

# Microsatellite Instability and Intestinal Tumorigenesis

Meena A. Prasad and Barbara Jung

**Abstract** Microsatellites are simple, repetitive sequences of DNA present throughout the genome prone to DNA replication errors. The mismatch repair machinery is crucial to maintaining genetic integrity. When mismatch repair is defective, microsatellite instability occurs, resulting in deregulated tissue growth and tumorigenesis. Although defunct mismatch repair proteins can occur sporadically, these defects are the hallmark of Lynch syndrome, a hereditary colon cancer syndrome due to germline mutations in the mismatch repair genes *MLH1*, *MSH2*, *MSH6* or *PMS2*. Loss of mismatch repair protein function results in an intestinal and extra-intestinal tumor phenotype. This chapter reviews the mechanisms involved in the mismatch repair process, mouse models of defective mismatch repair, genetic versus sporadic microsatellite instability, methods of testing for microsatellite instability, and the clinical implications of microsatellite instability in the large bowel, small bowel and stomach.

**Keywords** Colorectal cancer · Gastrointestinal cancer · Intestinal tumorigenesis · Microsatellite instability · Mismatch repair · Lynch syndrome

## 1 Microsatellite Instability and Mismatch Repair

### 1.1 Microsatellites

Microsatellites are simple, repetitive sequences of DNA that are found throughout the genome and are prone to DNA replication errors [1]. They are typically composed of 1–6 base pair units that may be repeated up to 100 times. Given their

---

B. Jung (✉)

Department of Medicine, Division of Gastroenterology and Hepatology, University of Illinois at Chicago, 840 South Wood Street, 738A CSB, Chicago, IL 60612, USA  
e-mail: [bjung@uic.edu](mailto:bjung@uic.edu)

M. A. Prasad

Division of Gastroenterology and Hepatology, Northwestern University Feinberg School of Medicine, 676 N. St Clair Street, Suite 1400, Chicago, IL 60611, USA  
e-mail: [MeenaAPrasad@gmail.com](mailto:MeenaAPrasad@gmail.com)

repetitive nature, they are prone to errors that can occur during DNA replication [2]. The DNA mismatch repair (MMR) system recognizes and repairs base-pair mismatches that occur during DNA replication.

## 1.2 Mismatch Repair

MMR is a complex process dependent upon MMR proteins and multiple other proteins that interact with DNA. During DNA replication, DNA polymerase is prone to errors in incorporating the correct number of bases in long repetitive DNA sequences, such as microsatellites. These errors occur due to slippage during replication of such a repetitive sequence [3]. MMR was first recognized in yeast with the key genes involved termed *Mut S* homologue (*MSH*), *Mut L* homologue (*MLH*) and *post-meiotic segregation* (*PMS*) and numbered *MSH1* through *MSH6*, *MLH1* through *MLH3* and *PMS1* [3]. Some of the key proteins involved in human MMR include MutL homolog 1 (*MLH1*), MutS homolog 2 (*MSH2*), MutS homolog 6 (*MSH6*), and postmeiotic segregation increased 2 (*PMS2*).

In mammals, there are three different MutS homologs (*MSH2*, *MSH3* and *MSH6*) and four different MutL homologs (*MLH1*, *PMS1*, *PMS2*, *MLH3*) [4]. MMR involves formation of heterodimers of key MMR proteins. The *MSH2* and *MSH6* proteins associate as a heterodimer (termed MutS $\alpha$ ), which then constitutes a sliding clamp on the newly synthesized strand of DNA and recognizes base-base mispairs as well as 1 base pair insertion/deletion mutations (see Fig. 1) [5]. *MSH2* also forms a heterodimer with *MSH3* (termed MutS $\beta$ ), which recognizes 1–4 base pair insertion/deletion mutations. When *MSH2* heterodimer recognizes a DNA base pair mismatch, it recruits the *MLH1*-*PMS2* heterodimer (termed MutL $\alpha$ ) to assist in repairing the mismatch [4]. The interactions of the *MLH1*-*PMS2* heterodimer with MutS $\alpha$  (*MSH2*-*MSH6*) and MutS $\beta$  (*MSH2*-*MSH3*) are critical for activation of subsequent MMR steps, including excision of the DNA strand with the mismatched base(s) and its resynthesis [4]. The precise mechanism of repair remains unknown [6]. Of note, MutL homolog 3 (*MLH3*) also forms a heterodimer with *MLH1* (termed MutL $\gamma$ ) and participates in mismatch repair. In cases where a loss of *PMS2* is present, function can be partially compensated by *MLH3* [3]. However, loss of *MLH3* has not been implicated in microsatellite instability and clinical pathology. *MLH1* also forms a heterodimer with *PMS1* (termed MutL $\beta$ ), but has little or no role in mismatch repair [7].

If MMR proteins (specifically *MSH2*, *MSH6*, *MLH1* and *PMS2*) do not correct errors from slippage during replication, a temporary insertion-deletion loop is created. This may result in microsatellite instability manifested by frame shift mutation and possible downstream nonsense mutation with production of a truncated, nonfunctional protein if the mutation is located in a coding region [3].

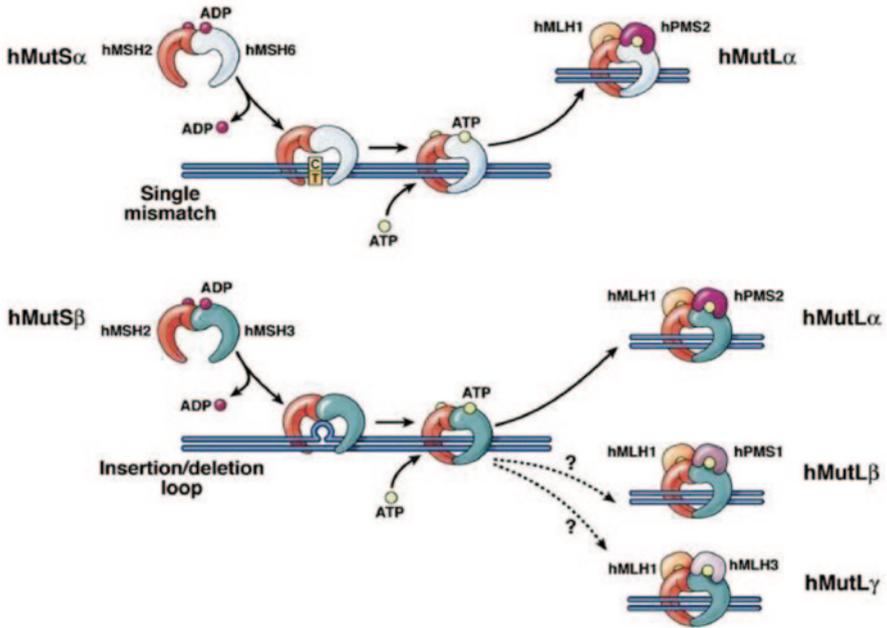


Fig. 1 The DNA MMR system

### 1.3 Microsatellite Instability

Microsatellite instability (MSI) is defined as a change of any length of DNA due to either insertion or deletion of repeating units in a microsatellite within a tumor when compared to normal tissue [2]. Typically MSI manifests itself in frame shifts that the defective MMR system is unable to correct [8]. MSI tumors rarely have large-scale genomic alterations but rather harbor mutations in coding microsatellite sequences in *TGFBR2*, *BAX*, *BRAF* and *ACVR2* [9]. Mutations in *APC*, *KRAS*, *P53* and *PIK3CA* are less common in MSI tumors [10].

## 2 Mouse Models of Mismatch Repair

Mouse models with inactivating mutations in MMR genes have allowed the methodical study of the individual function of each gene and their respective contribution to cancer susceptibility. In addition, specific loss of function of a gene may approximate the clinical phenotype. However, unlike Lynch syndrome (LS) patients with genetic defects in MMR, heterozygous MMR mutant mice do not develop early onset tumors. This may be due to the shorter life span of mice associated with a lower likelihood of a second hit induced somatic loss of the wild-type

MMR allele leading to tumor development. Only homozygous MMR mutant mice are prone to developing early tumors including gastrointestinal cancer, skin cancer and lymphoma. One limitation of mouse models to study MMR is that due to the development of aggressive lymphomas, the study of spontaneous GI tumors is difficult. Patients with LS (and monoallelic mutations) do not typically develop early lymphoma, but the rare patients with biallelic mutations in *MSH2*, *MSH6*, *MLH1* and *PMS2* have a severely reduced life span due to hematologic, brain and gastrointestinal malignancies similar to the knockout mice [11]. Interestingly, MMR deficient mice do not specifically develop colorectal cancer (CRC), but predominantly small bowel tumors [4].

## 2.1 MutS Homolog Deficient Mice

Of the three MutS homologs (*MSH2*, *MSH6*, *MHS3*), only deficiency in *MSH2* and *MSH6* demonstrated a strong tumor predisposition phenotype (summarized in Table 1).

In particular, the phenotypes are as follows:

## 2.2 *MSH2*

*Msh2* deficient mice display severely reduced survival with 50% mortality by 6 months and 100% mortality by 11–12 months. Phenotypically, these mice have a strong predisposition to malignancy with lymphomas observed in 80% and intestinal neoplasms observed in 70% of knockout mice 6 months or older and skin neoplasms (sebaceous gland tumors and squamous cell carcinoma) in 7% [12]. MSI is more common in carcinomas than in adenomas, but uncommon in normal tissue. Cell extracts from these mice are unable to repair single-base mismatches or insertion/deletion loops of one to four nucleotides, supporting the

**Table 1** Mouse lines with MutS MMR gene mutations

Genotype	Fertility M/F	Mononucleotide MSI panel	Dinucleotide MSI panel	Tumor incidence	Tumor spectrum
<i>Msh2</i> <sup>-/-</sup>	+/+	High	High	High	Lymphoma, GI, skin, other
<i>Msh3</i> <sup>-/-</sup>	+/+	Moderate	High	High	GI tumors
<i>Msh6</i> <sup>-/-</sup>	+/+	None	Low	High	Lymphoma, GI, other
<i>Msh3</i> <sup>-/-</sup> <i>Msh6</i> <sup>-/-</sup>	+/+	High	High	High	Lymphoma, GI, skin, other
<i>Msh4</i> <sup>-/-</sup>	-/-	N/A	N/A	None	None
<i>Msh5</i> <sup>-/-</sup>	-/-	N/A	N/A	None	None

concept that loss of *Msh2* in mice results in complete MMR deficiency. Clinically, LS patients with *MSH2* deficiency have earlier onset and more aggressive malignancy phenotype.

Due to the aggressive nature of lymphomas affecting *Msh2* deficient mice, the study of GI tumors is challenging as they often do not live long enough to develop GI tumors. To avoid early death caused by tumorigenesis in extraintestinal organs, conditional *Msh2* knockout mice with intestine specific gene inactivation have been generated. These mice carry an *Msh2<sup>loxP</sup>* allele that results in somatic deletion of *Msh2* in the intestinal epithelium by expression of the *Villin-Cre* recombinase transgene. Tumorigenesis is isolated to the small intestine and formation of 1–2 adenomas and/or adenocarcinomas is observed within the first year of life. This *Msh2*-deficient model, while slower, provides a more useful means of studying intestinal tumorigenesis [13].

### 2.3 *MSH6*

Knockout of *Msh6* in mice led to the search for and later implication of *Msh6* as a late-onset cancer susceptibility gene. Phenotypically, these mice have a tumor spectrum phenotype similar to *Msh2* deficient mice, but the onset of tumor development is delayed and the mice survive up to 16 months. Molecular analysis of these mice revealed a partial repair defect with defective repair of single nucleotide mismatches but intact repair of 1, 2 and 4 nucleotide insertion/deletion mismatches [14]. As a result, these mice accumulate base substitutions rather than the frame shift mutations that occur in the absence of *Msh2*. Due to preservation of insertion/deletion mismatches, the tumors in *Msh6* deficient mice do not display MSI phenotype. In concordance with the *Msh6* deficient mouse model, LS patients with germ line mutation in *MSH6* develop GI cancers at age 60 and older with variable MSI phenotypes [15]. Interestingly, both the female carriers of *MSH6* germ line mutation and mice with global *Msh6* deficiency develop endometrial cancer [16].

### 2.4 *MSH3*

*Msh3* deficient mice neither have a reduced survival or cancer phenotype, likely due to only moderate defects in MMR. These mice do develop tumors later in life, but the incidence is comparable to wild-type mice. The cell extracts from these mice show impaired repair of insertion/deletion mutations of 1–4 bases, but efficient repair of single-base substitution mutations, likely because of functional Msh2-Msh6 complexes. However, when *Msh3* deficiency is combined with *Msh6* deficiency via crosses, there is a complete loss of repair activity in cell extracts and the tumor predisposition phenotype is similar to that in *Msh2* deficiency. This suggests a functional redundancy between *MSH2-MSH6* and *MSH2-MSH3* complexes and indicates that *MSH3* is not implicated in LS [4, 17].

## 2.5 *MutL Homolog Deficient Mice*

Of the four different MutL homologs (MLH1, PMS1, PMS2, MLH3), only mice deficient in *Mlh1* or *Pms2*, which together form the MutL $\alpha$  heterodimer, demonstrate a strong tumor predisposition phenotype (summarized in Table 2).

## 2.6 *MLH1*

Mice deficient in *Mlh1* have reduced survival, a strong cancer phenotype similar to *Msh2* deficient mice and a tumor spectrum that includes T-cell lymphoma, intestinal adenoma/adenocarcinoma and skin neoplasia [4]. Typically these mice live up to 12 months. Extracts from these mice display complete MMR deficiency with MSI high tumors. This suggests that *Mlh1* is essential for MMR.

## 2.7 *PMS2*

*Pms2* deficient mice harbor a cancer predisposition phenotype that is less severe and different in its tumor spectrum compared to *Mlh1*, *Msh2* or *Msh6* mutant mice. These mice develop lymphomas and sarcomas, but not intestinal tumors [18]. They also exhibit increased mutation frequencies at mononucleotide repeat tracts, with an increase in mutation frequency up to three times lower when compared to *Mlh1* deficient mice [13, 19].

## 2.8 *MLH3*

Although MLH3 deficiency in humans is not known to result in a predisposition to malignancy, *Mlh3*-deficient mice develop adenomas/adenocarcinomas in their small intestines as well as lymphomas, basal cell carcinoma of the skin and other

**Table 2** Mouse lines with exonuclease and MutL MMR gene mutations

Genotype	Fertility M/F	Mononucleotide MSI panel	Dinucleotide MSI panel	Tumor incidence	Tumor spectrum
<i>Mlh1</i> <sup>-/-</sup>	-/-	High	High	High	Lymphoma, GI, skin, other
<i>Pms1</i> <sup>-/-</sup>	+/+	Low	Low	None	None
<i>Pms2</i> <sup>-/-</sup>	-/+	High	High	High	Lymphoma and sarcoma
<i>Mlh3</i> <sup>-/-</sup>	-/-	Low	Low	N/A	N/A
<i>Exo1</i> <sup>-/-</sup>	-/-	High	Low	Moderate	Lymphoma

tumors. These mice also display MSI at mononucleotide repeat sequences in their genomic DNA, but to a lesser extent than *Pms2* [13].

## **2.9 Exonuclease 1**

Exonuclease 1 (Exo1)-deficient mice do display a significant increase in GI tumors and therefore it was concluded that this gene might only rarely be involved in LS. Loss of Exo1 function causes defects in the repair of single base mismatches and 1 base insertion/deletion loops but does not affect repair of 2 base insertion/deletion loops. This mouse line has an increased predisposition to lymphoma late in life. Tumor tissues in these mice display MSI at mononucleotide repeat markers. When compared with MMR defects in *Msh2* or *Mlh1* deficient mice, Exo1 deficiency results in only a partial repair defect, suggesting that other redundant exonucleases with unknown identity are present [13] (Table 2).

## **2.10 Homozygous MMR Mutant Mice with Heterozygous Adenomatous polyposis coli (APC) Mutation**

Many colorectal cancers with MSI also carry somatic mutations in the *APC* gene, which implies that loss of APC function is critical for tumorigenesis. Mice with homozygous mutation of *Msh2*, *Msh6*, *Mlh1* or *Pms2* and heterozygous germ-line mutation of *Apc* develop tumors predominantly in the intestinal tract. Phenotypically, in *Apc*-mutant mice, loss of *Msh2* or *Mlh1* dramatically increases intestinal tumor burden while loss of *Msh6* or *Pms2* causes a milder intestinal tumor. In contrast, loss of *Msh3* does not result in any significant increase in tumors. Overall, the level of tumor burden correlates with the type of MMR defects in MMR deficient mice [20–23].

# **3 Microsatellite Instability in Human Cancers**

## **3.1 Lynch Syndrome or germline mutations in MMR genes**

LS is an autosomal dominant disease characterized by MSI high tumors of the colon and other organs. Previously patients with LS were identified by the Amsterdam I criteria, which required fulfillment of the following diagnostic criteria: 3 or more relatives with CRC, 2 or more successive generations affected, 1 or more relatives diagnosed before age 50 [24]. In 1998, this criterion was broadened to include the extracolonic tumors, known as Amsterdam II. Amsterdam II criteria requires fulfillment of the following diagnostic criteria: 3 or more relatives with a LS associated

cancer (CRC, endometrial, small bowel, ureter or renal pelvis), 2 or more successive generations affected, 1 or more relatives diagnosed before age 50, 1 should be a first-degree relative of the other two, familial adenomatous polyposis should be excluded, and tumors should be verified by pathologic examination [25]. In 1997, the National Cancer Institute (NCI) developed the Bethesda guidelines in order to identify individuals to undergo MSI testing and genetic testing for LS, which were subsequently revised in 2004 [26]. The revised Bethesda guidelines recommend genetic testing for MSI high tumors in the following situations: CRC diagnosed in a patient age less than 50, presence of synchronous or metachronous CRC or other LS associated tumor regardless of age, CRC with high MSI histology diagnosed in a patient less than 60 years old, CRC diagnosed in one or more first degree relatives with a LS associated tumor with one cancer diagnosed at age less than 50, CRC diagnosed in two or more first-degree or second-degree relatives with LS associated tumors regardless of age (summarized in Table 3).

LS is now defined by the presence of a germline mutation in a DNA MMR gene. Approximately 15–20% of colorectal cancers are related to a familial cancer syndrome with 3% of all CRC associated with LS and a germline mutation in a MMR protein [6]. LS is characterized by an increased risk of CRC (lifetimes risk: 54–74% males and 30–52% females) and earlier age of onset (mean age 42–61) [27]. Females with LS have a 28–68% lifetime risk for endometrial cancer and modestly increased risks for other cancers including gastric, small bowel, ovarian, urinary tract, pancreatic, biliary, brain and sebaceous gland tumors (summarized in Table 4) [27].

**Table 3** Clinical criteria to identify patients with Lynch syndrome

Amsterdam I Criteria
Amsterdam II Criteria
Revised Bethesda Guidelines
3 or more relatives with CRC
2 or more successive generations affected
1 or more relatives diagnosed before age 50 years
3 or more relatives with a LS associated cancer (CRC, endometrial, small bowel, ureter or renal pelvis)
2 or more successive generations affected
1 or more relatives diagnosed before age 50
1 should be a first-degree relative of the other two
Familial adenomatous polyposis should be excluded
Tumors should be verified by pathologic examination
CRC diagnosed in a patient age less than 50
Presence of synchronous or metachronous CRC or other LS associated tumor regardless of age
CRC with high MSI histology diagnosed in a patient less than 60 years old
CRC diagnosed in one or more first degree relatives with a LS associated tumor with one cancer diagnosed at age less than 50
CRC diagnosed in two or more first-degree or second-degree relatives with LS associated tumors regardless of age

**Table 4** MMR gene mutation and clinical human phenotype

Gene affected	Age onset	Tumor spectrum/risk
<i>MLH1</i> + <i>PMS2</i>	Early	CRC Endometrial cancer Any LS-associated cancer
<i>MSH2</i> + <i>MSH6</i>	Early	Endometrial cancer CRC Any LS-associated cancer
<i>MSH6</i>	Late	Endometrial cancer CRC Any LS-associated cancer
<i>PMS2</i>	Late	CRC 15–20% Endometrial cancer 15% Any LS-associated cancer 25–32% [27]

The MMR genes implicated in LS include *MLH1*, *MSH2*, *MSH6*, and *PMS2*; however, not all causative gene mutations are known. More recently, deletions in the 3' region of the gene *EPCAM* have been identified as a cause of constitutional epigenetic inactivation of *MSH2*. Of note, cancer risks and age of onset vary depending on which gene is affected. Lower risk of malignancy and later age of onset has been noted for individuals with *MSH6* and *PMS2* mutations [28, 29]. LS patients with endometrial cancer tend to have *MSH2* and *MSH6* gene mutations [30]. The contribution of each MMR protein to LS is 32% for *MLH1*, 38% for *MSH2*, 14% for *MSH6*, and 15% for *PMS2* [31].

As discussed previously, monoallelic inactivation of a MMR gene results in a normal phenotype but with an increased risk of developing cancer. Although LS is inherited in an autosomal dominant manner with a single mutated allele present in the germline of an affected individual, a second hit on the other allele is required before the defect in MMR occurs. Thus, MSI is not found in normal tissue unless biallelic germline mutation is present, but this is exceedingly rare. When a second alteration occurs that inactivates the wild-type allele in a cell of a target tissue, there is loss of DNA MMR activity, resulting in MSI [5]. Inactivation of the remaining wild-type allele can occur by a variety of mechanisms including deletion, gene conversion, methylation or point mutation.

Biallelic germline mutation of MMR genes has been documented in a rare childhood cancer syndrome [32]. The clinical phenotype of these children includes development of brain tumors, leukemias and lymphomas in the first decade of life. Those children who survive these initial malignancies may later develop gastrointestinal tumors, typically multiple adenomatous polyps that evolve into CRC and small bowel cancer, as well as café-au-lait macules. In patients with biallelic MMR gene mutations with gastrointestinal cancers, only 14% harbor *MLH1* or *MSH2* gene mutations and more than half biallelic *PMS2* mutations. Accordingly, parents of these children are obligate carriers of a heterozygous MMR gene mutation and require LS cancer screening. Siblings have a 25% chance of also having biallelic MMR gene mutations and a 50% chance of being heterozygous for either MMR gene mutation [32].

MMR gene mutations associated with LS include point mutations, large genomic deletions or rearrangements. Large genomic deletions can be challenging to detect since they are missed by qualitative standard sequencing technology given the potential for loss of the area's primer binding for amplification in the affected allele, resulting in amplification of the intact wild type allele only. Next generation sequencing, however, which allows for much more overlap and even assessment of mosaics, may detect such lesions.

### 3.2 *Sporadic MSI*

MSI is the hallmark feature of LS, but can occur in sporadic CRC as well. The most common etiology of sporadic MSI is acquired promoter hypermethylation of *MLH1* that occurs in a background of global hypermethylation of gene promoters known as CpG island methylator phenotype (CIMP). Methylation of the CpG site on the 3' end of the *MLH1* promoter, close to the start codon, mediates somatic gene silencing of *MLH1*. When CpG sites on the promoter regions of both copies of *MLH1* are hypermethylated, transcription factor binding is blocked and MLH1 expression is lost in epithelial cells, resulting in genomic MSI [33]. *MLH1* promoter hypermethylation is highly associated with the V600E mutation (c.1799T>A) in the *BRAF* oncogene that occurs in 40–50% of sporadic MSI CRC. *BRAF* mutations lead to constitutive activity of a kinase involved in response to growth signals. *BRAF* mutation is typically absent in LS associated tumors and helps to distinguish between sporadic and germline MSI [6]. Hypermethylation of the *MLH1* gene promoter is a somatic change associated with advanced age, female sex and proximal colon location. Sessile serrated adenomatous polyps may be the precursor lesions for the development of colon cancer in this pathway [33].

### 3.3 *Constitutional Epimutations*

Recently, cases of constitutional *MLH1* and *MSH2* epimutations have been reported in patients with normal genetic testing for MMR proteins, but meeting clinical criteria for LS.

In cases of constitutional *MLH1* epimutations, the promoter of one of the *MLH1* alleles is methylated in the germline, resulting in transcriptional silencing of this allele in non-neoplastic tissue [34]. Typically, an *MLH1* epimutation is identified when MLH1 and PMS2 are proteins that are not expressed, so these proteins are absent on immunohistochemistry (IHC) but methylation testing is positive for tumor tissue and adjacent normal tissue, no *BRAF* mutation is identified and no germline mutation identified with genetic testing. *MLH1* epimutations are reversible between generations and thus display non-Mendelian inheritance [34]. However, dominant transmission of *MLH1* epimutation has been documented [35].



<http://www.springer.com/978-3-319-19985-6>

Intestinal Tumorigenesis  
Mechanisms of Development & Progression  
Yang, V.W.; Bialkowska, A.B. (Eds.)  
2015, XI, 445 p., Hardcover  
ISBN: 978-3-319-19985-6