa Stage IIb Stage IIIa Stage IIIb Stage IIIc Stage IV	Patients with previous surgery were excluded.	Median age was 28 years (range 15-68).	Female
Campos	53	6 location unknown	Not mentioned	All patients in this study underwent surgery.	Average age was 40 years.	Not specified
Eigenbrod	1	1 liver	Stage II	Yes	51	Female
			Stage I Stage III Stage IV			
Panis	2	1 liver	Stages of 6 patients were given.	No	Average age was 40,5 years. Age	Female

				Stage I Stage II Stage III	1 2 3					of patient with metastasis were 28 (F) and 34 (M) years, respectively.	
Jang	23	11 location unknown	70% of patients with colon cancer had tumors confined to the bowel wall without nodal or distant metastases. The stage of the initial tumors of patients developing metastases was specified in 5 cases.	Stage II Stage III	4 1	All 11 patients underwent colectomy.	Mean age of patients diagnosed with CRC was 39 years.	Female			
Iizuka	1	1 lymph node	Advanced stage (stage IV)			Colectomy	41	Female			
Article	Number of carcinomas	Metastases described	Initial stage of colorectal tumors	Previous surgery before oophorectomy	Age at diagnosis	Gender					
Miyaki	1	1 ovary	Not mentioned	No	Not mentioned	Female					
Hosogi	1	1 ovary 1 lymph node	Stage IV (T4/N1-2/M1)	No	44	Female					
Donnellan	1	1 ovary 1 lung	Not mentioned	No	30	Female					

Overview of stage, previous surgery, age and gender of patients reported in table 2 in the manuscript.





## Chapter 3

# **Target-Enriched NGS Reveals Differences Between Primary And Secondary Ovarian Tumors in FFPE Tissue**

Stijn Crobach<sup>1</sup>, Dina Ruano<sup>1</sup>, Ronald van Eijk<sup>1</sup>, Gert Jan Fleuren<sup>1</sup>, Ivonne Minderhout<sup>2</sup>, Ronelle Snowdowne<sup>2</sup>, Carli Tops<sup>2</sup>, Tom van Wezel<sup>1</sup>  
and Hans Morreau<sup>1</sup>

<sup>1</sup>Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands

<sup>2</sup>Laboratory for Diagnostic Genome Analysis (LDGA), Leiden University Medical Center, Leiden, the Netherlands

**Abstract**

**Purpose:** Differentiating between primary endometrioid or mucinous ovarian tumors and secondary ovarian tumors (metastases) can be a challenge. Somatic mutation profiles of primary and secondary ovarian cancers were compared to investigate whether these profiles could aid in diagnosing ovarian tumors.

**Experimental design:** 115 cancer related genes were screened by target-enriched next generation sequencing (NGS) in formalin-fixed paraffin-embedded (FFPE) tumor tissue from 43 primary endometrioid and mucinous ovarian carcinomas and 28 proven colorectal cancer (CRC) metastases to the ovary. Validation was performed by High-Resolution Melting Curve Analysis (HRMCA) and Sanger sequencing. LOH and promoter hypermethylation of *APC* were also studied.

**Results:** *TP53*, *NOTCH1*, *PIK3CA* and *FAT4* were the most frequently mutated genes in the primary ovarian tumors. *APC*, *TP53*, *KRAS* and *FAT4* mutations were the most common mutations in ovarian CRC metastases. An inactivating *APC* mutation was found in 4.7% of primary ovarian tumors (2/43; CI -1,6–10,9%). In contrast, inactivating *APC* mutations were identified in 71% of CRC metastases (20/28; CI 55-88%) ( $p < 0.001$ ; sensitivity 71.4%; CI 51,1-86,0%, specificity 95.4%; CI 82,9-99,1%). LOH and *APC* promoter hypermethylation did not differ significantly between the primary and secondary ovarian tumors. *NOTCH1* mutations were observed specifically in primary ovarian tumors, though at a low frequency, but not in metastases (6/41; 14.6%; CI 3,8-25,4%).

**Conclusion:** *APC* mutation analysis can be used to differentiate primary endometrioid and mucinous ovarian tumors from CRC metastases to the ovary.

## Introduction

Ovarian tumors can be subdivided into primary (85%) and secondary tumors (metastases; 15%). For therapeutic and prognostic reasons, it is important to be able to distinguish between these two categories.[1] Roughly half of the primary tumors that give rise to ovarian metastases are located in the gastro-intestinal tract (GIT).[2, 3] Although these are mostly colorectal cancers, tumors in the appendix, stomach or pancreas can also metastasize to the ovary. These metastasizing primary GIT tumors are mostly conventional adenocarcinomas and are often easily recognized as such on basis of histology and immunohistochemical stainings. However, some of the primary ovarian malignancies such as endometrioid and mucinous adenocarcinomas can show extensive histological and immunohistochemical similarities to primary GIT tumors, making a final diagnosis difficult.[4, 5] Other frequently observed subtypes, such as serous papillary carcinomas of the ovary do not pose such problems.

Size, bilaterality and immunohistochemical stainings of ovarian lesions can be used to differentiate between primary and secondary tumors.[6-11] Nevertheless, the tools that are currently available cannot always discriminate adequately and present a risk of equivocal characterization of ovarian tumors. Therefore, finding molecular information that further characterizes some of these ovarian lesions would be helpful in daily practice. The COSMIC database, currently the foremost database of (somatic) mutation profiles [12], shows that similar cancer-driving mutations can be found in several types of cancer, although these mutations are often observed at different frequencies.[13, 14]

In this study, we focused on ovarian metastases of CRC. Based on mutation profiles described in the literature and the COSMIC database, *APC* appears to be the best candidate for discriminating between primary tumors versus metastases of CRC.[4, 15, 16]

115 cancer-related genes (including *APC* and *CTNNB1*) were analyzed in FFPE material, that is used in diagnostics on a daily basis, by target enriched NGS.[17] Also, methylation and loss of heterozygosity (LOH) of *APC* were investigated. In this study, we show that the detection of inactivating *APC* mutations can help to distinguish primary from secondary ovarian tumors.

## Methods

### Tissue Samples

Ovarian metastases of CRC were obtained from the archives of the LUMC Pathology Department (period 1985-2010; n=10) and from PALGA (the nationwide Dutch network and registry of histopathology and cytopathology; n=18).[18] FFPE tissue was available for all 28 ovarian metastases of CRC, fresh frozen for 6 metastases. Next, 35 endometrioid and 7 mucinous primary ovarian cancers were selected, because these subtypes can pose diagnostic problems based on their histologic similarity with CRC metastases to the ovary. One case showed a mixed type histology consisting of an endometrioid tumor with clear cell elements. Both frozen and FFPE tissue was available for all the primary ovarian tumors.

### Medical Consent

The present study falls under approval by the Medical Ethical Committee of the LUMC (protocol P01-019). Informed consent was obtained according to protocols approved by the LUMC Medical Ethical Committee (02-2004). The patient samples were handled according to the medical ethics guidelines described in the Code Proper Secondary Use of Human Tissue established by the Dutch Federation of Medical Sciences (<http://www.federa.org/codes-conduct>; accessed at January 2013).

### Tissue (Micro)-Dissection and DNA Extraction

On basis of pathological examination of hematoxylin and eosin stained slides tumor tissue containing > 50% tumor cells was selected from FFPE-material. Tumor tissue was taken using 0.6-2.0 mm tissue cylinder, depending on the tumor volume (Beecher Instruments, USA). In case tumor fields were too small for punching micro-dissection was performed as follows: 5-10x 10  $\mu$ M slides were cut and stained with hematoxylin only. Eosin staining was omitted to preserve the integrity of the DNA. After staining slides were visualized using an inverted microscope, and microdissected using a sharp-pointed knife. In case frozen tissue was used, tumor enrichment was achieved by removing non-tumorous tissue as much as possible after frozen section analysis. DNA was isolated using the Nucleospin Tissue kit (Bioke, The Netherlands). After isolation the DNA was dissolved in 50-100  $\mu$ l water. The concentration of the DNA ranged from 10-200 ng/ $\mu$ l, for microdissected tumors with dispersed tumor fields and solid tumors, respectively.

### Targeted sequencing in FFPE tissue

To screen 115 genes, a HaloPlex custom made target enrichment kit was designed (Agilent Technologies, Santa Clara, CA). Genes were selected based on the literature

(Supplemental Table 1). [19, 20] Library preparation was performed, with 250 ng as input, according to the standard protocol (Agilent Technologies, Santa Clara, CA). After equimolar pooling, the samples were sequenced on an Illumina HiSeq 2000 sequencer (Illumina, San Diego, USA). An average coverage of 359x was achieved (range 15-1972x; median 355x).

### **Data Analysis**

Adaptors, barcodes and enzyme footprints were removed from the sequenced reads using SureCall software (Technologies, Santa Clara, CA), after which the reads were aligned using BWA.[21] The GATK was used for realignment around indels and base quality recalibration.[22] Removal of duplicates is not necessary in HaloPlex hybridization-extension experiments, used to capture the target DNA regions. SNP and indel calling was carried out using varScan 2 software using the following arguments: minimum read depth = 8, minimum number of reads with the alternative allele = 2, minimum base quality = 20, and minimum variant allele frequency = 0.10. See Supplemental Table 2A and 2D for details on total reads and coverage.

Variants were functionally annotated using ANNOVAR. [23, 24] To remove false positives, three strategies were used. First, the variants with supported reads in only one direction where the wild-type allele had more than 2 reads in both directions were removed. Second, the variants with a frequency higher than 90% in an extended cohort of 102 patients were considered to be false variants and were removed. Third, the variants in known duplicated genomic regions were removed from downstream analysis. Matched normal tissue was not available for many of the samples; therefore variants that were present in the 1000 Genomes project (<http://www.1000genomes.org/>; data from April 2012) or in the NHLBI Exome Sequencing Project (<http://evs.gs.washington.edu/EVS/>; data from January 2013) were removed. We focused on splicing and exonic variants (excluding synonymous). After all the steps described above, a total of 301 variants was obtained.

### **Mutation Detection in Formalin-Fixed Paraffin-Embedded Tissue**

Primer sequences are listed in Table 1. All PCR reactions were performed using a 25 µl final reaction volume, containing 12.5 µl SYBR Green (BioRad), 9.9 µl H<sub>2</sub>O, 0.6 µl primer mix (Eurofins) and 2 µl diluted DNA (approximately 5 ng/µl). For *CTNNB1*, 1.5 µl of primer mix was used. After denaturation at 95°C for 5 min, the templates were amplified for 40 cycles (95°C for 10 s, 60°C for 10 s, and 72°C for 10 s) followed by a 10-min extension at 72°C. A melting curve was obtained to evaluate the quality of the PCR products. After purifying the PCR products using a Qiagen MinElute 96 UF PCR Purification Kit, the PCR amplicons were sequenced by a com-

mercial entity using standard forward and reverse M13 primers (Macrogen). The sequence trace files were analyzed using Mutation Surveyor DNA Variant Analysis software (version 3.97, SoftGenetics).

### **APC Mutation Detection in Primary Ovarian Frozen Tumor Tissue**

A standard PCR protocol, as described above, was used for amplification. High Resolution Melting Curve Analysis (HRMCA) was performed using a LightScanner system (BioFire Diagnostics). Abnormal melting curves were additionally analyzed by Sanger sequencing and analyzed with Seqscape v2.6.

### **Detection of LOH**

Six polymorphisms within *APC* (exon 11: c.1458T>C; exon 15: c.4479G>A, c.5034G>A, c.5268G>T, c.5465A>T and c.5880A>G) were used for probe-based melting curve analysis. To detect the loss of one of the alleles, peak heights of heterozygous alleles were compared for both normal and tumor tissue (Supplemental Figure 1A). Furthermore, the PCR products of the primary ovarian tumors were sequenced to confirm the observed LOH (Supplemental Figure 1B-1C).

### **Methylation analysis of APC**

The promoter region 1A of *APC* was analyzed by methylation-specific PCR (MSP) and bisulfite sequencing. Sodium bisulfite modification of 500 ng DNA was carried out using the EZ DNA methylation kit (Zymo Research, Orange, CA, USA). The amplification efficiency was determined by quantitatively comparing PCR products at specific melting temperatures. To determine whether this read-out was justified, several samples were put on agarose gel. In cases in which the MSP was not conclusive, bisulfite sequencing was performed (primer sequences are listed in Table 1).

### **Immunohistochemistry**

Immunohistochemistry was performed as described previously.[25] The antibodies and dilutions that were used are as follows: caudal homeobox 2/CDX2 (DAKO Denmark; 1:80), paired box gene 8/PAX8 (Protein Tech USA; 1:3200), Cyto-keratin 7/CK7 (DAKO Denmark; 1:1600), Cyto-keratin 20/CK20 (DAKO Denmark; 1:800) and VIMENTIN (DAKO Denmark; 1:2000).

## Results

### Immunohistochemical profiles of primary and secondary ovarian tumors

Primary and secondary ovarian tumors were stained for caudal homeobox 2 (CDX2), paired box gene 8 (PAX8), Cyto-Keratin 7 (CK7), Cyto-Keratin 20 (CK20) and VIMENTIN. Numerous primary ovarian cancers in our cohort gave immunohistochemical staining results that might give rise to discussions about their primary origin. For example 12/39 gave positive CDX2 staining expected to stain positive in CRC. Similarly, not all CRC metastases to the ovary gave results that would support an unequivocal diagnosis (Supplemental Table 2B and 2E).

### Mutation rate per gene in primary and secondary ovarian tumors

Using NGS, the mutation rate of the 115 targeted genes (Supplemental Table 1) was studied in FFPE tissue of both primary ovarian tumors (n=43) and CRC metastases to the ovary (n=28). The NGS results of *APC* were validated using HRMCA and Sanger sequencing in frozen tumor tissue for all primary ovarian tumors and for 6 CRC metastases to the ovary. Because of the absence of frozen tissue of the remaining CRC metastases to the ovary, only mutations within the Mutation Cluster Region of *APC* (MCR; codon 1286-1513) were validated by Sanger sequencing in these cases. Besides, a subset of mutation calls in *TP53* and hotspot mutations in *KRAS*, *BRAF* and *PIK3CA* were confirmed by Sanger sequencing and TaqMan assays, respectively.

Missense and nonsense *TP53* (61%), missense *NOTCH1*, activating *PIK3CA* and missense *FAT4* (all 15%) mutations were the most frequently observed in mucinous and endometrioid primary ovarian tumors (Figure 1 shows absolute mutation numbers). Two primary ovarian cases were excluded from this analysis due to suboptimal tissue processing. In CRC metastases to the ovary, *TP53* was mutated in 64%, *KRAS* in 39% and *FAT4* in 29%. As expected, mutations in *APC* stood out in frequency (71%). The other genes in the selected panel showed low mutation rates. *KRAS* was mutated more often in the ovarian metastases than in the primary ovarian tumors (11 vs. 2 cases), and *NOTCH1* was mutated more frequently in the primary ovarian tumors (14.6%; CI 3,8-25,4%; 6 vs. 0 cases). Although both genes show a significant difference in mutation rate between the two groups, caution must be exercised in interpreting these results because of the low absolute number of mutated cases. Therefore, we focused on the gene that best discriminated between the two groups, *APC*, and included the analysis of *CTNNB1*.

**APC mutations in primary endometrioid and mucinous ovarian tumors**

First, the complete *APC* gene was analyzed by NGS in FFPE tissue. All 7 primary mucinous ovarian tumors were negative for *APC* mutations. In the remaining 36 endometrioid tumors, only 1 inactivating mutation (c.4403A>T; p.Lys1468X), located in the MCR was found (Table 2). The mutation-positive tumor came from a 69-year-old patient, with a unilateral presentation, a diameter of 30 cm and a mixed type histology (endometrioid and clear cell), without a CRC history.

In addition, in 5 individual cases missense variants in *APC* were identified, of which 2 were located in the 3' portion of *APC*, downstream of the MCR (Figure 2 and Supplemental Table 3). In the COSMIC database, in total three of these missense variants were reported (Supplemental Table 4). One variant (c.902C>T; p.Pro301Leu) turned out to be a somatic event after analysis of matching normal DNA. Another variant (c.3949G>C; p.Glu1317Gln) was also identified in normal DNA (germ line variant). Two variants with low coverage (c.6136G>A; p.Ala2046Thr and c.3007G>A; p.Gly1003Asp) could not be confirmed both in normal and tumor tissue, possibly due to detection limits using Sanger sequencing. For one variant (c.8162G>A; p.Arg2721His) the quality of the DNA was insufficient to reliably interpret the Sanger sequences. The pathogenicity of the germline missense variants is still unclear according to an in-house *APC* database used to screen suspected FAP-patients (details available on request). Somatic missense *APC* mutations might have a role in ovarian tumorigenesis.

To confirm the results of the NGS analysis of *APC*, HRMCA (that requires high quality DNA from frozen tissue) and Sanger sequencing were performed. In the validation experiment however, in 2 primary endometrioid ovarian tumors (4,7%; CI -1,6–10,9) extra inactivating *APC* mutations were found, despite sufficient coverage with NGS (Supplemental Table 3; patients 21 and 26). Both mutations were positioned in the 3' portion of the *APC* gene, outside of the MCR (c.4666dup; p.Thr1556fs and c.7356delA; p.Leu2452X). The former mutation was located in a repetitive DNA sequence and was detected in a bilateral ovarian tumor process. Immunohistochemical stains were concordant with a primary ovarian tumor (OC125/Ker7+ and CEA/Ker20-). There was no CRC history in this patient. The second deleterious *APC* mutation (c.7356delA; p.Leu2452X) was found in the ovarian tumor in which the p.Lys1468X mutation was found by NGS.

**APC mutations in ovarian metastases from colorectal tumors**

The *APC* gene was analyzed by NGS in ovarian CRC metastases in FFPE tissue. *APC* mutations were identified in 71% (20/28; CI 55-88%) of the ovarian metastases, a significant difference compared to the primary ovarian tumors (Table 3; p<0.001).



The positive and negative predictive value is 90.9% (CI 69,4-98,4%) and 83.7% (69,7-92,2%), respectively. In total, 29 damaging *APC* mutations were found in the ovarian metastases (Supplemental Table 5). The nonsense and frameshift *APC* mutations were located in the MCR of *APC* or upstream (5' portion) (Figure 2). Furthermore, one missense mutation was observed (c.4092T>A; p.Ser1364Arg). All MCR mutations were confirmed by Sanger sequencing, and no additional variants were detected. After analysis of matching normal DNA the missense variant turned out to be a somatic event of unknown pathogenicity.

In the 6 ovarian metastasis of CRC for which frozen tissue was available, the whole *APC* gene was also investigated with HRMCA and Sanger sequencing. All 5 *APC* mutations detected by NGS, of which 2 were located outside the MCR, were confirmed (Supplemental Table 5; cases 1-6). All the mutation-negative cases were indeed negative.

#### **LOH analysis of *APC* in primary and secondary ovarian tumors**

In CRC, the second hit in *APC* is preferentially copy-neutral LOH or a second somatic mutation. Surprisingly, LOH of *APC* in cases without a mutation was observed in 28% of the primary ovarian tumors (12/43; CI 14,5-41,3%; Table 2 and Supplemental Table 2C). In the ovarian CRC metastases, LOH of *APC* was observed in 32% (9/28; CI 14,8-49,4%; Table 3 and Supplemental Table 2F;  $p=0.79$ ). A further delineation of the mechanism underlying the detected LOH (being either physical loss or copy neutral LOH) could not be obtained due to the method that was used to detect LOH.

#### **Methylation status of *APC* in primary and secondary ovarian tumors**

Hypermethylation of the promoter region of *APC* is an alternative mechanism leading to gene silencing. [26] Promoter hypermethylation of *APC* is reported in 18% of sporadic CRCs and is associated with the loss of expression of the *APC* transcript. [27] Hypermethylation of the *APC* promoter was observed in 12% (5/43; CI 2,0-21,2%) and 25% (7/28; CI 8,9-41,0%), in respectively the primary ovarian tumors and CRC metastases to the ovary ( $p=0.19$ ). Although the hypermethylation of the promoter region of *APC* did not show differences between the primary and secondary ovarian tumors as a group, *APC* promoter hypermethylation was observed more often in the primary mucinous ovarian carcinomas than in the endometrioid subtype. Three of the 7 analyzed mucinous tumors (42,9%; CI 6,2-79,5%) but only 2 out of 36 (5,6%; CI -1,9-13%) endometrioid tumors showed hypermethylation of *APC* ( $p=0.014$ ). Analyses of 16 additional primary ovarian tumors showed hypermethylation only in the cases ( $n=3$ ) with mucinous histology. Larger cohorts may indicate that the hypermethylation of the promoter region of *APC* is a distinctive feature of primary mucinous carcinomas.

---

### **Synergies of APC-inactivation mechanisms**

Next, we investigated whether primary ovarian tumors and CRC metastases to the ovary showed different combinations of events leading to the inactivation of *APC* (mutation, LOH or methylation). In the primary ovarian tumors, only one case showed a combination of 2 events (e.g., LOH and methylation; Supplemental Table 2C). Only one type of inactivating events was observed in all the other cases.

In the set of CRC metastases to the ovary, a synergism of inactivation mechanisms was seen more frequently (Supplemental Table 2F). In 9 cases, 2 deleterious mutations were identified within the same tumor. Five out of 28 cases showed both a deleterious *APC* mutation and LOH; in 2 of these cases, 2 *APC* mutations were detected. Three cases showed a combination of an *APC* mutation and hypermethylation. Two cases showed LOH, promoter hypermethylation and a mutation in the MCR. In conclusion, the complete inactivation of *APC* is more prevalent in CRC metastases than in ovarian tumorigenesis.

### ***CTNNB1* (B-catenin) mutation analysis**

*CTNNB1* mutation analysis was included because *CTNNB1* mutations are present in about 25% of primary endometrioid ovarian tumors and in MSI-H colorectal tumors, which mostly do not metastasize.[28, 29] Exon 3 is the mutation hotspot region of *CTNNB1*. In the primary ovarian tumors, 2 activating *CTNNB1* missense mutations in exon 3 (2/43; 4.7%; CI -1,6–10,9%) and 5 additional missense mutations in the 3' part of the gene were identified by NGS (Figure 3). The pathogenicity of the missense mutations found in the 3' part of the gene is unknown. All these mutations were found in the endometrioid subtype (Supplemental Table 3 shows annotations). The 2 identified mutations in exon 3 were confirmed with Sanger sequencing, and the exon 3 mutation-negative cases showed indeed no alterations.

In the CRC metastases 3 (10,7%; CI 0,7-22,1%) missense mutations in the 3' portion of *CTNNB1* were identified by NGS (Figure 3) and Sanger sequencing of exon 3 confirmed the absence of mutations in exon 3. No significant difference in the mutation frequency of *CTNNB1* comparing primary versus secondary ovarian tumors was observed ( $p=0.64$ ).

## Discussion

In this study, we showed that target-enriched NGS of a set of 115 complete genes is feasible on fragmented DNA from FFPE tissue. Furthermore analysis of the complete *APC* gene can be used to differentiate between primary endometrioid- or mucinous ovarian tumors and metastases of colorectal carcinoma. Especially, in those cases in which immunohistochemical results cause discussion, *APC* mutation analysis can be of additional value.

The DNA sequencing of FFPE tissue was already routinely possible using Sanger sequencing. There are, however, limitations to the latter technique, when large genes (like *APC*) need to be screened. Pathogenic *APC* mutations found by NGS were confirmed by other means. [30] In two cases, however, 2 additional variants were found by classic DNA sequencing methods, despite sufficient coverage with NGS. One of these variants was located in a repetitive sequence of *APC* (3' portion). The analysis of repetitive sequences can be problematic in NGS approaches.[31] Besides, intra-tumor heterogeneity might alternatively explain this discrepancy.[32] The identification of a mutation in the mutation cluster region or in the 5' part (upstream MCR) of *APC* is a strong argument in favor of a colorectal origin in cases of ovarian tumors with a mucinous- or endometrioid type morphology. As in many of the diagnostic evaluations, however, such a finding has high odds ratios for a colorectal origin, but there will always be exceptions to this rule. In the work we now present, we did not address the presence and frequency of *APC* mutations of other cancers that can metastasize to the ovary, such as tumors from other locations within the GIT. *APC* mutations have been reported in 7% and 6% of primary stomach and pancreatic carcinomas, respectively.[12] In appendiceal mucinous tumors, *APC* alterations have been reported, but large series have not yet been examined.[33] In subsequent series we are planning to analyze ovarian metastases from primary tumors located outside the colorectum. In this way mutation profiles for ovarian metastases per primary tumor location can be generated.

So far, genomic profiles of primary and secondary ovarian tumors have been compared by mRNA expression profiles (reviewed by Stella et al [34-36]) and by karyotyping [37], needing frozen or fresh material, respectively. According to the literature, one *APC* mutation has been detected (out of a total of 25 investigated cases) in primary endometrioid ovarian carcinomas.[38] No other reports of *APC* mutations in endometrioid or mucinous ovarian tumors are mentioned in the COSMIC database.[12] Kelemen et al. suggested that mucinous tumors of the colorectum and the ovary show similarity in their mutational patterns.[4] However, only one article was cited that used

immunohistochemical staining of beta-catenin, an indirect way to test the deregulation of the Wnt-pathway.[39] *CTNNB1* exon 3 (the hotspot mutation region) mutations were identified in 4,7% of primary ovarian endometrioid tumors in our study. This is a somewhat low number compared to the reported percentage (26%).[12] Nevertheless, because no *CTNNB1* mutation in exon 3 was found in the colorectal metastases to the ovary, identifying a *CTNNB1* mutation may still be an argument in favor of an ovarian origin of the tumor. Furthermore, *CTNNB1* mutations were identified in the 3' part of the gene in both primary and secondary ovarian tumors, their pathogenicity is unknown. No apparent difference in mutation spectra was found for most genes studied in the current literature and in our analysis. This is not a surprise; there is wide overlap and redundancy in genes that drive several tumor types. Single genes that are informative on their own, such as *APC*, are extremely rare. It may be possible to develop a computational model that combines all (non-significant) odds-ratios to arrive at a certain diagnosis. The weighted net result of these odds-ratios can give a value between 0 and 1. A value close to 1 would suggest a metastasis, whereas a value close to 0 would indicate a primary tumor. Such an approach could also be valuable for determining the chemosensitivity or prognosis of a tumor based on mutational profiles. In summary, we have shown that mutation analysis of complete genes, such as *APC* in FFPE tissue is feasible and can be used to differentiate between primary endometrioid- or mucinous ovarian tumors and metastases of colorectal carcinoma.

## Acknowledgements

We thank Lysette van Bommel (Laboratory for Diagnostic Genome Analysis) for technical help with the analysis of *APC* in DNA from frozen tissue

## References

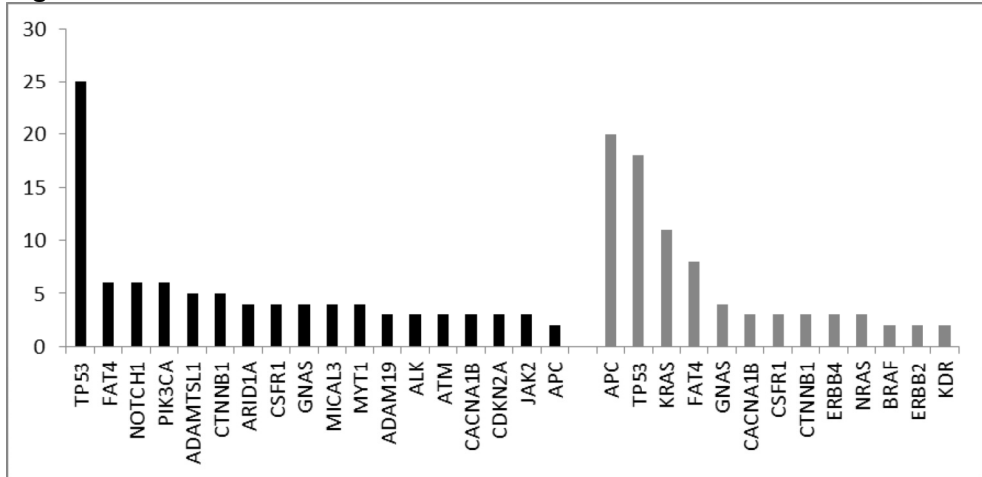
- [1] de Waal YR, Thomas CM, Oei AL, Sweep FC, Massuger LF. Secondary ovarian malignancies: frequency, origin, and characteristics. *Int J Gynecol Cancer* 2009;19:1160-5.
- [2] Moore RG, Chung M, Granai CO, Gajewski W, Steinhoff MM. Incidence of metastasis to the ovaries from nongenital tract primary tumors. *Gynecol Oncol* 2004;93:87-91.
- [3] Petru E, Pickel H, Heydarfadai M, Lahousen M, Haas J, Schaidler H, Tamussino K. Nongenital cancers metastatic to the ovary. *Gynecol Oncol* 1992;44:83-6.
- [4] Kelemen LE, Kobel M. Mucinous carcinomas of the ovary and colorectum: different organ, same dilemma. *Lancet Oncol* 2011.
- [5] Lee KR, Young RH. The distinction between primary and metastatic mucinous carcinomas of the ovary: gross and histologic findings in 50 cases. *Am J Surg Pathol* 2003;27:281-92.
- [6] Vang R, Gown AM, Barry TS, Wheeler DT, Yemelyanova A, Seidman JD, Ronnett BM. Cytokeratins 7 and 20 in primary and secondary mucinous tumors of the ovary: analysis of coordinate immunohistochemical expression profiles and staining distribution in 179 cases. *Am J Surg Pathol* 2006;30:1130-9.
- [7] Fraggetta F, Pelosi G, Cafici A, Scollo P, Nuciforo P, Viale G. CDX2 immunoreactivity in primary and metastatic ovarian mucinous tumours. *Virchows Arch* 2003;443:782-6.
- [8] Groisman GM, Meir A, Sabo E. The value of Cdx2 immunostaining in differentiating primary ovarian carcinomas from colonic carcinomas metastatic to the ovaries. *Int J Gynecol Pathol* 2004;23:52-7.
- [9] Khunamornpong S, Suprasert P, Pojchamarnwiputh S, Na CW, Settakorn J, Siriaunkgul S. Primary and metastatic mucinous adenocarcinomas of the ovary: Evaluation of the diagnostic approach using tumor size and laterality. *Gynecol Oncol* 2006;101:152-7.
- [10] Cathro HP, Stoler MH. Expression of cytokeratins 7 and 20 in ovarian neoplasia. *Am J Clin Pathol* 2002;117:944-51.
- [11] Ji H, Isacson C, Seidman JD, Kurman RJ, Ronnett BM. Cytokeratins 7 and 20, Dpc4, and MUC5AC in the distinction of metastatic mucinous carcinomas in the ovary from primary ovarian mucinous tumors: Dpc4 assists in identifying metastatic pancreatic carcinomas. *Int J Gynecol Pathol* 2002;21:391-400.
- [12] Bamford S, Dawson E, Forbes S, Clements J, Pettett R, Dogan A, Flanagan A, Teague J, Futreal PA, Stratton MR, Wooster R. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer* 2004;91:355-8.
- [13] Bell DA. Origins and molecular pathology of ovarian cancer. *Mod Pathol* 2005;18 Suppl 2:S19-S32.
- [14] Velculescu VE. Defining the blueprint of the cancer genome. *Carcinogenesis* 2008;29:1087-91.
- [15] Comprehensive molecular characterization of human colon and rectal cancer. *Nature*

---

2012;487(7407):330-7.

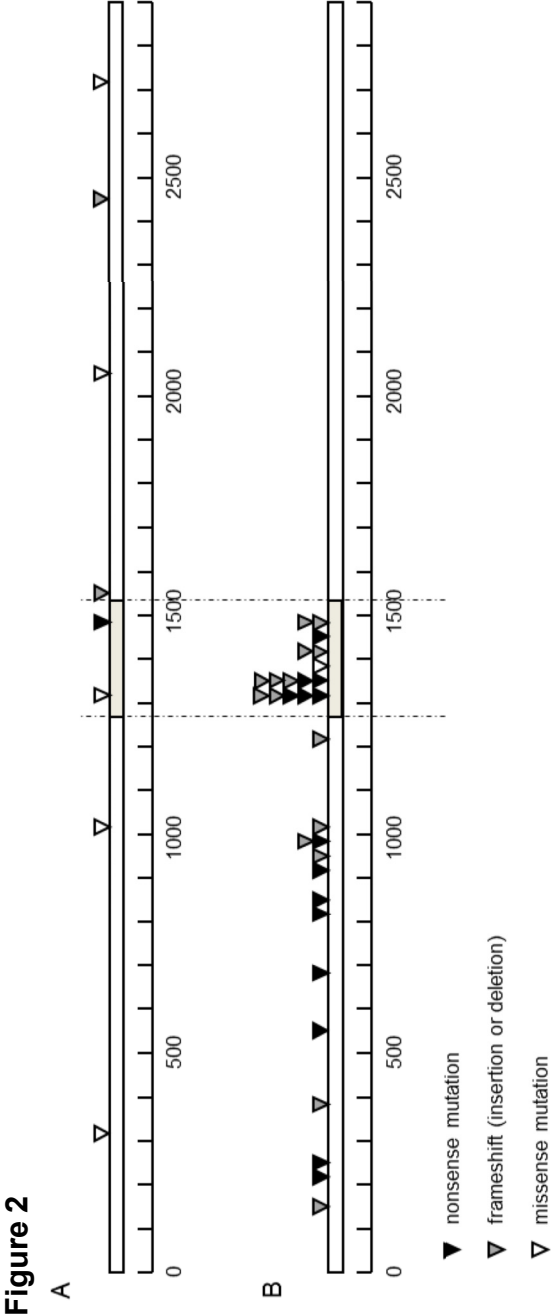
- [16] Miyoshi Y, Nagase H, Ando H, Horii A, Ichii S, Nakatsuru S, Aoki T, Miki Y, Mori T, Nakamura Y. Somatic mutations of the *APC* gene in colorectal tumors: mutation cluster region in the *APC* gene. *Hum Mol Genet* 1992;1:229-33.
- [17] Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ. Target-enrichment strategies for next-generation sequencing. *Nat Methods* 2010;7(2):111-8.
- [18] Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de Pol A, van Krieken JH, Meijer GA. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol* 2007;29:19-24.
- [19] Starr TK, Allaei R, Silverstein KA, Staggs RA, Sarver AL, Bergemann TL, Gupta M, O'Sullivan MG, Matise I, Dupuy AJ, Collier LS, Powers S, Oberg AL, Asmann YW, Thibodeau SN, Tessarollo L, Copeland NG, Jenkins NA, Cormier RT, Largaespada DA. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science* 2009;323(5922):1747-50.
- [20] Torkamani A, Schork NJ. Identification of rare cancer driver mutations by network reconstruction. *Genome Res* 2009;19(9):1570-8.
- [21] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25(14):1754-60.
- [22] DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del AG, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43(5):491-8.
- [23] Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;22(3):568-76.
- [24] Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38(16):e164.
- [25] Lips EH, van ER, de Graaf EJ, Doornebosch PG, de Miranda NF, Oosting J, Karsten T, Eilers PH, Tollenaar RA, van WT, Morreau H. Progression and tumor heterogeneity analysis in early rectal cancer. *Clin Cancer Res* 2008;14:772-81.
- [26] Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 2002;3(6):415-28.
- [27] Esteller M, Sparks A, Toyota M, Sanchez-Cespedes M, Capella G, Peinado MA, Gonzalez S, Tarafa G, Sidransky D, Meltzer SJ, Baylin SB, Herman JG. Analysis of adenomatous polyposis coli promoter hypermethylation in human cancer. *Cancer Res* 2000;60(16):4366-71.
- [28] Mirabelli-Primdahl L, Gryfe R, Kim H, Millar A, Luceri C, Dale D, Holowaty E, Bapat B,

- Gallinger S, Redston M. Beta-catenin mutations are specific for colorectal carcinomas with microsatellite instability but occur in endometrial carcinomas irrespective of mutator pathway. *Cancer Res* 1999;59:3346-51.
- [29] Lovig T, Meling GI, Diep CB, Thorstensen L, Norheim AS, Lothe RA, Rognum TO. APC and CTNNB1 mutations in a large series of sporadic colorectal carcinomas stratified by the microsatellite instability status. *Scand J Gastroenterol* 2002;37:1184-93.
- [30] Kerick M, Isau M, Timmermann B, Sultmann H, Herwig R, Krobitch S, Schaefer G, Verdorfer I, Bartsch G, Klocker H, Lehrach H, Schweiger MR. Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. *BMC Med Genomics* 2011;4:68.
- [31] Treangen TJ, Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet* 2011;13(1):36-46.
- [32] Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res* 2012;72(19):4875-82.
- [33] Shih IM, Yan H, Speyrer D, Shmookler BM, Sugarbaker PH, Ronnett BM. Molecular genetic analysis of appendiceal mucinous adenomas in identical twins, including one with pseudomyxoma peritonei. *Am J Surg Pathol* 2001;25(8):1095-9.
- [34] Stella GM, Senetta R, Cassenti A, Ronco M, Cassoni P. Cancers of unknown primary origin: current perspectives and future therapeutic strategies. *J Transl Med* 2012;10:12-0.
- [35] Meyniel JP, Cottu PH, Decraene C, Stern MH, Couturier J, Lebigot I, Nicolas A, Weber N, Fourchette V, Alran S, Rapinat A, Gentien D, Roman-Roman S, Mignot L, Sastre-Garau X. A genomic and transcriptomic approach for a differential diagnosis between primary and secondary ovarian carcinomas in patients with a previous history of breast cancer. *BMC Cancer* 2010;10:222.
- [36] Mhawech-Fauceglia P, Rai H, Nowak N, Cheney RT, Rodabaugh K, Lele S, Odunsi K. The use of array-based comparative genomic hybridization (a-CGH) to distinguish metastatic from primary synchronous carcinomas of the ovary and the uterus. *Histopathology* 2008;53:490-5.
- [37] Dionigi A, Facco C, Tibiletti MG, Bernasconi B, Riva C, Capella C. Ovarian metastases from colorectal carcinoma. Clinicopathologic profile, immunophenotype, and karyotype analysis. *Am J Clin Pathol* 2000;114:111-22
- [38] Wu R, Zhai Y, Fearon ER, Cho KR. Diverse mechanisms of beta-catenin deregulation in ovarian endometrioid adenocarcinomas. *Cancer Res* 2001;61:8247-55.
- [39] Chou YY, Jeng YM, Kao HL, Chen T, Mao TL, Lin MC. Differentiation of ovarian mucinous carcinoma and metastatic colorectal adenocarcinoma by immunostaining with beta-catenin. *Histopathology* 2003;43:151-6.

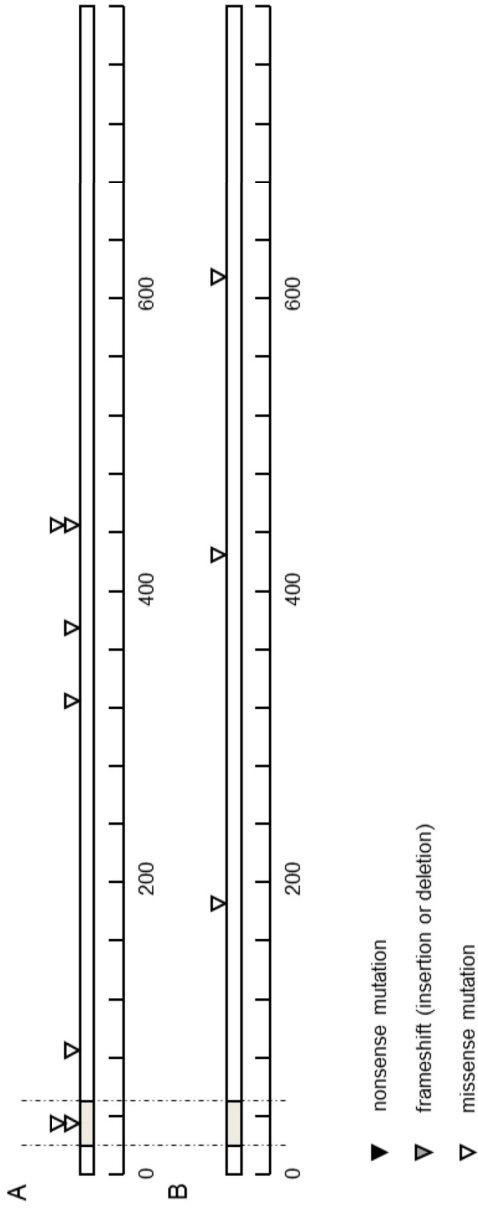
**Figure 1**

*Mutational profiles of primary endometrioid and mucinous ovarian tumors (in black) and ovarian metastases of CRC (in grey). Absolute number of mutations is given per gene. Genes with a low mutation rate (<3 for primary ovarian tumors; <2 for ovarian CRC metastases) are not shown, except for APC.*





Overview of APC mutations, locations are indicated with arrows, in (2A) primary ovarian tumors and (2B) ovarian metastases of colorectal tumors. The mutation cluster region (MCR; codon 1286-1513) is indicated with dotted lines.

**Figure 3**

Overview of CTNNB1 mutations, locations are indicated with arrows, in primary ovarian tumors (A) and in ovarian metastases of colorectal tumors (B). Exon 3 (codon 29-42) of CTNNB1 is indicated with dotted lines.

**Table 1**

<b>Sequencing primers for APC mutation cluster region (MCR; codon 1286-1513)</b>	
S1	5' - GAAATAGGATGTAATCAGACG - 3'      5' - CGCTCCTGAAGAAAAATCAAC - 3'
S2	5' - ACTGCAGGGTTCTAGTTATC - 3'      5' - AGCTGGCAATCGAACGACT - 3'
S3	5' - ACTTCTGTCAGTTCACCTTGATA - 3'      5' - ATTTT TAGGTACTTCTCGCTTG - 3'
S4	5' - AAACACCTCCACCACCTCC - 3'      5' - CATTATTCTTAATCCACATC - 3'
<b>Sequencing primers for CTNMB1 exon 3 (codon 29-42)</b>	
5' - GATTTGATGGAGTTGGACATGG - 3'	5' - TGTTCCTTGAGTGAAGGACTGAG - 3'
<b>Methylation specific PCR primers APC promoter 1A</b>	
Methylated	5' - TATTGCGGAGTGCGGGTC - 3'      5' - TCGAGGAACCTCCCGACGA - 3'
Unmethylated	5' - GTGTTTATTGTGGAGTGTGGTT - 3'      5' - CCAATCAACAAAATCCCAACA - 3'
<b>Bisulfite sequencing primers for APC promoter 1A</b>	
5' - GGGTTAGGGTTAGGTAGTTGT - 3'	5' - ACACCTCCATTCTATCTCCAATAAC - 3'

*Primer sequences. Overview of the sequencing primers for APC and CTNMB1. The primers for methylation-specific PCR and bisulfite sequencing of the promoter region 1A of APC are also shown.*

Table 2

	No.	APC mutation in MCR	APC mutation outside MCR	LOH APC	Methylation APC	CTNWB1 mutation exon 3	CTNWB1 mutation outside exon 3
endometrioid	35	0	1*	12	2	2	5
mucinous	7	0	0	0	3	0	0
mixed subtype	1	1*	1*	0	0	0	0
<b>total</b>	<b>43</b>	<b>2.3% (1/43)</b>	<b>4.7% (2/43)</b>	<b>27.9% (12/43)</b>	<b>11.6% (5/43)</b>	<b>4.7% (2/43)</b>	<b>11.6% (5/43)</b>

Overview of pathogenic APC mutations (frameshift or truncating mutations) in primary ovarian endometrioid and mucinous tumors. APC missense mutations were excluded from this table. In total, 43 primary ovarian tumors were analyzed using the HaloPlex system. \*In total, 3 APC mutations were found (in 2 unique cases); 2 of these mutations were located in the 3' portion of the gene (downstream of the MCR).

Table 3

	No.	APC mutation in MCR	APC mutation outside MCR	LOH APC	Methylation APC	CTNMB1 mutation exon 3	CTNMB1 mutation outside exon 3
mucinous	9	6	4**	2	2	0	2
non- mucinous	19	9	10**	7	5	0	1
<b>total</b>	<b>28</b>	<b>53.5% (15/28)</b>	<b>50.0% (14/28)</b>	<b>32.1% (9/28)</b>	<b>25% (7/28)</b>	<b>0.0% (0/28)</b>	<b>10.7% (3/28)</b>

Overview of pathogenic APC mutations (frameshift or truncating mutations) in ovarian metastases of colorectal tumors. APC missense mutations were excluded from this table. In total, 28 ovarian metastases were analyzed. \*\* All APC mutations found outside of the MCR were located in the 5' portion (upstream of the MCR) of the gene.

## Supplemental Figure 1A - 1C

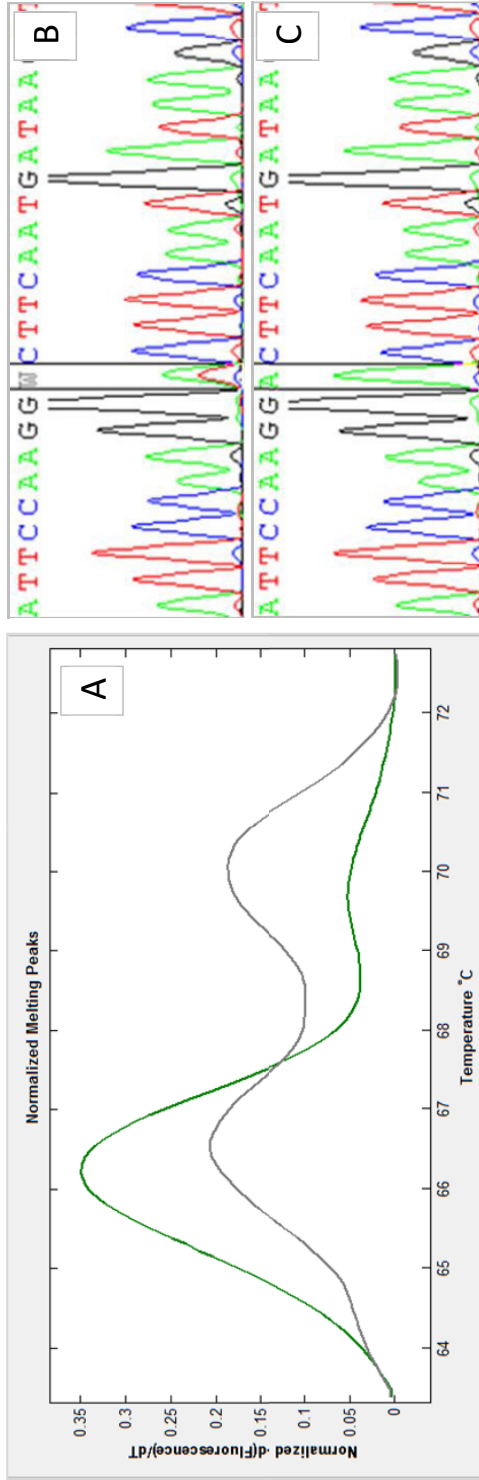


Illustration of High Resolution Melting Curve Analysis (HRMCA) and sequencing to detect Loss of Heterozygosity (LOH). (1A) An example (CRC metastasis case no. 1) of LOH detected by High Resolution Melting Curve Analysis (HRMCA) is shown. The normal heterozygote control at SNP c.5465A>T is shown in grey. As can be seen, two peaks are present, representing both alleles. The tumor material (in green) shows loss of one allele. (1B) Sequence generated by Sanger sequencing of the same material. In the upper panel the sequence of the heterozygote normal material is shown. As can be observed, there are two alleles present. (1C) In the panel at the bottom, the sequence of the tumor material is shown. In this sequence only one allele is present, due to loss of one of the alleles.

**Supplemental Table 1**

1	<i>ABCA1</i>	31	<i>CSF1R</i>	61	<i>KRAS</i>	91	<i>PRUNE2</i>
2	<i>ABL1</i>	32	<i>CTNNB1</i>	62	<i>LIMK2</i>	92	<i>PTEN</i>
3	<i>ADAM19</i>	33	<i>DSTYK</i>	63	<i>LOH12CR1</i>	93	<i>PTPN1</i>
4	<i>ADAMTSL1</i>	34	<i>EGFR</i>	64	<i>LTK</i>	94	<i>PTPN11</i>
5	<i>ADAMTSL3</i>	35	<i>EPHA3</i>	65	<i>MAP2K1</i>	95	<i>RB1</i>
6	<i>AKT1</i>	36	<i>ERBB2</i>	66	<i>MAP2K2</i>	96	<i>RET</i>
7	<i>AKT2</i>	37	<i>ERBB4</i>	67	<i>MAP2K4</i>	97	<i>RIOK3</i>
8	<i>ALK</i>	38	<i>FAT4</i>	68	<i>MDM2</i>	98	<i>SIAH1</i>
9	<i>ALPK1</i>	39	<i>FBXW7</i>	69	<i>MEN1</i>	99	<i>SMAD2</i>
10	<i>APC</i>	40	<i>FES</i>	70	<i>MET</i>	100	<i>SMAD3</i>
11	<i>ARID1A</i>	41	<i>FGFR1</i>	71	<i>MICAL3</i>	101	<i>SMAD4</i>
12	<i>ATM</i>	42	<i>FGFR3</i>	72	<i>MLH1</i>	102	<i>SMAD7</i>
13	<i>BAX</i>	43	<i>FGR</i>	73	<i>MLL3</i>	103	<i>SMARCB1</i>
14	<i>BMP2</i>	44	<i>FLT3</i>	74	<i>MMP2</i>	104	<i>SMO</i>
15	<i>BMPR1A</i>	45	<i>FOXO1</i>	75	<i>MMP9</i>	105	<i>SRC</i>
16	<i>BMPR2</i>	46	<i>GATA3</i>	76	<i>MSH2</i>	106	<i>SRGAP1</i>
17	<i>BRAF</i>	47	<i>GNAS</i>	77	<i>MSH6</i>	107	<i>STAB1</i>
18	<i>C11orf66</i>	48	<i>GUCY1A2</i>	78	<i>MUTYH</i>	108	<i>STK11</i>
19	<i>CACNA1B</i>	49	<i>HIF1A</i>	79	<i>MYC</i>	109	<i>SYNC</i>
20	<i>CACNA2D3</i>	50	<i>HOXA4</i>	80	<i>MYT1</i>	110	<i>TGFBR1</i>
21	<i>CASR</i>	51	<i>HOXB4</i>	81	<i>NEGR1</i>	111	<i>TGFBR2</i>
22	<i>CCNB2</i>	52	<i>HOXC4</i>	82	<i>NOTCH1</i>	112	<i>TIE1</i>
23	<i>CCND1</i>	53	<i>HOXD4</i>	83	<i>NRAS</i>	113	<i>TP53</i>
24	<i>CCNT2</i>	54	<i>HRAS</i>	84	<i>NTRK1</i>	114	<i>TP53BP1</i>
25	<i>CDC42BPA</i>	55	<i>IDH1</i>	85	<i>NTRK3</i>	115	<i>VHL</i>
26	<i>CDC73</i>	56	<i>JAK1</i>	86	<i>PANK4</i>		
27	<i>CDH1</i>	57	<i>JAK2</i>	87	<i>PARP1</i>		
28	<i>CDK4</i>	58	<i>JAK3</i>	88	<i>PDGFRA</i>		
29	<i>CDKN2A</i>	59	<i>KDR</i>	89	<i>PIK3CA</i>		
30	<i>COL3A1</i>	60	<i>KIT</i>	90	<i>PMS2</i>		

List of genes included in the custom-made gene panel (n=115).

Supplemental Table 2

Prim. Ovarium (2A)	total reads	mean coverage	%_bases_ above_1	%_bases_ above_1 0	%_bases_ above_20	%_bases_ above_30
1	562667999	541.33	95.4	82.4	75.7	71.2
2	379378269	364.99	90.8	74.4	66.8	62.2
3	356370116	342.86	87.2	69.6	62.9	58.6
4	96697971	93.03	74.6	55.3	46.3	40.6
5	39768331	38.26	63.9	39.9	30.3	24.7
6	61270131	58.95	62.3	42.8	34.2	29
7	407196076	391.75	92.7	79.9	73.2	68.5
8	420720024	404.76	90.8	75.8	69.2	65
9	378169419	363.83	84.9	71.5	65	60.9
10	280861817	270.21	96.6	85.2	77.7	72.6
11	463614399	446.03	96.3	85.7	79.2	74.8
12	526563591	506.59	93.6	81.6	75.1	70.7
13	152609160	146.82	96.7	84.2	75	68.2
14	94323066	90.75	86.7	64.1	55	48.7
15	73004285	70.24	66.2	46.1	37.6	32.3
16	365500760	351.64	91.6	77.4	70.7	65.8
17	210754662	202.76	67.3	56	49.3	44.9
18	530727173	510.6	97.4	90.3	84.6	80.5
19	77216295	74.29	62.1	43.1	35.1	30.4
20	160662204	154.57	85.1	67	58.6	52.8
21	415117066	399.37	98	90.8	85	80.5
22	297369341	286.09	96.7	84.4	75	68.1
23	164192207	157.97	63.7	48.3	40.5	35.5
24	259907149	250.05	96.5	85	77.6	72.3
25	149717992	144.04	89.6	68.3	59.3	53.6
26	551591937	530.67	93.5	81.7	75	70.4
27	19975283	19.22	57.3	28.1	19.3	14.6
28	474382629	456.39	79.7	67.8	61.6	57.7
29	158241257	152.24	95.4	78.8	69.5	62.9
30	89124782	85.74	95	78	68	60.6
31	110481983	106.29	71.9	51.4	42.9	37.4
32	321932494	309.72	89.1	73.1	66.1	61.7
33	29258865	28.15	71.5	42.3	31.6	24.7
34	345800030	332.69	83.2	68.4	61.2	56.6
35	203099255	195.4	90	70.6	62.4	57.3
36	212950285	204.87	78.2	60.8	53.4	48.3
37	433832794	417.38	92.5	80.2	73.9	69.7
38	443987429	427.15	93.3	79.6	72.8	68.4
39	81759525	78.66	75.5	53.9	45	39.7
40	149304509	143.64	85.4	64.5	55.7	50.3
41	380109542	365.69	96	84.5	76.9	72.3
42	458889329	441.49	97.7	90.5	84.8	80.1
43	112404076	108.14	69.5	50.2	42.5	37.4



Prim. Ovarium (2B)	Histologic subtype	CDX2 -	PAX8 +	CK7 +	CK20 -	VIM +
1	En	E	E	E	E	U
2	M	E	E	E	U	U
3	En	E	E	E	E	U
4	En	E	E	U	E	U
5	En	E	U	E	E	U
6	En	E	U	E	E	U
7	En	E	E	E	E	U
8	En	E	E	E	E	E
9	En	E	E	E	E	U
10	En	E	E	E	E	U
11	En	E	E	E	E	U
12	En	E	E	E	E	E
13	En	U	E	E	E	U
14	M	U	E	E	U	U
15	En	E	E	E	E	U
16	En	E	E	U	E	U
17	En	U	E	U	E	U
18	En	U	E	E	E	U
19	En	E	E	E	E	U
21	En+C	E	E	E	E	U
22	M	U	E	E	U	U
23	M	U	U	E	U	U
24	En	U	E	E	E	U
25	En	E	E	E	E	U
26	En	E	E	E	E	U
27	En	E	E	E	E	U
28	En	E	E	E	E	U
29	En	E	E	E	E	U
30	M	E	U	E	E	U
32	M	U	U	E	E	U
33	En	U	U	E	E	U
34	En	E	U	U	E	U
35	En	E	E	E	E	E
36	En	E	E	E	E	U
37	En	U	E	E	E	U
38	En	E	E	E	E	U
39	En	U	E	E	E	E
40	En	U	U	E	E	U
41	En	E	E	U	E	U

Prim. Ovarium (2C)		APC MUTATIONS				
Case	Histologic subtype	5'end (<codon 1286)	MCR (1286-1513)	3'end (>codon 1513)	LOH	Methylation
1	En	N	N	N	UK	N
2	M	N	N	N	N	Y
3	En	N	N	N	N	Y
4	En	N	N	N	Y	N
5	En	N	N	N	N	N
6	En	N	N	N	Y	N
7	En	N	N	N	Y	Y
8	En	N	N	N	Y	N
9	En	N	N	N	N	N
10	En	N	N	N	UK	N
11	En	N	N	N	N	N
12	En	N	N	N	N	N
13	En	N	N	N	Y	N
14	M	N	N	N	N	N
15	En	N	N	N	N	N
16	En	N	N	N	N	N
17	En	N	N	N	UK	N
18	En	N	N	N	UK	N
19	En	N	N	N	UK	N
20	En	N	N	N	Y	N
21	En+C	N	Y	Y	UK	N
22	M	N	N	N	UK	Y
23	M	N	N	N	N	N
24	En	N	N	N	UK	N
25	En	N	N	N	Y	N
26	En	N	N	Y	UK	N
27	En	N	N	N	UK	N
28	En	N	N	N	Y	N
29	En	N	N	N	Y	N
30	M	N	N	N	N	N
31	M	N	N	N	UK	N
32	M	N	N	N	UK	Y
33	En	N	N	N	Y	N
34	En	N	N	N	N	N
35	En	N	N	N	UK	N
36	En	N	N	N	N	N
37	En	N	N	N	N	N
38	En	N	N	N	N	N
39	En	N	N	N	N	N
40	En	N	N	N	N	N
41	En	N	N	N	Y	N
42	En	N	N	N	UK	N
43	En	N	N	N	Y	N
<b>n=43</b>		<b>0%</b>	<b>2.3%</b>	<b>4.7%</b>	<b>27.9%</b>	<b>11.6%</b>

CRC metastases (2D)	total reads	mean coverage	%_bases_ above_1	%_bases_ above_10	%_bases_ above_20	%_bases_ above_30
1	531435788	511.28	95.3	83.5	76.9	72.5
2	258895399	249.08	96	82.6	75.1	69.7
3	366236567	352.35	88.3	75.8	69.2	64.6
4	284445865	273.66	93.8	78.1	70.8	66.3
5	386080901	371.44	85.9	70	62.8	58
6	419398099	403.49	98.1	92.2	86.9	82.6
7	166188874	159.89	91.7	70.8	61.5	55.6
8	370868504	356.8	95.3	83.6	76.7	72
9	528614247	508.57	90.9	80.9	74.8	70.6
10	410658067	395.08	97.5	89	82.5	77.8
11	377617408	363.3	96.1	84.7	77.3	72.5
12	266186282	256.09	97.1	87.8	81.1	75.5
13	404134644	388.81	97	87.3	80.9	75.9
14	148608397	142.97	93.9	74.8	64.9	58.3
15	316669636	304.66	97	89.1	83.1	78.2
16	287800179	276.89	96.2	84.5	77.3	72.2
17	303684904	292.17	88.7	74.7	68.4	64
18	190242156	183.03	83.3	65.1	56.8	51.3
19	270909355	260.64	91.1	74.4	66.6	61.3
20	191464997	184.2	88.4	69.8	61.5	56.2
21	825364679	794.06	95.6	88.2	83	79.2
22	168253406	161.87	96.6	83.2	74.4	68.1
23	238299380	229.26	93.4	77.1	69.3	64
24	411089878	395.5	91.4	77.3	70.5	65.8
25	450911323	433.81	83.9	70.8	64.2	59.7
26	325136401	312.81	86.9	73.1	66.7	62.6
27	465015253	447.38	97.3	88.3	82	77.6
28	170601993	164.13	89.2	70.6	62.4	56.7

---

CRC Metastases (2E)	CDX2 +	PAX8 -	CK7 -	CK20 +	VIM -
1	E	E	E	E	E
2	U	E	E	E	E
3	E	E	E	E	E
4	E	U	E	U	E
5	E	E	E	U	E
7	E	U	E	E	E
8	E	E	E	U	E
9	E	E	E	E	E
10	E	E	E	U	E
11	E	E	E	E	E
12	E	U	E	E	E
13	E	E	E	E	E
14	E	E	E	U	E
15	E	E	E	E	E
16	E	E	E	E	E
17	E	E	E	E	E
18	E	E	E	E	E
19	E	E	E	U	E
20	E	E	E	U	E
21	E	E	E	E	E
22	E	E	E	E	E
23	E	E	E	E	E
24	U	E	E	U	E
25	U	E	E	E	E
26	E	E	E	U	E
27	E	U	E	U	E
28	E	E	E	E	E

CRC metastases (2F)	APC MUTATIONS				
	5'end (<codon 1286)	MCR (1286-1513)	3'end (>codon 1513)	LOH	Methylation
1	Y	N	N	Y	N
2	Y	Y	N	Y	N
3	N	Y	N	N/A	N
4	N	N	N	N/A	N
5	N	Y	N	N/A	N
6	N	N	N	N	Y
7	N	N	N	Y	N
8	N	N	N	N/A	N
9	Y	Y	N	N	Y
10	N	Y	N	N/A	Y
11	Y	N	N	N	N
12	Y	Y	N	Y	Y
13	N	Y	N	Y	N
14	Y	Y	N	N/A	N
15	Y	Y	N	N/A	N
16	Y	Y	N	N/A	N
17	N	N	N	N	N
18	N	N	N/A	Y	N
19	N	N	N	N/A	N
20	Y	N	N	N/A	Y
21	Y	Y	N	N/A	Y
22	Y	Y	N	N	N
23	N	Y	N	N/A	N
24	Y	N	N	Y	N
25	Y	Y	N	Y	N
26	N	Y	N	Y	Y
27	Y	Y	N	N/A	N
28	N	N	N	N	N
<b>n=28</b>	<b>46.4%</b>	<b>60.7%</b>	<b>0%</b>	<b>30.1%</b>	<b>25.0%</b>

Total reads, mean coverage and percentage of bases covered more than 1x, 10x, 20x and 30x is showed for (2A) primary ovarian tumors and (2D) CRC metastases to the ovary. Overview of staining results of 39 primary ovarian tumors en 27 proven CRC metastases to the ovaries in our series (2B and 2E). A small number of tumors (4 primary ovarian tumors and 1 CRC metastasis) was not included. The expected immunohistochemical profile of primary ovarian tumors is CK7+ , CK20- , CDX2- , PAX8+ and vimentin+, while the CRC metastases to the ovary are expected to show opposite staining results (E=expected; U=unexpected). Not all cases show a characteristic immunostaining profile that would lead to an unequivocal diagnosis. The numbers in the first column of Supplemental Table 2B and 2E correspond with the numbers in respectively Supplemental Table 4 and 6. En=endometrioid; M= mucinous; C=clearcell. Overview of APC-inactivation mechanisms for primary ovarian tumors (2C) and ovarian CRC metastases (2F). The Y (=yes) indicates alterations, and the N (=no) indicates no alteration. En = endometrioid; M = mucinous; C = clear cell; N/A = not analyzed.

Supplemental Table 3

No.	Histology	Screening method	Mutation outside MCR APC (5'end; <codon 1286)	Mutation within MCR APC (codon 1286-1513)	Mutation outside MCR APC (3'end; >codon 1513)	LOH	Methylation	B-catenin mutation (exon 3)	B-catenin mutation (outside exon 3)
1	E	HRMCA + Sanger	-	-	-	UK	N	-	N/A
2	M	HaloPlex + Sanger	-	-	-	N	Y	-	N/A
3	E	HRMCA + Sanger	-	-	-	N	Y	-	N/A
4	E	HRMCA + Sanger	-	c.3949G>C; p.E1317Q	-	Y	N	-	N/A
5	E	HRMCA + Sanger	-	-	c.6136G>A; p.A2046T	N	N	-	N/A
6	E	HRMCA + Sanger	-	-	-	Y	N	-	c.1346G>A; p.R449H
7	E	HaloPlex + Sanger	-	-	-	Y	Y	-	N/A
8	E	HaloPlex + Sanger	c.902C>T; p.P301L	-	-	Y	N	-	N/A



18	E		HaloPlex	-	-	-	c.8162G>A; p.R2721H	UK	N	-		c.1000G>A; p.E334K
			HRMCA + Sanger	-	-	-	-	UK	N	-		N/A
			HaloPlex	-	-	-	-			-		-
19	E		HRMCA + Sanger	-	-	-	-	UK	N	-		N/A
			HaloPlex	-	-	-	-			-		-
20	E		HRMCA + Sanger	-	-	-	-	Y	N	-		N/A
			HaloPlex	-	-	-	-			-		-
21	E+C		HRMCA + Sanger	-	-	-	c.7356delA; p.L2452X	UK	N	-		N/A
			HaloPlex	-	-	-	-			-		-
22	M		HRMCA + Sanger	-	-	c.4403A>T; p.K1468X	-	UK	Y	-		N/A
			HaloPlex	-	-	-	-			-		-
23	M		HRMCA + Sanger	-	-	-	-	N	N	-		N/A
			HaloPlex	-	-	-	-			-		-
24	E		HRMCA + Sanger	-	-	-	-	UK	N	-		N/A
			HaloPlex	-	-	-	-			-		-
25	E		HRMCA + Sanger	-	-	-	-	Y	N	-		N/A
			HaloPlex	-	-	-	-			-		-
26	E		HRMCA + Sanger	-	-	-	c.4666dup; p.T1556fs	UK	N	-		N/A
			HaloPlex	-	-	-	-			-		-
27	E		HRMCA +	-	-	-	-	UK	N	-		N/A







**Supplemental Table 4**

Case	Variant	Present in COSMIC?	Present in Tumor?	Present in Normal?
Primary ovarian tumors				
3	c.3949G>C;p.Glu1317Gln	Y	Y	Y
4	c.6136G>A;p.Ala2046Thr	Y	N	N
7	c.902C>T;p.Pro301Leu	N	Y	N
17	c.8162G>A;p.Arg2721His	N	UK	UK
28	c.3007G>A;p.Gly1003Asp	Y	N	N
CRC metastasis				
22	c.4092T>A;p.Ser1364Arg	Y	Y	N

Overview of the 6 missense variants in APC. The results of the Sanger sequences and the presence in the COSMIC database are given. Note that c.3949G>C;p.Glu1317Gln is mentioned in the COSMIC database but present in normal tissue. Y=yes; N=no.

Supplemental Table 5

No.	Screening Method	Mutation outside MCR APC (5'end; <codon 1286)	Mutation within MCR APC (codon 1286-1513)	Mutation outside MCR APC (3'end; >codon 1513)	LOH	Methylation	Mucinous	B-catenin mutation (exon 3)	B-catenin mutation (outside exon 3)
1	HRMCA + Sanger	c.637C>T; p.R213X	-	-	Y	N	N	-	-
	HALOplex	c.637C>T; p.R213X	-	-				-	-
2	Sanger	c.847C>T; p.R283X	c.4465_4466insA T; p.L1489fs	-	Y	N	N	-	c.1862T>G; p. L621R
	HALOplex	c.847C>T; p.R283X	c.4465_4466insA T; p.L1489fs	-				-	-
3	HRMCA + Sanger	-	c.4057G>T; p.E1353X	-	UK	N	Y	-	-
	HALOplex	-	c.4057G>T; p.E1353X	-				-	-
4	HRMCA + Sanger	-	-	-	UK	N	N	-	-
	HALOplex	-	-	-				-	-
5	HRMCA + Sanger	-	c.4010_4011delT G; p.L1337fs	-	UK	N	Y	-	-

	HALOplex	-	c.4010_4011delT G; p.L1337fs	-								-	-
6	HRMCA + Sanger	-	-	-	N	Y	N					-	-
	HALOplex	-	-	-								-	-
7	Sanger	N/A	-	N/A	Y	N	N					-	-
	HALOplex	-	-	-								-	-
8	Sanger	N/A	-	N/A	UK	N	N	Y				-	-
	HALOplex	-	-	-				Y				-	-
9	Sanger	N/A	c.4308delT; p.S1436fs	N/A	N	Y	N					-	-
	HALOplex	c.1192- 1193delAA; p.K398fs	c.4308delT; p.S1436fs	-								-	-
10	Sanger	N/A	c.3944C>A; p.S1315X	N/A	UK	Y	Y					-	c.1276A>G; p.N426D
	HALOplex	-	c.3944C>A; p.S1315X	-								-	-
11	Sanger	N/A	-	N/A	N	N	N					-	-
	HALOplex	c.3709delC; p.Q1237fs	-	-								-	-

12	Sanger	N/A	c.3948delT; p.A1316fs	N/A	Y	Y	Y	Y	-	-	-
	HALOplex	c.419-420insA; p.N140fs	c.3948delT; p.A1316fs	-					-	-	-
13	Sanger	N/A	c.3927_3931delA AAGA; p.E1309fs	N/A	Y	N	N	N	-	-	-
	HALOplex	-	c.3927_3931delA AAGA; p.E1309fs						-	-	-
14	Sanger	N/A	c.3949G>T; p.E1317X	N/A	UK	N	N	Y	-	-	-
	HALOplex	c.2701C>T; p.Q901X	c.3949G>T; p.E1317X	-					-	-	-
15	Sanger	N/A	c.4241_4242insT; p.V1414fs	N/A	UK	N	N	N	-	-	-
	HALOplex	c.2950G>T; p. E985X	c.4241_4242insT; p.V1414fs	-					-	-	-
16	Sanger	N/A	-	N/A	UK	N	N	Y	-	-	c.569G>A; p.R190H
	HALOplex	c.2943delC; p.P981fs		-					-	-	-
17	Sanger	N/A	-	N/A	N	N	N	N	-	-	-
	HALOplex	-		-					-	-	-

18	Sanger	N/A	-	N/A	Y	N	N	-	-
	HALOplex	-	-	-				-	-
19	Sanger	N/A	-	N/A	UK	N	N	-	-
	HALOplex	-	-	-				-	-
20	Sanger	N/A	-	N/A	UK	Y	N	-	-
	HALOplex	c.2626C>T; p.R876X	-	-				-	-
21	Sanger	N/A	c.4359delT; p.P1453fs	N/A	UK	Y	N	-	-
	HALOplex	c.2097G>A; p.W699X	c.4359delT; p.P1453fs	-				-	-
22	Sanger	N/A	c.4092T>A; p.S1364R / c.4348C>T; p. R1450X	N/A	N	N	N	-	-
	HALOplex	c.1660C>T; p.R554X	c.4092T>A; p.S1364R / c.4348C>T; p. R1450X	-				-	-
23	Sanger	N/A	c.4067C>A; p.S1356X	N/A	UK	N	N	-	-

	HALOplex	-		c.4067C>A; p.S1356X	-						-	-	
24	Sanger	N/A		-	N/A	Y	N	N			-	-	
	HALOplex	c.2882delA; p.N961fs		-	-						-	-	
25	Sanger	N/A		c.4462_4463insA T; p.L1488fs	N/A	Y	N	N			-	-	
	HALOplex	c.2413C>T; p.R805X		c.4462_4463insA T; p.L1488fs	-						-	-	
26	Sanger	N/A		c.4056_4057insT; p.E1353fs	N/A	Y	Y	N			-	-	
	HALOplex	-		c.4056_4057insT; p.E1353fs	-						-	-	
27	Sanger	N/A		c.3925G>T; p.E1309X	N/A	UK	N	Y			-	-	
	HALOplex	c.3067- 3068insA; p.T1023fs		c.3925G>T; p.E1309X	-						-	-	
28	Sanger	N/A		-	N/A	N	N	Y			-	-	
	HALOplex	-		-	-						-	-	
<b>Total</b>		<b>50.0% (14/28)</b>		<b>57.1% (16/28)</b>	<b>0.0% (0/28)</b>	<b>32.1% (9/28)</b>	<b>25.0% (7/28)</b>	<b>32.1% (9/28)</b>	<b>32.1%</b>	<b>0.0% (0/28)</b>	<b>0.0% (0/28)</b>	<b>10.7% (3/28)</b>	

Overview of APC and CTNNB1 mutations in ovarian CRC metastases. LOH could not be determined for all of the cases because of the homozygosity of the SNPs or because the normal tissue was not available (UK=unknown). The Y (=yes) indicates alterations, and the N (=no) indicates no alteration. N/A = not analyzed. Reference sequence NM\_001127510.



## Chapter 4

# **Somatic mutation profiles in primary colorectal cancers and matching ovarian metastases: Identification of driver and passenger mutations**

Stijn Crobach<sup>1</sup>, Dina Ruano<sup>1</sup>, Ronald van Eijk<sup>1</sup>, Melanie Schrumpf<sup>1</sup>, PALGA group<sup>2</sup>, Gertjan Fleuren<sup>1</sup>, Tom van Wezel<sup>1</sup> and Hans Morreau<sup>1</sup>

<sup>1</sup>Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands

<sup>2</sup>PALGA, the nationwide histopathology and cytopathology data network and archive

---

**Abstract**

The mutational profiles of primary colorectal cancers (CRCs) and corresponding ovarian metastases were compared. Using a custom-made next generation sequencing panel, 115 cancer-driving genes were analyzed in a cohort of 26 primary CRCs and 30 matching ovarian metastases (4 with bilateral metastases). To obtain a complete overview of the mutational profile, low thresholds were used in bioinformatics analysis to prevent low frequency passenger mutations from being filtered out. A subset of variants was validated using Sanger and/or hydrolysis probe assays. The mutational landscape of CRC that metastasized to the ovary was not strikingly different from CRC in consecutive series. When comparing primary CRCs and their matching ovarian metastases, there was considerable overlap in the mutations of early affected genes. A subset of mutations demonstrated less overlap, presumably being passenger mutations. In particular, primary CRCs showed a substantially high number of passenger mutations. We also compared the primary CRCs and matching metastases for stratifying variants of 6 genes (*KRAS*, *NRAS*, *BRAF*, *FBXW7*, *PTEN* and *PIK3CA*) that select for established (*EGFR* directed) or future targeted therapies. In a total of 31 variants 12 were not found in either of the two locations. Tumours thus differed in the number of discordant variants between the primary tumours and matching metastases. Half of these discordant variants were pathogenic class 4/5 variants. However, in terms of temporal heterogeneity, no clear relationship was observed between the number of discordant variants and the time interval between primary CRCs and the detection of ovarian metastases. This suggests that dormant metastases may be present from the early days of the primary tumours.

## Introduction

Next-generation sequencing (NGS) provides the ability to determine the mutational profiles of tumours in a rapid and cost-effective manner.[1, 2] Previous NGS experiments showed that distinct parts of the same tumour show different mutation profiles (spatial intra-tumour heterogeneity; ITH).[3, 4] Additionally, primary tumours and their metastases can differ in their mutational pattern, thereby showing temporal heterogeneity.[5, 6] Determining the concordance between primary tumours and metastases is of interest for choosing the optimal treatment, i.e., targeted therapies that are directed against variants present in the primary tumour but not in metastases will not be effective. Studies investigating the overlap and differences between the mutational profiles of primary tumours and matched metastases at specific locations are mostly lacking.

In this study, we selected colorectal tumours (CRCs) that metastasized to the ovaries. CRCs frequently metastasize to the liver and the lung, whereas ovarian metastases are sparse.[7] Ovarian metastasis occurs in approximately 3.4% of women diagnosed with a colorectal malignancy.[8] However, in up to 38% of cases, ovarian metastasis detection may precede the detection of the primary CRC.[9-11] In such cases, it is important for treatment strategies to recognize that the ovarian tumour is a metastasis and not a primary ovarian tumour. Extensive genomic profiling of CRCs and primary ovarian tumours has revealed a limited number of genes helpful in discriminating between these malignancies.[12]

Previous studies have primarily investigated mutational differences between CRCs and liver metastasis.[5, 13-15] The mutational status of *KRAS* showed high concordance between CRCs and metastases.[16, 17] Because the *KRAS* mutation status has predictive value for *EGFR*-mediated treatment inhibition, mutations in liver metastases of CRC were concluded to be reliable when predicting the effects of the targeted therapies.

In a limited number of studies, broader gene panels of 5 to more than 1,000 genes were studied in CRC metastases. Vermaat et al. studied 1,264 genes and showed a gain of 83 and loss of 70 potentially function-impairing variations between primary CRCs and liver metastases.[5] Vakiani et al. reported a higher frequency of *TP53* and a lower frequency of *BRAF* mutations in the liver metastases compared with the primary tumours.[13, 14] However, the

same mutations were identified in both the circulating tumour cells and the primary CRC tumour.[18] Goranova studied the mutation rate in 6 liver metastases and primary CRCs.[19] In contrast with the study by Vakiani et al., no discrepancies between the primary tumour and the metastases were detected for *TP53* and *KRAS*. However, fewer *APC* mutations were detected in the liver metastases. In summary, due to the limited number of studies and the few cases included in the studies, no clear overview of the complete mutational profile of (liver) metastases of CRC is currently available (see also Supplemental Table 4).[20]

Interestingly, CRCs positive for a *KRAS* mutation have a higher risk of metastasis to the lungs.[21, 22] Among other factors, varying mutational profiles of CRC may enable successful homing at specific locations (e.g. the ovaries). It is currently unknown whether mutation profiles differ according to metastatic location. Identifying such stratifying mutations could assist in clinical diagnostics.

Much remains unknown regarding the biology of the process of metastasis.[23] In cases of ovarian metastasis of CRC, dissemination through the lymph-vascular system or through direct peritoneal spreading are considered to be the first steps.[24] Next, circulating tumour cells in lymph or blood vessels require homing signals to settle at distant sites. Because primary tumours and metastases are clonally related, it is possible to study the overlap of mutations and the effect of analysis settings.

## Materials and methods

### Medical consent

The present study was approved by the Medical Ethical Committee of the Leiden University Medical Center (protocol P01-019). Informed consent was obtained according to protocols approved by the LUMC Medical Ethical Committee (02-2004). Patient samples were handled according to the medical ethics guidelines described in the Code for Proper Secondary Use of Human Tissue established by the Dutch Federation of Medical Sciences ([www.federa.org](http://www.federa.org); accessed July 2014).

### Sample selection and DNA isolation

Twenty-six colorectal cancers, all diagnosed as adenocarcinomas, that metastasized to the ovaries were selected together with their 30 matching ovarian metastases. In 4 cases, metastases to both ovaries were included. The samples were obtained from the archives of the LUMC Pathology Department (period 1985-2010;  $n = 13$ ) and from PALGA (the nationwide Dutch network and registry of histopathology and cytopathology;  $n = 13$ ).<sup>[25]</sup> The MMR proteins were not stained; however no class 5 (pathogenic) *CTNNB1* variants, which are characteristic of most sporadic MSI-H cancers, were found. Hereditary non-polyposis colorectal cancer (HNPCC) cases were not included. So most likely, the CRC cohort consisted primarily of microsatellite stable (MSS)-*BRAF* negative cases (24/26) and a subset of MSS-*BRAF* positive cases (2/26). The tissue taken for analysis was enriched for tumour cells after the evaluation of haematoxylin and eosin (H&E)-stained slides. Based on this evaluation, 0.6- or 2.0-mm tissue punches were taken from the selected tumour foci in the FFPE block using a tissue microarrayer (Beecher Instruments, Sun Prairie, WI, USA). In cases where the tumour cells were more dispersed, micro-dissection was performed on 10 unstained 10- $\mu$ m sections to achieve the highest tumour percentage (at least 50%). Prior to DNA isolation, the tissue was deparaffinized in xylene and washed in 70% ethanol. DNA was isolated using the NucleoSpin Tissue Genomic DNA Purification Kit (Machery-Nagel, Düren, Germany) according to the manufacturer's instructions.

### Construction of the gene list for target-enriched next-generation sequencing (NGS)

A gene list was compiled based on the most frequently mutated genes in COSMIC and the mutated genes lists described in the literature, resulting in

a selection of 115 genes targeting 0.015% of the human genome (486,013 bp).[26-28] See Supplemental Table 1 for an overview of the genes included.

### **Sample library preparation**

Library preparation was performed according to the HaloPlex protocol (Agilent, Santa Clara, CA, USA). In short, 225 ng of FFPE-DNA was fragmented using 8 pairs of restriction enzymes. Hereafter, the customized probe library was added and hybridized to the targeted fragments. Additionally, a sample barcode sequence was incorporated in this step. Next, the targeted fragments were purified and amplified. The enriched, barcoded samples were sequenced on an Illumina HiSeq 2000. See Supplemental Table 2 for the coverage numbers of the CRCs and the matching metastases to the ovaries.

### **Data analysis**

Adaptors, barcodes and enzyme footprints were removed from the sequenced reads using SureCall software (Agilent Technologies, Santa Clara, CA), after which the reads were aligned to the human genome (hg19) using the Burrows-Wheeler aligner (BWA, version 0.7.5a).[21] The Genome Analysis Toolkit (GATK, version 2.5) was used for realignment around the indels and base quality recalibration.[22] Duplicate removal was not performed due to the nature of hybridization-extension used to capture the target DNA regions. SNP and indel calling were carried out using VarScan software (version v2.3.6) with the following arguments: minimum read depth = 8, minimum number of reads with the alternative allele = 2, minimum base quality = 15, and minimum variant allele frequency = 0.10. VarScan somatic mode was used to analyze the primary vs. metastasis pairs.

Variants were functionally annotated using ANNOVAR.[23, 24] We then selected variants more likely to have a deleterious effect. This was achieved by focusing on splicing and exonic variants (excluding synonymous) and removing the variants that were present with a frequency higher than 1% in the 1000 Genomes project (<http://www.1000genomes.org/>; data from April 2012) and/or in the NHLBI Exome Sequencing Project (<http://evs.gs.washington.edu/EVS/>; data from January 2013) because they are more likely to be germline. Variants in only the primary tumour or the metastasis were visually inspected to identify false discordant calls, i.e., variants that are in fact present in both the primary tumour and the metastasis but that failed the 10% minimum threshold variant allele in one of the tissues.

### **Validation target-enriched next-generation sequencing**

Validation of the variants detected using target-enriched NGS was obtained via allele-specific qPCR of hotspot mutations and classic DNA Sanger sequencing. Allele-specific qPCR was performed to confirm the status of *KRAS*, *BRAF* and *PIK3CA* mutation hotspot loci.[29] *TP53* (exons 5-8) and *APC* (mutation cluster region; exon 15) were analyzed via Sanger DNA sequencing. Sequences were analyzed with Mutation Surveyor (Bioke, Leiden, The Netherlands).

---

## Results

### Patient characteristics

In total, 26 primary CRCs and 30 matching ovarian metastases were tested (in 4 cases, both left and right ovarian metastases were included). The average age at CRC diagnosis was 56 years (range 28-84). In 9 cases (35%), the ovarian metastases were synchronously present at the time of CRC diagnosis (cut-off point 6 months). In the other 17 cases, the ovarian metastases were diagnosed on average 2.9 years later (range 0.5-13). See Table 1 for details.

### Data Analysis

An average of 409 unfiltered variants were identified per sample with a standard deviation of 201. Regarding only variants more likely to have a deleterious effect (see Methods section), the numbers of variants per sample decreased to an average of 37 and a standard deviation of 30. The number of variants detected in this study was higher than the average number of variants observed in comparable sequencing efforts targeting a comparable number of bases.[30, 31] The variant calling parameters were set at relatively low values to prevent passenger mutations from being filtered out. However, because the primary CRCs and the ovarian metastases were analyzed in pairs, when in doubt the variant could be manually compared with the matched sample. If the tumours had been analyzed without matched metastases, the parameters would have been higher to remove false positive variants. For more details, see Table 2 and Supplemental Table 3 for all variants detected.

### Mutation profile of primary CRCs and their metastases to the ovaries

An overview of the number of gene variants per case, varying from 9 to 113, can be observed in Table 2 and Supplemental Table 3. Supplemental Figure 1A shows an overview of the number of variants per gene. *APC*, *FAT4*, *NOTCH1*, *CACNA1B*, *STAB1* and *TP53* showed the highest numbers of variants in the primary CRC samples. The most frequently affected gene was *APC*. Variants in *APC* were identified in 19 out of 26 samples (73%). *FAT4* showed variants in 15 out of 26 analyzed (58%). *NOTCH1* carried variants in 58%. In 54% of the cases, variants were observed in *CACNA1B*, *STAB1* and *TP53*.

In the COSMIC database, the *PIK3CA*, *FBXW7* and *SMAD4* genes were reported to be mutated in CRC with frequencies of ~23%, ~20% and ~26%, re-



spectively.[32] In this cohort, only 2 variants in *PIK3CA* (8%), 5 in *FBXW7* (15%) and 6 in *SMAD4* (23%) were found. To ensure that the low number of mutations in *PIK3CA* was not a sequencing artifact, the coverage was checked. The average coverage for *PIK3CA* was 155x; (median 113x; with a range of 1.4-1401). However, one of the hotspot positions in *PIK3CA* (E542) showed less coverage than other parts of *PIK3CA*. Only one sample showed more than 20 reads at position E542 in *PIK3CA*. To investigate this mutation hotspot position, a TaqMan assay was performed. An additional mutation (p.E542K, c.1624G>A) was found in two cases leading to a mutation frequency in 15% of *PIK3CA* (4/26). Thus, in this cohort of colorectal cancers, no differences were noted in mutation frequency for these driver genes. The most frequently affected genes in the ovarian metastasis were *APC*, *TP53*, *CACNA1B* and *FAT4* (Supplemental Figure 1B). Although the gene lists in Supplemental Figure 1A and 1B slightly appear to contradict the gene lists that are normally reported to be mutated in CRC, the census genes show comparable mutation frequencies.

### **Concordance analysis of genes that select for targeted therapy**

We compared the presence of stratifying mutations in the primary CRCs versus the ovarian metastases that select for established (EGFR directed) or future targeted therapies. These genes comprise *KRAS*, *NRAS*, *BRAF*, *FBXW7*, *PTEN* and *PIK3CA*. *MTOR*, *TSC1* and *TSC2* were not covered in our gene panel. *KRAS* was discordant in 3 of 12 mutated cases; *NRAS* was not discordant (0/4); *BRAF* was discordant in 4/6 mutated cases; *FBXW7* was discordant in 3/3 mutated cases; *PTEN* was discordant in 1/3 mutated cases and lastly *PIK3CA* was discordant in 1/3 mutated cases. Overall 9 gene variants (2x *KRAS*, 3x *BRAF*, 3x *FBXW7*, 1x *PIK3CA*) that were present in the primary CRCs were not found in the metastases. Three gene variants (of *KRAS*, *BRAF* and *PTEN*) that were identified in the metastases were not found in the primary CRCs. Half of the 12 discordant variants were class 4/5 pathogenic variants.

### **Concordance analysis between primary CRCs and matching ovarian metastases: Effect of time intervals**

All genes were analyzed for concordant and discordant variants. In cases of discordant variants, we investigated whether this was caused by an absence of the variant in the primary CRC or the metastasis. There were no variants that were called discordant due to an absence of reads in the matching sample at the position of that specific variant. The total number of discordant vari-

ants was 848. The average number of discordant variants per tumour pair was 28. Most discordant variants were caused by presence in the primary CRC tumours and absence in the metastases. Thus, primary CRCs demonstrate a large cohort of passenger mutations of which only a minor part is present in the matching metastases. The known driver genes were (as expected) mostly concordant. The numbers of discordant variants were more or less comparable between cases with the exception of one case (case number 9), which showed a remarkably high number of unique variants in the primary CRC (see Table 2).

Next, we plotted the time intervals between primary CRCs and metastases versus the amount of unique variants in the metastases. No correlation was observed (Figure 1). For example, case 8 had a long interval of 13 years between the primary tumour and the metastasis; however, only 3 of 16 discordant variants were unique to the metastasis in the ovary.

### **Separate ovarian metastases of the same primary tumour show evidence for different metastasizing patterns.**

In 4 cases, both left and right ovarian metastases of the same CRC were sequenced. Two of the cases (numbers 4 and 17) showed a limited number of additional variants that were shared by both metastases but were not present in the primary CRC. In these cases, both ovaries were likely affected by the same metastasizing clone. In the other two cases (9 and 15), the additional variants that were present in both metastases were not observed in the primary tumour and showed no overlap with each other. In these cases, the metastases to both ovaries are most likely to be independent events originating from different subclones with their own specific mutational profiles. See Supplemental Table 3 for details.

## Discussion and Conclusion

Using the analysis of 115 cancer-driving genes, we compared the mutation profiles of primary CRC and matching ovarian metastases. Mutations could be grossly classified into three categories: mutations that are (1) ubiquitous (present in both the primary tumour and the metastasis), (2) restricted to the primary tumour or (3) only found in the metastasis. We show that loosening filter settings and manual inspection of mutation positions reveal a substantially larger overlap in mutation profiles. Many (driver) mutations are present in both the primary tumour and the metastasis, although sometimes only in a limited number of tumour cells. This could explain the dissimilarities in the mutational status of *KRAS* and *EGFR* in CRCs and hepatic metastases reported earlier.[5]

Primary CRCs and their metastases showed considerable concordance for driver genes. In contrast to the classic driver genes, we identified a subset branch type of genes that displayed substantially less overlap.[33] The primary CRCs show substantially more passenger mutations than the ovarian metastases of CRC. It could be speculated that the large number of passenger mutations in the primary CRCs displays a large number of subclones that are spatially present. In this model certain subclones within primary tumours are most capable of homing into different target organs and even surviving adjuvant therapy. The other subclones do not contribute to the metastasizing process. Vignot et al. observed a similar pattern in lung tumours and their metastases.[6]

We analyzed the number of variants in a temporal context (with metastasis occurring synchronously or metachronously with intervals of up to 13 years). It is assumed that new mutations will arise as time passes between the detection of the primary tumour and the metastasis, leading to more discordant gene variants. For synchronous metastases, the mutation profile is expected to be a comparable reflection of the mutation profile of the primary tumour. However, no correlation between the number of variants and the time interval between primary CRC and matching metastasis was observed. Apparently, the underlying biology driving each individual tumour is more important than the actual intervals between the primary tumour and the metastasis in our cohort. Recent publications indicate differing mutational burdens in different cancer types.[34, 35] These differences can for example occur as a result of highly mutagenic influences (smoking, sunburn, asbestos, etc.), through the

inactivation of DNA repair systems or the activation of APOBEC deaminases (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like).[36] Interestingly, the primary CRC in case number 9 carried a remarkably high number of passenger mutations (see Table 2), possibly caused by an underlying hypermutability deficit. An explanation to keep in mind is that the input DNA could have been of poor quality, leading to false positive variants. This option appears less likely in our study because the DNA quality was checked at case selection. An explanation for comparable mutation profiles between the primary tumour and the metastasis in cases where there is a long period between primary and metastasis detection is that metastases arise early during the development of the disease and are dormant for a period before they present clinically.

Colorectal cancers have been extensively characterized at the molecular level. The genes most frequently mutated in colorectal cancer are *APC*, *TP53*, *KRAS*, *PIK3CA* and *SMAD4*. [37] All other genes are mutated in less than 10% of samples. In our series, *APC* and *TP53* are frequently mutated (in 73% and 58% of the cases). The prevalence of other mutations in our series is comparable with the mutational profiles described in the literature and the COSMIC database.[32] The initial low frequency of *PIK3CA* mutations could be attributed to a low coverage of one of the *PIK3CA* hotspots. After performing an additional TaqMan analysis, the mutation frequency of *PIK3CA* was 15%, equal to data in available databases and previous studies (on average 15%). [37-39]

The link between mutation profiles and metastasizing patterns has been analyzed before. *KRAS* mutations in CRC were found to be associated with lung metastases.[21] CRCs with a wild type *KRAS* status showed more frequent liver and distant lymph node metastases.[22] *KRAS* mutation status is not informative in predicting peritoneal or ovarian metastases. Additionally, *BRAF* mutations are claimed to correlate with higher rates of peritoneal metastasis, distant lymph node metastasis, and lower rates of lung metastasis.[40] However, the number of reports on this topic are limited; thus, the metastasizing pattern in correlation with specific mutation profiles of CRC is not completely understood. CRCs showing metastases to the ovaries did not show a specific profile. Besides, no specifically mutated gene in the metastases that could for example explain the 'homing capacity' of circulating tumour cells was identified.

The use of targeted therapy has become standard practice in advanced CRC; therefore it is important to determine the mutational landscape in various tumour locations within the same patient.[41] Depending on the type of targeted therapy, effectiveness will depend on the absence or presence of certain gene variants preferably in all tumour locations. We now present the first study that compares mutational profiles of ovarian metastases with their matching primary CRCs. Furthermore, druggable or stratifying mutations that select for targeted therapy can be present in one region of the tumour but absent in another, a phenomenon known as intra-tumour heterogeneity (ITH). In this study we compared the primary tumour with one metastatic site. However, only one region of both the primary tumours and the metastases were investigated. When analyzing 6 different genes that select for current or future targeted therapies we found remarkable differences in *KRAS*, *NRAS* and *BRAF* mutational status that can select for *EGFR* directed therapy.[42] We identified differences in 7 of 22 variants when comparing primary tumours and matching metastases. These variants were identified in 20 cases, as in two cases both *KRAS* and *BRAF* mutations were identified, probably in distinct clones. Five of these 7 were not found in the metastases, 2 of 7 not in the primary tumours. Two of these 7 variants were known pathogenic *KRAS* variants (class 5). Also gene variants (of *PIK3CA*, *PTEN*, *FBXW7*) that potentially could select for mTOR pathway(s) directed therapies showed remarkable differences. Will there be any benefit to a patient if only a minor fraction of the tumour mass carries a druggable mutation? Or what will be the benefit if the druggable mutation is present in the primary tumour, but not in the metastases? For CRC, data from actual studies testing these variables are mostly lacking. On the other hand previous studies have shown that *KRAS* pathogenic variations are often concordant between primary tumours and matching metastases. In our study *KRAS*, *NRAS* and *BRAF* mutational status did overlap between primary tumours and matching metastasis in 15/22 cases. In order to avoid the unrealistic goal of testing all tumour sites of every individual tumour it has been proposed that testing circulating free tumour DNA in plasma might be an alternative approach to pursue.

Limitations of our study are the restricted number of genes we investigated. Whole exome/genome sequencing might reveal differences that were not found with our 115-gene panel. Secondly, more extensive molecular characterization of tumours that also includes the analysis of transcriptome, methylome, microRNA, and proteome profiles could potentially show alterations that would explain why a small subset of CRCs show metastasis to ovarian sites.

Finally, the comparison of metastases to both left and right ovaries in individual cases revealed mutations that were shared by both metastases but were not identified within the matching primary tumours in two cases. The latter could suggest that separate metastases of the same primary tumour can have more overlap with each other than with the primary tumour.

In conclusion, this study showed a high concordance rate between CRCs and corresponding ovarian metastases for driver genes but less overlap for passenger genes. Although gene variants currently known to be clinically relevant were largely concordant between primary CRCs and matching metastases to the ovaries, there was a subset of cases that showed differences. The clinical relevance of mutations that are present in only a small percentage of tumour cells needs to be clarified. The number of discordant variants could likely be better explained by intra-tumour characteristics than by the time interval between the primary tumour and metastasis. CRCs metastasizing to the ovaries did not show a specific mutation profile in comparison to consecutive series of CRC, nor did the ovarian metastases.

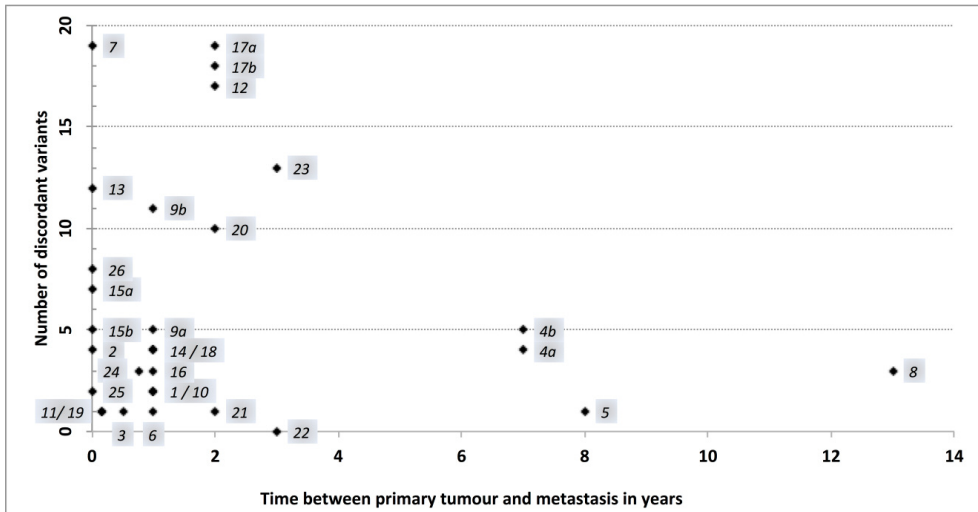
## References

- [1] Mamanova L, Coffey AJ, Scott CE, et al. Target-enrichment strategies for next-generation sequencing. *Nat Methods* 2010;**7(2)**:111-8.
- [2] Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet* 2010;**11(1)**:31-46.
- [3] Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012;**366(10)**:883-92.
- [4] Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res* 2012;**72(19)**:4875-82.
- [5] Vermaat JS, Nijman IJ, Koudijs MJ, et al. Primary colorectal cancers and their subsequent hepatic metastases are genetically different: implications for selection of patients for targeted treatment. *Clin Cancer Res* 2012;**18(3)**:688-99.
- [6] Vignot S, Frampton GM, Soria JC, et al. Next-generation sequencing reveals high concordance of recurrent somatic alterations between primary tumor and metastases from patients with non-small-cell lung cancer. *J Clin Oncol* 2013;**31(17)**:2167-72.
- [7] Patanaphan V, Salazar OM. Colorectal cancer: metastatic patterns and prognosis. *South Med J* 1993;**86(1)**:38-41.
- [8] Lewis MR, Deavers MT, Silva EG, et al. Ovarian involvement by metastatic colorectal adenocarcinoma: still a diagnostic challenge. *Am J Surg Pathol* 2006;**30**:177-84.
- [9] de Waal YR, Thomas CM, Oei AL, et al. Secondary ovarian malignancies: frequency, origin, and characteristics. *Int J Gynecol Cancer* 2009;**19**:1160-5.
- [10] Judson K, McCormick C, Vang R, et al. Women with undiagnosed colorectal adenocarcinomas presenting with ovarian metastases: clinicopathologic features and comparison with women having known colorectal adenocarcinomas and ovarian involvement. *Int J Gynecol Pathol* 2008;**27**:182-90.
- [11] Petru E, Pickel H, Heydarfadai M, et al. Nongenital cancers metastatic to the ovary. *Gynecol Oncol* 1992;**44**:83-6.
- [12] Integrated genomic analyses of ovarian carcinoma. *Nature* 2011;**474**:609-15.
- [13] Vakiani E, Janakiraman M, Shen R, et al. Comparative genomic analysis of primary versus metastatic colorectal carcinomas. *J Clin Oncol* 2012;**30(24)**:2956-62.
- [14] Baas JM, Krens LL, Guchelaar HJ, et al. Concordance of predictive markers for EGFR inhibitors in primary tumors and metastases in colorectal cancer: a review. *Oncologist* 2011;**16(9)**:1239-49.
- [15] Han CB, Li F, Ma JT, et al. Concordant KRAS mutations in primary and metastatic colorectal cancer tissue specimens: a meta-analysis and systematic review. *Cancer Invest* 2012;**30(10)**:741-7.

- 
- [16] Miglio U, Mezzapelle R, Paganotti A, et al. Mutation analysis of KRAS in primary colorectal cancer and matched metastases by means of highly sensitivity molecular assay. *Pathol Res Pract* 2013;**209(4)**:233-6.
- [17] Knijn N, Mekenkamp LJ, Klomp M, et al. KRAS mutation analysis: a comparison between primary tumours and matched liver metastases in 305 colorectal cancer patients. *Br J Cancer* 2011;**104(6)**:1020-6.
- [18] Heitzer E, Auer M, Gasch C, et al. Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res* 2013;**73(10)**:2965-75.
- [19] Goranova TE, Ohue M, Shimoharu Y, et al. Dynamics of cancer cell subpopulations in primary and metastatic colorectal tumors. *Clin Exp Metastasis* 2011;**28(5)**:427-35.
- [20] Baldus SE, Schaefer KL, Engers R, et al. Prevalence and heterogeneity of KRAS, BRAF, and PIK3CA mutations in primary colorectal adenocarcinomas and their corresponding metastases. *Clin Cancer Res* 2010;**16(3)**:790-9.
- [21] Tie J, Lipton L, Desai J, et al. KRAS mutation is associated with lung metastasis in patients with curatively resected colorectal cancer. *Clin Cancer Res* 2011;**17(5)**:1122-30.
- [22] Kim MJ, Lee HS, Kim JH, et al. Different metastatic pattern according to the KRAS mutational status and site-specific discordance of KRAS status in patients with colorectal cancer. *BMC Cancer* 2012;**12**:347-12.
- [23] Brabletz T. To differentiate or not—routes towards metastasis. *Nat Rev Cancer* 2012;**12(6)**:425-36.
- [24] Sakakura C, Hagiwara A, Kato D, et al. Manifestation of bilateral huge ovarian metastases from colon cancer immediately after the initial operation: report of a case. *Surg Today* 2002;**32**:371-5.
- [25] Casparie M, Tiebosch AT, Burger G, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol* 2007;**29**:19-24.
- [26] Starr TK, Allaei R, Silverstein KA, et al. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science* 2009;**323(5922)**:1747-50.
- [27] Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. *Science* 2007;**318(5853)**:1108-13.
- [28] Torkamani A, Schork NJ. Identification of rare cancer driver mutations by network reconstruction. *Genome Res* 2009;**19(9)**:1570-8.
- [29] van ER, Licht J, Schrupf M, et al. Rapid KRAS, EGFR, BRAF and PIK3CA mutation analysis of fine needle aspirates from non-small-cell lung cancer using allele-specific qPCR. *PLoS One* 2011;**6(3)**:e17791.
- [30] Brannon AR, Vakiani E, Sylvester BE, et al. Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic colorectal cancer lesions. *Genome Biol* 2014;**15**:454-0454.



- [31] Han SW, Kim HP, Shin JY, et al. Targeted sequencing of cancer-related genes in colorectal cancer using next-generation sequencing. *PLoS One* 2013;**8**:e64271.
- [32] Bamford S, Dawson E, Forbes S, et al. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer* 2004;**91**:355-8.
- [33] Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res* 2012;**72**:4875-82.
- [34] Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature* 2013;**500**:415-21.
- [35] Jia P, Pao W, Zhao Z. Patterns and processes of somatic mutations in nine major cancers. *BMC Med Genomics* 2014;**19**:7:11.
- [36] Saraconi G, Severi F, Sala C, et al. The RNA editing enzyme APOBEC1 induces somatic mutations and a compatible mutational signature is present in esophageal adenocarcinomas. *Genome Biol* 2014;**15**:417-0417.
- [37] Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012;**487(7407)**:330-7.
- [38] Farina SA, Zeestraten EC, van WT, et al. PIK3CA kinase domain mutation identifies a subgroup of stage III colon cancer patients with poor prognosis. *Cell Oncol* 2011;**34**:523-31.
- [39] Reimers MS, Bastiaannet E, Langley RE, et al. Expression of HLA class I antigen, aspirin use, and survival after a diagnosis of colon cancer. *JAMA Intern Med* 2014;**174**:732-9.
- [40] Tran B, Kopetz S, Tie J, et al. Impact of BRAF mutation and microsatellite instability on the pattern of metastatic spread and prognosis in metastatic colorectal cancer. *Cancer* 2011;**117**:4623-32.
- [41] Martinez P, Birkbak NJ, Gerlinger M, et al. Parallel evolution of tumour subclones mimics diversity between tumours. *J Pathol* 2013;**230(4)**:356-64.
- [42] Van Emburgh BO, Sartore-Bianchi A, Di NF, et al. Acquired resistance to EGFR-targeted therapies in colorectal cancer. *Mol Oncol* 2014;**8**:1084-94.

**Figure 1**

Correlation between the number of discordant variants plotted against the time between the primary tumour and the metastases to the ovaries. The number of discordant variants is plotted on the y-axis (in bold). The time between the primary tumour and the metastases is plotted on the x-axis (in bold). The individual cases (1 to 26) are displayed in italics.

**Table 1**

Total number of patients	26
Age at colorectal cancer diagnosis	
mean	57
range	28-84
Synchronous tumours	9
Metachronous tumours	17
Time between primary tumour resection and resection of metastasis (years)	
mean	2.9
range	0.5-13

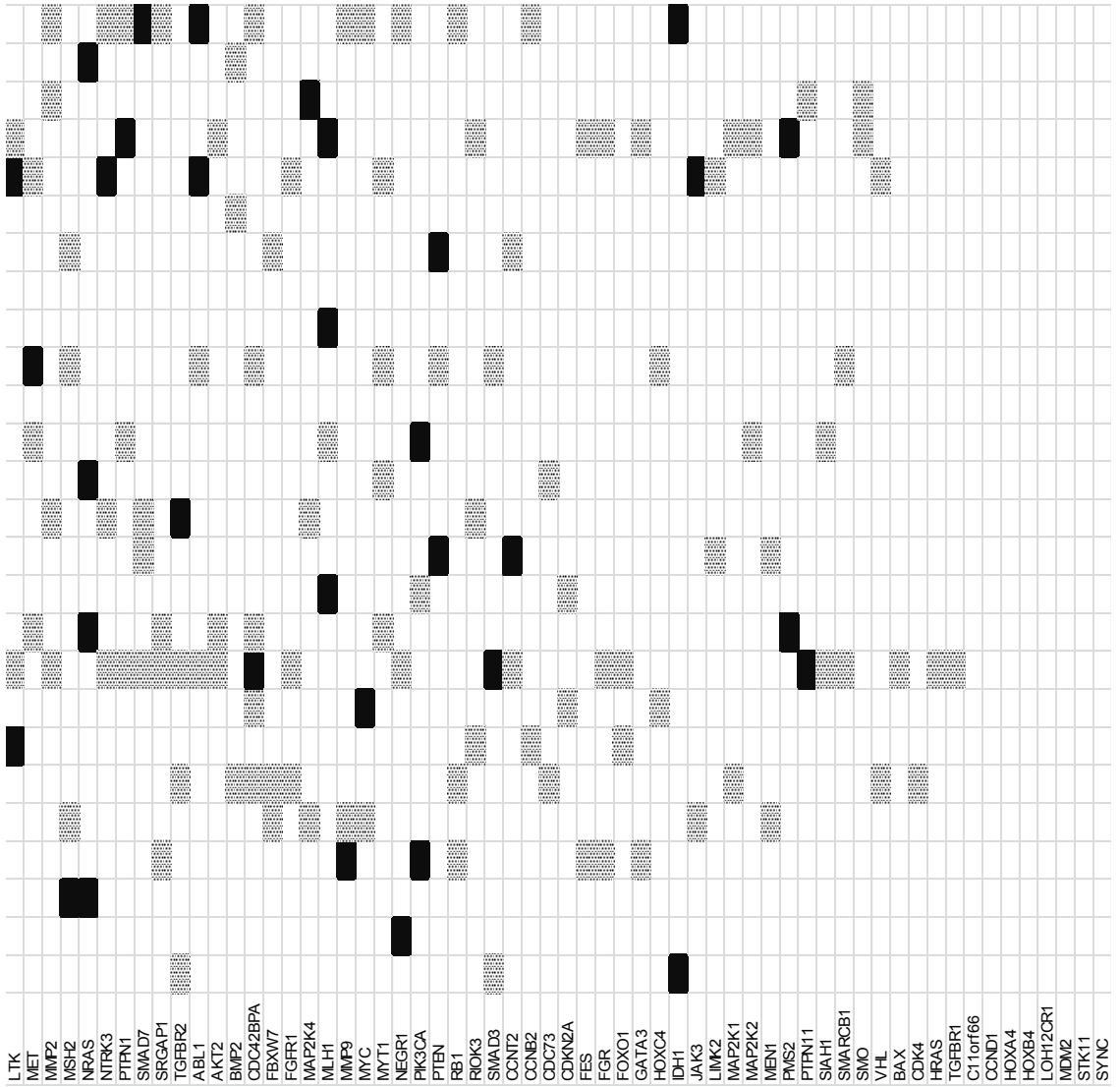
Overview of patient characteristics. Age, synchronous vs. metachronous tumours (cut-off at 6 months) and time between the primary CRC and the metastasis are shown.

**Table 2**

	Total	TotalSelected	Concordant	Discordant	Unique in CRC	Unique in metastasis
1	228	22	16	6	4	2
2	147	15	9	6	2	4
3	365	15	13	2	2	0
4 Left	788	20	11	9	6	3
4 Right	804	22	10	12	8	4
5	162	28	3	25	24	1
6	160	41	3	38	37	1
7	204	33	10	23	5	18
8	256	24	8	16	13	3
9 Left	854	113	12	101	96	5
9 Right	808	109	11	98	88	10
10	251	42	6	36	34	2
11	187	9	7	2	1	1
12	242	25	7	18	2	16
13	286	37	6	31	19	12
14	242	10	5	5	1	4
15 Left	1136	42	10	32	25	7
15 Right	1003	36	9	27	22	5
16	238	15	7	8	5	3
17 Left	892	56	10	46	27	19
17 Right	781	47	8	39	21	18
18	134	16	7	9	5	4
19	136	9	5	4	3	1
20	315	76	11	65	55	10
21	206	15	3	12	12	0
22	209	68	13	55	55	0
23	229	55	7	48	35	13
24	378	25	6	19	16	3
25	238	9	4	5	3	2
26	395	61	10	51	43	8
<b>Total</b>	<b>12274</b>	<b>1095</b>	<b>247</b>	<b>848</b>	<b>669</b>	<b>179</b>

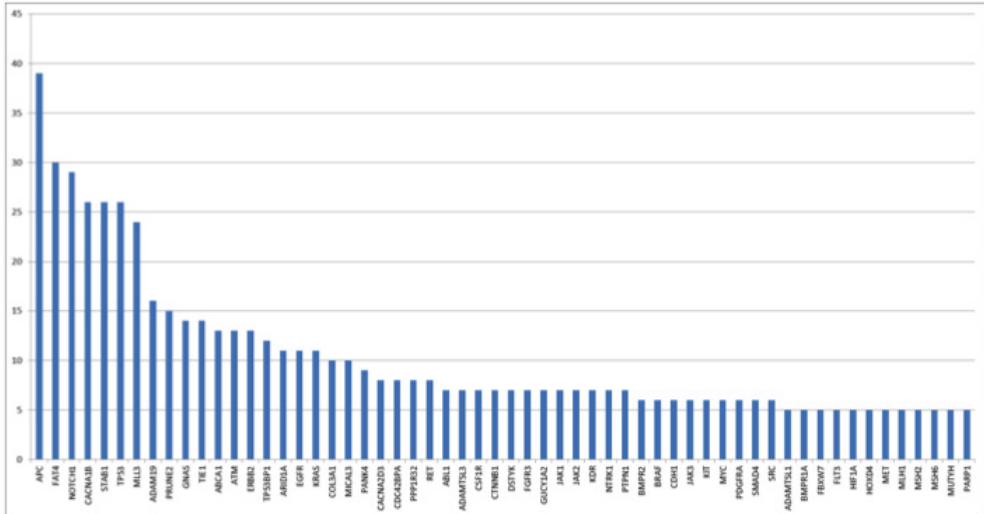
*Overview of paired analysis of primary CRCs and matching metastases to the ovaries. Shown are number of variants, number of selected variants (see “Data analysis in “Materials and methods”), number of concordant and discordant variants, number of unique variants in the primary CRCs and in the metastases to the ovaries.*





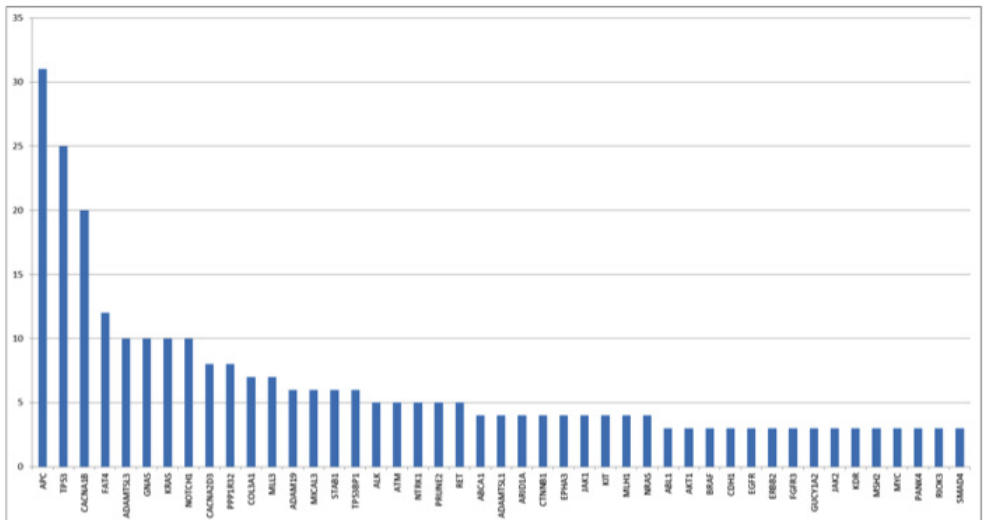
Concordant and discordant variants. Illustration of concordant and discordant variants. Black boxes indicate concordant variants (present in both the primary colorectal tumour as the metastasis to the ovary). Dotted boxes indicate discordant variants (present only in the primary colorectal tumour or in the metastasis to the ovary).

**Supplemental Figure 1A**



Overview of number of variants per gene in metastases to the ovaries from primary CRCs. Genes with 5 or more variants are displayed.

**Supplemental Figure 1B**



Overview of number of variants per gene in metastases to the ovaries from primary CRCs. Genes with 5 or more variants are displayed.

**Supplemental Table 1**

1	<i>ABCA1</i>	31	<i>CSF1R</i>	61	<i>KRAS</i>	91	<i>PRUNE2</i>
2	<i>ABL1</i>	32	<i>CTNNB1</i>	62	<i>LIMK2</i>	92	<i>PTEN</i>
3	<i>ADAM19</i>	33	<i>DSTYK</i>	63	<i>LOH12CR1</i>	93	<i>PTPN1</i>
4	<i>ADAMTSL1</i>	34	<i>EGFR</i>	64	<i>LTK</i>	94	<i>PTPN11</i>
5	<i>ADAMTSL3</i>	35	<i>EPHA3</i>	65	<i>MAP2K1</i>	95	<i>RB1</i>
6	<i>AKT1</i>	36	<i>ERBB2</i>	66	<i>MAP2K2</i>	96	<i>RET</i>
7	<i>AKT2</i>	37	<i>ERBB4</i>	67	<i>MAP2K4</i>	97	<i>RIOK3</i>
8	<i>ALK</i>	38	<i>FAT4</i>	68	<i>MDM2</i>	98	<i>SLAH1</i>
9	<i>ALPK1</i>	39	<i>FBXW7</i>	69	<i>MEN1</i>	99	<i>SMAD2</i>
10	<i>APC</i>	40	<i>FES</i>	70	<i>MET</i>	100	<i>SMAD3</i>
11	<i>ARID1A</i>	41	<i>FGFR1</i>	71	<i>MICAL3</i>	101	<i>SMAD4</i>
12	<i>ATM</i>	42	<i>FGFR3</i>	72	<i>MLH1</i>	102	<i>SMAD7</i>
13	<i>BAX</i>	43	<i>FGR</i>	73	<i>MLL3</i>	103	<i>SMARCB1</i>
14	<i>BMP2</i>	44	<i>FLT3</i>	74	<i>MMP2</i>	104	<i>SMO</i>
15	<i>BMPR1A</i>	45	<i>FOXO1</i>	75	<i>MMP9</i>	105	<i>SRC</i>
16	<i>BMPR2</i>	46	<i>GATA3</i>	76	<i>MSH2</i>	106	<i>SRGAP1</i>
17	<i>BRAF</i>	47	<i>GNAS</i>	77	<i>MSH6</i>	107	<i>STAB1</i>
18	<i>C11orf66</i>	48	<i>GUCY1A2</i>	78	<i>MUTYH</i>	108	<i>STK11</i>
19	<i>CACNA1B</i>	49	<i>HIF1A</i>	79	<i>MYC</i>	109	<i>SYNC</i>
20	<i>CACNA2D3</i>	50	<i>HOXA4</i>	80	<i>MYT1</i>	110	<i>TGFBR1</i>
21	<i>CASR</i>	51	<i>HOXB4</i>	81	<i>NEGR1</i>	111	<i>TGFBR2</i>
22	<i>CCNB2</i>	52	<i>HOXC4</i>	82	<i>NOTCH1</i>	112	<i>TIE1</i>
23	<i>CCND1</i>	53	<i>HOXD4</i>	83	<i>NRAS</i>	113	<i>TP53</i>
24	<i>CCNT2</i>	54	<i>HRAS</i>	84	<i>NTRK1</i>	114	<i>TP53BP1</i>
25	<i>CDC42BPA</i>	55	<i>IDH1</i>	85	<i>NTRK3</i>	115	<i>VHL</i>
26	<i>CDC73</i>	56	<i>JAK1</i>	86	<i>PANK4</i>		
27	<i>CDH1</i>	57	<i>JAK2</i>	87	<i>PARP1</i>		
28	<i>CDK4</i>	58	<i>JAK3</i>	88	<i>PDGFRA</i>		
29	<i>CDKN2A</i>	59	<i>KDR</i>	89	<i>PIK3CA</i>		
30	<i>COL3A1</i>	60	<i>KIT</i>	90	<i>PMS2</i>		

List of genes included in the custom-made gene panel (n=115). The total number of genes is shown.

**Supplemental Table 2**

Primary CRC (2A)	total	mean	%_bases - above_1	%_bases - above_10	%_bases - above_20	%_bases - above_30
1	480470949	462,25	97,3	87,9	81,4	77
2	195102371	187,7	91,1	73,8	66,4	61,5
3	2249426017	2164,12	99,1	97,5	96,4	95,7
4	160424949	154,34	87,4	67,4	59,3	53,9
5	158871721	152,85	86,2	67,5	59,3	53,7
6	69207190	66,58	68,1	45,7	37,5	32,8
7	306852498	295,22	89,9	70,5	62,7	58
8	486964922	468,5	89,5	77,6	71,1	67
9	115616104	111,23	70	55,7	47,6	42,5
10	91273070	87,81	77,7	57,2	48,3	42,1
11	251614742	242,07	93,3	75,6	67,6	62,3
12	275390044	264,95	95,7	82,7	75,4	70,4
13	308248474	296,56	83,6	66,4	59,1	54,5
14	309295455	297,57	95,2	81,8	74,2	69,1
15	408945444	393,44	84,5	69,8	63,5	59,3
16	462954757	445,4	93,3	79,9	72,9	68,4
17	136442316	131,27	78,9	60,8	53,1	47,9
18	407641847	392,18	89	73,9	67,3	62,9
19	117302213	112,85	94,7	77,6	67,6	60,7
20	126240021	121,45	66,4	50,4	42,8	37,9
21	462349947	444,82	83,1	64,9	58,2	53,9
22	13653814	13,14	51,8	24	16	11,8
23	173730165	167,14	78,6	65,7	57,8	53
24	551493045	530,58	88,8	76,8	70,8	66,7
25	346645995	333,5	96,4	84,7	78,2	73,4
26	460541334	443,08	78,6	66,6	61,1	57,5

Coverage for (2A) primary CRC tumours and (2B) metastases to the ovaries. Total reads, mean coverage and percentage of bases covering more than 1x, 10x, 20x and 30x is shown for (2A) primary CRCs and (2B) metastases to the ovaries.



Chapter 4

<b>CRC metastases to the ovary (2B)</b>	<b>total</b>	<b>mean</b>	<b>%_bases_ above_1</b>	<b>%_bases_ above_10</b>	<b>%_bases_ above_20</b>	<b>%_bases_ above_30</b>
1	258895399	249,08	96	82,6	75,1	69,7
2	284445865	273,66	93,8	78,1	70,8	66,3
3	419398099	403,49	98,1	92,2	86,9	82,6
4 Left	125729981	120,96	96,1	80,3	70,7	64,5
4 Right	168253406	161,87	96,6	83,2	74,4	68,1
5	166188874	159,89	91,7	70,8	61,5	55,6
6	238299380	229,26	93,4	77,1	69,3	64
7	528614247	508,57	90,9	80,9	74,8	70,6
8	410658067	395,08	97,5	89	82,5	77,8
9 Left	170601993	164,13	89,2	70,6	62,4	56,7
9 Right	151499411	145,75	87,4	67,3	58,9	53,4
10	370868504	356,8	95,3	83,6	76,7	72
11	531435788	511,28	95,3	83,5	76,9	72,5
12	366236567	352,35	88,3	75,8	69,2	64,6
13	386080901	371,44	85,9	70	62,8	58
14	404134644	388,81	97	87,3	80,9	75,9
15 Left	411089878	395,5	91,4	77,3	70,5	65,8
15 Right	213475530	205,38	91,9	74,1	65,7	60,2
16	316669636	304,66	97	89,1	83,1	78,2
17 Left	450911323	433,81	83,9	70,8	64,2	59,7
17 Right	122588632	117,94	87,1	65,7	56,6	50,9
18	148608397	142,97	93,9	74,8	64,9	58,3
19	287800179	276,89	96,2	84,5	77,3	72,2
20	303684904	292,17	88,7	74,7	68,4	64
21	377617408	363,3	96,1	84,7	77,3	72,5
22	270909355	260,64	91,1	74,4	66,6	61,3
23	191464997	184,2	88,4	69,8	61,5	56,2
24	825364679	794,06	95,6	88,2	83	79,2
25	465015253	447,38	97,3	88,3	82	77,6
26	325136401	312,81	86,9	73,1	66,7	62,6

Supplemental Table 3

	<b>gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
<b>1</b>	ADAM19	c.G379A	p.V127I	NM_033274	x	x
	APC	c.C847T	p.R283X	NM_001127510	x	x
	APC	c.4462_4463insTA	p.L1488fs	NM_001127510	x	x
	CASR	c.G2536A	p.V846I	NM_001178065	x	x
	COL3A1	c.G2035A	p.A679T	NM_000090	x	x
	COL3A1	c.G3626C	p.G1209A	NM_000090	x	x
	CTNNB1	c.T1862G	p.L621R	NM_001904	x	x
	GNAS	c.T689C	p.L230P	NM_001077490	x	x
	IDH1	c.C394T	p.R132C	NM_005896	x	x
	KRAS	c.G35T	p.G12V	NM_033360	x	x
	NOTCH1	c.A1045C	p.T349P	NM_017617	x	x
	NTRK1	c.C16T	p.R6W	NM_001012331	x	x
	PDGFRA	c.G769C	p.G257R	NM_006206	x	x
	PPP1R32	c.A422G	p.Q141R	NM_145017	x	x
	PPP1R32	c.C713A	p.T238N	NM_145017	x	x
	TP53	c.G396C	p.K132N	NM_001126114	x	x
	APC	c.G6914C	p.R2305T	NM_001127510	x	
	NTRK1	c.T1253G	p.L418R	NM_001012331	x	
	PARP1	c.G205A	p.V69I	NM_001618	x	
	TGFBR2	c.G1583A	p.R528H	NM_003242	x	
FAT4	c.T7886C	p.F2629S	NM_024582		x	
SMAD3	c.A1277T	p.X426L	NM_005902		x	
<b>2</b>	<b>gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
	ABCA1	c.G5563T	p.A1855S	NM_005502	x	x
	ADAMTSL3	c.A4547T	p.N1516I	NM_207517	x	x
	CACNA1B	c.A2986G	p.T996A	NM_000718	x	x
	GUCY1A2	c.G1975T	p.G659W	NM_001256424	x	x

MLL3	c.T13931G	p.V4644G	NM_170606	x	
MSH6	c.C1186G	p.L396V	NM_000179	x	
NEGR1	c.C94G	p.L32V	NM_173808	x	
NOTCH1	c.A1762C	p.T588P	NM_017617	x	
TP53	c.T325G	p.F109V	NM_001126114	x	
ARID1A	c.1057dupG	p.A353fs	NM_006015	x	
TIE1	c.C1765T	p.R589W	NM_005424	x	
FGFR3	c.C2423A	p.T808K	NM_001163213	x	
MICAL3	c.C4379A	p.P1460H	NM_015241	x	
MICAL3	c.C478T	p.R160C	NM_015241	x	
MICAL3	c.4495dupC	p.R1499fs	NM_015241	x	
<b>3 gene</b>	<b>cdNA</b>	<b>protein</b>	<b>transcript</b>	<b>M</b>	
ADAM19	c.G1303A	p.E435K	NM_033274	x	
FAT4	c.G13789T	p.D4597Y	NM_024582	x	
GNAS	c.T689C	p.L230P	NM_001077490	x	
GNAS	c.C1798G	p.R600G	NM_080425	x	
MSH2	c.A182G	p.Q61R	NM_000251	x	
MSH2	c.G2500A	p.A834T	NM_000251	x	
NOTCH1	c.A1762C	p.T588P	NM_017617	x	
NOTCH1	c.A1045C	p.T349P	NM_017617	x	
NRAS	c.C181A	p.Q61K	NM_002524	x	
NTRK1	c.T1253G	p.L418R	NM_001012331	x	
PPP1R32	c.C190A	p.Q64K	NM_145017	x	
PPP1R32	c.G582T	p.L194F	NM_145017	x	
TP53	c.G743A	p.R248Q	NM_001126114	x	
ARID1A	c.1057dupG	p.A353fs	NM_006015	x	
PPP1R32	c.1062dupC	p.P354fs	NM_145017	x	
<b>4 gene</b>	<b>cdNA</b>	<b>protein</b>	<b>transcript</b>	<b>M Left</b>	<b>M Right</b>
APC	c.C1660T	p.R554X	NM_001127510	x	x
APC	c.T4092A	p.S1364R	NM_001127510	x	x
APC	c.C4348T	p.R1450X	NM_001127510	x	x

CACNA1B	c.C501G	p.N167K	NM_000718	x	x
MMP9	c.C1199T	p.A400V	NM_004994	x	x
NOTCH1	c.A931C	p.T311P	NM_017617	x	x
NTRK1	c.G53A	p.G18E	NM_001012331	x	x
PIK3CA	c.T1031G	p.V344G	NM_006218	x	x
TP53	c.G404T	p.C135F	NM_001126114	x	x
ALK	c.G3082C	p.E1028Q	NM_004304	x	x
GNAS	c.T689C	p.L230P	NM_001077490	x	x
MLL3	c.G7388A	p.R2463H	NM_170606	x	x
ARID1A	c.1057dupG	p.A353fs	NM_006015	x	x
FES	c.C841T	p.R281W	NM_001143783	x	x
FGR	c.G715A	p.A239T	NM_001042729	x	x
GATA3	c.G295A	p.G99S	NM_001002295	x	x
HOXD4	c.G548A	p.R183H	NM_014621	x	x
MUTYH	c.A1634C	p.N545T	NM_001128425	x	x
NTRK1	c.G1828T	p.D610Y	NM_001012331	x	x
PIK3CA	c.T1031G	p.V344G	NM_006218	x	x
RB1	c.C1727A	p.S576X	NM_000321	x	x
FGFR3	c.C2204G	p.A735G	NM_001163213	x	x
KIT	c.A851G	p.D284G	NM_000222	x	x
TIE1	c.C242T	p.A81V	NM_005424	x	x
SRGAP1	c.C391A	p.Q131K	NM_020762	x	x
<b>5 gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
GNAS	c.C1268A	p.P423H	NM_001077490	x	x
PANK4	c.G1702A	p.D568N	NM_018216	x	x
TP53	c.C586T	p.R196X	NM_001126114	x	x
ADAM19	c.C2608T	p.R870C	NM_033274	x	x
ADAMTSL1	c.G3064A	p.D1022N	NM_001040272	x	x
ALPK1	c.G1423T	p.G475X	NM_025144	x	x
APC	c.G8162A	p.R2721H	NM_001127510	x	x
BMPR1A	c.G742T	p.V248L	NM_004329	x	x
COL3A1	c.C1112A	p.P371H	NM_000090	x	x

CSF1R	c.C2695A	p.H899N	NM_005211	x	
ERBB2	c.C121A	p.L41M	NM_001005862	x	
FBXW7	c.C174A	p.S58R	NM_018315	x	
FGFR3	c.G755A		NM_001163213	x	
JAK3	c.G1918T	p.D640Y	NM_000215	x	
KDR	c.C1136A	p.A379E	NM_002253	x	
MAP2K4	c.C859T	p.R287C	NM_003010	x	
MEN1	c.C1453T	p.R485W	NM_130803	x	
MICAL3	c.C4374A	p.S1458R	NM_015241	x	
MLL3	c.C6029A	p.P2010Q	NM_170606	x	
MMP9	c.G742A	p.G248S	NM_004994	x	
MSH6	c.C3013T	p.R1005X	NM_000179	x	
MUTYH	c.C346T	p.R116W	NM_001128425	x	
MYC	c.G949T	p.V317F	NM_002467	x	
MYC	c.G1009A	p.A337T	NM_002467	x	
STAB1	c.C5794A	p.H1932N	NM_015136	x	
TIE1	c.G2335T	p.A779S	NM_005424	x	
TP53BP1	c.G4792T	p.E1598X	NM_005657	x	
MSH2	c.C1322T	p.T441I	NM_000251	x	
<b>6 gene</b>	<b>cdNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
CACNA1B	c.A2986G	p.T996A	NM_000718	x	x
TP53	c.G733A	p.G245S	NM_001126114	x	x
TP53BP1	c.T880C	p.S294P	NM_005657	x	x
ADAM19	c.C2506A	p.P836T	NM_033274	x	
ALK	c.C2039G	p.T680R	NM_004304	x	
ATM	c.A3008T	p.Q1003L	NM_000051	x	
ATM	c.C3059T	p.T1020I	NM_000051	x	
BMP2	c.C77T	p.P26L	NM_001200	x	
CACNA1B	c.C6394T	p.R2132C	NM_000718	x	
CACNA1B	c.3825dupG	p.V1275fs	NM_000718	x	
CDC42BPA	c.G4636A	p.E1546K	NM_003607	x	
CDC42BPA	c.T3344C	p.V1115A	NM_003607	x	

CDC73	c.A238T	p.T80S	NM_024529	x	
CDH1	c.G371A	p.R124H	NM_004360	x	
CDK4	c.G365A	p.R122H	NM_000075	x	
COL3A1	c.G2374A	p.G792R	NM_000090	x	
CSF1R	c.G2843T	p.S948I	NM_005211	x	
DSTYK	c.G1184T	p.R395L	NM_015375	x	
DSTYK	c.A1180G	p.T394A	NM_015375	x	
ERBB4	c.C2566A	p.H856N	NM_005235	x	
ERBB4	c.G868A	p.V290I	NM_005235	x	
FAT4	c.C5513T	p.T1838I	NM_024582	x	
FAT4	c.G9697T	p.V3233L	NM_024582	x	
FBXW7	c.T1072A	p.S358T	NM_018315	x	
FBXW7	c.G678T	p.Q226H	NM_018315	x	
FGFR1	c.C350T	p.S117L	NM_001174064	x	
FLT3	c.T2083G	p.C695G	NM_004119	x	
GNAS	c.G3068T	p.R1023L	NM_080425	x	
JAK1	c.G535A	p.E179K	NM_002227	x	
JAK2	c.G3188A	p.R1063H	NM_004972	x	
MAP2K1	c.G602A	p.R201H	NM_002755	x	
MICAL3	c.C2233T	p.R745W	NM_015241	x	
MLL3	c.G8263A	p.D2755N	NM_170606	x	
MLL3	c.C7384T	p.Q2462X	NM_170606	x	
PRUNE2	c.G5416T	p.A1806S	NM_015225	x	
PRUNE2	c.G3184A	p.G1062S	NM_015225	x	
RB1	c.G1618A	p.G540S	NM_000321	x	
TGFBR2	c.G610A	p.G204S	NM_003242	x	
TIE1	c.C2693T	p.P898L	NM_005424	x	
VHL	c.T407G	p.F136C	NM_000551	x	
ERBB4	c.T3305C	p.F1102S	NM_005235	x	
<b>7 gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
ADAMTSL3	c.C2872T	p.R958W	NM_207517	x	x
ADAMTSL3	c.A4547T	p.N1516I	NM_207517	x	x

8	gene	cDNA	protein	transcript	P	M
APC	c.T4308A	p.S1436R		NM_001127510	x	x
APC	c.1192_1193del	p.398_398del		NM_001127510	x	x
APC	c.4307delG	p.S1436fs		NM_001127510	x	x
CACNA2D3	c.G1333A	p.D445N		NM_018398	x	x
CDH1	c.C1792T	p.R598X		NM_004360	x	x
KRAS	c.G35A	p.G12D		NM_033360	x	x
LTK	c.G728A	p.R243Q		NM_002344	x	x
TP53	c.G818A	p.R273H		NM_001126114	x	x
FAT4	c.G5621T	p.G1874V		NM_024582	x	x
FAT4	c.C5890T	p.L1964F		NM_024582	x	x
FAT4	c.G12772A	p.V4258I		NM_024582	x	x
APC	c.G4307A	p.S1436N		NM_001127510	x	x
NOTCH1	c.G4144A	p.E1382K		NM_017617	x	x
ADAMTSL3	c.C1439A	p.P480H		NM_207517	x	x
ADAMTSL3	c.C4942T	p.R1648W		NM_207517	x	x
BMP2	c.G1165T	p.E389X		NM_001204	x	x
CACNA2D3	c.G1700A	p.R567Q		NM_018398	x	x
CACNA2D3	c.G1708A	p.V570M		NM_018398	x	x
CACNA2D3	c.G1994A	p.R665H		NM_018398	x	x
CCNB2	c.C978A	p.N326K		NM_004701	x	x
CDH1	c.C1115A	p.P372H		NM_004360	x	x
COL3A1	c.C3189A	p.D1063E		NM_000090	x	x
FAT4	c.G2280T	p.G760H		NM_024582	x	x
FOXO1	c.C14T	p.P5L		NM_002015	x	x
KDR	c.G2837A	p.R946H		NM_002253	x	x
MLL3	c.C5935A	p.P1979T		NM_170606	x	x
PPP1R32	c.C216A	p.N72K		NM_145017	x	x
PRUNE2	c.G3413T	p.W1138L		NM_015225	x	x
RIOK3	c.G977A	p.R326H		NM_003831	x	x
STAB1	c.G7060A	p.D2354N		NM_015136	x	x
TIE1	c.G2489T	p.S830I		NM_005424	x	x

APC	c.A3173G	p.D1058G	NM_001127510	x	
APC	c.C3944A	p.S1315X	NM_001127510	x	
CTNNB1	c.A1276G	p.N426D	NM_001904	x	
FAT4	c.C905T	p.T302M	NM_024582	x	
KRAS	c.G35T	p.G12V	NM_033360	x	
MYC	c.C1088T	p.S363L	NM_002467	x	
SMAD4	c.552_553insCCACCAAGTA	p.H184fs	NM_005359	x	
STAB1	c.C1909G	p.L637V	NM_015136	x	
ALK	c.C32T	p.P11L	NM_004304	x	
ATM	c.C2941T	p.R981C	NM_000051	x	
BMPR1A	c.C760T	p.R254C	NM_004329	x	
CDC42BPA	c.C1054T	p.R352W	NM_003607	x	
CDC42BPA	c.G694T	p.V232F	NM_003607	x	
CDKN2A	c.C295T	p.R99C	NM_058195	x	
FGFR3	c.G2035A	p.V679I	NM_001163213	x	
JAK3	c.C424T	p.R142C	NM_000215	x	
MYC	c.C977T	p.A326V	NM_002467	x	
NOTCH1	c.G5422A	p.D1808N	NM_017617	x	
SMAD7	c.C1049T	p.P350L	NM_005904	x	
STAB1	c.G7258A	p.A2420T	NM_015136	x	
TIE1	c.C202T	p.R68C	NM_005424	x	
ARID1A	c.4684_4685del	p.1562_1562del	NM_006015	x	
CACNA1B	c.G2713T	p.E905X	NM_000718	x	
HOXC4	c.C646T	p.R216X	NM_014620	x	
<b>9 gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M Left</b> <b>M Right</b>
ALK	c.A3089C	p.H1030P	NM_004304	x	x
CACNA1B	c.C501G	p.N167K	NM_000718	x	x
CACNA1B	c.G3457A	p.E1153K	NM_000718	x	x
FGFR3	c.G2278A	p.D760N	NM_001163213	x	x
RET	c.C950T	p.T317M	NM_020975	x	x
TP53	c.652_654del	p.218_218del	NM_001126114	x	x
AKT1	c.C1424T	p.S475L	NM_001014432	x	x



ALK	c.G1756A	p.A586T	NM_004304	x	x
BMPR2	c.C673T	p.R225C	NM_001204	x	x
CACNA2D3	c.G1309A	p.V437M	NM_018398	x	x
EPHA3	c.C2182T	p.R728X	NM_005233	x	x
PARP1	c.G1687A	p.V563I	NM_001618	x	x
CDC42BPA	c.C2651T	p.S884L	NM_003607	x	x
MICAL3	c.C2689T	p.R897C	NM_015241	x	x
PTPN11	c.C1241T	p.T414M	NM_002834	x	x
SMAD3	c.C238T	p.R80W	NM_005902	x	x
STAB1	c.G3839A	p.R1280H	NM_015136	x	x
ABCA1	c.G5347A	p.V1783M	NM_005502	x	x
ABCA1	c.C3084A	p.S1028R	NM_005502	x	x
ABL1	c.C2191T	p.R731C	NM_007313	x	x
ABL1	c.C2243T	p.T748M	NM_007313	x	x
ADAM19	c.G1817A	p.R606Q	NM_033274	x	x
ADAM19	c.G1036A	p.E346K	NM_033274	x	x
ADAMTSL1	c.C3669A	p.F1223L	NM_001040272	x	x
ADAMTSL3	c.C502A	p.Q168K	NM_207517	x	x
ADAMTSL3	c.G1895A	p.R632Q	NM_207517	x	x
AKT1	c.C1408T	p.P470S	NM_001014432	x	x
AKT2	c.G1249A	p.V417M	NM_001626	x	x
AKT2	c.G268A	p.V90M	NM_001626	x	x
ALK	c.G1110T	p.E370D	NM_004304	x	x
APC	c.G8162A	p.R2721H	NM_001127510	x	x
ARID1A	c.G4624A	p.E1542K	NM_006015	x	x
ARID1A	c.G5725A	p.A1909T	NM_006015	x	x
BMPR2	c.G893T	p.W298L	NM_001204	x	x
BMPR2	c.G2076T	p.Q692H	NM_001204	x	x
BRAF	c.C1100T	p.P367L	NM_004333	x	x
CACNA1B	c.G1088A	p.R363Q	NM_000718	x	x
CACNA1B	c.G3763A	p.V1255I	NM_000718	x	x
CACNA1B	c.C5446T	p.Q1816X	NM_000718	x	x
CASR	c.G1683T	p.R561S	NM_001178065	x	x

CDH1	c.G2336A	p.R779Q	NM_004360	x
CDH1	c.G2494A	p.V832M	NM_004360	x
CDKN2B	c.C244T	p.R82X	NM_004936	x
CSF1R	c.G2335A	p.V779M	NM_005211	x
CSF1R	c.C110T	p.T37M	NM_005211	x
CTNNB1	c.C548A	p.A183D	NM_001904	x
CTNNB1	c.G1564A	p.A522T	NM_001904	x
DSTYK	c.C1813T	p.R605W	NM_015375	x
EGFR	c.C549A	p.H183Q	NM_005228	x
EGFR	c.G1522T	p.A508S	NM_005228	x
EGFR	c.G1648T	p.V550L	NM_005228	x
EPHA3	c.G544T	p.A182S	NM_005233	x
ERBB2	c.G380A	p.R127Q	NM_001005862	x
ERBB2	c.G1211A	p.R404Q	NM_001005862	x
ERBB2	c.G1625A	p.G542D	NM_001005862	x
ERBB2	c.G3053A	p.R1018H	NM_001005862	x
ERBB2	c.G3460A	p.V1154I	NM_001005862	x
ERBB2	c.G3481A	p.V1161M	NM_001005862	x
FAT4	c.G13885A	p.A4629T	NM_024582	x
FBXW7	c.G188A	p.G63D	NM_018315	x
FGFR3	c.G580A	p.E194K	NM_001163213	x
FGFR3	c.G713A	p.R238Q	NM_001163213	x
FGR	c.C18A	p.C6X	NM_005248	x
FOXO1	c.G641T	p.R214L	NM_002015	x
GNAS	c.C727T	p.R243W	NM_001077489	x
HOXD4	c.G161A	p.R54Q	NM_014621	x
HRAS	c.C382T	p.R128W	NM_176795	x
JAK3	c.C1744T	p.R582W	NM_000215	x
KRAS	c.C378A	p.D126E	NM_033360	x
LTK	c.C971T	p.A324V	NM_002344	x
MLL3	c.C11651A	p.T3884K	NM_170606	x
MLL3	c.G10156T	p.D3386Y	NM_170606	x
MLL3	c.C9967A	p.H3323N	NM_170606	x

MLL3	c.C8290A	p.L2764I	NM_170606	x
MSH6	c.C3013T	p.R1005X	NM_000179	x
NEGR1	c.C412A	p.P138T	NM_173808	x
NOTCH1	c.A7502T	p.Q2501L	NM_017617	x
NOTCH1	c.G5632A	p.G1878R	NM_017617	x
NOTCH1	c.G3224A	p.W1075X	NM_017617	x
NOTCH1	c.G1309A	p.E437K	NM_017617	x
NTRK3	c.G1729A	p.E577K	NM_001007156	x
NTRK3	c.C442T	p.Q148X	NM_001007156	x
PANK4	c.C1805A	p.S602Y	NM_018216	x
PARP1	c.C193T	p.R65W	NM_001618	x
PMS2	c.G632A	p.R211Q	NM_000535	x
PPP1R32	c.C737A	p.T246N	NM_145017	x
PRUNE2	c.G8889T	p.L2963F	NM_015225	x
PTPN1	c.C617T	p.P206L	NM_002827	x
PTPN1	c.C1145T	p.A382V	NM_002827	x
RET	c.C1102T	p.R368C	NM_020975	x
SIAH1	c.C721T	p.R241X	NM_003031	x
SMAD4	c.G404A	p.R135Q	NM_005359	x
SMARCB1	c.C197T	p.S66L	NM_003073	x
SMARCB1	c.C801A	p.N267K	NM_003073	x
SRC	c.C128T	p.S43L	NM_005417	x
SRC	c.G829T	p.G277C	NM_005417	x
SRC	c.G1261T	p.A421S	NM_005417	x
SRGAP1	c.C974T	p.A325V	NM_020762	x
SRGAP1	c.G2546A	p.R849Q	NM_020762	x
STAB1	c.G4100A	p.C1367Y	NM_015136	x
STAB1	c.G4580A	p.R1527H	NM_015136	x
STAB1	c.G5242A	p.G1748S	NM_015136	x
TGFBR1	c.G899T	p.R300M	NM_001130916	x
TGFBR2	c.G610A	p.G204S	NM_003242	x
TIE1	c.C1777T	p.R593W	NM_005424	x
TP53BP1	c.G5872T	p.G1958X	NM_005657	x

TP53BP1	c.G4538C	p.G1513A	NM_005657	x	
TP53BP1	c.C3754T	p.R1252C	NM_005657	x	
TP53BP1	c.G490T	p.A164S	NM_005657	x	
ADAMTSL3	c.G1262A	p.R421H	NM_207517	x	
COL3A1	c.G1715A	p.R572Q	NM_000090	x	
MDM2	c.C748A	p.Q250K	NM_002392	x	
MSH6	c.C3226T	p.R1076C	NM_000179	x	
MYC	c.G997T	p.D333Y	NM_002467	x	
BAX	c.G579A	p.W193X	NM_004324		x
BRAF	c.G1385A	p.R462K	NM_004333		x
CCNT2	c.C676A	p.P226T	NM_058241		x
FGFR1	c.C2369T	p.P790L	NM_001174064		x
MMP2	c.G1942A	p.V648M	NM_004530		x
NOTCH1	c.G4861A	p.G1621S	NM_017617		x
NTRK3	c.C1552T	p.R518C	NM_001007156		x
SMAD2	c.C340T	p.R114C	NM_001003652		x
SMAD4	c.C884T	p.P295L	NM_005359		x
SMAD7	c.C973T	p.R325W	NM_005904		x

10	gene	cDNA	protein	transcript	P	M
	APC	c.G3916T	p.E1306X	NM_001127510	x	x
	FAT4	c.G6608A	p.R2203Q	NM_024582	x	x
	HIF1A	c.A1949G	p.E650G	NM_001530	x	x
	MLL3	c.C12118A	p.L4040I	NM_170606	x	x
	NRAS	c.C181A	p.Q61K	NM_002524	x	x
	TP53	c.583dupA	p.I195fs	NM_001126114	x	x
	ABCA1	c.G1882A	p.V628I	NM_005502	x	
	ADAMTSL1	c.G2807A	p.R936H	NM_001040272	x	
	AKT2	c.G253A	p.E85K	NM_001626	x	
	ALK	c.C550T	p.R184X	NM_004304	x	
	ALK	c.C193T	p.R65C	NM_004304	x	
	ALPK1	c.C3314A	p.P1105H	NM_025144	x	
	ARID1A	c.C5626A	p.P1876T	NM_006015	x	

ATM	c.C2074T	p.R692C	NM_000051	x
CACNA2D3	c.G2123A	p.W708X	NM_018398	x
CDC42BPA	c.G420T	p.L140F	NM_003607	x
CDKN2B	c.C244T	p.R82X	NM_004936	x
CTNNB1	c.G635A	p.R212H	NM_001904	x
ERBB4	c.C738A	p.C246X	NM_005235	x
FAT4	c.C1668G	p.D556E	NM_024582	x
FAT4	c.C10775T	p.P3592L	NM_024582	x
FAT4	c.C11374T	p.R3792W	NM_024582	x
FGFR3	c.G1253A	p.R418H	NM_001163213	x
JAK1	c.C525A	p.D175E	NM_002227	x
KIT	c.G25A	p.D9N	NM_000222	x
MET	c.A1948G	p.S650G	NM_001127500	x
MUTYH	c.C1547T	p.P516L	NM_001128425	x
MYT1	c.G286T	p.D96Y	NM_004535	x
NOTCH1	c.G5885A	p.R1962H	NM_017617	x
NOTCH1	c.G4699A	p.E1567K	NM_017617	x
NOTCH1	c.C3394A	p.R1132S	NM_017617	x
NOTCH1	c.G3100A	p.G1034S	NM_017617	x
NOTCH1	c.G2983A	p.G995S	NM_017617	x
NOTCH1	c.C339A	p.N113K	NM_017617	x
NTRK1	c.C1079T	p.T360M	NM_001012331	x
PRUNE2	c.G1193T	p.S398I	NM_015225	x
SMAD4	c.G1393A	p.V465M	NM_005359	x
SRGAP1	c.C1912A	p.L638I	NM_020762	x
STAB1	c.G4078A	p.E1360K	NM_015136	x
TIE1	c.G1354T	p.A452S	NM_005424	x
CACNA1B	c.G2422T	p.D808Y	NM_000718	x
STAB1	c.G5615A	p.R1872H	NM_015136	x
<b>11 gene</b>	<b>cdNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>
APC	c.C637T	p.R213X	NM_001127510	x
JAK2	c.A1936C	p.N646H	NM_004972	x

KRAS	c.G38A	p.G13D	NM_033360	x	x
MLH1	c.G2146A	p.V716M	NM_000249	x	x
PRUNE2	c.C1132G	p.L378V	NM_015225	x	x
TP53	c.G796A	p.G266R	NM_001126114	x	x
TP53BP1	c.C5242T	p.R1748C	NM_005657	x	x
PIK3CA	c.G2593A	p.G865S	NM_006218	x	x
KDR	c.55delG	p.A19fs	NM_002253	x	x
<b>12</b>	<b>gene</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
ADAM19	c.C187T	p.L63F	NM_033274	x	x
APC	c.G4057T	p.E1353X	NM_001127510	x	x
CCNT2	c.C1984T	p.R662W	NM_001241	x	x
KRAS	c.G35T	p.G12V	NM_033360	x	x
MUTYH	c.A536G	p.Y179C	NM_001128425	x	x
PRUNE2	c.C3313T	p.R1105W	NM_015225	x	x
PTEN	c.C388T	p.R130X	NM_000314	x	x
ATM	c.G1010A	p.R337H	NM_000051	x	x
PANK4	c.G1628A	p.R543H	NM_018216	x	x
ADAM19	c.G1835A	p.R612Q	NM_033274	x	x
ADAMTSL3	c.G4030T	p.V1344F	NM_207517	x	x
AKT1	c.C1436T	p.T479M	NM_001014432	x	x
BRAF	c.G1271A	p.R424Q	NM_004333	x	x
CACNA1B	c.G566A	p.R189Q	NM_000718	x	x
CSF1R	c.G2068A	p.G690S	NM_005211	x	x
EGFR	c.C2481A	p.Y827X	NM_005228	x	x
EPHA3	c.C2395T	p.R799C	NM_005233	x	x
GNAS	c.C854T	p.S285F	NM_080425	x	x
GNAS	c.C1684T	p.R562C	NM_080425	x	x
LIMK2	c.C583T	p.R195X	NM_016733	x	x
MEN1	c.C1456T	p.R486W	NM_130803	x	x
SMAD2	c.C169T	p.R57X	NM_001003652	x	x
SMAD4	c.C1067T	p.P356L	NM_005359	x	x
SMAD7	c.C1121T	p.A374V	NM_005904	x	x

13	gene	cDNA	protein	transcript	P	M
STAB1		c.G16A	p.G6S	NM_015136		x
	<b>gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
	APC	c.4010_4011del	p.1337_1337del	NM_001127510	x	x
	NOTCH1	c.G5215A	p.V1739M	NM_017617	x	x
	NTRK1	c.G2321A	p.R774Q	NM_001012331	x	x
	SMAD4	c.G1082A	p.R361H	NM_005359	x	x
	TGFBR2	c.G5C	p.G2A	NM_003242	x	x
	TP53	c.C1009T	p.R337C	NM_001126114	x	x
	ABCA1	c.G5950A	p.V1984I	NM_005502	x	
	ABCA1	c.G4844A	p.R1615Q	NM_005502	x	
	ADAM19	c.G1531A	p.A511T	NM_033274	x	
	ADAM19	c.C181A	p.H61N	NM_033274	x	
	APC	c.G4349A	p.R1450Q	NM_001127510	x	
	BMPR1A	c.C760T	p.R254C	NM_004329	x	
	CACNA1B	c.G3917A	p.G1306E	NM_000718	x	
	CACNA1B	c.C6017T	p.A2006V	NM_000718	x	
	CASR	c.G854A	p.R285Q	NM_001178065	x	
	CDH1	c.G1423A	p.V475M	NM_004360	x	
	COL3A1	c.G1787A	p.R596Q	NM_000090	x	
	JAK1	c.G1028A	p.R343Q	NM_002227	x	
	KDR	c.G3288A	p.W1096X	NM_002253	x	
	MAP2K4	c.G810T	p.M270I	NM_003010	x	
	MICAL3	c.C367T	p.R123C	NM_015241	x	
	MLL3	c.G13199A	p.R4400Q	NM_170606	x	
	PANK4	c.G502A	p.V168M	NM_018216	x	
	PRUNE2	c.G3674T	p.R1225I	NM_015225	x	
	SMAD7	c.G265T	p.E89X	NM_005904	x	
	ADAMTSL1	c.G3268A	p.E1090K	NM_001040272		x
	ADAMTSL1	c.C4545A	p.C1515X	NM_001040272		x
	ADAMTSL1	c.G4676A	p.R1559Q	NM_001040272		x
	CACNA2D3	c.C824T	p.A275V	NM_018398		x
	COL3A1	c.G187A	p.D63N	NM_000090		x

CSF1R	c.G1660A	p.E554K	NM_005211	X
JAK1	c.G1595A	p.R532H	NM_002227	X
MMP2	c.G1942A	p.V648M	NM_004530	X
NTRK3	c.G80A	p.G27D	NM_001007156	X
PANK4	c.G1658A	p.R553Q	NM_018216	X
RET	c.C691T	p.R231C	NM_020975	X
RIOK3	c.G1255A	p.V419I	NM_003831	X

## 14

gene	cDNA	protein	transcript	M
APC	c.3921_3925del	p.1307_1309del	NM_001127510	X
EPHA3	c.C2330G	p.A777G	NM_005233	X
JAK1	c.A2578T	p.K860X	NM_002227	X
NRAS	c.C181A	p.Q61K	NM_002524	X
TP53	c.G273A	p.W91X	NM_001126114	X
CDC73	c.C787T	p.R263C	NM_024529	X
ARID1A	c.G28A	p.A10T	NM_006015	X
CDC73	c.C1563A	p.Y521X	NM_024529	X
MLL3	c.T5863G	p.S1955A	NM_170606	X
NOTCH1	c.T1034C	p.F345S	NM_017617	X

## 15

gene	cDNA	protein	transcript	M Left	M Right
ALK	c.A3089C	p.H1030P	NM_004304	X	X
CACNA1B	c.C501G	p.N167K	NM_000718	X	X
FAT4	c.T8651A	p.L2884H	NM_024582	X	X
GNAS	c.C676T	p.R226C	NM_016592	X	X
KRAS	c.G35C	p.G12A	NM_033360	X	X
PARP1	c.G2656A	p.V886M	NM_001618	X	X
PPP1R32	c.G347A	p.R116Q	NM_145017	X	X
TP53	c.C817T	p.R273C	NM_001126114	X	X
TP53	c.G473A	p.R158H	NM_001126114	X	X
TP53BP1	c.C2930A	p.A977D	NM_005657	X	X
ADAM19	c.G349A	p.G117S	NM_033274	X	X
ADAMTSL3	c.G4535A	p.R1512Q	NM_207517	X	X



CDC73	c.C314T	p.S105L	NM_024529	x
COL3A1	c.C2173A	p.Q725K	NM_000090	x
COL3A1	c.C3155A	p.P1052H	NM_000090	x
EGFR	c.C2288A	p.A763D	NM_005228	x
FAT4	c.G8560T	p.V2854L	NM_024582	x
FLT3	c.G2572A	p.E858K	NM_004119	x
GNAS	c.G3068T	p.R1023L	NM_080425	x
KDR	c.C481A	p.L161I	NM_002253	x
MAP2K2	c.C893T	p.P298L	NM_030662	x
MAP2K2	c.C806T	p.P269L	NM_030662	x
MEN1	c.C1450T	p.R484W	NM_130803	x
MET	c.G3608T	p.G1203V	NM_001127500	x
MICAL3	c.G3036T	p.E1012D	NM_015241	x
MICAL3	c.G1710T	p.L570F	NM_015241	x
MLH1	c.C1757A	p.A586D	NM_000249	x
MLL3	c.G6197A	p.R2066Q	NM_170606	x
MSH2	c.C2579T	p.S860L	NM_000251	x
MYC	c.G939T	p.K313N	NM_002467	x
NOTCH1	c.G7426A	p.V2476M	NM_017617	x
NOTCH1	c.G3262A	p.G1088S	NM_017617	x
NOTCH1	c.G2068A	p.G690R	NM_017617	x
PDGFRA	c.G255T	p.L85F	NM_006206	x
PDGFRA	c.C272T	p.S91L	NM_006206	x
PRUNE2	c.C3320T	p.T1107M	NM_015225	x
PRUNE2	c.G6106T	p.D2036Y	NM_015225	x
PTPN1	c.C505T	p.R169X	NM_002827	x
SHAH1	c.C77T	p.R3C	NM_003031	x
STAB1	c.G7258A	p.A2420T	NM_015136	x
SYNC	c.G1378T	p.G460C	NM_001161708	x
CCNT2	c.A854G	p.N285S	NM_058241	x
CDC42BPA	c.G523T	p.E175X	NM_003607	x
GATA3	c.C407A	p.A136D	NM_001002295	x
KDR	c.C543A	p.S181R	NM_002253	x

MLL3	c.G11473A	p.A3825T	NM_170606	x	
MMP2	c.G1144A	p.D382N	NM_004530	x	
NOTCH1	c.G4336A	p.E1446K	NM_017617	x	
ADAMTSL1	c.C4897A	p.Q1633K	NM_001040272		x
ERBB4	c.A2491G	p.M831V	NM_005235		x
ERBB4	c.A1801C	p.N601H	NM_005235		x
HOXD4	c.G202A	p.G68R	NM_014621		x
SMAD2	c.G1202T	p.G401V	NM_001003652		x
<b>16 gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>M</b>	
ADAM19	c.C2080T	p.P694S	NM_033274	x	
APC	c.G2950T	p.E984X	NM_001127510	x	
APC	c.4241dupT	p.V1414fs	NM_001127510	x	
CACNA1B	c.G4166A	p.R1389H	NM_000718	x	
FAT4	c.T5539C	p.S1847P	NM_024582	x	
KRAS	c.G38A	p.G13D	NM_033360	x	
TP53	c.G524A	p.R175H	NM_001126114	x	
APC	c.G1085C	p.G362A	NM_001127510	x	
FAT4	c.C1897T	p.R633C	NM_024582	x	
HOXD4	c.G268A	p.A90T	NM_014621	x	
MLL3	c.G7544C	p.R2515T	NM_170606	x	
PANK4	c.G2147A	p.R716H	NM_018216	x	
ATM	c.G3682T	p.E1228X	NM_000051	x	
CACNA1B	c.G2685T	p.E895D	NM_000718	x	
ERBB4	c.A3005C	p.K1002T	NM_005235	x	
<b>17 gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>M Left</b>	<b>M Right</b>
ALK	c.A3089C	p.H1030P	NM_004304	x	x
APC	c.C2413T	p.R805X	NM_001127510	x	x
APC	c.4462_4463insTA	p.L1488fs	NM_001127510	x	x
CACNA1B	c.C501G	p.N167K	NM_000718	x	x
MET	c.C2962T	p.R988C	NM_001127500	x	x
MLL3	c.C12655G	p.L4219V	NM_170606	x	x

TP53	c.G826C	p.A276P	NM_001126114	X	X
ABCA1	c.6203delT	p.L2068fs	NM_005502	X	X
STAB1	c.G6442A	p.A2148T	NM_015136	X	X
TP53BP1	c.G3146A	p.R1049Q	NM_005657	X	X
FAT4	c.G6569A	p.R2190H	NM_024582	X	X
ADAMTSL3	c.C2812T	p.R938X	NM_207517	X	X
AKT1	c.C1108T	p.R370C	NM_001014432	X	X
ARID1A	c.C5372A	p.S1791X	NM_006015	X	X
CACNA1B	c.G52A	p.G18R	NM_000718	X	X
CACNA1B	c.C2758T	p.R920W	NM_000718	X	X
CACNA1B	c.G4123A	p.V1375M	NM_000718	X	X
CDC42BPA	c.A116G	p.Y39C	NM_003607	X	X
COL3A1	c.C4275A	p.S1425R	NM_000090	X	X
CSF1R	c.G2762A	p.R921Q	NM_005211	X	X
FAT4	c.T5276C	p.I1759T	NM_024582	X	X
FAT4	c.G6569A	p.R2190H	NM_024582	X	X
FES	c.G2206T	p.E736X	NM_001143783	X	X
GUCY1A2	c.G1982A	p.R661Q	NM_001256424	X	X
HOXC4	c.G260T	p.G87V	NM_014620	X	X
JAK1	c.G3190A	p.V1064I	NM_002227	X	X
JAK1	c.G323A	p.R108Q	NM_002227	X	X
KDR	c.G3428A	p.W1143X	NM_002253	X	X
MEN1	c.C1450T	p.R484W	NM_130803	X	X
MMP2	c.G1472A	p.R491Q	NM_004530	X	X
MSH2	c.G469T	p.G157C	NM_000251	X	X
MUTYH	c.G1318T	p.G440W	NM_001128425	X	X
PRUNE2	c.C3320T	p.T1107M	NM_015225	X	X
RET	c.G2935A	p.E979K	NM_020975	X	X
SMAD4	c.G1060A	p.V354M	NM_005359	X	X
SRC	c.G659T	p.R220L	NM_005417	X	X
STAB1	c.G4031A	p.G1344E	NM_015136	X	X
STK11	c.C16G	p.P6A	NM_000455	X	X
TIE1	c.C202T	p.R68C	NM_005424	X	X

ATM	c.T3401C	p.M1134T	NM_000051	x	x
FAT4	c.C13223A	p.P4408H	NM_024582	x	x
ABCA1	c.G6181A	p.G2061S	NM_005502	x	
ADAMTSL3	c.C3485T	p.S1162L	NM_207517	x	
APC	c.4724delT	p.L1575fs	NM_001127510	x	
CACNA1B	c.C2800T	p.R934C	NM_000718	x	
CACNA1B	c.G4306A	p.E1436K	NM_000718	x	
COL3A1	c.C2296A	p.P766T	NM_000090	x	
COL3A1	c.C2512A	p.P838T	NM_000090	x	
DSTYK	c.C1774T	p.R592W	NM_015375	x	
FLT3	c.C1159T	p.R387X	NM_004119	x	
HOXD4	c.G61A	p.E21K	NM_014621	x	
HOXD4	c.G187A	p.G63R	NM_014621	x	
JAK1	c.G415A	p.E139K	NM_002227	x	
MYC	c.G1096T	p.E366X	NM_002467	x	
NOTCH1	c.G7426A	p.V2476M	NM_017617	x	
RET	c.C1094T	p.S365L	NM_020975	x	
RIOK3	c.G1372A	p.A458T	NM_003831	x	
TP53	c.G998A	p.R333H	NM_001126114	x	
ABL1	c.C1777T	p.R593X	NM_007313	x	x
ADAMTSL3	c.C340T	p.R114W	NM_207517	x	x
ADAMTSL3	c.C2433A	p.H811Q	NM_207517	x	x
AKT1	c.A380G	p.D127G	NM_001014432	x	x
ALPK1	c.G1543T	p.D515Y	NM_025144	x	x
CDH1	c.C1517A	p.T506K	NM_004360	x	x
CSF1R	c.G1819A	p.E607K	NM_005211	x	x
HIF1A	c.G1507A	p.D503N	NM_001530	x	x
JAK2	c.G508T	p.V170L	NM_004972	x	x
NEGR1	c.G670T	p.A224S	NM_173808	x	x
NOTCH1	c.G4906A	p.E1636K	NM_017617	x	x
PARP1	c.G1195A	p.G399R	NM_001618	x	x
PTEN	c.G151T	p.D51Y	NM_000314	x	x
SMAD3	c.C5T	p.S2L	NM_005902	x	x

x  
x

NM\_005902  
NM\_003073

p.W30L  
p.R190W

c.G89T  
c.C568T

SMAD3  
SMARCB1

<b>18</b>	<b>gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
	APC	c.C2701T	p.G901X	NM_001127510	x	x
	APC	c.G3949T	p.E1317X	NM_001127510	x	x
	FAT4	c.C13921A	p.Q4641K	NM_024582	x	x
	KRAS	c.G35T	p.G12V	NM_033360	x	x
	MLH1	c.A1852G	p.K618E	NM_000249	x	x
	MLH1	c.A1853C	p.K618T	NM_000249	x	x
	TP53	c.G646A	p.V216M	NM_001126114	x	x
	ALK	c.C32T	p.P11L	NM_004304	x	
	BRAF	c.G998T	p.S333I	NM_004333	x	
	CACNA1B	c.G415A	p.G139R	NM_000718	x	
	CACNA1B	c.G3949A	p.E1317K	NM_000718	x	
	FAT4	c.G6025A	p.E2009K	NM_024582	x	
	ATM	c.G5432A	p.C1811Y	NM_000051		x
	EPHA3	c.G1615T	p.V539F	NM_182644		x
	FAT4	c.G3719T	p.G1240V	NM_024582		x
<b>19</b>	<b>gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
	ADAMTSL1	c.G2465A	p.G822D	NM_001040272	x	x
	APC	c.2941delC	p.P981fs	NM_001127510	x	x
	CTNNB1	c.G569A	p.R190H	NM_001904	x	x
	JAK2	c.G3188A	p.R1063H	NM_004972	x	x
	KRAS	c.G35T	p.G12V	NM_033360	x	x
	APC	c.C4132T	p.Q1378X	NM_001127510	x	
	CACNA1B	c.C2755T	p.H919Y	NM_000718	x	
	NOTCH1	c.C5920T	p.Q1974X	NM_017617	x	
	CACNA1B	c.G1628A	p.R543Q	NM_000718		x
<b>20</b>	<b>gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
	ABCA1	c.G4402A	p.D1468N	NM_005502	x	x

ATM	c.C2074T	p.R692C	NM_000051	x	x
CACNA2D3	c.G994A	p.A332T	NM_018398	x	x
EGFR	c.G2030A	p.R677H	NM_005228	x	x
GUCY1A2	c.G1982A	p.R661Q	NM_001256424	x	x
KIT	c.A848G	p.N283S	NM_000222	x	x
PANK4	c.G2275A	p.G759S	NM_018216	x	x
PTEN	c.G625A	p.G209R	NM_000314	x	x
RET	c.C1573T	p.R525W	NM_020975	x	x
TP53	c.G733A	p.G245S	NM_001126114	x	x
TP53BP1	c.C3712T	p.R1238C	NM_005657	x	x
ABCA1	c.C3556A	p.L1186I	NM_005502	x	x
ABL1	c.C2041T	p.P681S	NM_007313	x	x
ADAM19	c.G796A	p.G266R	NM_033274	x	x
AKT2	c.G1318A	p.D440N	NM_001626	x	x
AKT2	c.G1303A	p.D435N	NM_001626	x	x
AKT2	c.G535A	p.A179T	NM_001626	x	x
ALPK1	c.C298T	p.L100F	NM_025144	x	x
ALPK1	c.C2015T	p.T672I	NM_025144	x	x
ALPK1	c.G2295T	p.Q765H	NM_025144	x	x
APC	c.3072delA	p.P1024fs	NM_001127510	x	x
ARID1A	c.G4564A	p.A1522T	NM_006015	x	x
ATM	c.G341A	p.R114K	NM_000051	x	x
ATM	c.A3008T	p.Q1003L	NM_000051	x	x
CACNA1B	c.T4219C	p.F1407L	NM_000718	x	x
CACNA2D3	c.C1211A	p.A404E	NM_018398	x	x
CCNT2	c.C1933T	p.R645W	NM_001241	x	x
CDH1	c.G1360A	p.V454I	NM_004360	x	x
CSF1R	c.C2344A	p.R782S	NM_005211	x	x
DSTYK	c.C2125T	p.R709X	NM_015375	x	x
DSTYK	c.C1555T	p.Q519X	NM_015375	x	x
EGFR	c.G2509A	p.D837N	NM_005228	x	x
EPHA3	c.G2158T	p.V720F	NM_005233	x	x
ERBB2	c.G784A	p.G262S	NM_001005862	x	x

ERBB2	c.G1211A	p.R404Q	NM_001005862	x
ERBB2	c.G3145A	p.E1049K	NM_001005862	x
FAT4	c.C1085T	p.S362L	NM_024582	x
FAT4	c.C8999T	p.T3000M	NM_024582	x
FAT4	c.G12503A	p.R4168H	NM_024582	x
FAT4	c.G14075A	p.C4692Y	NM_024582	x
FBXW7	c.C859T	p.R287X	NM_018315	x
GNAS	c.C2524T	p.R842C	NM_080425	x
GUCY1A2	c.G2057A	p.R686H	NM_001256424	x
GUCY1A2	c.G2027A	p.S676N	NM_001256424	x
GUCY1A2	c.G1007A	p.S336N	NM_001256424	x
JAK1	c.G1501A	p.E501K	NM_002227	x
JAK2	c.A652G	p.I218V	NM_004972	x
JAK3	c.C3124T	p.R1042W	NM_000215	x
KDR	c.G2413A	p.V805I	NM_002253	x
KIT	c.G532A	p.A178T	NM_000222	x
MLL3	c.C13886T	p.S4629L	NM_170606	x
MLL3	c.G4742A	p.S1581N	NM_170606	x
MLL3	c.6186_6187insC	p.A2063fs	NM_170606	x
MSH2	c.G583T	p.G195X	NM_000251	x
MYT1	c.G2030A	p.R677H	NM_004535	x
NOTCH1	c.G4663A	p.E1555K	NM_017617	x
NOTCH1	c.G3862A	p.V1288I	NM_017617	x
PANK4	c.G1507A	p.V503M	NM_018216	x
PANK4	c.G601A	p.V201M	NM_018216	x
PDGFRA	c.2453delA	p.D818fs	NM_006206	x
PTPN1	c.C333A	p.N111K	NM_002827	x
SMAD4	c.C565T	p.R189C	NM_005359	x
STAB1	c.G3325A	p.G1109S	NM_015136	x
STAB1	c.G4615A	p.E1539K	NM_015136	x
TIE1	c.C2116T	p.R706C	NM_005424	x
TP53	c.C1097T	p.S366F	NM_001126114	x
ADAMTSL3	c.C265T	p.R89W	NM_207517	x

APC	c.G4349A	p.R1450Q	NM_001127510	x
ARID1A	c.C2596T	p.R866W	NM_006015	x
ARID1A	c.G6259A	p.G2087R	NM_006015	x
CACNA1B	c.C6394T	p.R2132C	NM_000718	x
CDH1	c.G1501A	p.V501M	NM_004360	x
ERBB2	c.G773A	p.R258Q	NM_001005862	x
NOTCH1	c.G1453C	p.V485L	NM_017617	x
RET	c.C1094T	p.S365L	NM_020975	x
STAB1	c.G6260A	p.R2087H	NM_015136	x

21 gene	cdna	protein	transcript	P	M
APC	c.3709delC	p.Q1237fs	NM_001127510	x	x
MICAL3	c.G4529A	p.R1510Q	NM_015241	x	x
TP53	c.A659G	p.Y220C	NM_001126114	x	x
ATM	c.G949A	p.D317N	NM_000051	x	
ATM	c.A967T	p.I323L	NM_000051	x	
BMP2	c.C250T	p.R84C	NM_001200	x	
EGFR	c.C2325A	p.C775X	NM_005228	x	
FAT4	c.A5923T	p.I1975F	NM_024582	x	
FAT4	c.C7319T	p.A2440V	NM_024582	x	
FAT4	c.G8774A	p.G2925E	NM_024582	x	
FLT3	c.C2251A	p.Q751K	NM_004119	x	
FLT3	c.G606T	p.Q202H	NM_004119	x	
HIF1A	c.C1489T	p.Q497X	NM_001530	x	
PRUNE2	c.C7163T	p.S2388L	NM_015225	x	
STAB1	c.G3034A	p.G1012S	NM_015136	x	

22 gene	cdna	protein	transcript	P	M
ABL1	c.C2269T	p.R757W	NM_007313	x	x
CACNA1B	c.C6143T	p.S2048L	NM_000718	x	x
CACNA2D3	c.G1196A	p.R399Q	NM_018398	x	x
EGFR	c.C1591T	p.R531X	NM_005228	x	x
GNAS	c.C2611T	p.R871C	NM_080425	x	x



JAK3	c.A2333G	p.N778S	NM_000215	x
JAK3	c.C2141T	p.T714M	NM_000215	x
KIT	c.C2540T	p.T847M	NM_000222	x
LTK	c.C1304G	p.S435X	NM_002344	x
MLL3	c.G9755A	p.R3252H	NM_170606	x
NTRK3	c.C1375T	p.R459W	NM_001007156	x
SRC	c.C1022T	p.T341M	NM_005417	x
TP53	c.293delC	p.P98fs	NM_001126114	x
ABCA1	c.C4462T	p.Q1488X	NM_005502	x
ABCA1	c.C3166A	p.L1056M	NM_005502	x
ABL1	c.C2347T	p.R783W	NM_007313	x
ADAM19	c.G2269A	p.G757R	NM_033274	x
ADAM19	c.G1748A	p.R583Q	NM_033274	x
ADAMTSL1	c.C2318T	p.S773L	NM_001040272	x
ALK	c.C4162T	p.Q1388X	NM_004304	x
ARID1A	c.A3112T	p.M1038L	NM_006015	x
ARID1A	c.C4894A	p.P1632T	NM_006015	x
ATM	c.C2738A	p.T913N	NM_000051	x
BMPR2	c.C1585T	p.R529C	NM_001204	x
CASR	c.C35G	p.A12G	NM_001178065	x
COL3A1	c.G2155A	p.A719T	NM_000090	x
CSF1R	c.G1897A	p.E633K	NM_005211	x
DSTYK	c.C1060T	p.R354C	NM_015375	x
DSTYK	c.2038delC	p.L680fs	NM_015375	x
EGFR	c.G1300A	p.G434S	NM_005228	x
EGFR	c.T1737G	p.C579W	NM_005228	x
EPHA3	c.A488G	p.K163R	NM_005233	x
ERBB2	c.G3262A	p.V1088I	NM_001005862	x
ERBB4	c.G488A	p.W163X	NM_005235	x
FGFR1	c.C1625T	p.T542M	NM_001174064	x
FGFR1	c.A771C	p.K257N	NM_001174064	x
FGFR1	c.G646C	p.D216H	NM_001174064	x
FGFR3	c.G2420A	p.R807Q	NM_001163213	x

FLT3	c.T103G	p.C35G	NM_004119	x	
GNAS	c.C923T	p.P308L	NM_080425	x	
KDR	c.G2806A	p.V936I	NM_002253	x	
LIMK2	c.A988G	p.K330E	NM_016733	x	
MET	c.C1234T	p.R412C	NM_001127500	x	
MET	c.A1241T	p.D414V	NM_001127500	x	
MICAL3	c.C5735T	p.S1912L	NM_015241	x	
MICAL3	c.C3692A	p.P1231Q	NM_015241	x	
MICAL3	c.G3559A	p.A1187T	NM_015241	x	
MLL3	c.A7168G	p.I2390V	NM_170606	x	
MSH6	c.A203G	p.K68R	NM_000179	x	
MUTYH	c.226delG	p.V76fs	NM_001128425	x	
MYT1	c.G2224A	p.E742K	NM_004535	x	
MYT1	c.C2831T	p.P944L	NM_004535	x	
NOTCH1	c.G6328A	p.V2110M	NM_017617	x	
NOTCH1	c.G1270A	p.E424K	NM_017617	x	
PDGFRA	c.1733delT	p.M578fs	NM_006206	x	
PRUNE2	c.A8199T	p.K2733N	NM_015225	x	
PRUNE2	c.G4460A	p.R1487Q	NM_015225	x	
PRUNE2	c.C4259A	p.S1420Y	NM_015225	x	
RET	c.C913T	p.P305S	NM_020975	x	
RET	c.G1588T	p.E530X	NM_020975	x	
SMAD2	c.G1291T	p.V431L	NM_001003652	x	
STAB1	c.G4732A	p.G1578S	NM_015136	x	
STAB1	c.G5450A	p.R1817Q	NM_015136	x	
STAB1	c.G5489A	p.R1830Q	NM_015136	x	
TIE1	c.C10T	p.R4W	NM_005424	x	
TP53	c.A1103C	p.H368P	NM_001126114	x	
VHL	c.C394T	p.Q132X	NM_000551	x	
VHL	c.T407G	p.F136C	NM_000551	x	
<b>23 gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
FAT4	c.C10847T	p.T3616M	NM_024582	x	x

JAK2	c.G143A	p.G48E	NM_004972	x
MLH1	c.G2146A	p.V716M	NM_000249	x
PTPN1	c.G71A	p.R24Q	NM_002827	x
PTPN1	c.C1117T	p.R373W	NM_002827	x
TP53	c.C742T	p.R248W	NM_001126114	x
TP53BP1	c.G4346A	p.R1449Q	NM_005657	x
ABCA1	c.G5572A	p.V1858M	NM_005502	x
ADAM19	c.G2071A	p.D691N	NM_033274	x
AKT2	c.G1441A	p.E481K	NM_001626	x
APC	c.3856delG	p.E1286fs	NM_001127510	x
ARID1A	c.G2196T	p.Q732H	NM_006015	x
BMPR2	c.A1400G	p.K467R	NM_001204	x
BMPR2	c.C1606T	p.R536C	NM_001204	x
BRAF	c.C1672T	p.R558X	NM_004333	x
CACNA1B	c.C1612T	p.P538S	NM_000718	x
CACNA2D3	c.G1906A	p.E636K	NM_018398	x
CTNNB1	c.C470A	p.T157K	NM_001904	x
EGFR	c.G967T	p.V323F	NM_005228	x
ERBB2	c.G2020A	p.G674R	NM_001005862	x
ERBB2	c.G3667A	p.V1223M	NM_001005862	x
FES	c.C868T	p.R290W	NM_001143783	x
FGR	c.G1090A	p.A364T	NM_001042729	x
GATA3	c.G428A	p.G143E	NM_001002295	x
GNAS	c.G3068T	p.R1023L	NM_080425	x
GUCY1A2	c.G1991A	p.R664H	NM_001256424	x
LTK	c.C2032T	p.R678C	NM_002344	x
MAP2K1	c.G649A	p.D217N	NM_002755	x
MAP2K2	c.C806T	p.P269L	NM_030662	x
MLL3	c.G4855A	p.A1619T	NM_170606	x
MSH6	c.C3202T	p.R1068X	NM_000179	x
NOTCH1	c.G4823A	p.R1608H	NM_017617	x
PARP1	c.C700A	p.L234I	NM_001618	x
RET	c.G3106T	p.E1036X	NM_020975	x

RET	c.C3233T	p.T1078M	NM_020975	X	
RIOK3	c.G944A	p.R315H	NM_003831	X	
STAB1	c.G1897A	p.V633M	NM_015136	X	
STAB1	c.G2896A	p.A966T	NM_015136	X	
STAB1	c.G4474A	p.G1492R	NM_015136	X	
STAB1	c.G4745A	p.R1582H	NM_015136	X	
STAB1	c.G7654A	p.D2552N	NM_015136	X	
TIE1	c.C3040T	p.R1014C	NM_005424	X	
ADAMTSL3	c.G1190A	p.S397N	NM_207517	X	
AKT1	c.G49A	p.E17K	NM_001014432	X	
APC	c.C2626T	p.R876X	NM_001127510	X	
ERBB2	c.G3053A	p.R1018H	NM_001005862	X	
GUCY1A2	c.G2260A	p.G754S	NM_001256424	X	
JAK1	c.C3410T	p.S1137F	NM_002227	X	
KRAS	c.G38A	p.G13D	NM_033360	X	
MAP2K2	c.C1027G	p.Q343E	NM_030662	X	
MICAL3	c.C4064T	p.T1355M	NM_015241	X	
PDGFRA	c.C2760A	p.H920Q	NM_006206	X	
PRUNE2	c.C170A	p.P57Q	NM_015225	X	
SMAD2	c.G680T	p.S227I	NM_001003652	X	
SMO	c.C77T	p.P26L	NM_005631	X	
<b>24 gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>	<b>M</b>
APC	c.G2097A	p.W699X	NM_001127510	X	X
APC	c.4359delT	p.P1453fs	NM_001127510	X	X
BRAF	c.G1405A	p.G469R	NM_004333	X	X
MAP2K4	c.A362C	p.H121P	NM_003010	X	X
NTRK1	c.C152T	p.T51I	NM_001012331	X	X
TP53	c.C916T	p.R306X	NM_001126114	X	X
BMPR1A	c.C934T	p.H312Y	NM_004329	X	X
BRAF	c.G716A	p.R239Q	NM_004333	X	X
COL3A1	c.C1247A	p.P416Q	NM_000090	X	X
HIF1A	c.G1388A	p.R463Q	NM_001530	X	X

HOXD4	c.G259A	p.G87S	NM_014621	x		
KRAS	c.G40A	p.V14I	NM_033360	x		
MICAL3	c.G4529A	p.R1510Q	NM_015241	x		
MICAL3	c.G4102A	p.V1368I	NM_015241	x		
MLL3	c.12039delA	p.V4013fs	NM_170606	x		
MMP2	c.C1195T	p.L399F	NM_004530	x		
NOTCH1	c.G2668A	p.G890S	NM_017617	x		
PTPN11	c.C1190T	p.T397M	NM_002834	x		
SMO	c.C1639T	p.R547C	NM_005631	x		
STAB1	c.C2063A	p.P688H	NM_015136	x		
TIE1	c.G3326A	p.R1109H	NM_005424	x		
TP53BP1	c.G4193A	p.R1398H	NM_005657	x		
CACNA1B	c.G28A	p.G10S	NM_000718	x		x
CACNA1B	c.C2722T	p.R908X	NM_000718	x		x
ERBB2	c.G3481A	p.V1161M	NM_001005862	x		x
<b>25 gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>		<b>M</b>
APC	c.G3925T	p.E1309X	NM_001127510	x		x
APC	c.3067dupA	p.D1022fs	NM_001127510	x		x
NRAS	c.A182G	p.Q61R	NM_002524	x		x
TP53	c.G976T	p.E326X	NM_001126114	x		x
BMP2	c.C110T	p.S37L	NM_001200	x		
FAT4	c.G3346T	p.G1116W	NM_024582	x		
MLL3	c.G11984A	p.R3995Q	NM_170606	x		
ATM	c.G3619A	p.E1207K	NM_000051			x
MICAL3	c.C4376A	p.P1459Q	NM_015241			x
<b>26 gene</b>	<b>cDNA</b>	<b>protein</b>	<b>transcript</b>	<b>P</b>		<b>M</b>
ABL1	c.G2173A	p.G725S	NM_007313	x		x
ABL1	c.C2243T	p.T748M	NM_007313	x		x
ADAM19	c.C2080T	p.P694S	NM_033274	x		x
APC	c.4055dupT	p.V1352fs	NM_001127510	x		x
BRAF	c.C2044T	p.R682W	NM_004333	x		x

GNAS	c.C1798G	p.R600G	NM_080425	x	x
IDH1	c.G103A	p.V35M	NM_005896	x	x
KIT	c.G157A	p.E53K	NM_000222	x	x
SMAD7	c.C281T	p.A94V	NM_005904	x	x
TP53	c.G743A	p.R248Q	NM_001126114	x	x
ABCA1	c.C5914A	p.L1972I	NM_005502	x	x
APC	c.G933T	p.K311N	NM_001127510	x	x
APC	c.C1991A	p.T664N	NM_001127510	x	x
ATM	c.C2494T	p.R832C	NM_000051	x	x
ATM	c.C9076T	p.L3026F	NM_000051	x	x
BMIPR1A	c.C571A	p.R191S	NM_004329	x	x
CACNA1B	c.G331A	p.V111M	NM_000718	x	x
CACNA1B	c.C5536T	p.P1846S	NM_000718	x	x
CACNA2D3	c.G2429A	p.R810Q	NM_018398	x	x
CCNB2	c.G367A	p.E123K	NM_004701	x	x
CDC42BPA	c.G4978A	p.G1660R	NM_003607	x	x
COL3A1	c.G1270A	p.A424T	NM_000090	x	x
FAT4	c.G13138A	p.G4380R	NM_024582	x	x
HIF1A	c.C421A	p.H141N	NM_001530	x	x
HIF1A	c.G734A	p.R245Q	NM_001530	x	x
HOXD4	c.G614A	p.R205Q	NM_014621	x	x
JAK1	c.G361A	p.D121N	NM_002227	x	x
JAK2	c.G2840A	p.R947Q	NM_004972	x	x
JAK2	c.C3337A	p.R1113S	NM_004972	x	x
KDR	c.G1963A	p.V655M	NM_002253	x	x
KIT	c.G2059T	p.D687Y	NM_000222	x	x
MLL3	c.C6041A	p.A2014E	NM_170606	x	x
MLL3	c.G4644T	p.L1548F	NM_170606	x	x
MMP2	c.G1201A	p.A401T	NM_004530	x	x
MMP9	c.G742A	p.G248S	NM_004994	x	x
MYC	c.C1159T	p.R387W	NM_002467	x	x
NEGR1	c.C461A	p.T154N	NM_173808	x	x
NOTCH1	c.G2353A	p.G785S	NM_017617	x	x

NTRK3	c.C2143A	p.P715T	NM_001012338	x
PANK4	c.G679A	p.A227T	NM_018216	x
PDGFRA	c.C272T	p.S91L	NM_006206	x
PRUNE2	c.C7678T	p.H2560Y	NM_015225	x
PRUNE2	c.G5399T	p.W1800L	NM_015225	x
PTPN1	c.C595T	p.R199X	NM_002827	x
RB1	c.C2321A	p.T774N	NM_000321	x
SRC	c.C1259T	p.T420M	NM_005417	x
SRGAP1	c.G2167A	p.G723S	NM_020762	x
STAB1	c.G4537A	p.E1513K	NM_015136	x
STAB1	c.G4615A	p.E1539K	NM_015136	x
STAB1	c.G4961A	p.R1654Q	NM_015136	x
TIE1	c.C65T	p.A22V	NM_005424	x
TIE1	c.C2458T	p.R820W	NM_005424	x
TIE1	c.G3030T	p.K1010N	NM_005424	x
CACNA1B	c.G2741A	p.G914D	NM_000718	x
CTNNB1	c.G1544A	p.R515Q	NM_001904	x
FAT4	c.T233dupC	p.N2411fs	NM_024582	x
PIK3CA	c.G2975A	p.R992Q	NM_006218	x
PRUNE2	c.C8794T	p.R2932W	NM_015225	x
SIAH1	c.C643T	p.R215C	NM_003031	x
SMO	c.C20T	p.A7V	NM_005631	x
TGFBR1	c.G533A	p.R178H	NM_001130916	x

List of mutations in primary CRC tumours and metastases to the ovaries; and mutation profiles. In the first column the variants present in the primary colorectal tumour are shown. In the second (and third) column the variants present in the metastases to the ovaries are shown.

Supplemental Table 4

Study	Number of cases included	Number of genes analysed	Tissue(s) examined	APC	TP53	PIK3CA	KRAS	BRAF
Baldus (2010)	20	3	liver lymph nodes	NA	NA	Less frequent in metastases	Less frequent in metastases	No difference
Goranova (2011)	6	3	liver	Less frequent in metastases	No difference	NA	No difference	NA
Vermaat et al. (2012)	21	1264	liver	NA	NA	More frequent in metastases	More frequent in metastases	More frequent in metastases
Vakiani (2012)	84	5	liver (78%) lung soft tissue brain ovary lymph nodes	NA	More frequent in metastases*	Less frequent in metastases	More frequent in metastases	Less frequent in metastases*
Heitzer (2013)	2	3	CTCs liver brain	NA	More frequent in metastases	NA	NA	NA

\*Significant differences

*Overlap and differences in mutation profiles of primary CRCs and matching metastases. A number of studies comparing the mutational profile of primary CRCs and matching metastases are shown. Due to the low number of studies a clear pattern cannot be observed.*



**Supplemental Table 5**

Gene	Codon	Amino Acid	Pathogenicity class*
KRAS	c.378C>A	p.D126E	3
KRAS	c.38G>A	p.G13D	5
KRAS	c.40G>A	p.V14I	5
BRAF	c.1100C>T	p.P367L	3
BRAF	c.1271G>A	p.R424Q	3
BRAF	c.998G>T	p.S333I	2
BRAF	c.1672C>T	p.R558X	1
FBXW7	c.174C>A	p.S58R	2
FBXW7	c.1072T>A	p.S358T	4
FBXW7	c.859C>T	p.R287X	5
PIK3CA	c.2593G>A	p.G865S	4
PTEN	c.151G>T	p.D51Y	4

List of discordant variants detected in KRAS, BRAF, FBXW7 and PTEN and their pathogenicity class. \*<http://www.interactive-biosoftware.com/alamut-visual/>



## Chapter 5

### **Next generation sequencing using the HaloPlex targeting method in formalin-fixed paraffin-embedded (FFPE) material.**

Stijn Crobach<sup>1</sup>, Dina Ruano<sup>1</sup>, Ronald van Eijk<sup>1</sup>, Melanie Schruppf<sup>1</sup>, Hans Morreau<sup>1</sup> and Tom van Wezel<sup>1</sup>

*<sup>1</sup>Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands.*

## **Abstract**

Next generation sequencing (NGS) is the current standard method for somatic variant detection in molecular tumor pathology. Despite of the fragmented nature of formalin-fixed paraffin embedded (FFPE) material, NGS is suitable in a diagnostic setting with FFPE-DNA as input. A large number of targeted sequencing approaches are available, where the capture can be done using polymerase chain reactions (PCR), hybridization or circularization reactions. In this study we show that a circularization based approach (HaloPlex), followed by sequencing on Illumina HiSeq is successful for targeted sequencing of DNA from FFPE material. Detected variants were validated with a PCR-based targeted enrichment method (Ion AmpliSeq) followed by sequencing on an Ion PGM sequencer. A high concordance between the detected variants in HaloPlex and AmpliSeq capture was observed. Discordant variants could largely be explained by (subtle) setting differences in the analysing pipeline. Thus, an optimal bioinformatics pipeline analysis that has to be adjusted to the chosen platform is crucial for correct detection of variants. Input from distinct DNA isolations can explain discordant sequencing variants, emphasizing the presence of tumor intra-heterogeneity (ITH).

## Introduction

Somatic gene variant profiling of large gene sets by next generation sequencing (NGS), instead of analysis of single genes by Sanger sequencing, is the current standard in daily clinical practice.[1, 2] The mutational status of genes is decisive for the choice of targeted therapies and can be useful in primary diagnostic analysis. As example, success of reaction towards EGFR inhibitors is seen in lung adenocarcinomas with activating *EGFR* variants. [3] The amount of genes known to be involved in responses to targeted therapies increases rapidly. For CRC, not only *KRAS*, but also pathogenic *NRAS*, *PIK3CA*, *BRAF* gene variants were shown to be involved. [4-6] Furthermore, for other malignancies like melanomas and gastro-intestinal stromal tumors (GIST) targeted therapies directed at respectively *BRAF/RAS* and *c-KIT* are now available.[7, 8] Next, sensitivity for radio- and chemotherapy might be correlated with mutational profiles.[9, 10] So, mutational profiles of large gene sets are becoming more important for clinical decision making and probably also for statements about prognoses.

Next generation sequencing (NGS) methodologies enable high throughput sequencing, resulting in large amounts of data. As only a limited number of genes is involved in treatment responses, targeted sequencing is the preferred method in diagnostic settings.[11-13] For research questions whole exome or whole genome analyses remain valuable. Several target enrichment strategies based on distinct methodologies are available (Table 1). The approaches can be based on hybridization, circularization or PCR.[11] Hybridization is regarded as the preferred method for large regions, while PCR is suitable for targeting a limited amount of genes. The coverage obtained with hybridization is in general more homogeneous than with other techniques. Drawbacks are the need for relative large amounts of DNA input and for additional equipment like array plates in contrast to in solution captures. Advantages of using PCR as enrichment technique is that also DNA of average quality is suitable. Furthermore, PCR based techniques show relatively few off-target reads. A challenge using PCR is that coverage can differ between separate amplicons. Circularization techniques are very specific in targeting, although low coverage uniformity can be noted with such approach (Table 1). For use in a diagnostic setting the performance of different targeting techniques has to be examined with DNA isolated from formalin fixed paraffin embedded tissue (FFPE) as input, as this is used in daily diagnostics in pathology.

In this research we tested a circularization reaction (HaloPlex), that was validated with a PCR based approach (Ion AmpliSeq). The former technique is based on the digestion of DNA with different sets of restriction enzymes (<http://www.agilent.com>;

accessed January, 2016), after which regions of interest are captured in circular DNA fragment (circularization).[12] To validate the results generated by the HaloPlex target enrichment, the Ion AmpliSeq Cancer Panel was used. The Ion AmpliSeq Cancer Panel is a PCR based technique with an input of only 10 ng of FFPE-DNA (<http://www.lifetechnologies.com>; accessed January, 2016).

Multiple sequencing devices that are updated in a high rate, are available for sequencing targeted DNA.[9, 13] Most devices are based on optical read-outs as the result of incorporation of fluorescent nucleotides (Table 2). [11, 14] A non-optical method is semiconductor sequencing that measures hydrogen ions that are released during polymerization of DNA.

To implement NGS for clinical purposes, validation experiments of a chosen analysis method are necessary. [15-17] In this paper tested the HaloPlex targeted enrichment method, and validated it with the Ion AmpliSeq targeted enrichment protocol. We show that both methodologies deliver correct mutation data in FFPE material and thus can be used for clinical purposes.

## Materials and methods

### Ethics Statement

All samples were handled according to the medical ethical guidelines described in the Code Proper Secondary Use of Human Tissue established by the Dutch Federation of Medical Sciences ([www.federa.org](http://www.federa.org)). Accordingly to these guidelines all human material used in this study has been anonymized. Because of this anonymization procedure individual patients' permission is currently not needed.

### Case selection and DNA isolation

FFPE tumor blocks of colorectal cancers and matching (ovarian) metastases were selected (8 pairs; 16 tumors). The tissue used for DNA isolation was enriched for tumor cells. Based on a hematoxylin and eosin (H&E)-stained slide, 0.6-mm tissue punches were taken using a tissue microarrayer (Beecher Instruments, Sun Prairie, WI, USA). DNA from both the tumors was isolated. Prior to DNA isolation, the FFPE-tissue was deparaffinized in xylene and washed in 70% ethanol. In half of the cases (n=8) DNA was isolated using the NucleoSpin Tissue Genomic DNA Purification kit (Machery-Nagel, Düren, Germany) according to the manufacturer's instructions. The other cases (n=8) were isolated using a fully automated nucleic acid purification method produced by Siemens.[18] In 6 cases the same DNA isolation was used for both the HaloPlex as the Ampliseq target enrichment. In 10 cases separate DNA isolates were used. Samples were selected after the quality of the DNA was tested by PCR. Base pair products of 150-, 255-, 343-, and 511-bases were sequenced. In case small base pair products (150 and 225 bp) were not generated, the sample was excluded from further analysis.

### Construction of the target enrichment panels (HaloPlex and Ion AmpliSeq)

The HaloPlex (Agilent Technologies) gene panel was custom designed for genes relevant in CRC. The panel, targeting 115 genes, was constructed using gene lists described in literature.[19, 20] See Supplemental Table 1 for the complete gene list. The Ion AmpliSeq Cancer Hotspot Panel v2 consist of 50 oncogenes. See Supplemental Table 2 for the complete list of genes.

### Sample library preparation using HaloPlex and Ion AmpliSeq kits

The HaloPlex target enrichment system is a circularization based enrichment method (<http://www.agilent.com>; accessed January, 2016). The first step is the fragmentation of the input 225 ng DNA by a set of 8 different restriction enzymes. Next, targeted nucleic acid sequences are hybridized with oligonucleotide constructs called selectors. The selectors contain target-complementary end-sequences that are joined by

a general linking sequence and that act as ligation templates to direct the circularization of target DNA fragments. Also, in this step a sample specific barcode is added. The selectors are biotinylated and therefore the targeted fragments can be retrieved with magnetic streptavidin beads. Hereafter, the circular molecules are closed by ligation, which is only efficient for perfectly hybridized fragments, which makes the method theoretically very sensitive and specific. Only these circularized targets are then amplified in a multiplex PCR, using one universal PCR primer pair that is specific for the general linking sequence in the selectors. The total size of the targeted region is 486013 bp divided over 115 cancer related genes (Supplemental Table 1).

The Ion AmpliSeq target enrichment kit is using a multiplex PCR reaction by which the regions of interest are first amplified. Next, the primer sequences of the PCR products that are specifically designed for this purpose are partially degraded by a FuPa reagent. Hereafter, adapters and sample specific barcodes are added to the PCR product by a ligation reaction. Finally an emulsion PCR is performed after which the samples are sequenced. The total size of the Ion AmpliSeq panel is 22027 bp divided over 50 cancer related genes (Supplemental Table 2).

### **High Throughput Sequencing using Illumina HiSeq and Ion PGM**

The libraries generated using the HaloPlex target enrichment kit, were sequenced on a Illumina HiSeq 2000 sequencer (ServiceXS, Leiden). Sequencing for libraries prepared with the Ion AmpliSeq target enrichment kit was performed on a Ion Torrent PGM sequencer (Thermo Fischer) using the Ion PGM 200 sequencing kit according to the manufacturer's instructions.

### **Data Analysis**

HaloPlex data was analyzed as previously described.[21] In short, the adaptors, barcodes and enzyme footprints were removed from the sequenced reads using Sure-Call software (Agilent Technologies, Santa Clara, CA), after which the reads were aligned to the human genome (hg19) using the Burrows-Wheeler aligner (BWA, version 0.7.5a).[22] The Genome Analysis Toolkit (GATK, version 2.5) was used for realignment around the indels and base quality recalibration.[23] SNP and indel calling were carried out using VarScan software (version v2.3.6) with the following arguments: minimum read depth = 8, minimum number of reads with the alternative allele = 2, minimum base quality = 15, and minimum variant allele frequency = 0.10.

Variants were functionally annotated using ANNOVAR.[24, 25] We then selected variants more likely to have a deleterious effect. This was achieved by focusing on splicing and exonic variants (excluding synonymous) and removing the variants that were present with a frequency higher than 1% in the 1000 Genomes project (<http://www.1000genomes.org/>; data from April 2012) and/or in the NHLBI Exome



Sequencing Project (<http://evs.gs.washington.edu/EVS/>; data from January 2013) because they are more likely to be germline in origin.

For AmpliSeq data, reads were mapped to the human reference genome (hg19) using the TMAP software with default parameters (<https://github.com/iontorrent/TS>). Subsequently variant calling was done using the Ion Torrent specific caller, Torrent Variant Caller (TVC), using the recommended Variant Caller Parameter for Cancer Hotspot Panel v2.

HaloPlex and AmpliSeq target regions were intersected resulting in 197 regions covering a total of 20294 bp captured by both panels. Variants in present in only AmpliSeq or HaloPlex were visually inspected using IGV, to identify false positive or negative calls.

### **Allele specific qPCR**

As extra validation step of the NGS results, allele-specific qPCR was performed as described previously.[26] Seven variants of *KRAS* (p.G12C, p.G12R, p.G12S, p.G12V, p.G12A, p.G12D, p.G13D), 2 variants of *EGFR* (p.L858R and exon 19 deletion), one variant of *BRAF* (p.V600E) and 3 variants of *PIK3CA* (p.E542K, p.E545K, p.H1047R) were analyzed.

---

## Results

### HaloPlex target enrichment.

HaloPlex target enrichment has been optimized for high quality DNA. To test whether the protocol also works with low quality FFPE-DNA as input, we processed 16 tumor FFPE-DNA samples. The samples consisted of 16 tumor samples. The regions of interest were enriched using a customized designed panel consisting of 115 genes. The panel consisted of colon cancer driver genes and currently relevant genes for clinical treatment decisions (*BRAF*, *EGFR*, *KRAS* and *PIK3CA*) were included. The total number of reads generated was  $8.4 \times 10^7$  of which  $6.2 \times 10^7$  were aligned (73.8%). Of the aligned reads  $5.9 \times 10^7$  were on target (95.2%). In Supplemental Figure 1A an overview is given for the number of reads per sample (range  $1.4 \times 10^6$  –  $25.0 \times 10^6$ ). One sample (no. 7) generated substantially more reads in comparison to the other samples,  $25.0 \times 10^6$  reads compared to an average number of  $5.3 \times 10^6$  reads. The average coverage was 452x (range 15-1972; Supplemental Figure 1B). In total 1962 regions were captured, of which 19 (~1%) showed a coverage of 10 reads or less.

### Validation with Ion AmpliSeq target enrichment.

To validate the HaloPlex results, the same 16 patients were analyzed with the Ion AmpliSeq target enrichment followed by Ion PGM sequencing. In 6 cases the same DNA was used, in the remaining 10 cases new DNA isolations were obtained. The total number of reads generated was  $9.2 \times 10^6$ ; of which  $8.6 \times 10^6$  were aligned (93.5%). The majority of the aligned reads (89%) was on target. All the 16 samples produced a comparable number of reads (range  $3.9 \times 10^5$  –  $7.5 \times 10^5$ ; average  $5.8 \times 10^5$ ). The average coverage per sample was 1873 (range 1357-2658; Supplemental Figure 2A). All amplicons showed sufficient coverage (1951x, range 61 – 11170; Supplemental Figure 2B).

### Haloplex and AmpliSeq comparison

To produce a reliable comparison between the two targeting techniques, the targeted regions that showed overlap between the two techniques were analyzed in more detail. Almost all regions covered with the Ion AmpliSeq panel were also targeted with customized HaloPlex panel. In total 197 separate DNA fragments (20294 bp) were overlapping in both target enrichment techniques. Figure 1 shows the coverage per targeted region for each sample, showing a higher number of reads in the AmpliSeq data, although in a comparable pattern. Figure 2 shows the average coverage per targeted region, also showing a similar image in the two targeting techniques. In the HaloPlex approach these 197 targeted regions showed an average coverage of 578x

(range 3-3774). In total 25 variants were detected in 16 cases (~2 variants per case). See Supplemental Table 4 for the complete overview of detected variants with the HaloPlex targeting technique. With the Ampliseq method an average coverage of 1951x (range 61-11170) was achieved in the 197 overlapping DNA regions. In total 44 variants were detected in 16 cases (~3 variants per case) in the 197 overlapping regions. In Supplemental Table 4 a complete overview of detected variants is shown.

The concordance rate of variants in the overlapping regions that were covered by both panels was 56%. In total 44 variants were detected of which 25 were found with both methodologies.

The 19 discordant variants were all detected with the Ion AmpliSeq panel, and initially not found with the HaloPlex panel. Nine of the 19 discordant variants could be detected in the HaloPlex data after manually curating the sequence reads. These 9 variants were filtered out using the standard analysis, due to low numbers of (mutant) reads. In 2 cases no reads were present at the specific positions in the HaloPlex experiments. Eight variants (divided over 3 cases) that were detected by Ion AmpliSeq target enrichment were, although with sufficient reads on target, not detected in the HaloPlex target experiment. The discrepancy could be explained by the use of separate DNA-isolations, as these discordant results were only detected as different DNA isolations of the same tumor was used. So, possibly the detected differences are best explained by intra-tumor heterogeneity (ITH).

### **Mutation profile of hotspots in KRAS, BRAF, EGFR and PIK3CA**

The coverage of *KRAS*, *BRAF*, *EGFR* and *PIK3CA* was comparable with the other genes in both target enrichment approaches used (Supplemental Figure 3A and 3B). However, when looking to coverage of specific base positions, it was noted that one of the *PIK3CA* hotspots (p.E542) had a very low coverage in the HaloPlex panel. In three cases this specific variant was not detected with the HaloPlex target enrichment because of low numbers of (mutant) reads that not reached the thresholds. In one case no reads were present at the *PIK3CA* p.E542 hotspot position.

In total 13 mutations were called in *KRAS*, *BRAF*, *EGFR*, and *PIK3CA*. Six of these variant were concordant. Of the seven discordant variants, 5 were detected by manually looking into the sequence data. Two variants were not detected, although there was sufficient coverage. As stated above, these discrepancies can be explained by ITH. Mutations calls were validated using a hydrolysis probe assay, showing no discordant results.

Interestingly, while manually checking the reads of the Ampliseq experiment in some cases a very small number of reads (varying from 1 to 7) carrying a pathogenic variant at known hotspot locations were observed (Supplemental Table 3). These variants

had a frequency lower than the thresholds and were not detected by using the bioinformatics pipeline. Samples analyzed with the Ion AmpliSeq panel showed this phenomenon more often, possibly because of the total number of reads was higher in comparison with the HaloPlex experiment. The hotspots in *EGFR* and *BRAF* do not show this phenomenon. Whether these reads present true low frequent mutations or are a technical artifact cannot be defined on these numbers. *APC* variants, showing no discordances, were detected in half of the cases (Supplemental Table 4).

## Discussion and Conclusion

Next generation sequencing (NGS) is the standard method for mutation screening in diagnostic pathology.[1] DNA isolated from formalin-fixed paraffin-embedded material (FFPE), has proven to be suitable for high-throughput approaches.[15] Targeted sequencing generates high coverage of the genes of interest.[27] Several target enrichment strategies based on hybridization, circularization or PCR are available.[11] For large regions hybridization is advised, while for small regions PCR-based techniques are preferred. Targeting methodologies using a circulation approach in combination with FFPE DNA as input material are not frequently used.[12]

An advantage of HaloPlex is the large number of probes targeting the genes of interest. Due to overlapping probes that target DNA regions, failing of one or more probes per region still can result in successful enrichment. In our experiments with FFPE material, the coverage varied considerably between samples. The low DNA quality of FFPE material might have caused this difference in sequencing efficiency between the samples. In general the genes of interest were successfully targeted (about 95% of the aligned reads were on the target regions), however some targets were not captured or showed a very low number of reads impairing their analysis. In total 1962 regions were captured, of which 19 (~1%) showed insufficient coverage (<10 reads). The smaller Ion AmpliSeq panel showed a more evenly distribution among the samples, although the coverage of the 197 overlapping regions showed a comparable pattern in the two targeting techniques (Figure 1 and 2). The overlapping regions showed a coverage of respectively 578x and 1951x in the HaloPlex and AmpliSeq targeted approach.

Bioinformatics pipelines are crucial for data interpretation. Several variants were initially not detected in our HaloPlex data analysis. These pathogenic variants were only detected in a retrospective analysis of the HaloPlex data. As larger regions of the DNA are targeted that show more variation in the number of reads per region, it is more difficult to construct an optimal pipeline. On one hand a pipeline should not have loose settings that result in false positive results, on the other hand too strict settings might not reveal variants with only a limited number of mutant reads. The HaloPlex panel resulted in regions with low coverage leading to undetected, but true, variants.

Several companies already offer analyzing software, that only need sequencing data as input. In this way all data is analyzed in a standardized method. However, the bioinformatics pipeline settings cannot be changed. Continuous validation and ad-

adjustments of the bioinformatics pipelines is necessary to get the most reliable results. As example, adding genes to the targeted panel might influence the processing of the DNA, needing thorough validation. Also, the implementation of newer sequencing devices that can deliver result of more patient samples in a shorter time span, needs to be validated. So, as the developments in the field of molecular diagnostics follow each other at high rate, continuous control experiments have to be performed.

The smaller Ion AmpliSeq panel (22.027 bp versus 486.013 bp targeted in the HaloPlex experiment) seems to be the preferred method in a clinical setting, as robust data could be delivered. Because only a limited amount of base pairs is targeted, a solid multiplex PCR reaction is present.

It is expected that the speed of sequencing will increase spectacular. Targeted approaches could potentially become unnecessary. In that case whole exome, or whole genome could potentially be performed on any patient sample. Only the genes of interest need then inspection. The other sequence data can be stalled for a renewed analyses.

## References

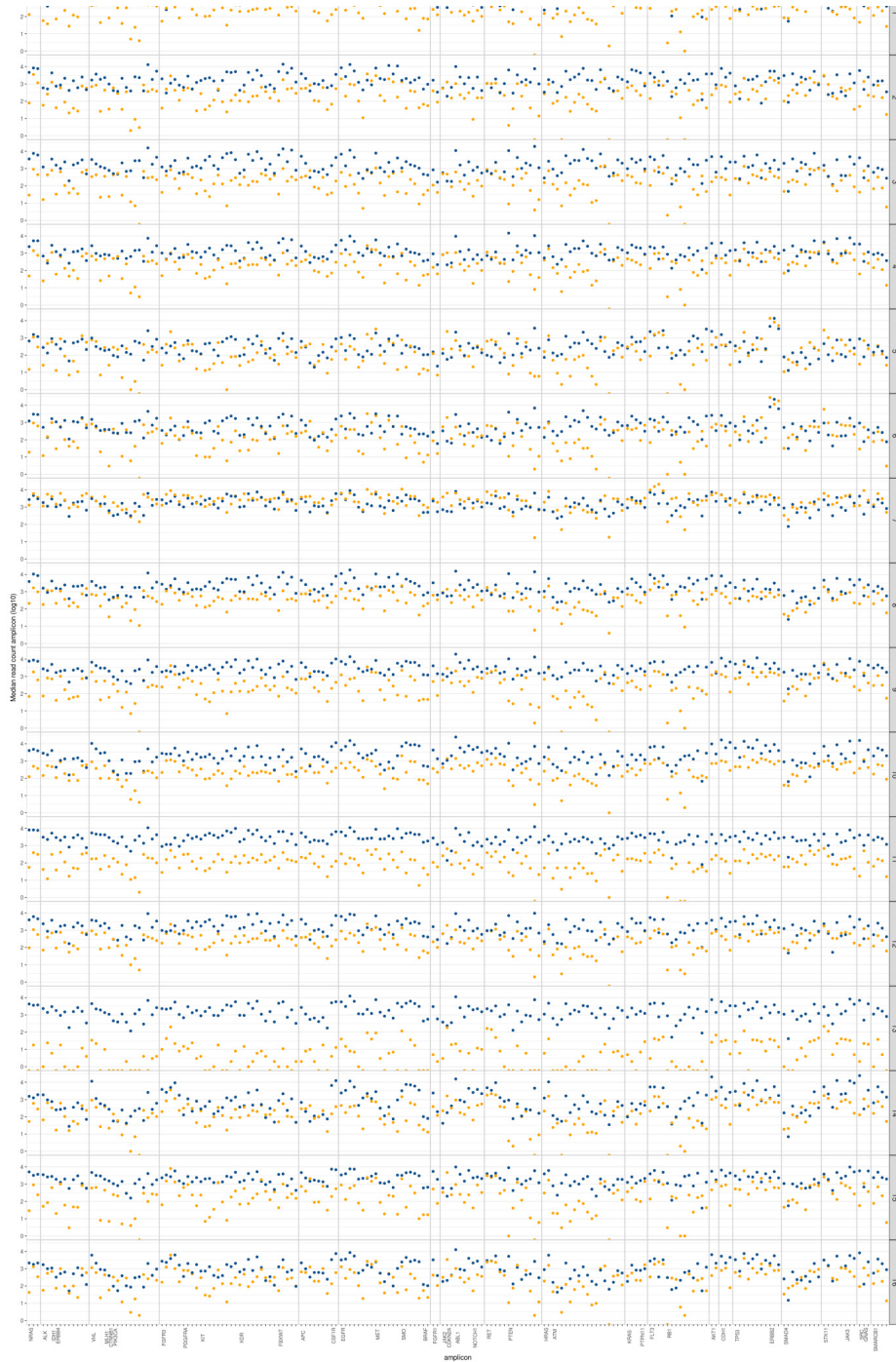
1. Natrajan, R. and J.S. Reis-Filho, *Next-generation sequencing applied to molecular diagnostics*. *Expert. Rev. Mol. Diagn*, 2011. **11**(4): p. 425-444.
2. Korf, B.R. and H.L. Rehm, *New approaches to molecular diagnosis*. *JAMA*, 2013. **309**(14): p. 1511-1521.
3. Wong, R. and D. Cunningham, *Using predictive biomarkers to select patients with advanced colorectal cancer for treatment with epidermal growth factor receptor antibodies*. *J. Clin. Oncol*, 2008. **26**(35): p. 5668-5670.
4. Di, N.F., et al., *Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer*. *J. Clin. Oncol*, 2008. **26**(35): p. 5705-5712.
5. Sartore-Bianchi, A., et al., *PIK3CA mutations in colorectal cancer are associated with clinical resistance to EGFR-targeted monoclonal antibodies*. *Cancer Res*, 2009. **69**(5): p. 1851-1857.
6. De, R.W., et al., *Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis*. *Lancet Oncol*, 2010. **11**(8): p. 753-762.
7. Kim, E.J. and M.M. Zalupski, *Systemic therapy for advanced gastrointestinal stromal tumors: beyond imatinib*. *J. Surg. Oncol*, 2011. **104**(8): p. 901-906.
8. Kudchadkar, R.R., et al., *Targeted therapy in melanoma*. *Clin. Dermatol*, 2013. **31**(2): p. 200-208.
9. Stratton, M.R., *Exploring the genomes of cancer cells: progress and promise*. *Science*, 2011. **331**(6024): p. 1553-1558.
10. Gonzalez-Angulo, A.M., B.T. Hennessy, and G.B. Mills, *Future of personalized medicine in oncology: a systems biology approach*. *J. Clin. Oncol*, 2010. **28**(16): p. 2777-2783.
11. Morey, M., et al., *A glimpse into past, present, and future DNA sequencing*. *Mol Genet Metab*, 2013. **110**(1-2): p. 3-24.
12. Johansson, H., et al., *Targeted resequencing of candidate genes using selector probes*. *Nucleic Acids Res*, 2011. **39**(2): p. e8.
13. Metzker, M.L., *Sequencing technologies - the next generation*. *Nat. Rev. Genet*, 2010. **11**(1): p. 31-46.
14. Mardis, E.R., *Next-generation sequencing platforms*. *Annu Rev Anal Chem (Palo Alto Calif)*, 2013. **6**: p. 287-303.
15. Hadd, A.G., et al., *Targeted, high-depth, next-generation sequencing of cancer genes in formalin-fixed, paraffin-embedded and fine-needle aspiration tumor specimens*. *J. Mol. Diagn*, 2013. **15**(2): p. 234-247.
16. Gerlinger, M., et al., *Intratumor heterogeneity and branched evolution revealed by multi-region sequencing*. *N. Engl. J. Med*, 2012. **366**(10): p. 883-892.
17. Hosen, N., et al., *The Wilms' tumor gene WT1-GFP knock-in mouse reveals the dynamic*

- regulation of WT1 expression in normal and leukemic hematopoiesis*. *Leukemia*, 2007. **21**(8): p. 1783-1791.
18. van Eijk, R., et al., *Assessment of a fully automated high-throughput DNA extraction method from formalin-fixed, paraffin-embedded tissue for KRAS, and BRAF somatic mutation analysis*. *Exp Mol Pathol*, 2013. **94**(1): p. 121-5.
  19. Starr, T.K., et al., *A transposon-based genetic screen in mice identifies genes altered in colorectal cancer*. *Science*, 2009. **323**(5922): p. 1747-1750.
  20. Torkamani, A. and N.J. Schork, *Identification of rare cancer driver mutations by network reconstruction*. *Genome Res*, 2009. **19**(9): p. 1570-1578.
  21. Crobach, S., et al., *Target-enriched next-generation sequencing reveals differences between primary and secondary ovarian tumors in formalin-fixed, paraffin-embedded tissue*. *J. Mol. Diagn*, 2015. **17**(2): p. 193-200.
  22. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform*. *Bioinformatics*, 2009. **25**(14): p. 1754-1760.
  23. DePristo, M.A., et al., *A framework for variation discovery and genotyping using next-generation DNA sequencing data*. *Nat. Genet*, 2011. **43**(5): p. 491-498.
  24. Koboldt, D.C., et al., *VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing*. *Genome Res*, 2012. **22**(3): p. 568-576.
  25. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data*. *Nucleic Acids Res*, 2010. **38**(16): p. e164.
  26. van, E.R., et al., *Rapid KRAS, EGFR, BRAF and PIK3CA mutation analysis of fine needle aspirates from non-small-cell lung cancer using allele-specific qPCR*. *PLoS. One*, 2011. **6**(3): p. e17791.
  27. Han, S.W., et al., *Targeted sequencing of cancer-related genes in colorectal cancer using next-generation sequencing*. *PLoS. One*, 2013. **8**(5): p. e64271.

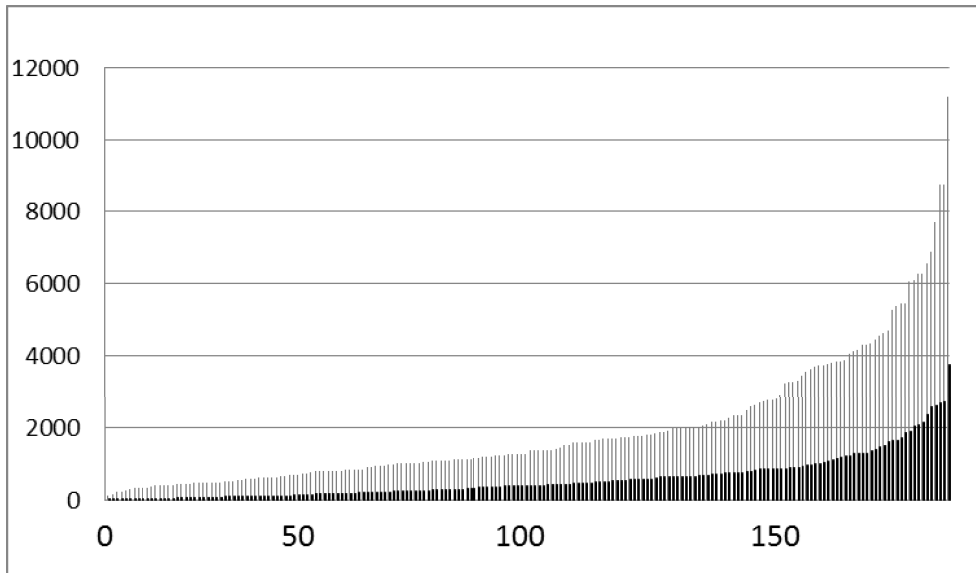
Coverage (y-axis) for shared targeted regions between the Ampliseq (blue dots) and Haloplex panel (yellow dots) for each sample. Targeted genes are shown on the x-axis. ►



Figure 1



**Figure 2**



*Overview of coverage (y-axis) of 197 overlapping DNA regions (x-axis) for both targeting techniques in grey (Ampliseq) and black (HaloPlex).*

**Table 1**

<b>Pro's and Con's</b>	<b>Examples</b>
<b>Hybrid Capture</b>	<b>Hybrid Capture</b>
+ Able to capture a large number of targets	Sureselect (Agilent)
+ Homogeneous coverage	SeqCap EZ (Roche NimbleGen)
- Large input of DNA is necessary	FlexGen (FlexSelect)
- Additional equipment is needed in case of using microarrays	Mybaits (Mycroarray)
<b>Circularization</b>	<b>Circularization</b>
- Low coverage uniformity	HaloPlex (Agilent)
	Single molecule Molecular Inversion Probes (smMIP)
<b>PCR</b>	
+ High quality DNA is not required	<b>PCR</b>
+ High enrichment ratio with few off-targets	Ion Ampliseq (Life Technologies)
- Coverage uniformity might be less	SequalPrep (Life Technologies)
	Seqtarget System (Qiagen)
	Acces Array System (Fluidigm)
	Thunderstorm/RDT 1000 System (Raindance)
	Truseq (Illumina)

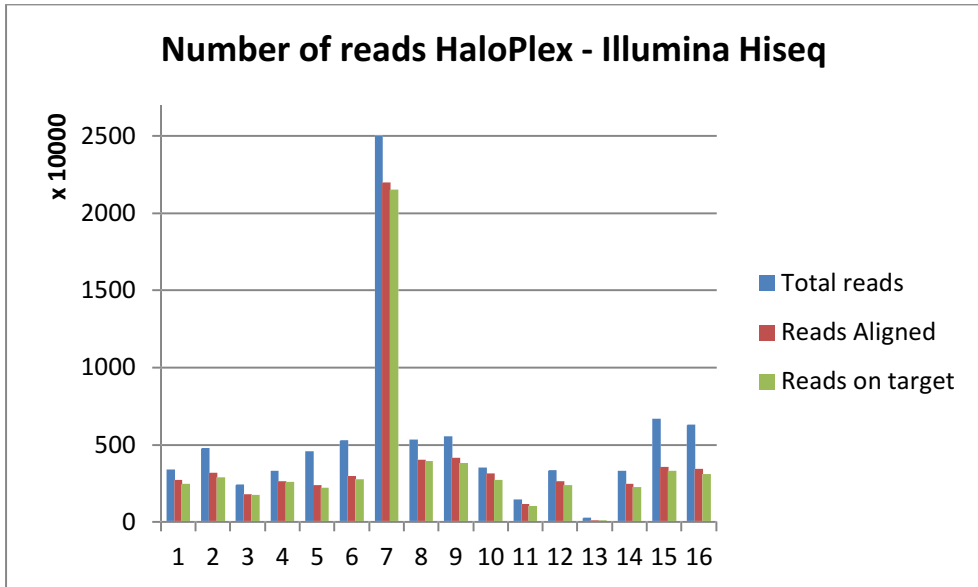
*Overview of Pro's and Con's of different target enrichment*

**Table 2**

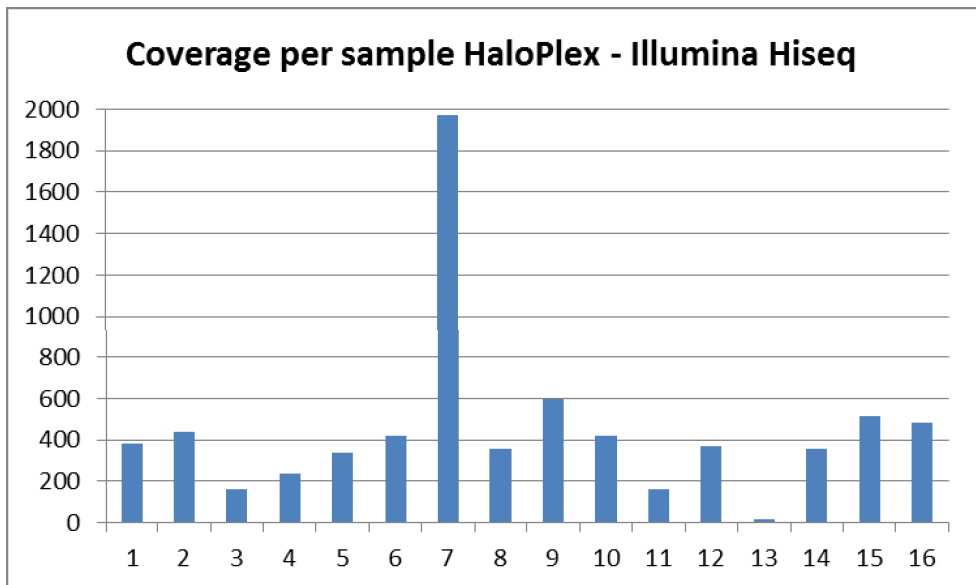
<b>Target Enrichment Strategy</b>	<b>Sequencing Platform</b>
<b>Hybrid Capture</b>	<b>Optical methods</b>
Sureselect (Agilent)	<u>Reversible Dye Terminator</u>
SeqCap EZ (Roche NimbleGen)	MiSeq (Illumina)
FlexGen (FlexSelect)	HiSeq (Illumina)
Mybaits (Microarray)	Genome Analyzer IIX (Illumina)
<b>Circularization</b>	<u>Single Molecule Real-time (SMRT)</u>
HaloPlex (Agilent)	PabBio RS (Pacific Biosciences)
Single molecule Molecular Inversion Probes (smMIP)	Oxford Nanopore sequencing (Oxford Nanopore)
<b>PCR</b>	<u>Pyrosequencing</u>
Ion AmpliSeq (Life Technologies)	454 (Roche)
SequalPrep (Life Technologies)	GS FLX Titanium (Roche)
Seqtarget System (Qiagen)	
Acces Array System (Fluidigm)	<u>Oligonucleotide Probe Ligation</u>
Thunderstorm/RDT 1000 System (Raindance)	Solid 4 (Life Technologies)
Truseq (Illumina)	Complete Genomics (BGI)
	<u>DNA Nanoball sequencing</u>
	Complete genomics
	<b>Non-optical methods</b>
	<u>Semiconductor sequencing</u>
	Ion PGM (Life Technologies)
	Ion Proton (Life Technologies)

*Overview of target enrichment strategies and sequencing platforms.*

Supplemental Figure 1A en 1B

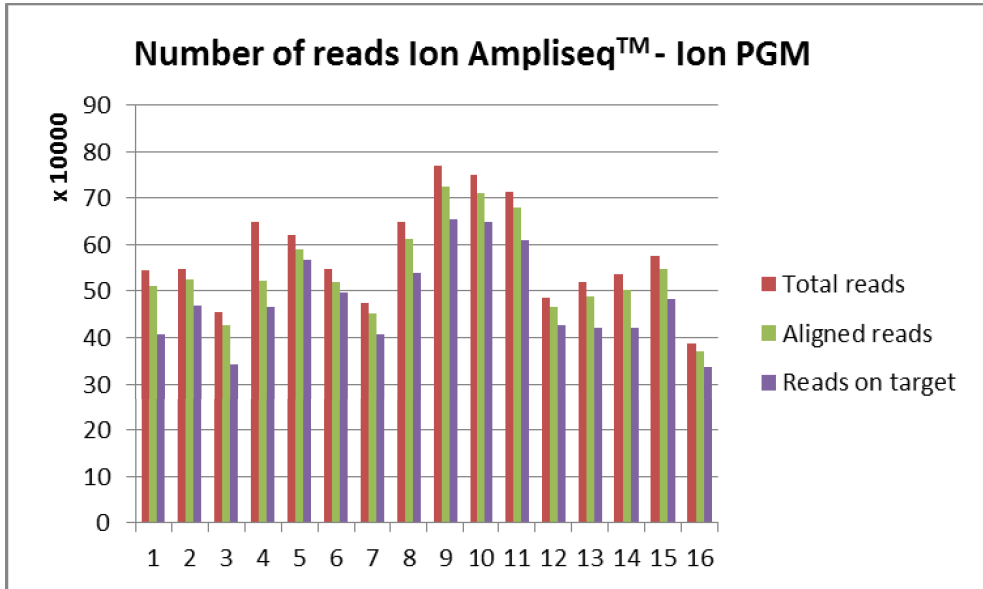


Total number of reads, reads aligned and reads on target are shown.

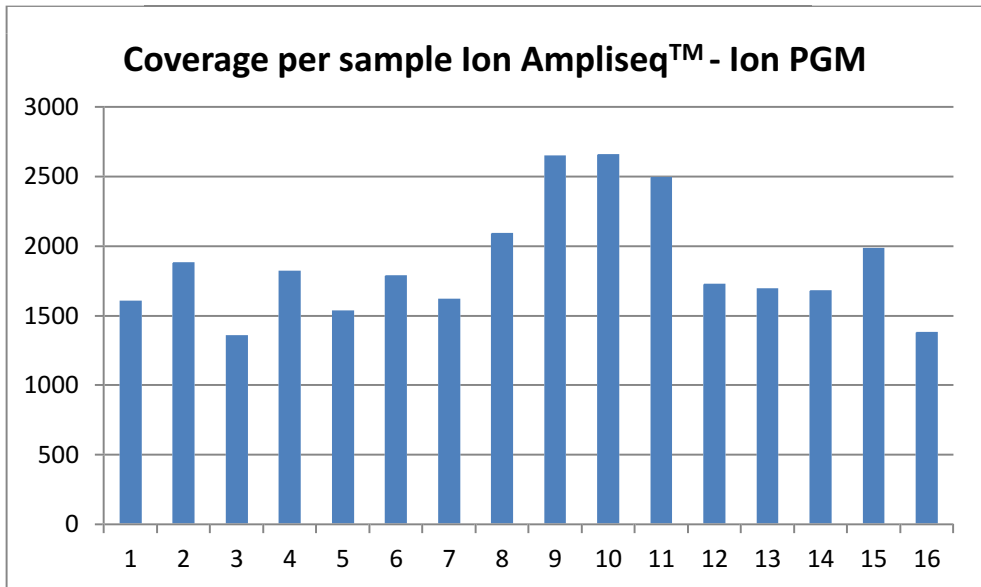


Average coverage per sample.

Supplemental Figure 2A en 2B

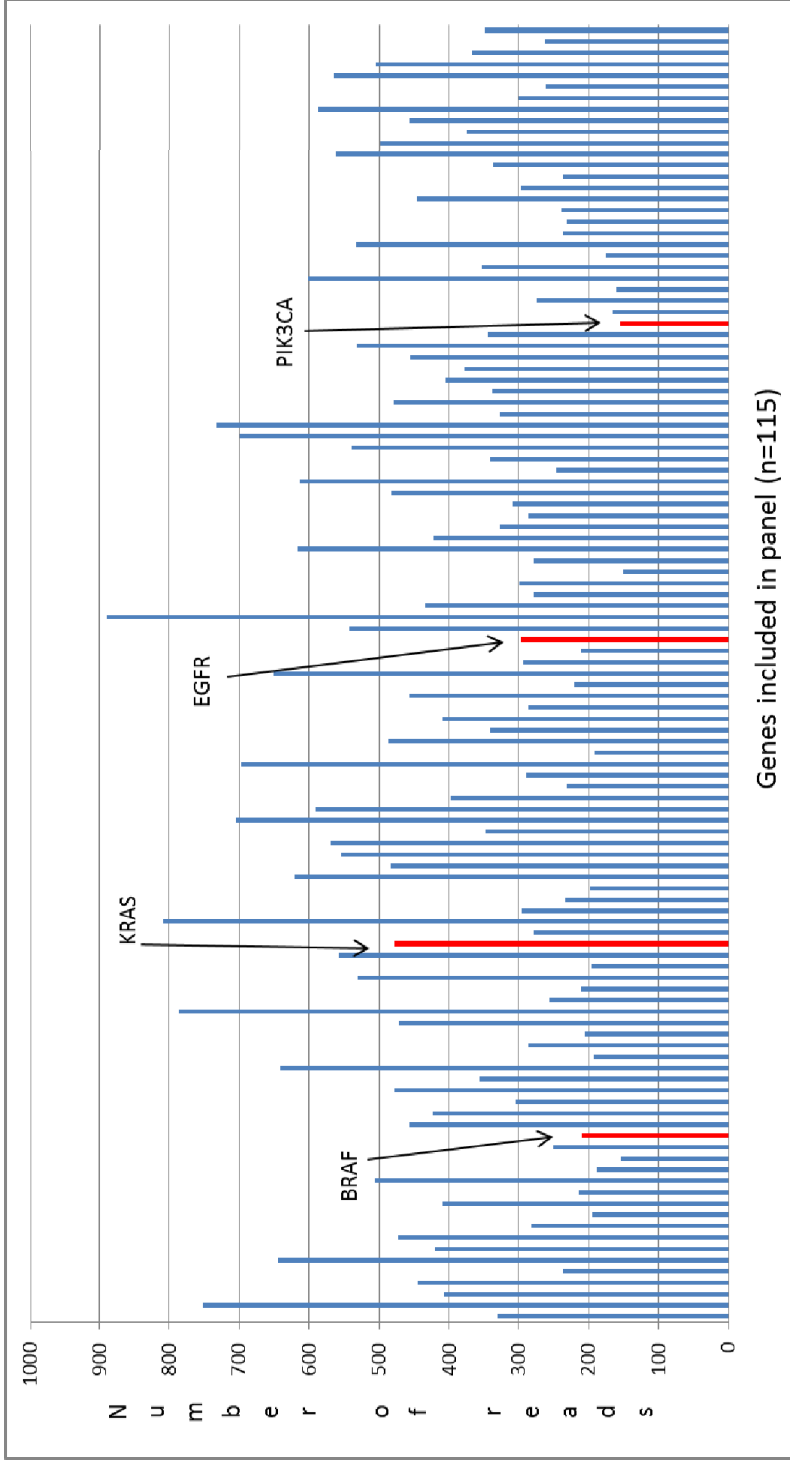


Total number of reads, reads aligned and reads on target are shown.



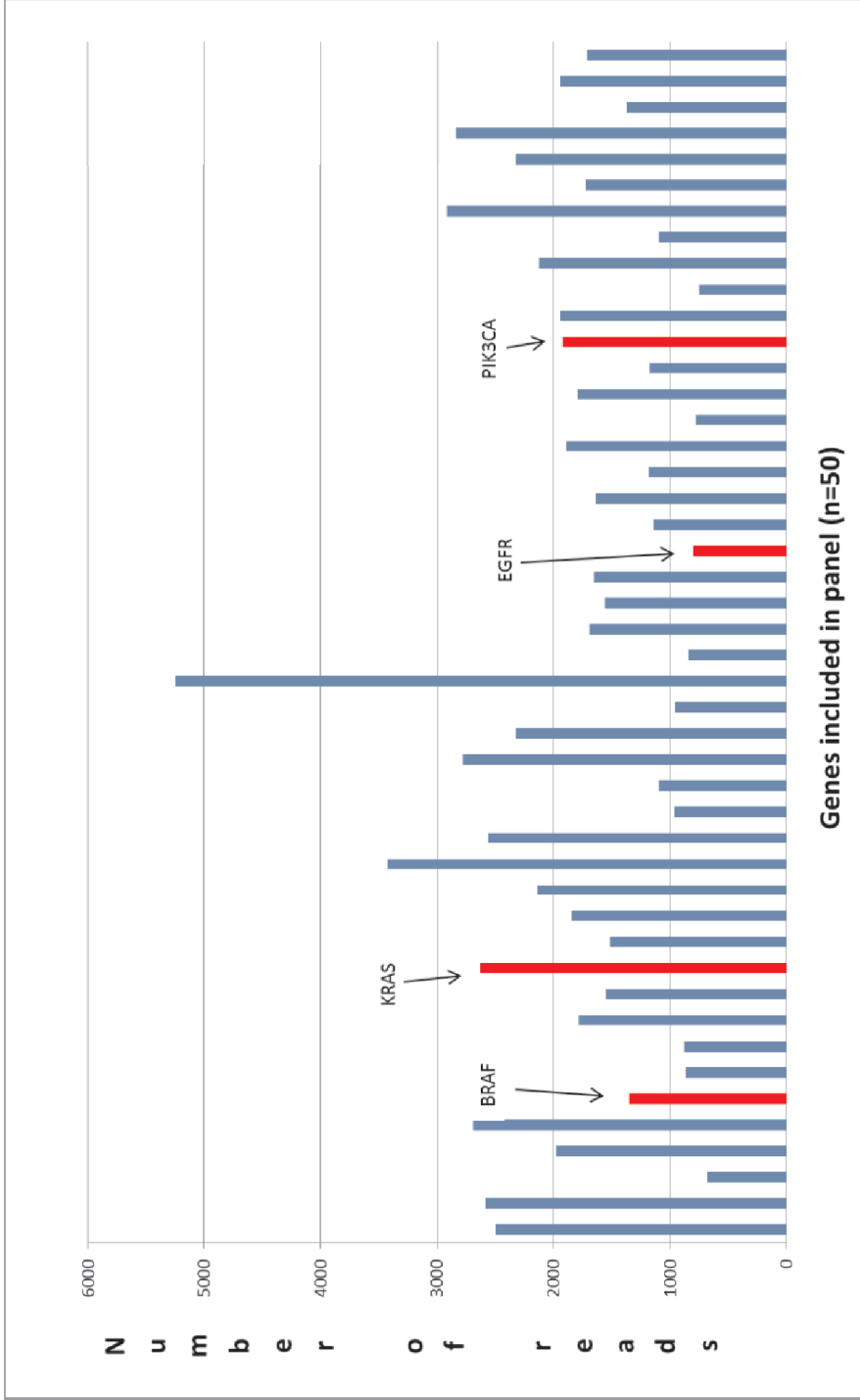
Average coverage per sample.

Supplemental Figure 3A Coverage of clinically relevant genes



Average coverage per targeted gene (n=115) in the custom made HaloPlex panels is shown. In red the clinically relevant genes.

Supplemental Figure 3B Coverage of clinically relevant genes



Average coverage per targeted gene (n=50) in the Ion AmpliSeq™ Cancer Hotspot Panel v2. In red clinically relevant genes.



**Supplemental Table 1**

1	ABCA1	31	CSF1R	61	KRAS	91	PRUNE2
2	ABL1	32	CTNNB1	62	LIMK2	92	PTEN
3	ADAM19	33	DSTYK	63	LOH12CR1	93	PTPN1
4	ADAMTSL1	34	EGFR	64	LTK	94	PTPN11
5	ADAMTSL3	35	EPHA3	65	MAP2K1	95	RB1
6	AKT1	36	ERBB2	66	MAP2K2	96	RET
7	AKT2	37	ERBB4	67	MAP2K4	97	RIOK3
8	ALK	38	FAT4	68	MDM2	98	SHAH1
9	ALPK1	39	FBXW7	69	MEN1	99	SMAD2
10	APC	40	FES	70	MET	100	SMAD3
11	ARID1A	41	FGFR1	71	MICAL3	101	SMAD4
12	ATM	42	FGFR3	72	MLH1	102	SMAD7
13	BAX	43	FGR	73	MLL3	103	SMARCB1
14	BMP2	44	FLT3	74	MMP2	104	SMO
15	BMPR1A	45	FOXO1	75	MMP9	105	SRC
16	BMPR2	46	GATA3	76	MSH2	106	SRGAP1
17	BRAF	47	GNAS	77	MSH6	107	STAB1
18	C11orf66	48	GUCY1A2	78	MUTYH	108	STK11
19	CACNA1B	49	HIF1A	79	MYC	109	SYNC
20	CACNA2D3	50	HOXA4	80	MYT1	110	TGFBR1
21	CASR	51	HOXB4	81	NEGR1	111	TGFBR2
22	CCNB2	52	HOXC4	82	NOTCH1	112	TIE1
23	CCND1	53	HOXD4	83	NRAS	113	TP53
24	CCNT2	54	HRAS	84	NTRK1	114	TP53BP1
25	CDC42BPA	55	IDH1	85	NTRK3	115	VHL
26	CDC73	56	JAK1	86	PANK4		
27	CDH1	57	JAK2	87	PARP1		
28	CDK4	58	JAK3	88	PDGFRA		
29	CDKN2A	59	KDR	89	PIK3CA		
30	COL3A1	60	KIT	90	PMS2		

Overview of genes targeted in the HaloPlex™ panel. Genes targeted are colon cancer driver genes (all CCDS inclusive 30 bp intronic on 5' and 3' side). The average fragment length after digestion is +/- 100 bp. The total number of targeted regions is 1958. Total target region size is 486013 bp. DNA-input required is 225 ng.

---

**Supplemental Table 2**

1	ABL1	26	IDH2
2	AKT1	27	JAK2
3	ALK	28	JAK3
4	APC	29	KDR
5	ATM	30	KIT
6	BRAF	31	KRAS
7	CDH1	32	MET
8	CDKN2A	33	MLH1
9	CSF1R	34	MPL
10	CTNNB1	35	NOTCH1
11	EGFR	36	NPM1
12	ERBB2	37	NRAS
13	ERBB4	38	PDGFRA
14	EZH2	39	PIK3CA
15	FBXW7	40	PTEN
16	FGFR1	41	PTPN11
17	FGFR2	42	RB1
18	FGFR3	43	RET
19	FLT3	44	SMAD4
20	GNA11	45	SMARCB1
21	GNAQ	46	SMO
22	GNAS	47	SRC
23	HNF1A	48	STK11
24	HRAS	49	TP53
25	IDH1	50	VHL

*Overview of genes targeted in the Ion AmpliSeq™ Cancer Hotspot Panel v2 (n=50). Number of amplicons is 207. The number of COSMIC mutations covered with this panel is 2800. DNA-input required is 10 ng.*



**Supplemental Table 3**

Case	KRAS				EGFR						BRAF			
	HaloPlex		Ampliseq		HaloPlex			Ampliseq			HaloPlex			
	1		1		2	3		2	3		4			
1	19	47	966	4040	0	500	0	222	0	982	0	1235	0	84
2	13	24	2572	5589	0	730	0	114	0	864	0	1272	0	65
3	0	29	10	2310	0	229	0	43	0	557	0	555	0	31
4	0	65	11	2183	0	663	0	87	0	1138	0	1023	0	56
5	0	23	11	589	0	308	0	19	0	97	0	111	0	15
6	0	61	14	1431	0	424	0	39	0	217	0	217	0	5
7	0	697	10	2278	1	3415	1	2300	1	1053	0	1014	0	1362
8	0	130	9	3240	0	803	0	400	0	1403	0	1538	0	156
9	51	91	839	2018	0	946	0	252	0	2523	0	2802	0	47
10	36	93	537	1115	0	601	0	222	0	2629	0	2718	0	81
11	11	64	1461	3288	0	142	0	107	0	2660	0	2562	0	53
12	32	54	1172	2508	0	493	0	240	0	1035	0	937	0	62
13	0	0	279	1052	0	24	0	1	0	1128	0	1147	0	0
14	0	14	224	254	0	336	0	113	0	1115	0	460	0	16
15	11	13	679	1100	0	591	0	85	0	2850	0	1949	0	18
16	30	57	438	825	0	747	0	116	0	1183	0	576	0	55

Number of WT and mutant reads for hotspot locations in KRAS, EGFR, BRAF and PIK3CA. 1= KRAS p.12; 2 = EGFR p.L858R; 3 = EGFR del19; p.747-753; 4 = BRAF p.V600E; 5 = PIK3CA p.E542K; 6= PIK3CA p.E545K; 7= PIK3CA p.H1047R.

BRAF				PIK3CA																
HaloPlex		Ampliseq		HaloPlex							Ampliseq									
4		4		5		6		7			5		6		7					
0	84	0	1100	0	5	0	360	0	878	3	2355	6	2355	2	2200					
0	65	0	515	0	3	0	437	0	906	0	2225	0	2225	0	1807					
0	31	0	503	0	0	0	399	0	544	550	2831	5	2582	0	1748					
0	56	0	382	0	2	0	409	0	635	0	1465	0	1465	1	1139					
0	15	0	106	0	0	0	284	0	374	1	501	2	501	0	310					
0	5	0	228	0	0	0	352	0	507	1	1272	9	1272	2	699					
0	1362	0	500	0	139	0	1493	0	3862	0	1048	1	1048	2	1244					
0	156	0	729	0	11	0	444	0	678	0	1730	0	1730	0	1635					
0	47	0	1231	0	0	0	275	0	535	3	2476	11	2476	1	1662					
0	81	0	526	0	4	0	152	0	324	1	928	0	928	1	856					
0	53	0	1542	0	3	0	113	0	296	0	3301	1	3301	3	2383					
0	62	1	484	0	5	0	233	0	447	0	1304	2	1304	0	1452					
0	0	0	441	0	0	0	5	0	6	15	1842	3	1842	0	1000					
0	16	0	104	0	0	0	73	0	285	2	285	0	285	0	264					
0	18	0	596	0	0	0	195	0	518	7	1655	0	1655	1	941					
0	55	0	124	0	2	0	134	0	746	0	286	0	286	0	414					

## Supplemental Table 4

Case 1 Primary colon tumor	HALO_Illumina	Ampliseq_Torrent
APC:NM_001127510:exon17:c.G4057T:p.E1353X	present	present
KRAS:NM_033360:exon2:c.G35T:p.G12V	present	present
PTEN:NM_000314:exon5:c.C388T:p.R130X	present	present
CCNT2:NM_001241:exon10:c.C1984T:p.R662W	present	not included in amliseq panel
ATM:NM_000051:exon8:c.G1010A:p.R337H	present after manual inspection	present
<b>Case 3 Primary colon tumor</b>		
ABCA1:NM_005502:exon41:c.G5563T:p.A1855S	present	not included in amliseq panel
TP53:NM_001126114:exon4:c.T325G:p.F109V	present	present
GUCY1A2:NM_000855:exon7:c.G1882T:p.G628W	present	not included in amliseq panel
<b>Case 5 Primary colon tumor</b>		
TP53:NM_001126112:exon10:c.C1009T:p.R337C	present	present
APC:NM_001127510:exon17:c.4010_4011del:p.1337_1337del	present	not included in amliseq panel
MICAL3:NM_001122731:exon3:c.C367T:p.R123C	present	not included in amliseq panel
TGFBR2:NM_001024847:exon1:c.G5C:p.G2A,	present	not included in amliseq panel
<b>Case 7 Primary colon tumor</b>		
NRAS:NM_002524:exon3:c.C181A:p.Q61K	present	present
MSH2:NM_000251:exon15:c.G2500A:p.A834T	present	not included in amliseq panel
FAT4:NM_024582:exon17:c.G13789T:p.D4597Y	present	not included in amliseq panel
TP53:NM_001126114:exon7:c.G743A:p.R248Q	present after manual inspection	present
<b>Case 9 Primary colon tumor</b>		
KRAS:NM_033360:exon2:c.G38A:p.G13D	present	present
TP53:NM_001126114:exon5:c.G524A:p.R175H	present	present
FAT4:NM_024582:exon3:c.T5539C:p.S1847P	present	not included in amliseq panel
APC:NM_001127510:exon11:c.G1085C:p.G362A	present	not included in amliseq panel
APC:NM_001127510:exon17:c.G2950T:p.E984X	present	not included in amliseq panel
APC:NM_001127510:exon17:c.4240_4241insT:p.V1414fs	present	not included in amliseq panel
CDH1:NM_004360:exon3:c.T208C:p.S70P	not detected	present
<b>Case 11 Primary colon tumor</b>		
CTNNB1:NM_001098209:exon5:c.G569A:p.R190H	present	not included in amliseq panel
KRAS:NM_033360:exon2:c.G35T:p.G12V	present after manual inspection	present
APC:NM_001127510:exon17:c.C4132T:p.Q1378X	present after manual inspection	present
<b>Case 13 Primary colon tumor</b>		
TP53:NM_001126114:exon4:c.293delC:p.P98fs	present	present
CSF1R:NM_005211:exon8:c.A1085C:p.H362P	present	not included in amliseq panel
KRAS:NM_033360:exon2:c.G34T:p.G12C	no reads present	present
<b>Case 15 Primary colon tumor</b>		
KRAS:NM_033360:exon2:c.G35C:p.G12A	present	present
TP53:NM_001126114:exon8:c.C817T:p.R273C	present	present
TP53:NM_001126114:exon5:c.G473A:p.R158H	present	present
GNAS:NM_016592:exon1:c.C676T:p.R226C	present	not included in amliseq panel
PARP1:NM_001618:exon19:c.G2656A:p.V886M	present	not included in amliseq panel
KDR:NM_002253:exon4:c.C481A:p.L161I	present	not included in amliseq panel
PIK3CA:NM_006218:exon10:c.C1636A:p.Q546K	present after manual inspection	present

Variants detected with the HaloPlex and AmpliSeq panel are shown. Discordant variants are highlighted in grey.

Case 2 Ovarian metastasis	HALO_Illumina	Ampliseq_Torrent
APC:NM_001127510:exon17:c.G4057T:p.E1353X	present	present
KRAS:NM_033360:exon2:c.G35T:p.G12V	present	present
PTEN:NM_000314:exon5:c.C388T:p.R130X	present	present
SMAD4:NM_005359:exon9:c.C1067T:p.P356L	present	present
CCNT2:NM_001241:exon10:c.C1984T:p.R662W	present	not included in amliseq panel
GNAS:NM_080425:exon1:c.C1684T:p.R562C	present	not included in amliseq panel
BRAF:NM_004333:exon10:c.G1271A:p.R424Q	present	not included in amliseq panel
HRAS:NM_176795:exon3:c.G239A:p.C80Y	not detected	present
FLT3:NM_004119:exon16:c.G2038A:p.A680T	not detected	present
FGFR3:NM_022965:exon16:c.C2012T:p.S671L	not detected	present
PDGFRA:NM_006206:exon15:c.G2071A:p.D691N	not detected	present
NOTCH1:NM_017617:exon34:c.C7394T:p.P2465L	not detected	present
<b>Case 4 Ovarian metastasis</b>		
ABCA1:NM_005502:exon41:c.G5563T:p.A1855S	present	not included in amliseq panel
TP53:NM_001126114:exon4:c.T325G:p.F109V	present	present
SMAD2:NM_001003652:exon2:c.C5T:p.S2L	present	not included in amliseq panel
PIK3CA:NM_006218:exon10:c.G1624A:p.E542K	no reads present	present
<b>Case 6 Ovarian metastasis</b>		
TP53:NM_001126115:exon6:c.C613T:p.R205C	present	present
TGFBR2:NM_001024847:exon1:c.G5C:p.G2A	present	not included in amliseq panel
APC:NM_001127510:exon17:c.4010_4011del:p.1337_1337del	present	not included in amliseq panel
<b>Case 8 Ovarian metastasis</b>		
NRAS:NM_002524:exon3:c.C181A:p.Q61K	present	present
MSH2:NM_000251:exon15:c.G2500A:p.A834T	present	not included in amliseq panel
FAT4:NM_024582:exon17:c.G13789T:p.D4597Y	present	not included in amliseq panel
TP53:NM_001126114:exon7:c.G743A:p.R248Q	present after manual inspection	present
<b>Case 10 Ovarian metastasis</b>		
KRAS:NM_033360:exon2:c.G38A:p.G13D	present	present
TP53:NM_001126114:exon5:c.G524A:p.R175H	present	present
ERBB4:NM_001042599:exon25:c.A3005C:p.K1002T	present	not included in amliseq panel
FAT4:NM_024582:exon3:c.T5539C:p.S1847P	present	not included in amliseq panel
APC:NM_001127510:exon17:c.G2950T:p.E984X	present	not included in amliseq panel
APC:NM_001127510:exon17:c.4240_4241insT:p.V1414fs	present	not included in amliseq panel
CDH1:NM_004360:exon3:c.T208C:p.S70P	not detected	present
<b>Case 12 Ovarian metastasis</b>		
CTNNB1:NM_001098209:exon5:c.G569A:p.R190H	present	not included in amliseq panel
KRAS:NM_033360:exon2:c.G35T:p.G12V	present after manual inspection	present
APC:NM_001127510:exon17:c.2941delC;p.P981fs	present	not included in amliseq panel
EGFR:NM_005228:exon15:c.1783_1784insC;p.C595fs	not detected	present
<b>Case 14 Ovarian metastasis</b>		
TP53:NM_001126114:exon4:c.293delC:p.P98fs	present	present
CSF1R:NM_005211:exon8:c.A1085C:p.H362P	present	not included in amliseq panel
ATM:NM_000051:exon55:c.G8146A:p.V2716I	present	not included in amliseq panel
KRAS:NM_033360:exon2:c.G34T:p.G12C	present after manual inspection	present
<b>Case 16 Ovarian metastasis</b>		
KRAS:NM_033360:exon2:c.G35C:p.G12A	present	present
TP53:NM_001126114:exon8:c.C817T:p.R273C	present	present
TP53:NM_001126114:exon5:c.G473A:p.R158H	present	present
GNAS:NM_016592:exon1:c.C676T:p.R226C	present	not included in amliseq panel
APC:NM_001127510:exon17:c.2880delA:p.S960fs	present	not included in amliseq panel
CSF1R:NM_005211:exon8:c.A1085C:p.H362P	present	not included in amliseq panel
FAT4:NM_024582:exon9:c.T8651A:p.L2884H	present	not included in amliseq panel
PIK3CA:NM_006218:exon10:c.C1636A:p.Q546K	present after manual inspection	present





## Chapter 6

### **Excluding Lynch syndrome in a female patient with metachronous DNA mismatch repair deficient colon- and ovarian cancer**

*Stijn Crobach<sup>1</sup>, Anne M.L. Jansen<sup>1</sup>, Marjolein J.L. Ligtenberg<sup>2</sup>, Marije Koopmans<sup>3</sup>, Maartje Nielsen<sup>3</sup>, Frederik J. Hes<sup>3</sup>, Juul T. Wijnen<sup>4</sup>, Winand N.M. Dinjens<sup>5</sup>, Tom van Wezel<sup>1</sup> & Hans Morreau<sup>1</sup>*

<sup>1</sup>Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands

<sup>2</sup>Department of Human Genetics and Department of Pathology, Radboud university medical center, Nijmegen, the Netherlands.

<sup>3</sup>Department of Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands

<sup>4</sup>Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands

<sup>5</sup>Department of Pathology, Erasmus MC Cancer Institute, Rotterdam, the Netherlands

## **Abstract**

Patients synchronously or metachronously presenting with ovarian and colon cancer can pose diagnostic challenges. A primary colon carcinoma can metastasize to one or both ovaries, two independent primary tumors can arise or an ovarian carcinoma can metastasize to the colon. Clinical and immunohistochemical characterization can aid the diagnosis. Recently, we reported that in difficult cases finding pathogenic APC variants supports a colonic origin.

In this case report we describe the clinical history of a female patient suspected for Lynch syndrome. She was diagnosed with a bilateral ovarian cancer at age 44, followed by the detection of a colon carcinoma 12.5 months later. Lesions of both sites showed a DNA mismatch repair deficiency with immunohistochemical loss of MLH1 and PMS2 expression without MLH1 promoter hypermethylation. In absence of germline MMR gene variants identical somatic MLH1 and CTNNB1 gene variants were found, indicating a clonal relation. MMR germline mosaicism was made unlikely by ultra deep sequencing of the MLH1 variant in DNA isolated from normal mucosa, blood, urine and saliva. Although initially being suspect for Lynch syndrome it was eventually concluded that a metachronously diagnosed colon carcinoma that metastasized to both ovaries was most likely.

## Introduction

In this report we describe a female patient diagnosed with bilateral endometrioid carcinoma of the ovaries at the age of 44. One year later an adenocarcinoma of the colon was detected. The discovery of the colon carcinoma created doubt about the primary origin of the ovarian tumors. Besides, because the patient met the Amsterdam/Bethesda revised criteria, Lynch syndrome (LS) was suggested.

The ovaries can be affected by metastases from several primary tumor sites.[1] Most metastases originate from the gastrointestinal tract, with the colon as most frequent primary location. However, primary ovarian tumors are more common than ovarian metastases; 85% versus 15%.[2] Since subtypes of primary ovarian cancers (especially endometrioid and mucinous adenocarcinomas) can show overlapping histological and immunohistochemical features with gastrointestinal tumor metastases, it can be difficult to discriminate these.[3, 4] A combined analysis of clinical and molecular features can be helpful in correctly diagnosing these tumors. Reanalysis of this case revealed both macroscopic and microscopic evidence for a colonic origin of the ovarian tumors. This thought was supported by up-to-date extensive molecular analyses that showed a clonal relationship between both tumors. Lynch syndrome, including DNA mismatch repair gene mosaicism, was ruled out.

---

## Materials and Methods

### Immunohistochemistry

Immunohistochemistry was performed as previously described.[8] The antibodies and dilutions that were used are as follows: MSH2 (1:25; DAKO Santa Clara, United States), MSH6 (1:400; GeneTex Irvine, United States), PMS2 (1:80; DAKO Santa Clara, United States) and MLH1 (1:40; DAKO Santa Clara, United States), CDX2 (1:80; DAKO Santa Clara, United States), keratin-7 (1:400; DAKO Santa Clara, United States), keratin-20 (1:200; DAKO Santa Clara, United States), ER (1:40; DAKO Santa Clara, United States), PR (1:400; DAKO Santa Clara, United States) and vimentin (1:1000; DAKO Santa Clara, United States).

### Methylation specific assay

The promoter region of *MLH1* was analyzed by methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) as previously described.[9]

### Microsatellite instability (MSI) analysis

Microsatellite analysis was performed using five mononucleotide microsatellite markers as previously described.[10]

### Germline analysis

Germline analysis of *MLH1*, *PMS2*, *MHS2* and *MSH6* variant was performed on DNA isolated from lymphocytes from a blood sample using standard procedures including the analysis for large deletions/duplications by the multiplex ligation-dependent probe amplification (MRC Holland, the Netherlands).

### Somatic and mosaicism analysis

Somatic mutation analysis of *MLH1* was performed using a laboratory developed multiplex AmpliSeq based NGS protocol followed by confirmation of detected mutations by Sanger sequencing.

Additional analysis of somatic variations was performed on DNA isolated using a fully automated DNA extraction procedure. The concentration of DNA was measured using a fluorometer (Qubit dsDNA HS, Thermo Fischer Scientific, Waltman MA USA). The amplicon library for targeted sequencing was constructed using AmpliSeq Cancer Hotspot Panel v2. This panel consists of a single primer pool and is designed to detect somatic cancer hotspot pathogenic variants in 207 amplicons covering 50 cancer related genes, including genes as *APC*, *KRAS*, *TP53*, *SMAD4* that are often altered in colorectal cancer. The whole *APC* gene was analyzed in a separate analysis as in the cancer hotspot panel only the mutation cluster region of *APC* is covered. Mo-

saicism analysis of the identified *MLH1* variant was performed by using a panel covering *MSH2*, *MSH6*, *PMS2*, *MLH1*, *POLD1* and *POLE*. Libraries were prepared with 10 ng of genomic DNA, and each sample was uniquely barcoded using IonXpress barcodes (Life Technologies). Next-generation sequencing was carried out according to the Ion Proton protocol.

### **Bioinformatic analysis**

The unaligned BAM file generated by the Proton sequencer were mapped against the human reference genome (GRCh37/hg19) using the TMAP 5.0.7 software with default parameters (<https://github.com/iontorrent/TS>). Subsequently variant calling was done using the Ion Torrent specific caller, Torrent Variant Caller (TVC)-5.0.2, using the recommended Variant Caller Parameter for Cancer Hotspot Panel v2. Variant interpretation was done using Geneticist Assistant (Softgenetics) which assigns Functional Prediction, Conservation scores and Disease associated information to each variant ([http://softgenetics.com/GeneticistAssistant\\_2.php](http://softgenetics.com/GeneticistAssistant_2.php)). Once pathogenicity is assigned to a variant, the same pathogenicity is automatically attributed the next time the variant is observed. Integrative Genomics Viewer (IGV) was used for visually inspecting variants (doi: 10.1093/bib/bbs017). The analysis of the complete *APC* gene was performed as described previously.[11] LOH was analyzed by comparison of variant and wild type DNA reads of the NGS results.

---

## Case Report

We describe a female patient with a family history of ovarian cancer (one sister at the age of 56 years), breast cancer (one sister at the age of 59 years), colon cancer (patient's mother at the age of 80 years) and a (non melanoma) skin cancer (the sister diagnosed with breast cancer). The index patient had one hyperplastic polyp removed from the rectum at the age of 43. Aged 44, she was diagnosed with bilateral endometrioid carcinoma of the ovaries with focally mucinous differentiation (Figure 1), clinical stage 1B according to the FIGO staging system. Surgery was performed and she was treated with adjuvant chemotherapy comprising a regimen of cyclophosphamide and carboplatin. At age 45, 12.5 months later, she was diagnosed with an adenocarcinoma of the colon, treated by a left-sided hemicolectomy. Based on these clinical records the patient met the Amsterdam/Bethesda revised criteria. Patient has remained disease-free until the age of 64. However, the discovery of the (mucinous) colon carcinoma showing partly a similar morphology as the ovarian tumors (Figure 1), created doubt about the primary origin of the ovarian tumors. Lynch syndrome (LS), in which independent ovarian and colon tumors had developed, was suggested.

Immunohistochemistry testing of the MMR proteins MLH1, PMS2, MSH2 and MSH6 of the colonic and ovarian tumors showed DNA MMR deficiency with loss of expression of MLH1 and PMS2. Microsatellite instability (MSI) testing using mononucleotide microsatellite markers showed an MSI-H phenotype. A sporadic origin of these MMR deficient tumors due to *MLH1* promoter hypermethylation was excluded. Our patient was subsequently referred to a clinical geneticist. However LS could not be confirmed after negative lymphocyte DNA testing of *MLH1*, *PMS2*, *MHS2* and *MSH6* for germline pathogenic variants. Also germline testing of *BRCA1* and *BRCA2* in two sisters of the patient was negative.

Reevaluation of the metachronously diagnosed colon tumor confirmed the primary origin in the colon as the bulk of the tumor was bulging in the colonic lumen. Furthermore, the serosal lining was unaffected. Immunohistochemical stainings of the ovarian tumor showed a phenotype compatible with a metastasis from a colon tumor (keratin-7 negative / keratine-20 and CDX-2 positive). ER, PR and vimentin were also negative. However, ovarian tumors with mucinous differentiation can show a wide variety of keratine-7/keratin-20 immunoprofile patterns, and should be interpreted with caution.[12]

Somatic testing (Table 1) of *MLH1* showed an identical *MLH1* class 5 pathogenic variant (c.1624C>T, p.(Gln542\*)) in both colon and ovarian tumors. Next, loss of het-

erozygosity for *MLH1* was shown by absence of the WT(wild type)-allele. We also somatically tested the complete *APC* gene for pathogenic variants in these lesions, as finding of pathogenic *APC* variants in ovarian neoplasms would point at a colonic origin of the lesions. No *APC* variants were found, however an identical activating class 5 *CTNNB1* pathogenic variant (c.134C>T; p.(Ser45Phe)) was identified, the molecular alternative for Wnt-pathway activation (Figure 2A). Finding identical *MLH1* and *CTNNB1* variants would suggest a clonal relation between the colon and ovarian tumor. Additionally, a class 5 pathogenic *TP53* variant (c.1024C>T, p.(Arg342\*)) was detected in the colon tumor, but absent in the ovarian tumor (Figure 2B). In order to estimate putative germline mosaicism we performed ultra-deep sequencing of the *MLH1* (c.1624C>T, p.(Gln542\*)) variant in DNA isolated from normal colonic mucosa, saliva, blood and urine. All isolates showed sufficient (>10K) coverage, but showing no presence of the *MLH1* variant, rendering germline mosaicism unlikely. It was concluded that a metachronously diagnosed colorectal tumor that metastasized to both ovaries was the most likely diagnosis.

---

## Discussion

In the current report we address a remarkable clinical dilemma once metachronous ovarian and colon tumors are diagnosed and the possibility of a Lynch syndrome needs to be answered. The female patient we now present with bilateral ovarian cancer was treated as having primary bilateral ovarian cancer. However, only 12.5 months after the first diagnosis the true primary origin of these lesions was questioned with the resection of a DNA mismatch repair deficient left sided colon cancer. After reevaluation and molecular analysis a clonal relation was identified between the ovarian and colonic lesions. As MMR deficient cancers mostly lack distant metastases possibly due to the interaction with the immune system, it is noteworthy that this DNA MMR deficient colon cancer probably metastasized to the ovaries.[13] .

About 15% of all ovarian tumors turn out to be metastases.[2] Histological parameters are not always sufficient to discriminate between a primary tumor and/or metastasis. Nowadays, molecular analysis can be a helpful tool to make this distinction in selected cases.[5] Inactivating *APC* pathogenic variants are almost exclusively found in colon tumors. Thus, the presence of a pathogenic variant is a strong argument for a primary colon tumor.[3, 5, 14] In our patient no pathogenic variants in the *APC* gene were found, but an identical activating *CTNNB1* variant was present in both ovary and colon tumors. As *CTNNB1* variants are very rare in colon carcinomas, this might suggest the ovarian tumor as the primary origin.[6] On the other hand activating *CTNNB1* pathogenic variants are often found in colon cancer associated with DNA mismatch repair deficiency.[7, 15, 16] Only an incidental report of metastatic mismatch repair deficient colon carcinoma to the ovaries is described.[17] In previous published research we did not find any *CTNNB1* pathogenic variants in MMR proficient colorectal metastases to the ovary.[5] With respect to the ovary, *CTNNB1* pathogenic variants have mainly been found in endometrioid ovarian cancers. [6] However, the histopathological findings in our case do not suggest metastases from the ovary since the colonic tumor was located at the luminal site. In case of a metastasis the bulk of the tumor would have been present on the serosal site. Besides, ovarian cancers metastasizing to the colon, and morphologically mimicking a primary colon tumor are probably very rare. Furthermore in case of bilateral ovarian tumors the odds favor metastases from a primary tumor elsewhere in the body.

In our patient the same somatic *MLH1* pathogenic variant and concomitant loss of heterozygosity of the wild type allele was present in the ovarian and colon cancer. As the detected *MLH1* variant was not found by deep sequencing of DNA isolated from normal mucosa, saliva, blood and urine a germline mosaicism was rendered unlikely.

Somatic *MLH1* pathogenic variants in sporadic tumors are mainly associated with



gastrointestinal tumors.[6, 18] *MLH1* pathogenic variants are not commonly found in ovarian cancer, although one study found *MLH1* pathogenic variants in 8,7% epithelial ovarian cancer.[19] Usually, *TP53* pathogenic variants occur early in the evolutionary development of a tumor. Our patient's tumors showed in two tumors *CTNNB1* as well as *MLH1* pathogenic variants, but only in the colon tumor a *TP53* pathogenic variant was identified. The presence of this variant can be explained by tumor progression within the primary colon tumor. Apparently in this case, the pathogenic *TP53* variant is not present in the metastasizing clone. Such spatial differences in mutation profiles within a tumor are known as intra-tumor heterogeneity.

In summary, we discuss the clinical dilemma with metachronous diagnosed bilateral mismatch repair deficient ovarian and colon cancer harboring a pathogenic MMR variant. In our case Lynch syndrome as well as a postzygotic somatic mutation leading to mosaicism of multiple normal tissues are very unlikely. Molecular analysis showed a clonal relationship between the ovarian and colon tumors with histopathological analysis suggesting the colon tumor being the primary tumor.

## **Acknowledgements**

We thank M.A. Wijngaarden for help with writing and critically reading the manuscript.

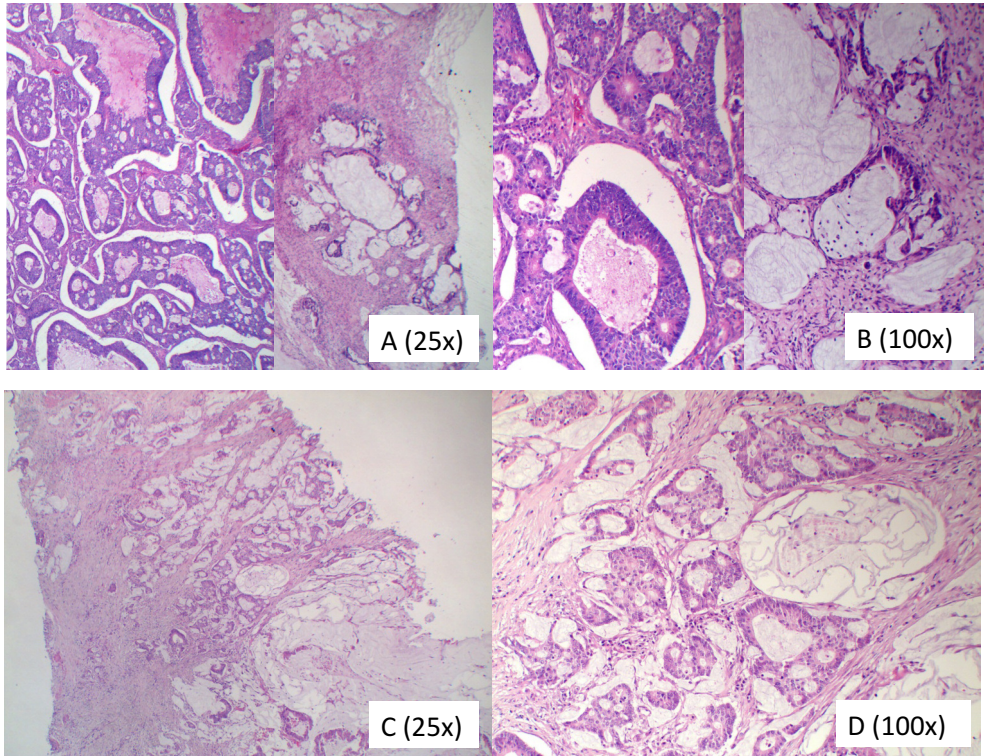
---

## References

1. Khunamornpong, S., et al., *Primary and metastatic mucinous adenocarcinomas of the ovary: Evaluation of the diagnostic approach using tumor size and laterality*. Gynecol. Oncol, 2006. **101**(1): p. 152-157.
2. de Waal, Y.R., et al., *Secondary ovarian malignancies: frequency, origin, and characteristics*. Int. J. Gynecol. Cancer, 2009. **19**(7): p. 1160-1165.
3. Kelemen, L.E. and M. Kobel, *Mucinous carcinomas of the ovary and colorectum: different organ, same dilemma*. Lancet Oncol, 2011.
4. Lee, K.R. and R.H. Young, *The distinction between primary and metastatic mucinous carcinomas of the ovary: gross and histologic findings in 50 cases*. Am. J. Surg. Pathol, 2003. **27**(3): p. 281-292.
5. Crobach, S., et al., *Target-enriched next-generation sequencing reveals differences between primary and secondary ovarian tumors in formalin-fixed, paraffin-embedded tissue*. J. Mol. Diagn, 2015. **17**(2): p. 193-200.
6. Bamford, S., et al., *The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website*. Br. J. Cancer, 2004. **91**(2): p. 355-358.
7. Lovig, T., et al., *APC and CTNNB1 mutations in a large series of sporadic colorectal carcinomas stratified by the microsatellite instability status*. Scand. J. Gastroenterol, 2002. **37**(10): p. 1184-1193.
8. Lips, E.H., et al., *Progression and tumor heterogeneity analysis in early rectal cancer*. Clin Cancer Res, 2008. **14**(3): p. 772-81.
9. Homig-Holzel, C. and S. Savola, *Multiplex ligation-dependent probe amplification (MLPA) in tumor diagnostics and prognostics*. Diagn Mol Pathol, 2012. **21**(4): p. 189-206.
10. Boland, C.R., et al., *A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer*. Cancer Res, 1998. **58**(22): p. 5248-57.
11. Jansen, A.M., et al., *Distinct Patterns of Somatic Mosaicism in the APC Gene in Neoplasms From Patients With Unexplained Adenomatous Polyposis*. Gastroenterology, 2017. **152**(3): p. 546-549 e3.
12. Cathro, H.P. and M.H. Stoler, *Expression of cytokeratins 7 and 20 in ovarian neoplasia*. Am. J. Clin. Pathol, 2002. **117**(6): p. 944-951.
13. de Miranda, N.F., et al., *Role of the microenvironment in the tumorigenesis of microsatellite unstable and MUTYH-associated polyposis colorectal cancers*. Mutagenesis, 2012. **27**(2): p. 247-53.
14. Miyoshi, Y., et al., *Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene*. Hum. Mol. Genet, 1992. **1**(4): p. 229-233.
15. Johnson, V., et al., *Exon 3 beta-catenin mutations are specifically associated with col-*

- orectal carcinomas in hereditary non-polyposis colorectal cancer syndrome. *Gut*, 2005. **54**(2): p. 264-7.
16. Mirabelli-Primdahl, L., et al., *Beta-catenin mutations are specific for colorectal carcinomas with microsatellite instability but occur in endometrial carcinomas irrespective of mutator pathway*. *Cancer Res*, 1999. **59**(14): p. 3346-3351.
  17. Ongom, P.A., et al., *Metastatic colorectal carcinoma mimicking primary ovarian carcinoma presenting as 'giant' ovarian tumors in an individual with probable Lynch syndrome: a case report*. *J Med Case Rep*, 2013. **7**: p. 158.
  18. Herfarth, K.K., et al., *Mutations in MLH1 are more frequent than in MSH2 in sporadic colorectal cancers with microsatellite instability*. *Genes Chromosomes Cancer*, 1997. **18**(1): p. 42-9.
  19. Kim, Y.M., et al., *Analysis and comparison of somatic mutations in paired primary and recurrent epithelial ovarian cancer samples*. *PLoS One*, 2014. **9**(6): p. e99451.

**Figure 1**



*Figure 1 shows the histological picture of the ovarian tumor (A and B) and the colon tumor (C and D). In Figure A en B both endometrioid and mucinous parts of the ovarian tumor are shown*

Figure 2

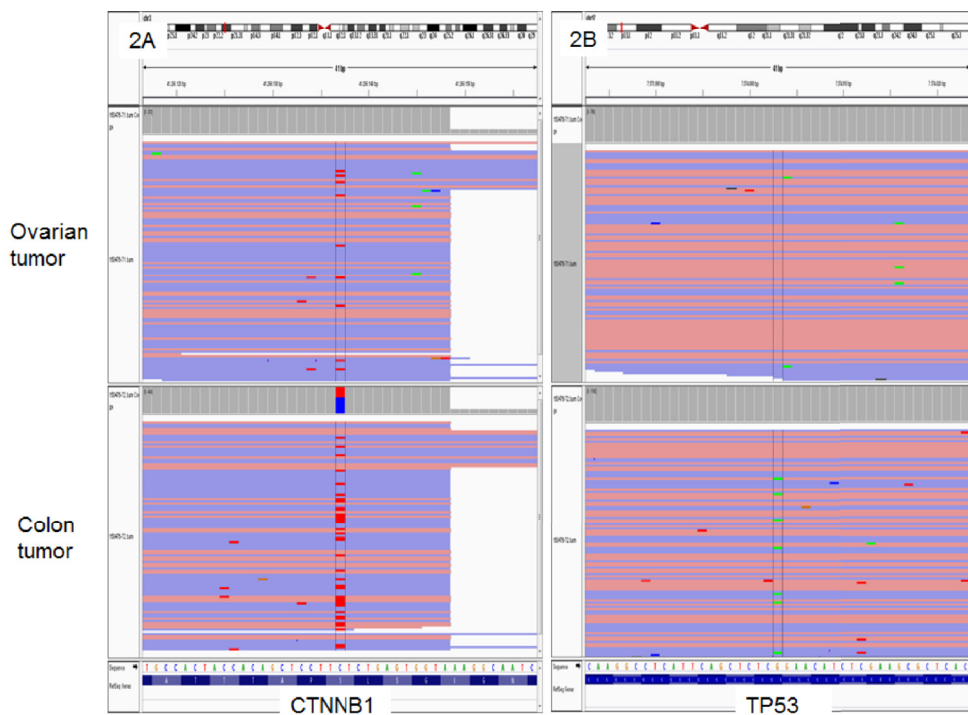


Figure 2A shows the reads including the pathogenic CTNNB1 variant that is present in both the ovarian and the colon tumor.

Figure 2B shows the reads including the pathogenic TP53 variant that is present in the colon tumor, but not in the ovarian tumor.

**Table 1**

<b>Gene</b>	<b>Ovary (T%: &gt; 60%)</b>	<b>Colon (T%: &gt; 50%)</b>	<b>Leucocytes</b>
<i>TP53</i>	No pathogenic variant	c.1024C>T, p.(Arg342*) / 6,1% mutant reads	
<i>MLH1</i>	c.1624C>T, p.(Gln542*) / 76% mutant reads LOH Loss of expression by IHC No promoter hypermethylation	c.1624C>T, p.(Gln542*) / 52% mutant reads LOH Loss of expression by IHC No promoter hypermethylation	No pathogenic variant
<i>CTNNB1</i>	c.134C>T, p.(Ser45Phe) / 9,2% mutant reads	c.134C>T, p.(Ser45Phe) 38% mutant reads	
<i>PMS2</i>	Loss of expression by IHC	Loss of expression by IHC	No pathogenic variant
<i>MSH2</i>	Normal expression by IHC	Normal expression by IHC	No pathogenic variant
<i>MSH6</i>	Normal expression by IHC	Normal expression by IHC	No pathogenic variant
<i>APC</i>	No pathogenic variant	No pathogenic variant	

LOH = loss of heterozygosity IHC = immunohistochemical staining T% = tumor cell percentage

*Table 1 shows an overview of the detected pathogenic variants, methylation assays and immunohistochemical staining results of mismatch repair genes in one of the ovarian tumors, the colon tumor and DNA isolated from blood.*

## Chapter 7

### **Concluding remarks and future directions**

This thesis describes the application of next-generation sequencing to the optimal diagnosis of ovarian tumors. For ovarian tumors, knowledge about their tissue of origin is crucial for therapeutic choices and statements regarding prognosis. The origin can be difficult to determine in a subset of cases.[1, 2] Macroscopic and histologic examination complemented by immunohistochemistry can, in some situations, be insufficient to identify the origin of a tumor process.[3, 4] Antibodies for novel targets are being developed but immunohistochemistry can still be difficult to interpret, not adhering to unmistakable guidelines. Therefore, additional (molecular) techniques are needed to help identify the origin of a tumor.

In chapter two, we describe a series of colorectal cancer (CRC) and duodenal cancer cases with ovarian metastases. In a cohort of 30 ovarian CRC metastases, four patients with Familial Adenomatous Polyposis (FAP) were identified. This enrichment of FAP patients in a series of CRCs is striking since the estimated incidence of FAP CRC is far below 1% of all CRCs.[5] The expected incidence of FAP CRC that metastasizes to the ovaries would thus be almost zero. The necessity of combined surgical resection of the primary CRC and bilateral oophorectomy is a matter of debate. Especially in female FAP CRC patients, bilateral oophorectomy during surgery should be discussed.

In chapter 3, we used a custom-made next-generation sequencing (NGS) panel, including 115 cancer-driving genes to screen formalin-fixed paraffin-embedded (FFPE) tumor tissue from 43 primary endometrioid and mucinous ovarian carcinomas. We then compared them to 28 proven ovarian colorectal cancer (CRC) metastases. The mutations were validated by high-resolution melting curve analysis (HRMCA) and Sanger sequencing. Furthermore, loss of heterozygosity (LOH) and promoter hypermethylation of APC were also studied. In the primary ovarian tumors, *TP53*, *NOTCH1*, *PIK3CA* and *FAT4* were the most frequently mutated genes. In the ovarian metastases, *APC*, *TP53*, *KRAS* and *FAT4* mutations were the most common mutations in ovarian CRC metastases. Inactivating *APC* mutations were identified in 71% of CRC metastases, which is in contrast to 4.7% identified in primary ovarian tumors. LOH and *APC* promoter hypermethylation did not differ significantly between the primary and secondary ovarian tumors. It can be concluded that *APC* mutation analysis can be used to differentiate primary endometrioid and mucinous ovarian tumors from ovarian CRC metastases.

In chapter four, the gene panel from chapter 3 was used to analyze the mutational profiles of primary colorectal cancers (CRCs) and the corresponding ovarian metastases. We compared 26 primary CRCs and 30 matching ovarian metastases (4 with



bilateral metastases). Low thresholds were used in bioinformatics analysis to prevent low frequency passenger mutations from being filtered out. Sanger and/or hydrolysis probe assays were used to validate a subset of variants. No striking differences were observed between the mutational landscape of CRC that metastasized to the ovary and CRC in consecutive series.[6] There was considerable overlap in the mutations of early-affected genes when comparing primary CRCs and their matching ovarian metastases. A subset of mutations, presumably passenger mutations, demonstrated less overlap. In particular, primary CRCs showed a substantially high number of supposed passenger mutations. We also compared the primary CRCs and matching metastases for stratifying variants of 6 genes (*KRAS*, *NRAS*, *BRAF*, *FBXW7*, *PTEN* and *PIK3CA*) that select for established (*EGFR*-directed) or future targeted therapies. Out of a total of 31 variants, 12 were not found in either of the two locations. Therefore, the number of discordant variants between the primary tumors and their matching metastases differed. Half of these discordant variants were pathogenic variants. In terms of temporal heterogeneity, no clear relationship was observed between the number of discordant variants related to the time interval between primary CRCs and the detection of ovarian metastases. Dormant metastases may therefore be present from the early days of the primary tumors.

In chapter 5, the Haloplex targeting technique used for next-generation sequencing (NGS) was studied and validated with another amplicon based targeted approach (Ion Ampliseq).[7, 8] NGS has been proven to be successfully applicable to fragmented formalin-fixed paraffin-embedded (FFPE) tissue.[9] A large number of targeted sequencing approaches are offered based on different principles, such as polymerase chain reaction (PCR), hybridization or circularization.[10] We showed that a circularization-based approach (HaloPlex), followed by sequencing on Illumina HiSeq, is successful for targeted sequencing of DNA from FFPE material. Detected variants were validated with a PCR-based targeted enrichment method (Ion AmpliSeq), followed by sequencing on an Ion PGM sequencer. A high concordance rate between the detected variants from different sample preparation techniques and sequencing methods was observed. The discordant variants could largely be explained by (subtle) setting differences in the analysis pipeline. Thus, optimal bioinformatics analysis is crucial for the correct detection of variants. In addition, tumor intra-heterogeneity (ITH) resulting in different DNA isolates can cause discordant sequencing results.

The case report in chapter 6 describes a female patient suspected of having Lynch syndrome. She was diagnosed with bilateral ovarian cancer at age 44, followed by the detection of a colon carcinoma 12.5 months later. Patients with synchronous or

metachronous ovarian and colon cancers pose diagnostic challenges.[11] Primary colon carcinomas can metastasize to one or both ovaries, two independent primary tumors can arise or an ovarian carcinoma can metastasize to the colon. Clinical and immunohistochemical characterization can be helpful. In this case, lesions of both sites showed a DNA mismatch repair deficiency with immunohistochemical expression loss of MLH1 without *MLH1* promoter hypermethylation. In the absence of germ line MMR gene variants, identical somatic *MLH1* and *CTNNB1* gene variants were found, indicating a clonal relation. MMR germ line mosaicism was ruled out by ultra-deep sequencing of the *MLH1* variant in DNA isolated from normal mucosa, blood, urine and saliva. Although initially suspected of having Lynch syndrome, it was eventually concluded that a metachronously metastasized colorectal tumor to both ovaries was most likely. This report illustrates the diagnostic dilemmas that can be encountered in solving suspected Lynch syndrome cases.

### **Next-generation sequencing (NGS)**

The last decade brought enormous technical developments in the field of molecular diagnostics. Sanger sequencing, a labor-intensive technique in which limited regions of a gene can be analyzed, has been replaced by next-generation sequencing (NGS), in which multiple genes can be examined in parallel.[10, 12] Next, the development of targeted therapies directed at specific genetic variants has taken an important place in the treatment of oncologic patients.[13] The detection of specific genetic variants is decisive in choosing the optimal treatment strategy. With extensive molecular profiling, the number of molecular targets that can be treated by specific agents is increasing. Therefore, accurate and fast methods are needed to detect targetable variants.[14]

Currently, targeted techniques in which only a selection of genes is investigated is the preferred method in a clinical setting, as the gene targeting and the subsequent bioinformatic analysis is less complex.[12, 15] Still, targeted sequencing devices that decrease sequencing time and reduce costs are desired. Most likely, whole-genome approaches will eventually become more important in the future.[16] A complete overview of most genomic variants, without first identifying the DNA of interest, can be obtained in this way.

Currently, several sequencing platforms are available, all of which have their pros and cons.[12] The development of these platforms resulted in the generation of longer sequencing reads (now generally several hundreds of base pairs), making genome assembly easier. A recent development in the field of NGS is the generation of even longer reads (thousands of base pairs).[12] These long reads are helpful in analyzing repetitive elements and complex structural variants.

The recent single molecule sequencing platforms overcame the problems generated by amplification of the input DNA.[12] As very limited input material is needed and the sequencing speed is faster than in the existing techniques, single molecule sequencing holds great potential for the future.[17] This is especially true as the input of tumor DNA might be decreasing with the increased use of neoadjuvant therapies (e.g., breast, rectum, esophagus and stomach cancers). Additionally, the applicability of sequencing techniques with freely circulating plasma-derived DNA as input, which will become more important in the near future, also needs to be investigated. Finally, the focus of attention may be shifting to sequencing non-coding regions, as recent research has shown that non-coding regions are involved in carcinogenesis.[18]

### **Bioinformatic analysis**

NGS produces enormous amounts of data. Bioinformatics analysis is necessary to obtain trustworthy data.[19] A list of pathogenic variants should not be contaminated by false-positive or false-negative findings. False-positive findings can be caused by technical issues that can differ by platform. Another problem is deciding whether a variant is non-pathogenic and present in the population or whether it is a true pathogenic variant responsible for tumor development. False-negative findings can be caused by suboptimal filter setting in a bioinformatic pipeline. As whole-genome sequencing is becoming more important, bioinformatic analysis and data storage infrastructure have to be able to cope with this increasing amount of data.[16] Next, new sequencing platforms can offer benefits with respect to the speed and the cost of sequencing. However, specific artifacts potentially linked to new sequencing techniques must be examined.

The refinement of sequencing detection techniques leads to problematic decision making about thresholds. The previously used Sanger sequencing was much less sensitive than the current NGS techniques. With Sanger sequencing, a somatic genetic variant was detectable once it was present in lesions with >20% of tumor cells. With the current techniques, somatic variants present in 0.1% of the cells can be detected. Whether the variant profile of such a small percentage of tumor cells should be decisive in choosing treatment strategies is unknown. For example, if 0.1% of the tumor cells of a colon carcinoma carry a *KRAS* mutation, it is unknown whether that patient should be excluded from targeted therapy or if it is still effective to target the other 99.9% of the tumor cells.

A positive effect of the high sensitivity of NGS techniques is the ability to detect variants in samples with a low tumor percentage. Previously, variants were often not discovered in samples with only a limited number of tumor cells. However, whether these

negative findings were false-negative observations could not be determined. Currently, the absence of variants after NGS analysis with sufficient read depth can be stated with much more certainty.

However, with a low input of DNA, the number of separate DNA molecules is also reduced. Sequencing artifacts caused by the NGS protocol, due to, for example, preferential amplification or cytosine deamination, can result in false-positive and false-negative results. Single molecule molecular inversion probes (smMIP) can help with detecting those sequencing artifacts.[20, 21]

### **Defining the origin of tumors**

With the introduction of NGS, the generation of tumor-specific mutation profiles, correlated to the primary tumor site, was anticipated. If each tumor type showed a characteristic mutation profile, these profiles could be used to define tumor origins. However, extensive sequencing projects (e.g., the TCGA project) have revealed that most tumor types share similar driver variants, in addition to a wide spectrum of passenger variants that are present in subclones within the tumor [6, 22]. In general, there are no specific mutation profiles linked to the origin of a tumor. However, according to the clinical setting, variants in specific genes can be informative.[23]

Somatic mutation detection is only one component of the molecular profile of a tumor. Copy number alterations can be detected by NGS or by array technology.[24]

Several array approaches that can help identify the primary location of a tumor have been described. Some of these tests are commercially available.[25] A more functional read-out of somatic variants and gene fusions can be obtained by transcriptome sequencing that covers the complete set of RNA proteins. Finally, epigenetic modifications influencing the expression levels of genes could be informative. Integration of all these techniques will produce a more complete molecular profile of tumors. Most likely, better predictions can be made about the origin of cancer metastasis in cases of unknown primary tumors based on these integrated profiles.

Nevertheless, despite the application of currently available tools (imaging, immunohistochemistry, and molecular analysis) approximately 3% of carcinomas cannot be traced back to their origin.[26] One may argue that defining the origin may be less important as potential targeted therapies can be applied irrespective of the location of the tumor (“basket studies”). The effectiveness of such an approach is currently being investigated. However, molecular analysis should be considered in the context of the complete clinical picture. For instance, *BRAF* inhibition alone can be an effec-

tive treatment for disseminated melanoma, but not for a metastasized colon carcinoma.[27]

### **Targeting therapies**

Clinical trials have shown that targeting a specific genetic variant can result in spectacular results. However, the effect of these therapies often only lasts for a limited period (several months). The idea that the targeting of (e.g., inhibition of) a single driver gene of a tumor could be enough to stop the growth of a malignant process has been proven to be too simplistic. Escape mechanisms are activated in tumor cells that enable their continuous growth.[28] A combination of targeted therapies aimed at more than one pathway has been shown to be more effective. For example, a combined therapy of *BRAF*- and *MEK*-inhibition in melanoma with *BRAF* mutations showed an improved response in comparison with *BRAF*-targeted single therapy.[29]

### **Intra-tumor heterogeneity (ITH)**

Another challenge with targeted therapies is that the genetic variants that are targeted can be present in only a portion of the tumor cells, reflecting intra-tumor heterogeneity (ITH). Tumors arise initially from a single cell in which driver mutations create a growth advantage. Further progression of the tumor leads to the accumulation of additional mutations, which are present in subclones of the tumor. ITH develops over time (temporal heterogeneity) and space (spatial heterogeneity), leading to a tumor consisting of a heterogeneous population of cells.[30] Targeted therapies that are directed against specific genetic variants will not be effective in tumor cells not carrying those specific variants. To get an up-to-date and complete overview of the genetic changes within a tumor, multiple samples over time and from different regions should be taken. Such an approach is impractical. Currently, the choices of targeted therapies are based on a single biopsy at one moment in time. That molecular profile is not an up-to-date and complete reflection of the mutation profile of the tumor. Liquid biopsies (the analysis of tumor DNA in systemic circulation) may overcome these problems. However, at this moment, it is not known whether the subclones that are present in the tumor contribute equally to circulating DNA.

Although targeted therapies may so far not have resulted in effective and sustained results in tumors, personalized care will remain an important treatment module in oncology. Standard use of liquid biopsies, simultaneous treatment with multiple drugs and expanding the repertoire of agents directed to targetable variants will be seen in the near future. Another important development in cancer treatment is immunotherapy.[31] Immunotherapy can be categorized into two types. The first type is active immunotherapy, in which the immune system of the patient is directed to the tumor

cells. The second type, passive immunotherapy, involves enabling existing immune responses to attack tumor cells or injecting antibodies or T-cells.

An active response of the immune system directed to the tumor cells can be achieved by injecting tumor specific antigens, as is done with the BCG (*Bacillus Calmette-Guérin*) treatments for bladder cancer. Antigen vaccines are now also developed for, among others, melanoma, colon cancer and leukemia.

Administration of antibodies is now possible for certain forms of lymphoma (anti-CD20), colorectal carcinoma (anti-EGFR / anti-VEGF) and gastric/breast carcinoma (anti-Her2neu).

Another option to create an anti-tumor response by the immune system is the injection of activated dendritic cells expressing tumor antigens.[32] These dendritic cells will provoke an immune response in lymph nodes. Successful trials have been performed for melanoma, renal cell carcinoma, lymphoma and glioblastoma.

Tumor infiltrating lymphocytes (TILs) are often present in tumor tissue but are not very effective because of the immunosuppressive environment that is created by the tumor cells.[33] Harvesting TILs from a patient, expanding their numbers, activating them and returning them into the patient can result in a tumor cell-directed immune response. Immunomodulation is possible by anti-CTL antigen-4 (CTLA-4) antibodies that neutralizes the inhibiting activity on T-cells.[34, 35] Anti-PD1 antibodies have a comparable working mechanism.

## References

1. Lewis, M.R., et al., *Ovarian involvement by metastatic colorectal adenocarcinoma: still a diagnostic challenge*. Am. J. Surg. Pathol, 2006. **30**(2): p. 177-184.
2. Prat, J., *Ovarian carcinomas, including secondary tumors: diagnostically challenging areas*. Mod. Pathol, 2005. **18 Suppl 2**: p. S99-111.
3. Baker, P.M. and E. Oliva, *Immunohistochemistry as a tool in the differential diagnosis of ovarian tumors: an update*. Int. J. Gynecol. Pathol, 2005. **24**(1): p. 39-55.
4. Dennis, J.L., et al., *Markers of adenocarcinoma characteristic of the site of origin: development of a diagnostic algorithm*. Clin. Cancer Res, 2005. **11**(10): p. 3766-3772.
5. Galiatsatos, P. and W.D. Foulkes, *Familial adenomatous polyposis*. Am. J. Gastroenterol, 2006. **101**(2): p. 385-398.
6. *Comprehensive molecular characterization of human colon and rectal cancer*. Nature, 2012. **487**(7407): p. 330-337.
7. Dahl, F., et al., *Multiplex amplification enabled by selective circularization of large sets of genomic DNA fragments*. Nucleic Acids Res, 2005. **33**(8): p. e71.
8. Singh, R.R., et al., *Clinical validation of a next-generation sequencing screen for mutational hotspots in 46 cancer-related genes*. J Mol Diagn, 2013. **15**(5): p. 607-22.
9. Hadd, A.G., et al., *Targeted, high-depth, next-generation sequencing of cancer genes in formalin-fixed, paraffin-embedded and fine-needle aspiration tumor specimens*. J. Mol. Diagn, 2013. **15**(2): p. 234-247.
10. Mamanova, L., et al., *Target-enrichment strategies for next-generation sequencing*. Nat. Methods, 2010. **7**(2): p. 111-118.
11. Kelemen, L.E. and M. Kobel, *Mucinous carcinomas of the ovary and colorectum: different organ, same dilemma*. Lancet Oncol, 2011.
12. Goodwin, S., J.D. McPherson, and W.R. McCombie, *Coming of age: ten years of next-generation sequencing technologies*. Nat Rev Genet, 2016. **17**(6): p. 333-51.
13. Huang, M., et al., *Molecularly targeted cancer therapy: some lessons from the past decade*. Trends Pharmacol Sci, 2014. **35**(1): p. 41-50.
14. Xuan, J., et al., *Next-generation sequencing in the clinic: promises and challenges*. Cancer Lett, 2013. **340**(2): p. 284-95.
15. Kerick, M., et al., *Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity*. BMC. Med. Genomics, 2011. **4**: p. 68.

16. Nakagawa, H., et al., *Cancer whole-genome sequencing: present and future*. *Oncogene*, 2015. **34**(49): p. 5943-50.
17. Shen, T., et al., *Clinical applications of next generation sequencing in cancer: from panels, to exomes, to genomes*. *Front Genet*, 2015. **6**: p. 215.
18. Chmielecki, J. and M. Meyerson, *DNA sequencing of cancer: what have we learned?* *Annu Rev Med*, 2014. **65**: p. 63-79.
19. Alioto, T.S., et al., *A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing*. *Nat Commun*, 2015. **6**: p. 10001.
20. Hiatt, J.B., et al., *Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation*. *Genome Res*, 2013. **23**(5): p. 843-54.
21. Chen, G., et al., *Cytosine deamination is a major cause of baseline noise in next-generation sequencing*. *Mol Diagn Ther*, 2014. **18**(5): p. 587-93.
22. Vogelstein, B., et al., *Cancer genome landscapes*. *Science*, 2013. **339**(6127): p. 1546-1558.
23. Crobach, S., et al., *Target-enriched next-generation sequencing reveals differences between primary and secondary ovarian tumors in formalin-fixed, paraffin-embedded tissue*. *J. Mol. Diagn*, 2015. **17**(2): p. 193-200.
24. Horlings, H.M., et al., *Gene expression profiling to identify the histogenetic origin of metastatic adenocarcinomas of unknown primary*. *J. Clin. Oncol*, 2008. **26**(27): p. 4435-4441.
25. Economopoulou, P., et al., *Cancer of Unknown Primary origin in the genomic era: Elucidating the dark box of cancer*. *Cancer Treat Rev*, 2015. **41**(7): p. 598-604.
26. *Comprehensive molecular portraits of human breast tumours*. *Nature*, 2012. **490**(7418): p. 61-70.
27. Holderfield, M., et al., *Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond*. *Nat Rev Cancer*, 2014. **14**(7): p. 455-67.
28. Prahallad, A., et al., *Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR*. *Nature*, 2012. **483**(7387): p. 100-3.
29. Flaherty, K.T., et al., *Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations*. *N Engl J Med*, 2012. **367**(18): p. 1694-703.
30. Swanton, C., *Intratumor heterogeneity: evolution through space and time*. *Cancer Res*, 2012. **72**(19): p. 4875-4882.
31. Scott, A.M., J.D. Wolchok, and L.J. Old, *Antibody therapy of cancer*. *Nat Rev Cancer*, 2012. **12**(4): p. 278-87.
32. Palucka, K. and J. Banchereau, *Cancer immunotherapy via dendritic cells*. *Nat Rev Cancer*, 2012. **12**(4): p. 265-77.



33. Vivier, E., et al., *Targeting natural killer cells and natural killer T cells in cancer*. Nat Rev Immunol, 2012. **12**(4): p. 239-52.
34. Khattak, M., et al., *Targeted therapy and immunotherapy in advanced melanoma: an evolving paradigm*. Ther Adv Med Oncol, 2013. **5**(2): p. 105-18.
35. Pardoll, D.M., *The blockade of immune checkpoints in cancer immunotherapy*. Nat Rev Cancer, 2012. **12**(4): p. 252-64.



# Chapter 8

## English Summary

Patients presenting with ovarian tumors may pose diagnostic challenges. Primary ovarian tumors should be distinguished from metastases to the ovaries from other primary tumors. Treatment choices and statements about prognosis are different for localised or metastatic disease. Diagnostic workup including imaging and histologic examination of tumors is not always successful in deciphering the origin of tumors. Therefore, additional characterisation of tumors at a molecular level might be helpful.

**Chapter 2** describes the observation of an increased number of metastases to the ovaries in case of primary colorectal or duodenal carcinomas in Familial Adenomatous Polyposis (FAP) patients. A pathology database search of cases with metastases to the ovaries resulted in an enrichment of FAP patients. This is surprising as the incidence of FAP is far less than 1% in consecutive CRC cohorts.

**Chapter 3** describes the comparison between the mutational profiles of primary ovarian tumors versus secondary ovarian tumors (metastases). As primary and secondary tumors can be difficult to distinguish from each other on histologic and immunohistochemical stainings, the mutation profiles of both tumor types were investigated. A comparable incidence of mutations in known driver genes was identified in both groups. However, a small number of genes (e.g. *APC* and *CTNNB1*) can be used to differentiate between primary and secondary ovarian tumors.

**Chapter 4** describes the comparison between the mutation profiles of primary colorectal tumors and matching metastases to the ovaries. There is a large overlap in driver variations between primary tumors and metastases. However, a large number of passenger mutations show less overlap. Multiple subclones are present in the primary CRC tumor and only a limited number in the metastases. The time between the detection of primary tumors and the metastases was not related to mutational differences.

**Chapter 5** describes a comparison between different targeting techniques in the workflow of next generation sequencing (NGS). At this moment, only a limited amount of genes is decisive for targeted treatment decisions. Therefore, genes of interest

are captured preceding sequencing. These targeting methods are based on different principles, each with their own pros and cons.

**Chapter 6** describes a patient that was first diagnosed with bilateral primary ovarian tumor. However, almost a year later the patient was treated for a colorectal tumor, causing doubt about the origin of the ovarian tumor. Interestingly, tumors of both sites showed an MSI-H phenotype, normally indicative of a low metastatic potential. Extensive molecular analysis showed a clonal relationship between the two processes, concluding that the patient suffered from metastasized colon cancer to the ovaries.

**Chapter 7** presents the concluding remarks and future perspectives. Next generation sequencing has changed the daily practice in pathology. Molecular profiling can be used to improve primary diagnostics, but is also needed for detection of targetable variants. The fast development of this field will lead to ongoing changes in the daily practice of molecular pathology.

# Nederlandse samenvatting

Patiënten met ovariumtumoren kunnen een diagnostische uitdaging vormen. Primaire ovariumtumoren moeten worden onderscheiden van ovariële uitzaaiingen van andere primaire tumoren. De behandeling en de prognose zijn namelijk verschillend in het geval van een gelokaliseerde of een uitgezaaide ziekte. Ondanks uitgebreid beeldvormend en histologisch onderzoek is het niet altijd mogelijk om de oorsprong van de tumor met zekerheid aan te geven. Aanvullend moleculair onderzoek kan hierbij behulpzaam zijn.

**Hoofdstuk 2** beschrijft een toegenomen aantal ovariële uitzaaiingen afkomstig van een primaire tumor in de dikke darm of de dunne darm indien ontstaan in de context van familiale adenomateuze polyposis (FAP). Een database met patiënten met ovariële metastasen afkomstig van de tractus digestivus toonde een verrijking van het aantal verwachte FAP patiënten.

**Hoofdstuk 3** beschrijft de vergelijking van DNA variant profielen van primaire ovarium tumoren met uitzaaiingen van de dikke darm (secundaire tumoren). Onderzocht werd of aanvullend moleculair onderzoek (detectie van DNA varianten) behulpzaam is bij dit diagnostische probleem. Een beperkte set van genen (bijv. *APC* en *CTNNB1*) kunnen worden gebruikt om een onderscheid te maken tussen primaire ovariumtumoren en uitzaaiingen van dikke darm tumoren.

**Hoofdstuk 4** beschrijft de vergelijking van de mutatieprofielen van primaire colorectale tumoren en de bijbehorende uitzaaiingen naar de eierstokken. Analyse van de DNA profielen geeft inzicht in de ontwikkeling van tumoren en de relatie met de uitzaaiingen. Een grote overeenkomst betreffende de sturende DNA varianten werd waargenomen tussen de primaire tumoren en de uitzaaiingen. Een groot aantal later ontstane varianten toonde een minder grote overeenkomst. De mate waarin de DNA variant profielen tussen de primaire en de uitzaaiingen verschilden was niet gerelateerd aan de tijd die was verstreken tussen het diagnosticeren van de primaire tumor en de metastasen.

**Hoofdstuk 5** beschrijft een vergelijking tussen verschillende manieren van verrijking van genen in de workflow van zogenaamde next generation DNA analyse (NGS). Op dit moment is slechts een beperkt aantal genen bepalend voor therapie keuzes in het kader van patiënt specifieke behandeling. Derhalve worden de genen die van belang zijn voor de behandeling verrijkt voorafgaande aan het sequencen. Al deze verrijkingsmethoden zijn gebaseerd op verschillende principes, elk met hun eigen voor- en nadelen.

**Hoofdstuk 6** beschrijft een patiënte die in eerste instantie werd gediagnosticeerd met tumoren in beide eierstokken. Echter, bijna een jaar later werd patiënte behandeld voor een colontumor. Deze opeenvolging van tumoren veroorzaakte twijfel over de oorsprong van de eierstoktumoren. Uitgebreide moleculaire analyse toonde een klonale relatie tussen de twee tumorprocessen. De uiteindelijke conclusie was dat patiënte leed aan een initieel onontdekte dikke darmtumor die uitzaaide naar de eierstokken. Hoewel eerst een onderliggende erfelijke ziekte de oorzaak leek, bleek dit niet het geval. Een gemetastaseerd coloncarcinoom naar beide ovaria de beste verklaring was voor de presentatie van de patiënte.

**Hoofdstuk 7** beschrijft de conclusie en de meest waarschijnlijke ontwikkelingen in de nabije toekomst. Next generation DNA analyse is onderdeel geworden van de dagelijkse praktijk in de pathologie. De zo gegenereerde moleculaire profielen kunnen worden gebruikt voor het stellen van de primaire diagnose, maar zijn ook nodig voor het bepalen van een effectieve patiënt specifieke behandeling. De snelle ontwikkelingen binnen dit vakgebied zullen leiden tot voortdurende veranderingen in de dagelijkse praktijk van de moleculaire diagnostiek.

## List of publications

Crobach S, Jansen AML, Ligtenberg MJL, Koopmans M, Nielsen M, Hes FJ, Wijnen JT, Dinjens WNM, van Wezel T, Morreau H.

Excluding Lynch syndrome in a female patient with metachronous DNA mismatch repair deficient colon- and ovarian cancer.

*Fam Cancer*. 2017 Nov 9. doi: 10.1007/s10689-017-0055-1. [Epub ahead of print]

Stephanie A Schubert, Dina Ruano, Fadwa A Elsayed, Arnoud Boot, Stijn Crobach, Arantza Farina Sarasqueta, Bruce Wolffenbuttel, Melanie M van der Klauw, Jan Oosting, Carli M Tops, Ronald van Eijk, Hans FA Vasen, Rolf HAM Vossen, Maartje Nielsen, Sergi Castellví-Bel, Clara Ruiz-Ponte, Ian Tomlinson, Malcolm G Dunlop, Pavel Vodicka, Juul T Wijnen, Frederik J Hes, Hans Morreau, Noel FCC de Miranda1 Rolf H Sijmons and Tom van Wezel.

Evidence for genetic association between chromosome 1q loci and predisposition to colorectal neoplasia.

*Br J Cancer*. 2017 Sep 5;117(6):876-884.

Voorzaat BM, van Schaik J, Crobach S, van Rijswijk CS, Rotmans JI.

Alpha-1 Antitrypsin Deficiency Presenting with MPO-ANCA Associated Vasculitis and Aortic Dissection.

*Case Rep Med*. 2017;2017:8140641. doi: 10.1155/2017/8140641. Epub 2017 Mar 6.

Jansen AM, Crobach S, Geurts-Giele WR, van den Akker BE, Garcia MV, Ruano D, Nielsen M, Tops CM, Wijnen JT, Hes FJ, van Wezel T, Dinjens WN, Morreau H. Distinct Patterns of Somatic Mosaicism in the APC Gene in Neoplasms From Patients With Unexplained Adenomatous Polyposis.

*Gastroenterology*. 2017 Feb;152(3):546-549

Crobach S, Ruano D, van Eijk R, Schrupf M, PALGA group, Fleuren G, van Wezel T and Morreau H. Somatic mutation profiles in primary colorectal cancers and matching ovarian metastases: Identification of driver and passenger mutations

*The Journal of Pathology: Clinical Research*. 2016 Apr 15;2(3):166-74

Berden AE, Beeres SL, Crobach S, Schaliij MJ, Rabelink TJ, Teng YK. Azathioprine-induced eosinophilic myocarditis in a patient with ANCA-associated vasculitis.

*Clin Exp Rheumatol*. 2016 May-Jun;34(3 Suppl 97):S146.

Boot A, van Eendenburg J, Crobach S, Ruano D, Speetjens F, Calame J, Oosting J, Morreau H, van Wezel T. Characterization of novel low passage primary and metastatic colorectal cancer cell lines.

*Oncotarget*. 2016 Feb 15. Mar 22;7(12):14499-509.

Hennink SD<sup>1</sup>, van der Meulen-de Jong AE<sup>1</sup>, Wolterbeek R<sup>1</sup>, Crobach S<sup>1</sup>, Becx MC<sup>1</sup>, Crobach WF<sup>1</sup>, van Haastert M<sup>1</sup>, Ten Hove WR<sup>1</sup>, Kleibeuker JH<sup>1</sup>, Meijssen MA<sup>1</sup>, Naggengast FM<sup>1</sup>, Rijk MC<sup>1</sup>, Salemans JM<sup>1</sup>, Stronkhorst A<sup>1</sup>, Tuynman HA<sup>1</sup>, Vecht J<sup>1</sup>, Verhulst ML<sup>1</sup>, de Vos Tot Nederveen Cappel WH<sup>1</sup>, Walinga H<sup>1</sup>, Weinhardt OK<sup>1</sup>, Westerveld D<sup>1</sup>, Witte AM<sup>1</sup>, Wolters HJ<sup>1</sup>, Cats A<sup>1</sup>, Veenendaal RA<sup>1</sup>, Morreau H<sup>1</sup>, Vasen HF<sup>2</sup>. Randomized Comparison of Surveillance Intervals in Familial Colorectal Cancer. *J Clin Oncol*. 2015 Dec 10;33(35):4188-93.

Crobach S and Morreau H. Diagnostically challenging cases: distinguishing primary from secondary ovarian tumours.

*Clinical laboratory international*. June 2015, page 36-38.

Crobach S, Ruano D, van Eijk R, Fleuren GJ, Minderhout I, Snowdowne R, Tops C, van Wezel T, Morreau H. Target-enriched next-generation sequencing reveals differences between primary and secondary ovarian tumors in formalin-fixed, paraffin-embedded tissue.

*J Mol Diagn*. 2015 Mar;17(2):193-200.

Spaans VM, Trietsch MD, Crobach S, Stelloo E, Kremer D, Osse EM, Haar NT, van Eijk R, Muller S, van Wezel T, Trimbos JB, Bosse T, Smit VT, Fleuren GJ. Designing a high-throughput somatic mutation profiling panel specifically for gynaecological cancers.

*PLoS One*. 2014 Mar 26;9(3):e93451. doi: 10.1371/journal.pone.0093451. eCollection 2014.

Crobach S, van Wezel T, Vasen HF, Morreau H. Ovarian metastases of colorectal and duodenal cancer in familial adenomatous polyposis.

*Fam Cancer*. 2012 Dec;11(4):671-3.

Vroegindewij E, Crobach S, Itoi M, Satoh R, Zuklys S, Happe C, Germeraad WT, Cornelissen JJ, Cupedo T, Holländer GA, Kawamoto H, van Ewijk W. Thymic cysts originate from Foxn1 positive thymic medullary epithelium.

*Mol Immunol*. 2010 Feb;47(5):1106-13.



## Curriculum vitae

Stijn Crobach werd geboren op 18 december 1983 te Maastricht. Hij groeide op in Melick en doorliep zijn middelbare school aan het Bisschoppelijk College Schöndeln te Roermond. In 2002 werd begonnen met de studie Geneeskunde aan de Rijksuniversiteit Leiden. In 2008 behaalde hij zijn artsexamen. Na een jaar op de onderzoeksafdeling Hematologie van het Erasmus Medisch Centrum te hebben gewerkt, is hij vanaf 2009 werkzaam als AIOSKO (arts in opleiding to specialist en klinisch onderzoeker) op de afdeling Pathologie van het Leids Universitair Medisch Centrum. Het promotie onderzoek werd begeleid door prof. dr. Morreau en dr. Tom van Wezel. Hij is getrouwd met Marjolein Wijngaarden, heeft een dochter Vera en een zoon Simon.

# Dankwoord

Velen hebben bijgedragen bij de totstandkoming van dit werk. Een aantal mensen zou ik speciaal willen bedanken.

Beste Hans, dank voor het jarenlange vertrouwen in mij. Een korte discussie met jou leert mij dat er nog altijd genoeg te leren valt! Beste Tom, dank voor het leveren van technische kennis, biologische knowhow en relativerende capaciteiten. Ook wil ik graag Gert Jan bedanken die bij het eerste begin van dit werk betrokken is geweest. Beste Dina, dank voor alle 0'en en 1'en op jouw beeldscherm die uiteindelijk leidden tot een overzichtelijk mutatie lijstje.

Beste Ronald, Melanie, Brendy, Ronelle, Ivonne en Carli, dank voor de praktische hulp bij het uitvoeren van de experimenten. Ook dank aan de studenten (Nesli en Monique) die geholpen hebben bij de uitvoering.

Beste Willem, dank dat je me op het vliegtuig naar Japan hebt gezet om te kunnen ervaren hoe leuk onderzoek is.

Daarnaast zou ik graag alle kamergenoten uit P1-39, aios, pathologen en andere medewerkers van de afdeling pathologie willen bedanken voor hun hulp en betrokkenheid.

Tevens dank aan mijn lieve familie en schoonfamilie.

Lieve pap en mam, dank voor alles wat jullie mij hebben meegegeven en voor alles dat jullie mogelijk hebben gemaakt. Zonder jullie steun was dit allemaal niet gelukt.

Lieve (dr.) Marjolein, eindelijk weer een gesprek op hetzelfde niveau... Het is heerlijk om met jou en de kinderen door het leven te gaan. Samen met jou ga ik vol vertrouwen de toekomst tegemoet.