CHAPTER 1

Introduction
GENERAL BACKGROUND

1.1. Lynch syndrome

Clinical definition
Lynch syndrome is an autosomal dominant inherited cancer-susceptibility syndrome. It is named after Dr. Henry Lynch, whose role was crucial in the clinical and scientific identification of the syndrome as an inherited and relatively frequent cause of colorectal and extra-colonic cancer. Lynch syndrome is also known as hereditary nonpolyposis colorectal cancer (HNPCC), a rather misleading name, as several cancers other than colorectal also belong to the disease spectrum. Lynch syndrome is the most common hereditary cause of colorectal cancer.

Long before the genetic mechanism underlying the disease was known, several major clinical features were described, and a first attempt to define a uniform set of minimal criteria for clinical diagnosis of Lynch syndrome based on family history was made in 1990, in a meeting of the International Collaborative Group on HNPCC (ICG-HNPCC), in Amsterdam (Vasen et al., 1991). These became known as the Amsterdam criteria (I). With time, and according to the new findings in the field, especially those related to the genetic basis of the disease, several refinements to this set of criteria were suggested, such as the Japanese, Mount Sinai, and Bethesda criteria (Fujita et al., 1996; Peltomaki et al., 2004; Umar et al., 2004). In 1999, the ICG-HNPCC proposed a new definition for HNPCC/Lynch syndrome and, with it, the revised Amsterdam criteria (II) (Vasen, 1999). (See Box.1)

Genetics of Lynch syndrome
The start of unravelling the genetic cause of Lynch syndrome was in 1993, when two major findings came together. One was the report of genetic instability associated with replication errors in microsatellite sequences in a large percentage of tumours from Lynch syndrome patients (Aaltonen et al., 1993; Ionov et al., 1993; Peltomaki et al., 1993a; Thibodeau et al., 1993). The other was the identification of two Lynch syndrome loci by linkage analysis, at chromosomes 2p and 3p (Lindblom et al., 1993; Peltomaki et al., 1993b).
Box 1. Details of the Amsterdam criteria for identifying Lynch syndrome families and the definition of the syndrome, by the 1999 International Collaborative Group on HNPCC (ICG-HNPCC) (Vasen et al., 1999).

**Amsterdam criteria II**

- At least three relatives should have histologically verified colorectal cancer, cancer of the endometrium, small bowel, ureter, or renal pelvis;
- One of them should be a first-degree relative of the other two;
- Familial adenomatous polyposis (FAP) should be excluded;
- At least two successive generations should be affected;
- In one of the relatives colorectal cancer should be diagnosed before 50 years of age.

**ICG-HNPCC definition of HNPCC/Lynch syndrome**

- Familial clustering of colorectal and/or endometrial cancer;
- Associated extra-colonic cancers: cancer of the stomach, ovary, ureter/renal pelvis, brain, small bowel, hepatobiliary tract, and skin (sebaceous tumours);
- Development of cancer at an early age;
- Development of multiple cancers;
- Features of colorectal cancer: (1) predilection for proximal colon; (2) improved survival; (3) multiple primary (synchronous/metachronous) colorectal cancers; (4) increased proportion of mucinous tumours, poorly differentiated tumours, and tumours with marked host-lymphocytic infiltration and lymphoid aggregation at the tumour margin;
- Features of colorectal adenoma: (1) the numbers vary from one to a few; (2) increased proportion of adenomas with a villous growth pattern and (3) probably rapid progression from adenoma to carcinoma;
- High frequency of MSI (MSI-H);
- Immunohistochemistry: loss of hMLH1, hMSH2, or hMSH6 protein expression;
- Germline mutation in MMR genes (hMSH2, hMLH1, hMSH3, hMSH6, hPMS1,hPMS2).

During 1994, the first germline mutations were found in two genes identified in those loci (\textit{MSH2} and \textit{MLH1}), both being human homologues of the well-known \textit{mutS} and \textit{mutL} mismatch repair (MMR) genes of bacteria and yeast. Thus, deficient DNA mismatch repair was identified as the cause of Lynch syndrome. This functional inactivation of the DNA MMR genes is due to germline mutations as the first hit (Fishel \textit{et al.}, 1993; Leach \textit{et al.}, 1993; Bronner \textit{et al.}, 1994; Papadopoulos \textit{et al.}, 1994), followed by somatic inactivation of the second allele as
the second hit (Hemminki et al., 1994; Lu et al., 1996). This second hit is usually a somatic mutation or loss of heterozygosity (LOH).

Germline mutations in the MLH1 and MSH2 genes form the vast majority of mutations found in Lynch syndrome cases (Peltomaki et al., 2004). Two other MMR genes – MSH6 and PMS2 – were later reported as being involved in the disease as well, since germline mutations are also found in a fraction of Lynch syndrome families (Berends et al., 2002; Hendriks et al., 2006). Germline deletion of the 3' exons of TACSTD1 can cause heritable somatic methylation and inactivation of the neighbouring MSH2 gene and thus Lynch syndrome (Kovacs et al., 2009; Ligtenberg et al., 2009). Also an interstitial deletion at 3p21.3 resulting in the genetic fusion of MLH1 and ITGA9 has been recently reported in a Lynch syndrome family, presumably defining a novel subclass of Lynch syndrome patients (Meyer et al., 2009). Several other genes, such as MLH3 and EXO1, also belong to the MMR pathway and these were therefore screened over the years as well. Germline mutations in MLH3 and EXO1 have been found (Wu et al., 2001a&b), but due to their low frequencies and type, mostly missense, they are not considered to be major players in Lynch syndrome (Hienonen et al., 2003; Jagmohan-Changur et al., 2003; Ou et al., 2008).

1.2. Mismatch repair and microsatellite instability

The MMR system is responsible for correcting errors that escape the activity of the polymerases during DNA replication. The system is able to correct mispaired nucleotides, as well as insertions and deletions loops (IDLs) that typically occur at short DNA tandem repeats - microsatellites. Therefore, when an MMR protein is inactivated, mutations will accumulate in those repeat sequences at a much higher rate (100- to 1000-fold) than that of spontaneous mutations in normal cells (Shibata et al., 1994). This phenomenon is referred to as “microsatellite instability” (MSI) (Ionov et al., 1993). It is easily recognized by decreased or increased lengths of the microsatellite, and therefore the detection of MSI became a key technique when searching for MMR-deficient tumours.

MSI was reported in Lynch syndrome patients in 1993, occurring in over 90% of tumours in those patients, and it was another important piece of the puzzle.
linking MMR deficiency and Lynch syndrome (Aaltonen et al., 1993; Ionov et al., 1993; Peltomaki et al., 1993a; Thibodeau et al., 1993). However, it is also found in a large proportion (15-25%) of sporadic tumours, not only of colorectal origin, but also in gastric and endometrial carcinomas (Boland et al., 1998). The underlying mechanism characterizing the sporadic forms of MSI tumours is also the functional inactivation of an MMR gene, namely MLH1, but in this case the bi-allelic inactivation of the gene is typically due to somatic promoter hypermethylation. MSI is a very early event in the tumorigenic process of tumours with MMR problems, as it has been detected in early lesions such as colorectal adenomas (Giuffrè et al., 2005).

1.3. Adenoma-carcinoma sequence

The adenoma-carcinoma sequence of colorectal cancer represents one of the best-known models of cancer development. Colorectal carcinomas arise through a multistep process, starting from early to high-grade dysplastic adenomas to carcinomas. This process of cancer development is basically caused by the progressive accumulation of genetic alterations in genes involved in cell growth, differentiation, proliferation, and apoptosis (Fearon & Vogelstein, 1990).

This accumulation of genetic alterations is thought to be due to genetic instability, in which several distinct forms can be distinguished. Those that are best-described are chromosomal instability (CIN) and microsatellite instability (MIN or MSI) (Royrvik et al., 2007). CIN is characterized by widespread chromosomal abnormalities such as aneuploidy and frequent loss of heterozygosity (LOH). MSI is caused by defects in the DNA mismatch repair (MMR) pathway, and is characterized by the accumulation of mutations in microsatellites (see above).

MSI is found in the very early stages of the adenoma carcinoma sequence, although generally in lower frequencies than in carcinomas. It is reported in about 1-2% of sporadic adenomas (Young et al., 1993; Iino et al., 1999; Loukola et al., 1999; Sugai et al., 2003) and in 10-90% of Lynch syndrome-associated adenomas (Aaltonen et al., 1994; Iino et al., 2000; Giuffrè et al., 2005). This wide range of MSI frequencies might be explained by the method of dissection used (laser microdissection vs. manual dissection) and it is related to the multi-clonality of the
tissue, i.e. different areas show different degrees of MSI and different degrees of
dysplasia (de Wind et al., 1998; Iino et al., 1999, 2000; Giuffrè et al. 2005;
Greenspan et al., 2007). In addition, there may be considerable variation in the
methods used to score MSI by different laboratories and between different
observers.

1.4. MSI detection

An international consensus panel of five microsatellite markers for detecting MSI
was proposed in 1997 (Boland et al., 1998) to facilitate the production of easily
comparable results, and this has become widely used. The panel includes two
mononucleotide markers (BAT-25 and BAT-26) and three dinucleotide markers
(D2S123, D5S346 and D17S250). Samples that are unstable for two or more of
these markers are designated MSI-high (MSI-H), while samples unstable for one
marker are MSI-low (MSI-L); samples that are stable for all the markers are
designated microsatellite stable (MSS). If it is necessary to distinguish between
MSI-L and MSS, then additional markers should be used (Boland et al., 1998). It is,
in fact, common that some labs use a different number of markers. In that case, a
sample is MSI-H if it is unstable for more than 30% of the markers used. More
recently, a pentaplex PCR assay for 5 mononucleotide markers was proposed
(Buhard et al., 2004). It includes the following markers: BAT-25, BAT-26, NR-21,
NR-22 and NR-24. The authors claim a sensitivity and specificity of 100%, and
suggest that the use of quasi-monomorphic mononucleotide repeats over
dinucleotide repeats is advantageous, as the latter are typically polymorphic and
more difficult to interpret. It is also believed that there is a greater sensitivity of
dinucleotides for MSI-L cases than for MSI-H cases (Hatch et al., 2005). In addition,
the use of mononucleotide markers might avoid needing normal tissue for
comparison in CRC cases. In fact, it has also been proposed that BAT-26 alone
and without normal matching mucosa might be sufficient for detecting MSI-H CRC
(Hoang et al., 1997; de la Chapelle, 1999). The above-mentioned pentaplex panel
has also been advised for endometrial carcinomas, although normal matching
mucosa DNA is in that case still recommended (Wong et al., 2006).
1.5. Target genes and tissue selection

Microsatellites are short DNA tandem repeat sequences spread throughout the genome, including non-coding and coding regions of genes. When MSI occurs in high frequencies in a coding sequence of a gene with important regulatory functions (involved in processes like apoptosis or proliferation for example), it is believed that such a gene, when impaired, contributes to development of cancer. These genes are generally called “target genes”. This is a rather simplistic definition; however, it has been controversial to agree on the criteria to define a real target gene (Woerner et al., 2001, 2003; Duval & Hamelin, 2002; Perucho, 2003). Due to the general lack of functional studies proving the true involvement of target genes in tumour development, these genes are generally classified as such based on a high mutation frequency. One major problem is the establishment of a valid cut-off value for the mutation frequency to separate real target genes from passengers or bystanders (those having the background mutations expected in an MMR-deficient context but not related with the progression to cancer). In 2002, Duval et al. proposed a cut-off frequency value of 10-15% and this has been used by other groups (Vilkki et al., 2002).

A number of target genes have been identified in MMR-deficient tumours, and these are thought to be key players in MSI-H tumorigenesis. Mutations were mostly searched for in MSI-H colorectal tumours, although now many of the genes have been screened in endometrial and gastric tumours as well (Duval et al., 1999; Schwartz et al., 1999; Duval et al., 2001; Vilkki et al., 2002; Royrvik et al., 2007). From several studies it became clear that there are target genes, such as BAX, commonly involved in MSI-H tumours of diverse origin, whereas others show considerable qualitative and quantitative differences between different tumour types, probably due to tissue-specific selection (Myeroff et al., 1995; Duval et al., 1999; Gurin et al., 1999; Schwartz et al., 1999; Semba et al., 2000; Duval et al., 2002a). It is also clear from these studies that more important genes remain to be found, especially in endometrial cancer. These tumours are subjected to less screening than colorectal ones, and only a few genes with a high mutation frequency have been found in them.
1.6. Endometrium

1.6.1. Histology and functional changes

The uterus is a hollow muscular, pear-shaped organ weighing 40-80 gram in a nonpregnant woman. The size of the uterus is highly variable as is demonstrated during pregnancy. There are two parts to the uterus: the main body, known as the corpus, and the lower part, which opens into the vagina, called the cervix. The wall of the uterus consists of three layers: different types of mucosa at the inner side; a thick muscular, highly vascularised part; and a thin layer of serosa covering the intraperitoneal part of the corpus. The cervical canal is covered by a single layer of cylindrical mucus secreting cells which extends into the underlying myocervix forming endocervical crypts. The inner lining of the corpus is called endometrium. The endometrium consists of a supportive stroma and an epithelial component the endometrial glands. The thickness and differentiation of the functional layer of the endometrium is highly regulated by the hormonal changes occurring during the menstrual cycle. The endometrial mucosa can be sub-divided in two areas related to those changes: a functional layer, adjacent to the cavity of the uterus, that is sloughed during menstruation and built up afterwards, and a basal inert layer, adjacent to the myometrium, that is not shed during the menstrual cycle and that functions as a regenerative zone for the functional layer. After menopause, in a low estrogenic situation, the endometrium consists of the basal layer only.

1.6.2. Endometrial cancer

Aetiology

Endometrial cancer (EC) is one of the most common types of gynaecological cancer in women worldwide. The highest incidence is found in North America, although the highest levels of mortality are in Eastern Europe. The incidence of endometrial cancer increases after menopause; approximately 75% of cases are diagnosed in postmenopausal women (Cancer Research UK website).

The major risk factor for endometrial cancer is the high, unopposed exposure to oestrogens (Sherman, 2000; Amant et al., 2005). Therefore, conditions
increasing the oestrogen levels are considered to increase the disease risk. These include for instance: early menarche, late menopause, nulliparity or low parity, and hormone replacement therapy (HRT) with exogenous oestrogen but without progesterone. Long-term use of tamoxifen, a drug used to treat breast cancer, also increases the risk for endometrial cancer (Polin & Ascher, 2008). Other risk factors for the disease include: a high-fat diet, obesity, hypertension, diabetes, age (more common after age of 50), personal history of breast, colorectal, or ovarian cancer and a family history of endometrial cancer or colon cancer (Lynch syndrome). Endometrial cancer risk has also been suggested to be increased in Cowden syndrome, caused by germline PTEN mutations. The use of oral combined contraceptives, on the other hand, is reported to reduce the risk of EC.

**Histopathological and molecular types of endometrial carcinomas**

Endometrial carcinomas are usually divided into two major groups that have different clinical and histological characteristics, as well as molecular differences (Emons et al., 2000; Lax et al., 2004; Ryan et al., 2005).

**Type I or oestrogen-dependent endometrioid carcinomas (EEC)**

Representing 80% of sporadic cases, this is the group of oestrogen-related tumours. They occur in both pre- and post-menopausal women, and their architectural features resemble endometrial glands. Tumours of this type are usually well differentiated (low grade) and confined to the uterus (low stage) and therefore the patient generally has a good prognosis. They are frequently preceded by endometrial hyperplasia.

**Type II or non-oestrogen-dependent ECs**

The tumours belonging to this group are unrelated to oestrogenic stimulation, and mainly occur in post-menopausal women. They display a more aggressive behaviour and poor prognosis. Frequently, by the time of diagnosis, the tumour has already spread outside the uterus. They are not usually preceded by hyperplasia, but originate from an atrophic endometrium instead. They are high-grade tumours with serous or clear-cell morphology.
In addition to the histopathological differences referred to above, there are also genetic differences between these two categories of endometrial carcinomas (Doll et al., 2007). In type I EC we can basically find mutations in PTEN (35-50%), K-Ras (15-30%) and β-catenin (25-40%) genes, and MMR defects, detected by high levels of MSI (20-40%). These characteristics are rarely seen in type II EC, which are characterized by high mutation frequencies on P53 gene (90%) and alterations on HER2/NEU and CDH1. The main type of genetic instability in this group is chromosomal instability (CIN), being aneuploidy and loss-of-heterozygosity (LOH) typical of EC type II. MSI (MIN) is extremely rare in these tumours (Emons et al., 2000; Lax et al., 2004; Ryan et al., 2005). Figure 1 shows the progression model of endometrial cancer proposed by Ryan et al. (2005).

**Figure 1.** Progression model of endometrial cancers type I and type II progression adapted from Ryan et al. (2005).
AIM AND OUTLINE OF THE THESIS

The main focus of this thesis was to understand the development of tumours that follow the MSI pathway. The study covered Lynch syndrome-associated tumours, with particular emphasis on colorectal and endometrial carcinomas and their sporadic counterparts.

Chapter 1 reviews the general background to Lynch syndrome, microsatellite instability, and the hereditary and sporadic cancers associated with this pathway.

Chapter 2 addresses how instability evolves along the adenoma-carcinoma sequence of colorectal cancer, and whether we are able to establish different profiles of MSI for hereditary and sporadic adenomas and carcinomas. Knowledge on this process might be helpful in understanding tumour development and in identifying Lynch syndrome patients in an easier and more specific way.

In chapter 3, we report on our comparison of the frequencies of instability of different types of microsatellites between colorectal and endometrial MSI-H tumours. In addition, we analyze features such as type (deletions/insertions) and size of microsatellite mutation for possible correlations with tissue specificity.

Chapter 4 describes our hunt for new genes involved in MSI-H endometrial tumours. It addresses the instability of mononucleotide repeats occurring in coding sequences. The aim of this work was to identify novel target genes that could explain MSI-H endometrial tumour development, and to unravel molecular pathways related to this type of cancer. We further wanted to determine whether the identified genes were also involved in colorectal and gastric tumours, and we speculate about the functional role of the proteins that are encoded by the genes we found mutated.

In chapter 5 we review and try to clarify the mechanisms linking hormones to cancer, and in particular how hormones can play a role in MSI tumorigenesis.

Finally, in chapter 6, the major findings of this project are discussed, conclusions are drawn and future perspectives are formulated.
REFERENCES


