The Doctoral School in Health Sciences (DSHealth)
Doctoral Programme in Biomedicine (DPBM)

Molecular classification of uterine leiomyomas by genome-wide methods

Miika Mehine

Medicum
Department of Medical and Clinical Genetics
Research Programs Unit
Genome-Scale Biology Research Program
Faculty of Medicine
University of Helsinki
Finland

ACADEMIC DISSERTATION

To be publicly discussed,
with the permission of the Faculty of Medicine, University of Helsinki,
in Haartman Institute, Lecture hall 2, Haartmaninkatu 3, Helsinki,
on the 21st of October 2016, at 12 noon.

Helsinki 2016

Supervised by

Academy Professor Lauri A. Aaltonen, M.D., Ph.D.
Department of Medical and Clinical Genetics, Research Programs Unit, Genome-Scale Biology Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Docent Esa Pitkänen, Ph.D.
Department of Medical and Clinical Genetics, Research Programs Unit, Genome-Scale Biology Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Docent Pia Vahteristo, Ph.D.
Department of Medical and Clinical Genetics, Research Programs Unit, Genome-Scale Biology Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Reviewed by

Associate Professor Ville Hietakangas, Ph.D.
Department of Biosciences, Institute of Biotechnology, University of Helsinki, Helsinki, Finland

Docent Merja Heinäniemi, Ph.D.
Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland

Official opponent

Professor Peter J. Campbell, Ph.D.
Wellcome Trust Sanger Institute, Cambridge, United Kingdom

ISBN 978-951-51-2487-6 (paperback)
ISBN 978-951-51-2488-3 (PDF)
http://ethesis.helsinki.fi/
Unigrafia Oy
Helsinki 2016

TABLE OF CONTENTS

ORIGINAL PUBLICATIONS **6**

ABBREVIATIONS **7**

ABSTRACT **8**

REVIEW OF THE LITERATURE **10**

1. TUMOR BIOLOGY 10
1.1 Tumor genetics *13*
1.1.1 Genomic instability 15
1.1.1.1 Chromothripsis 16
1.1.2 Driver and passenger mutations 17
1.1.2.1 Oncogenes 17
1.1.2.2 Tumor suppressor genes 18
1.1.3 Inherited tumor susceptibility 19
1.1.4 Tumor epigenetics 20
2. GENOME-WIDE METHODS FOR STUDYING TUMORIGENESIS 21
2.1 DNA microarray technology *21*
2.2 Massively parallel sequencing technology *22*
2.2.1 Sample processing and sequencing 22
2.2.2 Data analysis 23
3. UTERINE LEIOMYOMAS 24
3.1 Clinical features and prevalence *24*
3.2 Pathological classification of uterine smooth muscle tumors *26*
3.3 Diagnosis and clinical management *27*
3.4 Risk factors *28*
3.5 Etiology and pathogenesis *29*
3.6 Genetics *29*
3.6.1 Genetic predisposition 30
3.6.2 Genetic disorders associated with leiomyomas 30
3.6.2.1 Inactivation of *FH* 31
3.6.3 Common low-penetrance risk variants 32
3.6.4 Somatic chromosomal abnormalities 33
3.6.4.1 Chromosome 12q15 rearrangements and *HMGA2* 34
3.6.4.2 Chromosome 6p21 rearrangements and *HMGA1* 35
3.6.4.3 Chromosome 14q24 rearrangements and *RAD51B* 36
3.6.4.4 Chromosome 7q22 abnormalities 36
3.6.4.5 Chromosome 10q22 rearrangements and *KAT6B* 37

3.6.4.6 Rare chromosomal abnormalities and candidate genes 37
3.7 Clonality *38*
3.8 Animal models *38*
3.9 Signaling pathways *40*

AIMS OF THE STUDY 42

MATERIALS AND METHODS 43

1. STUDY MATERIAL AND ETHICAL ISSUES (I-IV) 43
2. HISTOPATHOLOGICAL EVALUATION (I-IV) 43
3. DNA, RNA, AND CDNA PREPARATION (I-IV) 44
4. GENOME-WIDE METHODS (I-IV) 44
4.1 Whole-exome sequencing (I) *44*
4.2. Whole-genome sequencing (II-IV) *45*
4.2.1 Detection of somatic substitutions and microindels (II-IV) 45
4.2.2 Detection of somatic copy-number alterations (II-IV) 46
4.2.3 Detection of chromosomal rearrangements (II-IV) 46
4.2.3.1 Detection of interconnected chromosomal rearrangements 46
4.3. Gene expression microarrays (I-IV) *47*
4.4 SNP arrays (III-IV) *48*
4.5 RNA sequencing (IV) *48*
5. SANGER SEQUENCING VALIDATION (I-IV) 49
6. *IN SILICO* PREDICTION TOOLS (I) 49

RESULTS 50

1. EXOME SEQUENCING REVEALS HIGH FREQUENCY OF *MED12* MUTATIONS (I) 50
2. CLONALLY RELATED LEIOMYOMAS ARE RELATIVELY COMMON (II-III) 51
2.1 Intratumor genetic heterogeneity and DEPDC5 mutations (III) *52*
3. WHOLE-GENOME SEQUENCING REVEALS COMPLEX CHROMOSOMAL REARRANGEMENTS (II) 53
4. CHROMOSOMAL ABNORMALITIES CREATE DRIVER CHANGES (II-IV) 54
4.1 Commonly deleted regions on chromosomes 7, 22 and 1 (II-IV) *57*
5. GENE EXPRESSION PROFILING REVEALS DISTINCT EXPRESSION PROFILES (I-II, IV) 60
5.1 Pathway enrichment analysis using differentially expressed genes (I, IV) *61*
5.2 Uniquely expressed genes in leiomyomas of different subtypes (II, IV) *62*
5.3 Downregulated genes by chromosome 7q22, 22q, and 1p deletions (IV) *65*

DISCUSSION 66

1. *MED12*: A KEY DRIVER IN LEIOMYOMAS 66
1.1 MED12 hotspot mutations in leiomyoma variants, leiomyosarcomas, and extrauterine leiomyomas *68*
1.2 MED12 hotspot mutations in other human neoplasms *70*
1.3 Other pathogenic MED12 mutations *72*
1.4 Possible mechanisms of tumorigenesis *72*
2. CLONALITY AND INTRATUMOR GENETIC HETEROGENEITY 75
2.1 Possible mechanisms of tumor spread *75*

2.2 Branched tumor evolution ... *76*
2.3 DEPDC5 as a secondary driver gene ... *77*
3. CHROMOTHRIPSIS IN LEIOMYOMAS ... 78
4. CHROMOSOMAL ABNORMALITIES AS DRIVERS ... 79
4.1 RAD51B as a putative driver gene ... *79*
4.2 CUX1 as a haploinsufficient tumor suppressor gene ... *80*
4.3 IRS4 as a driver of COL4A5-COL4A6 deletions ... *81*
4.4 PLAG1 as a putative driver gene ... *82*
4.5 Other candidate driver genes ... *82*
5. MOLECULAR CLASSIFICATION OF UTERINE LEIOMYOMAS ... 84
6. PUTATIVE BIOMARKERS AND DYSREGULATED PATHWAYS ... 86
6.1 Activation of Prolactin signaling ... *86*
6.2 Dysregulation of IGF-1 signaling ... *87*
6.3 Activation of the NRF2 pathway in leiomyomas of the FH subtype ... *87*

CONCLUDING REMARKS AND FUTURE PERSPECTIVES ... 90

ACKNOWLEDGEMENTS ... 91

REFERENCES ... 93

ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals (I-IV).

I Mäkinen N*, **Mehine M***, Tolvanen J, Kaasinen E, Li Y, Lehtonen HJ, Gentile M, Yan J, Enge M, Taipale M, Aavikko M, Katainen R, Virolainen E, Böhling T, Koski TA, Launonen V, Sjöberg J, Taipale J, Vahteristo P, Aaltonen LA. *MED12*, *the mediator complex subunit 12 gene*, is mutated at high frequency in uterine leiomyomas. *Science*. 2011, 334(6053):252-255.

II **Mehine M***, Kaasinen E*, Mäkinen N, Katainen R, Kämpjärvi K, Pitkänen E, Heinonen HR, Bützow R, Kilpivaara O, Kuosmanen A, Ristolainen H, Gentile M, Sjöberg J, Vahteristo P, Aaltonen LA. Characterization of Uterine Leiomyomas by Whole-Genome Sequencing. *New England Journal of Medicine*. 2013, 369(1):43-53.

III **Mehine M**, Heinonen HR, Sarvilinna N, Pitkänen E, Mäkinen N, Katainen R, Tuupanen S, Bützow R, Sjöberg J, Aaltonen LA. Clonally related uterine leiomyomas are common and display branched tumor evolution. *Human Molecular Genetics.* 2015, 24(15):4407-16.

IV **Mehine M**, Kaasinen E, Heinonen HR, Mäkinen N, Kämpjärvi K, Sarvilinna N, Aavikko M, Vähärautio A, Pasanen A, Bützow A, Heikinheimo O, Sjöberg J, Pitkänen E, Vahteristo P, Aaltonen LA. Integrated data analysis reveals uterine leiomyoma subtypes with distinct driver pathways and biomarkers. *Proceedings of the National Academy of Sciences*. 2016, 113(5):1315-20.

* Equal contribution

Publication I is included in the thesis of Netta Mäkinen (MED12: a novel player in uterine leiomyomas, Helsinki, 2015), and publication II is included in the thesis of Eevi Kaasinen (Large-scale data analysis to identify novel disease phenotypes and genes, Helsinki, 2014).

The publications are reprinted with the permission of the copyright holders.

ABBREVIATIONS

Abbreviation	Definition
2SC	S-(2-succino)-cysteine
3'UTR	three prime untranslated region
ADAM12	*ADAM metallopeptidase domain 12*
AJAP1	*adherens junctions associated protein 1*
AKR1B10	*aldo-keto reductase family 1 member B10*
ANOVA	analysis of variance
ATRX	*ATRX, chromatin remodeler*
ATS-DL	Alport syndrome with diffuse leiomyomatosis
bFGF	basic fibroblast growth factor
Bhd	Birt-Hogg-Dubé
bps	base pairs
BWA	Burrows-Wheeler Aligner
CARMN	*cardiac mesoderm enhancer-associated non-coding RNA*
CCND1	*cyclin D1*
CCR	complex chromosomal rearrangement
CDK6	*cyclin dependent kinase 6*
CDK8	*cyclin dependent kinase 8*
CG	Complete Genomics
CN-LOH	copy-neutral loss of heterozygosity
CNV	copy-number variation
COL12A1	*collagen type XII alpha 1 chain*
COL4A5	*collagen type IV alpha 5 chain*
COL4A6	*collagen type IV alpha 6 chain*
CIN	chromosomal instability
CUL1	*cullin 1*
CUL3	*cullin 3*
CUX1	*cut like homeobox 1*
dbSNP	The Single Nucleotide Polymorphism database
DEPDC5	*DEP domain containing 5*
ECM	extracellular matrix
EZH2	*enhancer of zeste 2 polycomb repressive complex 2 subunit*
FASN	*fatty acid synthase*
FDR	false discovery rate
FH	*fumarate hydratase*
FIMM	Institute of Molecular Medicine Finland
GWAS	genome-wide association study
HIF-1-α	hypoxia-inducible factor 1-alpha
HLRCC	hereditary leiomyomatosis and renal cell cancer
HMGA1	*high mobility group AT-hook 1*
HMGA2	*high mobility group AT-hook 2*
HPV	human papilloma virus
ICGC	International Cancer Genome Consortium
IDH	isocitrate dehydrogenase
IGF	insulin-like growth factor
IGF2BP2	*insulin like growth factor 2 mRNA binding protein 2*
IGFBP5	*insulin like growth factor binding protein 5*
IRS4	*insulin receptor substrate 4*
KAT6B	*lysine acetyltransferase 6B*
KEAP1	*kelch like ECH associated protein 1*
KEGG	kyoto encyclopedia of genes and genomes
KIF1B	*kinesin family member 1B*
LOH	loss of heterozygosity
MED12	*mediator complex subunit 12*
MLH1	*mutL homolog 1*
mTOR	mechanistic target of rapamycin
NF1	*neurofibromin 1*
NFE2L2	*nuclear factor, erythroid 2 like 2*
NPHP4	*nephrocystin 4*
NRF2	nuclear factor, erythroid 2 like 2
NIN	nucleotide instability
OMIM	Online Mendelian Inheritance in Man
PAPPA2	*pappalysin 2*
PDGFRB	*platelet derived growth factor receptor beta*
PI3K	phosphatidylinositol-3 kinase
PLAG1	*PLAG1 zinc finger*
PLAGL1	*PLAG1 like zinc finger 1*
PRL	*prolactin*
PRLHR	*prolactin releasing hormone receptor*
PyLT	polyomavirus large T antigen
RAD51B	*RAD51 paralog B*
RB1	*RB transcriptional corepressor 1*
REST	RE1 silencing transcription factor
SCID	severe combined immunodeficiency
SCNA	somatic copy-number alteration
SDH	succinate dehydrogenase
SFRP1	*secreted frizzled related protein 1*
SMARCB1	*SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1*
SNP	single nucleotide polymorphism
SNV	single nucleotide variant
STUMP	smooth muscle tumor of uncertain malignant potential
SV40	simian virus 40
TCGA	the cancer genome atlas
TERT	*telomerase reverse transcriptase*
TGF	transforming growth factor
TP53	*tumor protein p53*
Tsc2	*tuberous sclerosis 2*
VEGF	vascular endothelial growth factor
WIF1	*WNT inhibitory factor 1*
ZMAT3	*zinc finger matrin-type 3*
ZNHIT1	*zinc finger HIT-type containing 1*

ABSTRACT

Uterine leiomyomas, often called fibroids, are highly common tumors arising from smooth muscle cells of the myometrium. Whereas cancers have the ability to metastasize, leiomyomas are benign tumors that grow only locally. Nevertheless, leiomyomas frequently cause a variety of health complications, including abdominal pain, abnormal menstrual bleeding, and impaired fertility. Leiomyomas are the leading indication for hysterectomy worldwide, and pose a significant socio-economic impact. Despite their major public health impact, this disease attracts relatively little research.

Epidemiological and molecular studies have indicated that, in the etiology of leiomyomas, genetic factors play a central role. Early cytogenetic studies revealed that approximately half of all leiomyomas display non-random chromosomal abnormalities such as *high mobility group AT-hook 2* (*HMGA2*) gene translocations. Furthermore, family-based linkage studies revealed that germline mutations in the *fumarate hydratase* (*FH*) gene result in high penetrance susceptibility to uterine leiomyomas. Sporadic leiomyomas, however, rarely harbor *FH* mutations and the majority lack chromosomal abnormalities, suggesting that some driver genes remain undiscovered.

Recent advances in sequencing technologies have made it possible to examine tumor genomes on a previously unprecedented scale. The aim of this thesis was to characterize the molecular underpinnings of uterine leiomyomas by the use of genome-wide methods such as massively parallel sequencing technology and gene expression microarrays. Using exome sequencing, we discovered that 71% of leiomyomas display localized mutations in the *mediator complex subunit 12* (*MED12*) gene, making it their most commonly mutated gene. Furthermore, with whole-genome sequencing, we discovered that a subset of leiomyomas display highly complex chromosomal rearrangements, ones previously undetectable by conventional cytogenetic techniques. These rearrangements closely resembled chromothripsis, a phenomenon in which one or a few chromosomes are shattered into multiple pieces and randomly stitched together in a single event. We also found these events to have occurred multiple times, and some had resulted in genetic changes with a selective value, such as *collagen type IV alpha 5 chain* and *collagen type IV alpha 6 chain* (*COL4A5-COL4A6*) deletions.

Patients affected by leiomyomas frequently harbor multiple distinct tumor nodules. Whereas the majority of studies have proposed that each leiomyoma arises independently, we found some leiomyomas to display identical chromosomal abnormalities, suggesting a common clonal origin. Whole-genome sequencing of clonally related leiomyomas revealed intratumor genetic heterogeneity suggestive of a branching model of tumor growth.

Furthermore, we also discovered *DEP domain containing 5* (*DEPDC5*) as a novel tumor suppressor gene, acting as a secondary driver gene in a subset of leiomyomas.

Our integrative analyses demonstrated that specific genetic defects were the major determinants of expression changes in leiomyomas. Our observations indicate that at least four molecular subtypes exist: leiomyomas harboring a *MED12* hotspot mutation, *HMGA2* overexpression, *FH* inactivation, or *COL4A5-COL4A6* deletion. We also detected subtype-specific expression differences in key tumorigenic pathways, including Wnt/β-catenin, Prolactin, IGF-1, and NRF2 signaling.

Using genome-wide methods in this thesis work, we have discovered several novel molecular defects that underlie leiomyoma etiology. These studies emphasize the importance of stratification in leiomyoma research and offer a set of candidate biomarkers that may facilitate the molecular classification of uterine leiomyomas. Millions of women suffer from uterine leiomyomas, and the ability to classify each lesion should pave the way towards personalized treatments.

REVIEW OF THE LITERATURE

1. Tumor biology

The human body consists of trillions of cells that grow, divide, and die in an organized manner. A tumor, or a neoplasm, is an abnormal mass of cells that serves no biological purpose. Solid tumors, such as colon and breast neoplasms, can either be benign or malignant (cancer), whereas hematological malignancies such as leukemia are malignant neoplasms that do not produce a cell "mass". Benign tumors grow only locally, whereas malignant tumors have the ability to invade or spread to distant organs, a phenomenon known as metastasis. Metastases are the major cause of death from cancer,[1] making most benign tumors relatively harmless. Nevertheless, some benign tumor types, ones such as leiomyomas or prolactinomas, cause clinical complications by creating pressure upon adjacent organs or secreting abnormally high levels of hormones.[2, 3] Furthermore, some cancers arise from benign precursor lesions with varying malignant potential, often referred to as premalignant lesions.[4]

Cancer development (carcinogenesis) is a multistep process in which neoplastic cells gradually acquire new abilities in order to proliferate excessively, survive, and eventually spread.[5] Most cancers take several decades to develop, and aging is the greatest risk factor.[6] Cancer is not a single disease, but rather a general term referring to a heterogeneous group of disorders characterized by ten shared traits known as "the hallmarks of cancer" (Box 1).[7] The ability to invade and metastasize is the only hallmark that never applies to benign tumors.[7]

Several environmental factors, known as carcinogens, are directly involved in causing cancer. Two of the most widely acknowledged carcinogens include ultraviolet light[8] and tobacco smoke[9]. Tobacco smoking results in the death of 1 in 10 adults worldwide[10] and is the leading preventable cancer cause, but obesity is rapidly overtaking tobacco.[11] Infectious agents such as *Helicobacter pylori*[12] and human papilloma virus (HPV)[13] have also proven to cause cancer in humans.

It is overly simplistic to think of single causes of cancer, and a variety of both external as well as internal factors influence cancer risk. Indeed, while tumor formation is clearly influenced by environmental and lifestyle factors, internal factors such as inheritance are highly important as well.[14] Furthermore, a recent controversial study indicates that a significant proportion of cancers arise due to bad luck resulting from random, uncontrollable genetic errors rather than from hereditary or environmental factors.[15]

Box 1. The hallmarks of cancer

I. **Sustaining proliferative signaling**: The ability to continue to proliferate without a normally controlled signaling input.

II. **Evading growth suppressors**: The ability to evade growth suppressors that normally limit cell proliferation.

III. **Resisting cell death**: The ability to evade programmed cell death (=apoptosis) by circumventing critical apoptotic pathways.

IV. **Enabling replicative immortality**: The ability to maintain telomere length in order to evade cell death by telomere shortening.

V. **Inducing angiogenesis**: The ability to form new blood vessels that invade and nourishes the tumor with oxygen and other critical nutrients.

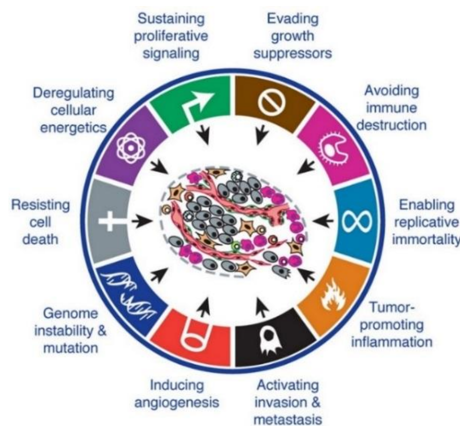
VI. **Reprogramming cellular metabolism**: The ability to adjust energy production in order to satisfy the increased rate of cell proliferation.

VII. **Avoiding immune destruction:** The ability to evade cell destruction by lymphocytes and macrophages.

VIII. **Activating invasion and metastasis:** The ability to escape from the primary growth site and attach to distant locations.

IX. **Genome instability and mutation:** The ability to generate genetic diversity in order to foster tumorigenesis and adapt to changing conditions.

X. **Tumor-promoting inflammation:** The ability to induce tumor-promoting inflammation that supplies the tumor microenvironment with factors that aid in tumor growth.



Modified from *Cell*, **144**, Hanahan and Weinberg, Hallmarks of Cancer: The Next Generation (2011), reprinted with the permission of Elsevier.

It is widely accepted that most, if not all, tumors arise due to genetic defects. Cancer is therefore called “a disease of the genome”. Indeed, the first step in tumorigenesis – known as tumor initiation – is triggered by genetic mutations.[16] Mutations can arise from physical or chemical agents known as mutagens, or spontaneously due to random errors in DNA replication, repair, or recombination. Tumor initiation is followed by tumor promotion in which existing tumors are stimulated to grow.[16] Common tumor promoters include hormones, particularly estrogens. These contribute to tumorigenesis by stimulating cell proliferation rather than inducing mutations. A single mutation is rarely sufficient for the development of a full-blown tumor, and additional genetic or non-genetic factors are required for the stepwise progression into a more malignant state. This last phase of tumorigenesis is known as tumor progression.[16]

Tumor masses are typically composed of both phenotypically and genetically distinct cell populations. Two models have aimed at explaining this tumor heterogeneity. The clonal evolution model states that mutated cells are selected and expanded based on their individual abilities.[17] Neoplastic cells that are selectively advantageous will generate more daughter cells than their competitor cells do and consequently become the more dominant cell population. This process is called clonal selection and is analogous to Darwinian natural selection, because neoplastic cells compete with each other for their survival.[17] Clonal selection continues to foster tumorigenesis in the process in which tumors become increasingly malignant.

The cancer stem cell model states that only specific cells with stem cell characteristics are responsible for the initiation and overgrowth of a tumor.[18] These cancer stem cells have an ability to self-renew and differentiate into different cell types.[18] The clonal evolution and the cancer stem cell models are not necessarily mutually exclusive, because the cancer stem cells may themselves undergo clonal evolution.[18] A recent study showed a strong correlation between the lifetime number of tissue-specific stem cell divisions and tissue-specific cancer risk, suggesting that a significant proportion of cancers arise due to random mutations in cancer-initiating stem cells.[15]

A tumor mass does not consist only of mutated tumor cells, but also the surrounding cellular environment in which the tumor grows. This tumor microenvironment is comprised of blood vessels, inflammatory cells, immune cells, connective tissue fibroblasts, and extracellular matrix (ECM) components.[19] Stromal cells can interact with the mutated tumor cells and thereby promote neoplastic transformation, support tumor growth and invasion, protect the tumor from host immunity, and foster therapeutic resistance.[19] The stromal cells of the tumor microenvironment are generally genetically normal, but may occasionally harbor a unique set of mutations,[20] suggesting that stromal cells could co-evolve with their neighboring neoplastic counterparts. The tumor microenvironment contributes in part to intratumor heterogeneity.[19]

1.1 Tumor genetics

A mutation refers to any change in a DNA sequence, and such a change can be harmful, neutral, or beneficial for the development of an organism.[21] Mutations that are inherited are known as germline mutations, whereas mutations that occur after conception are known as somatic mutations.[21] Mutation is an important process that creates genetic variation among species. Genetic variations that are common within a population (frequency ≥1%) are known as polymorphisms, with alternative forms of the same genes known as alleles.[21]

Mutations can be roughly divided into nucleotide-level (small-scale) or chromosome-level (large-scale) mutations due to differences in their size, in how they arise, and in how they are detected. Nucleotide-level mutations typically affect one or a few nucleotides and include substitutions, insertions, and deletions (Figure 1A). Insertions and deletions smaller than 50 nucleotides are collectively termed microindels. Mutations that affect only a single nucleotide are termed point mutations or single nucleotide variants (SNVs). Single nucleotide polymorphisms (SNPs) refer to point mutations that are common within a population and are the most common type of human genetic variation.[22] Sanger sequencing is the most widely used method for detecting nucleotide-level mutations.

Most eukaryotic genes are composed of exons and introns, reflecting the gene's protein-coding and non-coding sections. Proximal promoters, distal enhancers, silencers, and insulators are important non-coding sequences that control gene expression. Exons are surrounded by splice-site sequences that, during mRNA processing, instruct the removal of introns.[23] Exons are composed of codons: sequences of three nucleotides that encode for a specific amino acid or a stop signal during protein synthesis. Mutations that affect the protein-coding sequence are classified into silent (synonymous) mutations encoding for the same amino acid; missense (nonsynonymous) mutations encoding for a different amino acid; and nonsense mutations encoding for a stop codon that truncates the protein (Figure 1A). Frameshift mutations are indels of a small number that is not divisible by three, typically altering the reading frame of a gene and resulting in a truncated protein product. Splice-site mutations may result in exon skipping, in activation of cryptic splice sites, in creation of pseudo-exons within introns, or in intron retention (Figure 1A).[23]

Chromosome-level mutations, also called chromosomal abnormalities, include structural rearrangements such as amplifications: the copying of regions; deletions: the loss of regions; inversions: the reversal of regions; translocations: the interchange of regions between nonhomologous chromosomes; and insertions: the transfer of fragments from one region into another (Figure 1B). Complex chromosomal rearrangements involve more than two chromosomes or more than two DNA double-strand breakpoints.[24] An abnormal number of chromosomes is called aneuploidy, whereas a whole chromosome loss is called monosomy and a whole chromosome gain is called trisomy. Unlike nucleotide-level mutations, some chromosomal abnormalities are detectable microscopically by cytogenetic techniques such as G-banding or fluorescence in situ hybridization. Germline deletions or amplifications larger than 50 nucleotides are collectively termed copy-number variations (CNVs), whereas their somatic counterparts are termed somatic copy-number alterations (SCNAs).

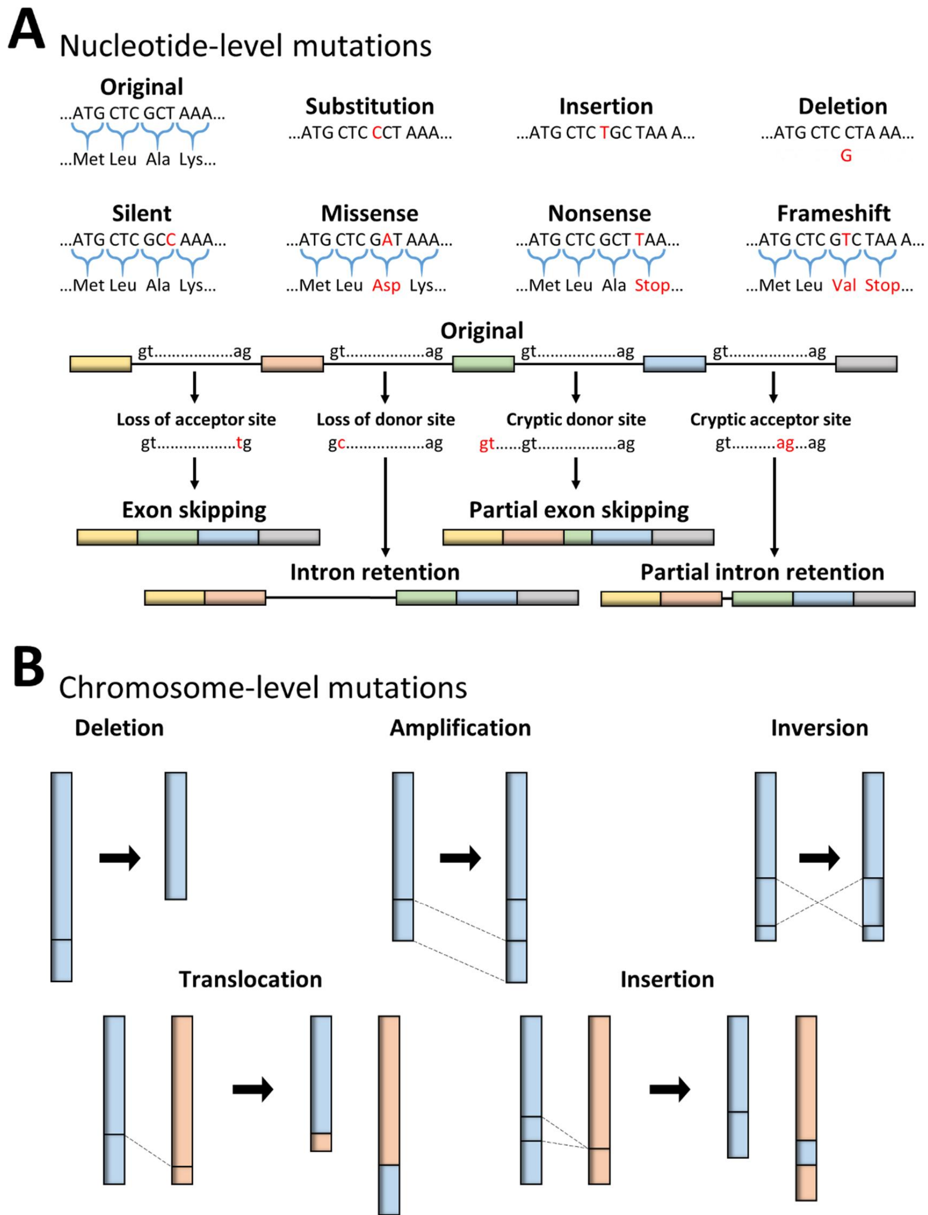


Figure 1. Types of mutations. Mutations can be classified into either nucleotide- or chromosome-level mutations due to differences in their size, in how they arise, and in how they are detected. **A**) Nucleotide-level mutations include substitutions, insertions, and deletions of one or a few nucleotides. Mutations located within the protein-coding sequences of genes are further divided into silent, missense, nonsense, or frameshift mutations, each reflecting differences in their encoded protein product. Splice-site mutations may also affect the protein-coding sequence of a gene, resulting in partial or complete intron retention or exon skipping. **B**) Chromosome-level mutations include amplifications, deletions, inversions, translocations, and insertions.

1.1.1 Genomic instability

Tumor genomes often exhibit a high rate of somatic mutations known as genomic instability.[25] The number of mutations often correlates with age,[26] suggesting that accumulation of spontaneous mutations or long-term exposure to specific mutagens may explain the high number of mutations frequently seen in tumors. Indeed, smokers' lung cancer genomes have approximately 10 times as many mutations as non-smokers'.[27] Prolonged exposure to radiation can also result in a high rate of DNA double-strand breaks.[28] Tumors rarely display genomic instability on both the small and large scale, indicating that distinct mutational processes generate different types of mutations.[25]

Chromosomal instability (CIN) is the most common type of genomic instability in tumors, and refers to a high rate of chromosome-level mutations.[25] CIN is subdivided into structural CIN referring to a high rate of rearrangements and numerical CIN referring to aneuploidy.[29] Although the underlying causes of CIN are still mostly unclear, at least five mechanisms are known to generate chromosomal rearrangements: non-homologous end joining, homologous recombination, microhomology-mediated replication-dependent recombination, long interspersed element 1 (LINE-1)-mediated retrotransposition, and telomere healing.[30] Indeed, mutations in *BRCA2* or genes encoding for the Fanconi anemia core complex result in structural CIN though defective repair of interstrand crosslinks and impaired homologous recombination.[31] Telomere shortening is another mechanism leading to structural CIN through repeated breakage-fusion-bridge cycles of unprotected sister chromatid ends.[32] Conversely, numerical CIN appears to result from weakened spindle-checkpoint signaling, excessive centrosomes, defects in chromatid cohesion, abnormal kinetochore-microtubule attachments, or increased spindle-microtubule dynamics.[33]

Nucleotide instability (NIN) refers to a high level of nucleotide-level mutations. Microsatellite instability is a type of NIN characterized by insertions or deletions of a small number of nucleotide repeats located within short repetitive sequences known as microsatellites.[34] Defects in the mismatch repair system give rise to this kind of instability, such as mutations in the DNA-repair gene *mutL homolog 1* (*MLH1*).[25] Mutations in the proofreading domains of DNA polymerases *POLE* and *POLD1* was recently associated with an extremely high rate of NIN, known as the hypermutator phenotype.[35]

Recent advances in sequencing technologies have made it possible to examine tumor genomes on a previously unprecedented scale.[36] This has led to the identification and characterization of several novel mutational signatures, ones displaying different combinations of mutation types.[37] Different types of mutational processes are likely to cause these distinct signatures. One of these signatures is termed Kataegis, which refers to a high rate of nucleotide-level mutations – mostly C>T transitions – that co-localize around chromosomal rearrangements.[38] These technologies have also uncovered chromosome-level mutations that have radically changed the view of how some chromosomal rearrangements arise. Indeed, some of these novel patterns challenge the conventional theory that cancer development is a gradual process of multiple cell divisions during which independent mutations accumulate.

1.1.1.1 Chromothripsis

In 2011, Stephens et al. described a novel form of CIN known as chromothripsis.[39] Chromothripsis is a phenomenon in which a single “catastrophic” event results in up to thousands of clustered rearrangements. Statistical modeling of complex rearrangements detected in cancer genomes led to the hypothesis that these were generated in a single event of local shattering followed by random reassembly.[39] The term chromothripsis is derived from the two Greek words: *chromos* for chromosome and *thripsis* for shattered into pieces.[39]

Rearrangements generated by chromothripsis have a few notable features. First, the double-strand breaks are confined to a single chromosome or to local regions over a few chromosomes. Second, the chromosomal fragments are randomly rearranged in all possible orientations. Third, some fragments are not incorporated into the newly formed chromosomes, thus resulting in two alternating copy-number states. Initially, the concept of chromothripsis attracted criticism, because no direct experimental evidence was available to support its existence.[40] Complex chromosomal rearrangements have occasionally been misinterpreted as chromothripsis, and so its hallmark features have recently been clarified in order to accurately infer chromothripsis from massively parallel sequencing data.[41]

Although first described in chronic lymphocytic leukemia, chromothripsis has since been established in a wide array of human cancers.[42] Some of these chromothripsis events have generated oncogenic fusion genes, disrupted tumor suppressor genes, or amplified oncogenes within double-minute chromosomes.[42] Interestingly, chromothripsis events can also occur in the germline of healthy individuals or can cause congenital disorders.[43] A recent study described a remarkable case in which a chromothripsis event spontaneously cured a patient of the immunodeficiency WHIM (warts, hypogammaglobulinemia, infections and myelokathexis) syndrome.[44] The chromothripsis event removed the defective gene (*CXCR4*) causing this syndrome, resulting in repopulation of the bone marrow by cells carrying this event.[44]

Following the initial discovery of chromothripsis, massively parallel sequencing technology have revealed a few similar “one-off” events, including chromoanasynthesis and chromoplexy. Chromoplexy is a phenomenon in which multiple adjacent double-strand breaks result in a complex chain of rearrangements across multiple chromosomes.[45] Chromoplexy differs from chromothripsis since it involves less breakpoints, more chromosomes, and it generally lacks alternating copy-number states.[45] Although the mechanism behind chromoplexy is still unknown, the breakpoints in chromoplexy are typically clustered within actively transcribed DNA and open chromatin.[45] Conversely, chromoanasynthesis results from serial, microhomology-mediated template switching during DNA replication.[46] Unlike chromothripsis, chromoanasynthesis is characterized by small-scale amplifications and retention of heterozygosity.[46] A common term proposed to describe all of these one-off events, independent of their underlying mechanisms, is chromoanagenesis.[47] The initial trigger of these events is still unknown, and the mechanism by which chromothripsis occurs was largely unclear until recently. Chromothripsis appears to be caused by missegregation of chromosomes during cell division into physically isolated micronuclei, inside of which the shattering and reassembly takes place.[48]

1.1.2 Driver and passenger mutations

Although tumor genomes may carry up to thousands of mutations, carcinogenesis directly involves only an estimated two to eight changes. Mutations that confer a selective growth advantage are known as "driver mutations", whereas mutations with no advantageous effect are known as "passenger mutations".[49] One of the central challenges in cancer research is to distinguish between these two. Another common challenge is to identify the "target genes" of large-scale chromosomal abnormalities – ones typically affecting numerous genes simultaneously. The key driver genes are typically mutated in a significantly higher proportion of tumors than would be expected by chance, making it possible to confidently identify some drivers from sequencing data alone. However, most cancer types display only a few "mountains" and numerous "hills", reflecting highly recurrently and less recurrently mutated genes.[49] The frequency does not always indicate drivers in the case of hills, because the mutational background rates between individual tumor genomes and regions may vary considerably. One effective way to recognize driver genes is to consider their mutational patterns. Indeed, all driver genes follow two main types of mutational patterns: gain-of-function mutations of oncogenes or loss-of-function mutations of tumor suppressor genes.

1.1.2.1 Oncogenes

A proto-oncogene is a gene that has the potential to drive tumorigenesis when activated by a mutation or expressed at exceptionally high levels.[49] When a proto-oncogene gains this ability, it is called an oncogene. Proto-oncogenes typically encode for proteins that normally function as growth factors, growth factor receptors, signal transducers, transcription factors, or chromatin remodelers.[50] In 1982, the very first human oncogene that is activated by somatic mutations was discovered, *HRAS*.[51]

Oncogenes are characteristically activated by gain-of-function mutations, such as point mutations, translocations, or amplifications.[50] Mutated oncogenes are dominant, because their activation requires only one single mutated copy. Missense mutations – the most common type of oncogenic mutation – typically alter or enhance a very specific activity of the encoded protein. These mutations are often located within specific regions known as mutational "hotspots". Although the majority of these hotspots lie in exons, recent studies indicate that oncogenic mutations occur in non-coding sequences as well. Indeed, missense mutations in the promoter of *telomerase reverse transcriptase* (*TERT*) – resulting in increased expression of telomerase – were recently discovered as the most common type of non-coding mutation in human cancers.[52] Chromosomal amplifications can also activate oncogenes by causing an increased gene dosage. Translocations can create oncogenic fusion proteins with a completely new function or a combined function of the two fusion partners. Translocations can also activate an oncogene by juxtaposing a proto-oncogene next to a regulatory element such as a highly active promoter or enhancer. Mutations in the three prime untranslated region (3'UTR) can activate an oncogene by disrupting the interaction between mRNAs and microRNAs.[53] MicroRNAs are small non-coding RNA molecules that negatively regulate the translation of proteins by binding to complementary sequences located at the 3' end of untranslated transcripts. In addition to these different types of mutations, retroviruses can insert an already activated oncogene into a host genome.[54]

1.1.2.2 Tumor suppressor genes

A tumor suppressor gene protects a cell against tumorigenesis.[49] The most commonly mutated tumor suppressor gene in human cancers is *tumor protein p53* (*TP53*).[55] This gene has numerous important functions in preventing tumorigenesis and is therefore known as "the guardian of the genome".[55] Tumor suppressor genes are generally classified into gatekeeper and caretaker genes.[56] Gatekeeper genes encode for proteins that directly limit cell division or stimulate apoptosis.[56] Loss of these genes results in uncontrolled cell growth. Caretaker genes, also known as stability genes, encode for proteins involved in maintenance of genomic stability.[56] Loss of these genes indirectly contributes to tumorigenesis by causing an increased mutation rate.[56] Mutations in caretaker genes themselves are therefore neither selectively advantageous nor required for tumorigenesis.[56] Consequently, caretaker gene mutations are rare in sporadic tumors and more often are the underlying cause of hereditary cancers.[56] Landscaper genes represent a third and less-common type of tumor suppressor.[57] The loss of these contributes to tumorigenesis by altering the stromal microenvironment to support the growth of neoplastic cells.[57]

According to Knudson's two-hit hypothesis, a selective growth advantage occurs only when both alleles of a tumor suppressor gene are inactivated (Figure 2).[58] Tumor suppressor genes are therefore recessive. However, exceptions include dominant negative mutations in which a mutant protein interferes with the action of the corresponding wild-type protein.[58] Another exception is haploinsufficiency, in which the loss of a single copy results in a gene dosage that is low enough to drive tumorigenesis.[58]

Tumor suppressor genes are typically inactivated by loss-of-function mutations, including nonsense or frameshift mutations.[49] Chromosomal rearrangements may also inactivate a tumor suppressor gene by disrupting the gene itself or separating it from essential regulatory elements.[49] The deletion of a gene and of its surrounding regions is known as loss of heterozygosity (LOH). The deleted region may also be replaced by a copy from the other homologous chromosome, resulting in no net change in copy number. This phenomenon is known as copy-neutral loss of heterozygosity (CN-LOH).[59]

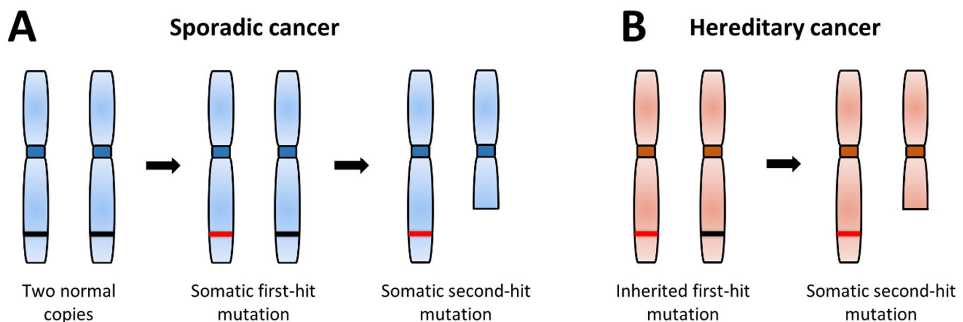


Figure 2. Knudson's two-hit model of tumorigenesis. According to Knudson's two-hit hypothesis, loss of both alleles of a tumor suppressor gene is a requirement for tumorigenesis. **A**) Tumor suppressor genes are often inactivated by a loss-of-function mutation in conjunction with a larger chromosomal deletion. **B**) In hereditary cancers, the "first-hit" is inherited in all cells of the body, leading to an earlier onset of multiple primary cancers.

1.1.3 Inherited tumor susceptibility

Tumor development is highly influenced by environmental and lifestyle factors, and the majority of tumors are sporadic. An estimated 5 to 10% of cancers are said to be hereditary, most of which result from germline mutations with high penetrance.[14] Penetrance refers to the proportion of individuals who carry a mutation that will give rise to a specific disease or trait. A hereditary cancer syndrome is a genetic disorder characterized by early onset of multiple primary tumors in several family members. The majority of the over 200 tumor syndromes described are inherited in an autosomal dominant manner with incomplete penetrance.[14] These syndromes characteristically result from germline mutations in tumor suppressor genes that conform to Knudson's two-hit model of tumorigenesis (Figure 2B). Family-based linkage studies have been highly successful in identifying such faulty genes, including *MLH1* in Lynch syndrome,[60] *RB transcriptional corepressor 1* (*RB1*) in hereditary retinoblastoma,[61] and *neurofibromin 1* (*NF1)* in neurofibromatosis type I.[62] However, the use of linkage studies is typically limited to highly penetrant syndromes (Figure 3).

During the previous decade, research focus has shifted from Mendelian (single-gene) models towards polygenic models of tumor susceptibility (Figure 3). The "common disease-common variant" model proposes that many common cancers arise due to a combined effect of multiple common low-penetrance variants.[63] Genome-wide association studies (GWAS) – in which thousands of individuals are investigated for associations between polymorphisms and disease phenotypes – have been successful in identifying such variants.[64] Understanding the molecular mechanisms behind such variants remains challenging because they are usually located in non-coding regions and are thus unlikely to be directly causal. Exceptions include a variant (rs6983267) located within an enhancer that directly regulates the activity of the proto-oncogene *MYC*.[65] The variants identified through GWAS and linkage studies explain only a small proportion of the entire estimated heritability of common cancers. This has led researchers to seek this "missing heritability".[64]

Approximately 15 to 20% of common cancers are considered "familial", which is defined as familial clustering of cancers.[14] In contrast to hereditary cancers, these cases follow no apparent inheritance pattern. The genetic background behind such cancer cases remains largely unexplained, and much of the latest research seek to understand them. Many of these cancers may arise due to rare variants with intermediate penetrance conferring a moderate cancer risk (Figure 3). Novel technologies combined with new strategies may have the potential to identify and characterize such complex patterns of inheritance.

Common risk variants may also be detectable by a novel approach known as admixture mapping.[66] With admixture mapping it is possible to detect disease-causing variants that differ in frequency between two historically separate populations.[66] This method is ideally performed in a population descended from two ancestral groups, each of which have been geographically isolated for tens of thousands of years: for example, African Americans with West African and European ancestry.[66]

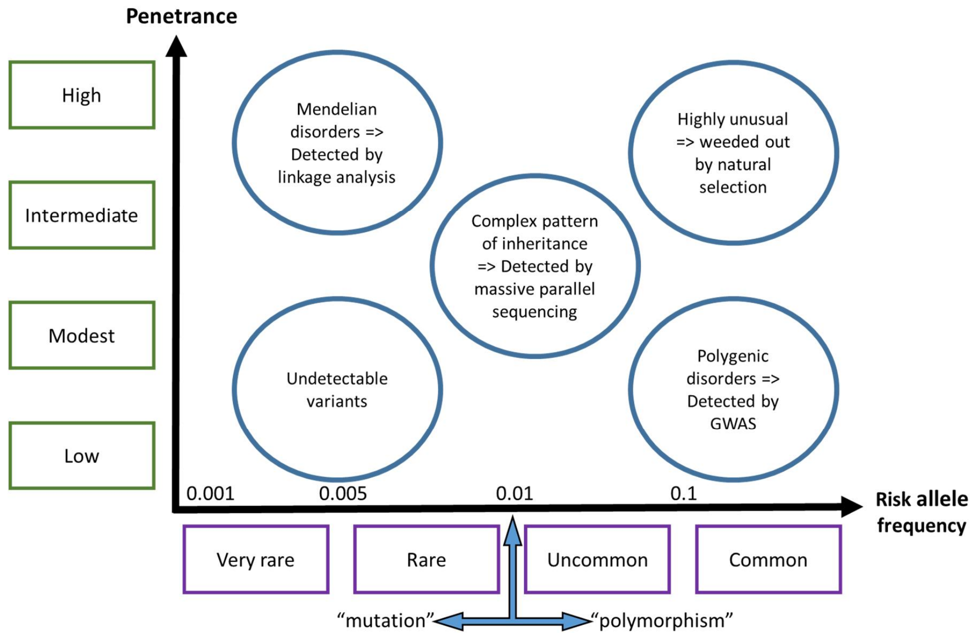


Figure 3. The relationship between penetrance and risk allele frequency. Linkage studies have been highly successful in identifying rare variants underlying Mendelian disorders. In contrast, genome-wide association studies have been successful in identifying common variants contributing to the risk for many common cancers. Familial aggregation studies indicate that a substantial fraction of the entire estimated heritability remains to be uncovered. This missing heritability may be explained by relatively rare variants with intermediate penetrance. Massively parallel sequencing combined with new strategies may have the potential to identify and characterize such complex patterns of inheritance.

1.1.4 Tumor epigenetics

Epigenetics is the study of cellular and physiological traits that are heritable by daughter cells and not caused by alterations to the DNA sequence.[67] Use of the term "epigenetics" to describe modifications that are not heritable remains controversial.[68] Several external and internal factors influence epigenetics, including age, the environment, and disease.[69] Epigenetic modifications, or "tags", include chemical alterations to the DNA strand or to proteins bound to it. These modifications may alter the chromatin structure of DNA and thereby alter gene expression. Epigenetics play a central role in nearly all cellular processes, including embryonic development and differentiation of pluripotent stem cells.[67] Furthermore, women carry two copies of the X chromosome, one of which is epigenetically silenced in each cell of the body. Epigenetics explains, in part, how cells carrying identical DNA differentiate into different tissues. In contrast to mutations, epigenetic alterations are reversible, and their status frequently differs from cell to cell. Only a few studies have been able to provide evidence for epigenetic tags that are transmittable to offspring.[70] The mechanisms by which these changes are transmitted are unclear, since the majority of epigenetic tags are erased after fertilization through a process called reprogramming.[71]

DNA methylation and histone modification are two universally accepted types of epigenetic mechanism. DNA methylation is a chemical process in which specific enzymes add methyl groups to DNA, predominantly at CpG sites where a cytosine is followed by a guanine.[72] DNA methylation can regulate the transcription of genes by preventing or allowing the binding of the basal transcriptional machinery or transcription factors to DNA.[72] Eukaryotic DNA is wrapped around special proteins called histones that form a large structure called the nucleosome. Histones play a role in maintaining chromatin structure, thereby regulating the transcription of genes.[67]

Epigenetic modifications may drive tumorigenesis, since tumors often exhibit a variety of abnormal epigenetic patterns not present in the normal tissue from which they arise.[73] Furthermore, hypomethylation can activate oncogenes, whereas hypermethylation can silence tumor suppressor genes.[73] Two different models aim to explain the mechanism by which methylation influences tumorigenesis. The stochastic model proposes that selectively advantageous changes occur by chance, whereas the instructive model proposes that DNA methylation is maintained by trans-acting mutations.[74] The role of epigenetic alterations in tumorigenesis is largely unclear, and epigenetics is currently one of the most rapidly expanding fields in cancer-related research.

2. Genome-wide methods for studying tumorigenesis

The field of molecular biology has undergone rapid changes during recent decades. Research has shifted from an analysis of a handful of molecules to simultaneous analysis of thousands of measurements on a genome-wide scale. Molecular biology has therefore become a data-intensive field requiring new considerations for experimental design, data analysis, and data visualization. Two of the most widely used high-throughput methods for studying molecular biology are microarrays and massively parallel sequencing.[36, 75] These technologies have revolutionized the study of human diseases, especially in the field of cancer. Indeed, cancers have historically been classified by their primary site of origin or by their histology, but these technologies have led to the identification and classification of molecular subtypes of cancers.

2.1 DNA microarray technology

DNA microarrays, also called gene-chips, were first introduced in 1995 as a tool for studying the expression of thousands of genes simultaneously.[75] This method was quickly adopted for other purposes as well, including detection of copy-number variations or SNP genotyping. Commercial- and custom-made microarrays are nowadays available from a variety of manufacturers, including Illumina, Affymetrix, and Agilent.[75] This technology is based on the hybridization of fluorescent-labeled molecules of interest to complementary probes that are attached on a surface consisting of silicon, nylon, or glass.[75] Microarrays can nowadays have up to millions of probes, consisting of oligonucleotides, cDNA, or cRNA.[75]

After hybridization, the fluorescent dye is excited with a laser, and the array images are quantified into numeric values.[75] One drawback of this method is that it does not allow for the identification of rearrangements such as translocations or inversions. Although microarrays are increasingly being replaced by massively parallel sequencing technology, the use of this technology remains particularly useful in gene expression profiling studies.

2.2 Massively parallel sequencing technology

Massively parallel sequencing, also known as next generation sequencing, refers to a novel sequencing approach in which millions to billions of DNA or RNA fragments are sequenced simultaneously.[36] Since its commercial launch in 2005, massively parallel sequencing has become widely available from several companies, including Illumina, Complete Genomics, Life Technologies, Roche 454, and Helicos BioSciences.[76] This technology has made it possible to generate high-resolution sequencing data in less time and at lower cost.[76] In comparison, the sequencing of the human genome, known as the Human Genome Project, was declared complete in April 2003 and ended up costing approximately $2.7 billion dollar by an approach known as shotgun sequencing.[77] These technologies have proven powerful for studying complex genetic diseases such as cancer.

Targeted sequencing is a method that combines targeted capture with massively parallel sequencing in order to explore regions of interest, typically exons.[76] Whole-exome sequencing refers to the sequencing of all known protein-coding regions of a genome, which constitutes approximately 1% of the human genome.[76] Whole-genome sequencing provides full coverage of the entire genome and allows for detection of chromosomal abnormalities with high precision.[76] This method is, however, costly for large sample sizes. This is particularly true for analysis of cancer genomes, since identification of somatic mutations requires sequencing of both tumor and normal tissues from the same patient.

Although this technology was initially used to study DNA, it has since been adopted for other purposes as well, such as investigating the transcriptome (RNA sequencing), the methylome (bisulfite sequencing), and interactions between DNA and transcription factors (chromatin immunoprecipitation sequencing).[36] Today, the cost of massively parallel sequencing is steadily decreasing and quickly overtaking Sanger sequencing as the gold standard sequencing technique. Since the sequencing of the first whole-cancer genome in 2008, this technology has led to the discovery of numerous cancer causing genes in a relatively short time.[78]

2.2.1 Sample processing and sequencing

All currently available sequencing platforms require preprocessing of DNA into a library that is suitable for sequencing. Although a wide variety of preparation protocols exists, each of them includes the fragmentation of DNA into an appropriate size followed by adapter ligation.[36] This is typically followed by PCR amplification in which sufficient quantities of

DNA are generated for the sequencing.[36] Illumina sequencing is currently the platform most widely used, and has a workflow that continues with immobilization of fragments onto a glass slide known as a flow cell, where each DNA molecule is sequenced in parallel.[36] This sequencing involves cycles of repeated addition of all four nucleotides, each labeled with a different dye.[36] Paired-end sequencing refers to the sequencing of both ends of a DNA fragment, thus generating higher quality and more easily alignable sequencing data.[36]

2.2.2 Data analysis

One current drawback from these technologies is that they generate relatively short reads, making genome assembly, alignment, and variant calling a difficult challenge that requires the development of novel bioinformatic tools and algorithms. Another drawback is that it produces gigabytes of data that are computationally intensive to analyze.

The first major step in the data analysis is alignment of the sequencing reads to a reference genome. This process is particularly time consuming and requires a significant amount of computing resources. Since a large part of the human genome consists repetitive regions, aligning short reads is a difficult challenge. Reads carrying large indels are also difficult to align and the PCR amplification step often results in overrepresentation of some DNA fragments. Duplicate read removal and local indel realignment are therefore often performed to refine alignments prior to variant calling.[79] Bowtie, Burrows-Wheeler Aligner (BWA), and Novoalign are some of the most widely used short read aligners.[80]

Once aligned, variant calling requires multiple different methods, each restricted to detecting a variant of a particular size. Indeed, substitutions, microindels, chromosomal rearrangements, and CNVs/SCNAs are typically detectable by different methods. Current methods for detecting substitutions show high sensitivity and specificity and include methods such as MuTect, VarScan 2, and Strelka.[81] On the other hand, detection of microindels and rearrangements are still problematic in both their sensitivity and specificity, requiring complex scoring and filtering strategies in order to achieve reliable calls. GATK Unified Genotyper, VarScan 2, SAMtools, and GTAK HaplotypeCaller include some of the most widely used tools for detecting microindels,[82] whereas BreakDancer, Pindel, Delly, and novoBreak[83] include some of the most widely used tools for detecting chromosomal rearrangements. Varscan 2, HMMCOPY, ExomeCNV, and ABSOLUTE are popular for detecting CNVs/SCNAs.[84] Variant calling is highly dependent on the sequencing depth coverage, and an average sequencing depth of 30x to 50x is often recommended for sequencing tumor samples.[85]

3. Uterine leiomyomas

3.1 Clinical features and prevalence

Leiomyomas are benign smooth muscle tumors of mesenchymal origin that can arise from any tissue containing smooth muscle cells, but most frequently occurs in the uterus, small bowel, or esophagus. Uterine leiomyomas originate from the muscle cells of the myometrium layer, and are classified into intramural, submucosal, and subserosal, according to their anatomic position within the uterus (Figure 4).[2] The ancient Greek physician Hippocrates had already recognized the existence of these uterine masses, and referred to these as "uterine stones".[86] The term "leiomyoma" will hereafter refer to uterine leiomyomas unless stated otherwise.

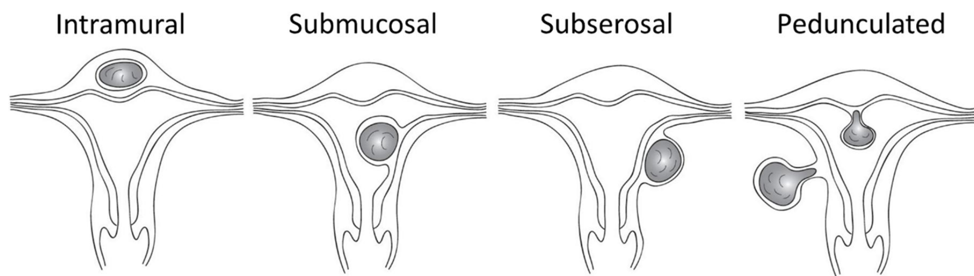


Figure 4. Anatomical classification of leiomyomas according to their position within the uterus. Leiomyomas most commonly grow within the smooth muscle layer of the uterine wall (intramural). They may occasionally grow towards the uterine cavity (submucosal) or towards the abdominal cavity (subserosal). Submucosal or subserosal leiomyomas may also grow on an elongated stalk (pedunculated). Leiomyomas can be single or multiple, and the spectrum of symptoms depends largely on their size, number, and location.

Leiomyomas are round, firm, sharply circumscribed and white-grayish neoplasms with a limited vascularization network. They are typically surrounded by a pseudocapsule that separates the neoplastic tissue from the myometrium.[87] These tumors are histologically composed of intersecting bundles of smooth muscle cells that resemble the normal myometrium. Leiomyomas are commonly called fibroids, owing to their distinctive features of fibrosis (excessive fibrous connective tissue).[88] Indeed, a particular hallmark of leiomyomas is the overproduction of extracellular matrix components, primarily consisting of collagens, fibronectins, and proteoglycans.[88] A major reason for their overgrowth is accumulation of this disorganized extracellular matrix that is exceptionally stable.[88]

Leiomyomas may grow as a single nodule or in clusters, averaging eight nodules per affected uterus.[89] Although they may grow to as large as 30 cm in diameter, the growth rate of leiomyomas is relatively slow, with an average increase of 9% in volume over a 6-month period.[90] However, they display significant variability in growth rate, with some lesions even regressing spontaneously.[90] Furthermore, despite having the same hormonal environment, leiomyomas within the same uterus often have different growth rates.[90]

Uterine leiomyomas are the most common neoplasms found in women of childbearing age. These tumors are typically detected during women's middle or late reproductive years, and they usually shrink after menopause. Approximately 70 to 80% of women will develop a leiomyoma by the age of 50, but the majority will be asymptomatic with the tumors remaining undiagnosed.[91] Nevertheless, approximately 25% of women with leiomyomas will present with symptoms,[89] making them one of the most common tumors afflicting women. Symptoms include pressure upon adjacent organs, abnormal menstrual bleeding and impaired fertility (Table 1).[2, 92] Most symptoms are reflective of leiomyoma size and location, but some symptoms may result from unknown biochemical alterations mediated by the tumor.

Table 1. Clinical presentation of uterine leiomyomas

Asymptomatic

Abnormal uterine bleeding

- Heavy or prolonged menstrual bleeding (menorrhagia)
- Painful menstruation (dysmenorrhea)
- Anemia

Pressure on the bladder

- Frequent urination
- Difficulty or inability to urinate
- Hydronephrosis

Pressure on the rectum

- Constipation
- Tenesmus

Obstetric complications

- Premature labor
- Spontaneous miscarriage
- Infertility

Pain symptoms

- Pelvic discomfort or pain
- Painful intercourse (dyspareunia)
- Lower back pain

Increased waist size

Ascites

Polycythemia

Data in part derived from *Reproductive Sciences*, **19**, Sabry M. et al., Medical Treatment of Uterine Leiomyoma (2012).

Because of their high morbidity, leiomyomas are the foremost indication for hysterectomy worldwide. Approximately 600,000 hysterectomies are performed yearly in the United States, making it the second most common surgical procedure for women after Caesarean section.[93] By the age of 60, more than one-third of all women will have had a hysterectomy, with 40% of these due to complications of leiomyomas.[93] The social and economic impact of this disease is therefore very high, in the United States costing an estimated 6 to 34 billion dollars per year.[94]

3.2 Pathological classification of uterine smooth muscle tumors

Smooth muscle tumors of the uterus consist of a heterogeneous group of mesenchymal tumors ranging from benign leiomyomas to high-grade malignant leiomyosarcomas. Several rare leiomyoma-like lesions should be distinguished from "conventional" leiomyomas, as they display unusual histopathology or distinct growth patterns. These leiomyoma variants account for approximately 10% of all leiomyomas (Table 2).[95]

Table 2. Classification of uterine smooth muscle tumors by the World Health Organization

Benign	Morphological features
Conventional leiomyoma	Bundles of smooth muscle cells with high extracellular content
Lipoleiomyoma	Admixture of mature adipocytes and smooth muscle cells separated by thin fibrous tissue
Epithelioid leiomyoma	Round to polygonal epithelial-like smooth muscle cells arranged in clusters or cords
Angioleiomyoma (Vascular Leiomyoma)	Presence of thick-walled vascular vessels
Myxoid leiomyoma	Presence of myxoid material between smooth muscle bundles
Hydropic leiomyoma	Presence of conspicuous zonal watery edema
Apoplectic leiomyoma	Presence of hemorrhage and edema within nodules of hypercellular smooth muscle cells
Unusual growth pattern	
Parasitic leiomyoma	A leiomyoma detached from the uterus and attached to peritoneal surfaces from which it derives its blood supply
Disseminated peritoneal leiomyomatosis	Smooth muscle nodules scattered over the peritoneal surfaces
Intravenous leiomyomatosis	Abnormal smooth muscle growth within vascular spaces from intrauterine venules to the right heart
Benign metastasizing leiomyoma	Smooth muscle tumors located at distant organs in women with a prior history of uterine surgery
Diffuse uterine leiomyomatosis	Almost complete replacement of the myometrium by innumerable poorly defined, confluent tumor nodules
Dissecting leiomyoma (Cotyledonoid leiomyoma)	Extrauterine bulbous growth continuous with a dissecting myometrial component
Intermediate	
Mitotically active leiomyoma	High mitotic index (>10 mitotic figures per 10 high-power fields), but lacking tumor cell necrosis and nuclear atypia
Cellular leiomyoma	Increased cellularity, but lacking tumor cell necrosis, nuclear atypia, and high mitotic index
Leiomyoma with bizarre nuclei (Atypical/Symplastic leiomyoma)	Presence of bizarre multinucleated smooth muscle cells (nuclear atypia), but lacking tumor cell necrosis and high mitotic index
Smooth muscle tumors of uncertain malignant potential (STUMP)	High mitotic index, increased nuclear atypia, and tumor cell necrosis, but not enough to meet the criteria of a leiomyosarcoma
Malignant	
Leiomyosarcoma	Presence of tumor cell necrosis, high mitotic index, nuclear atypia, and increased cellularity

Data in part derived from Kurman R.J. et al., *WHO Classification of Tumours of Female Reproductive Organs*. Fourth Edition (2014).

Smooth muscle tumors resembling uterine leiomyomas at the gross and histologic level, but present at unusual locations include disseminated peritoneal leiomyomatosis, intravenous leiomyomatosis, and benign metastasizing leiomyomas (Table 2).[96] The etiology of these conditions remains controversial, as they have been associated with a history of uterine surgery for leiomyomas.[96] Diffuse uterine leiomyomatosis is another usual growth pattern, one in which the uterus is symmetrically enlarged due to a complete replacement of the myometrium by innumerable poorly defined, confluent tumor nodules.[97] The existence of these quasi-malignant behaviors suggests that some leiomyoma cells have the potential to spread and attach to distant locations. Nevertheless, it is important to note that these variants are classified as clinically benign rather than premalignant.

The majority of uterine smooth muscle tumors are readily classified as either benign or malignant based upon their gross and histologic features. Malignant leiomyosarcomas are characterized by tumor cell necrosis, high mitotic activity, nuclear atypia, and increased cellularity.[98] Considerable debate still exists as to whether leiomyomas and leiomyosarcomas are part of the same disease continuum. Leiomyosarcomas are extremely rare and account for only 0.1% of uterine smooth muscle tumors, suggesting that leiomyosarcomas arise *de novo*.[99] The existence of leiomyoma variants displaying some, but not all, features of malignancy suggests, however, that leiomyosarcomas may develop from leiomyomas.[98] Furthermore, leiomyosarcomas may occasionally display areas that are histologically benign, supporting the view of a malignant transformation.[100] The distinction between leiomyosarcomas and some histologically intermediate variants tends to be challenging; smooth muscle tumors that cannot be diagnosed as either benign or malignant are sometimes classified as smooth muscle tumors of uncertain malignant potential (STUMP).[98] While it is conceivable that some leiomyomas may become malignant, what is still unclear is why the vast majority of these very common tumors have such a low malignant potential.

Although the uterus is the most common site of origin for both leiomyomas and leiomyosarcomas, both of these tumor types can arise from any tissue that contains smooth muscle cells. Uncommon sites of origin include the vulva, ovaries, bladder, urethra, abdominal cavities, kidneys, and skin.[101] Such extrauterine smooth muscle tumors are very rare, and present a greater diagnostic challenge.

3.3 Diagnosis and clinical management

Uterine leiomyomas are often diagnosed incidentally during routine pelvic examination, and confirmed by ultrasound.[102] Magnetic resonance imaging provides a more accurate, but also costly method for their detection and localizion.[102] In general, leiomyomas are treated only if they cause symptoms, and the treatment choice depends on patient age, general health, symptoms, reproductive desires, and leiomyoma size and location.[103] Leiomyomas can be treated surgically, medically, or with minimally invasive techniques.[103]

Hysterectomy, the surgical removal of the uterus, is currently the most common intervention for leiomyomas.[103] Myomectomy, the surgical removal of leiomyomas leaving the uterus intact, is the primary treatment choice for women who wish to preserve their fertility.[103] After myomectomy, up to 25% of women will experience recurrence, and 10% will require additional surgeries.[104] The general belief is that these recurrences represent newly developed leiomyomas rather than the same leiomyomas that grow back.

Open power morcellation refers to the dissection of large tissues into smaller pieces with an electric morcellator.[105] Although an efficient tool, this device has recently become a subject of controversy. Morcellation may accidentally spread tumor tissue around the peritoneal cavity or surgical ports. This may lead to postoperative growth of tumors at distant locations,

sometimes referred to as "morcellomas".[106] Furthermore, morcellation may spread unsuspected leiomyosarcomas, leading to worsened disease outcomes. An estimated 1/500 to 1/1000 hysterectomy specimens of a presumed leiomyoma will ultimately be confirmed as leiomyosarcoma.[107] The U.S. Food and Drug Administration have therefore issued a warning against the use of morcellators.[108] However, this tool is more cost-effective than its alternatives, with an estimated cost of 11 million dollars to prevent one unintentional case of disseminated leiomyosarcoma.[109]

At present, medical management of leiomyomas is very limited and mainly serves to temporarily relieve symptoms or to reduce the size of leiomyomas prior to surgery. The most common medications include non-steroidal anti-inflammatory drugs, oral contraceptive pills, selective progesterone receptor modulators, and gonadotropin-releasing hormone agonists.[92] Gonadotropin-releasing hormone agonists act on the pituitary gland to reduce the level of estrogen produced by the ovary, leading to a reduction in leiomyoma size by half within a 3-month period.[92] Nevertheless, leiomyomas typically regrow once treatment is stopped, and long-term use causes severe side effects associated with low estrogen. A better understanding of the molecular underpinnings may allow for identification of novel therapeutic targets that could effectively treat leiomyomas long-term.

3.4 Risk factors

Since leiomyomas are so common, it is reasonable to assume that many leiomyomas may share a common risk factor. Epidemiological and experimental studies have identified several potential risk factors for leiomyomas (Table 3), most of which are not yet fully understood and require further validation. Not only do these studies indicate that leiomyomas develop under unfavorable environmental conditions, but that inheritance is highly important as well.

Table 3. Risk factors for uterine leiomyomas

Risk	Risk factor	References	Risk	Risk factor	References
Increased	African American ethnicity	90, 91, 110	**Decreased**	Menopause	111, 112
	Positive family history	113-117		Increasing parity	118
	Early age of menarche	119		Fruit and vegetable intake	120, 121
	Age (reproductive years)	91, 122		Dairy intake	123
	Nulliparity	124		Green tea extract	125
	Hormone replacement therapy	112		Depot medroxyprogesterone acetate	126
	Thyroid disease	127		Tobacco smoking	128
	Hypertension	124, 129		Physical activity	130
	Polycystic ovary syndrome	131			
	Vitamin D deficiency	132	**Contradictory**	Oral contraceptives	128, 133
	Obesity	112, 134		Pelvic inflammatory disease	135, 136
	Psychological stress	137		Chlamydia	135, 136
	Red meat intake	120		Diabetes	124, 138
	Alcohol intake	139, 140			
	Caffeine intake	139			
	Radiation exposure	141			
	Diethylstilbestrol	142			

3.5 Etiology and pathogenesis

The etiology and pathogenesis of leiomyomas are still largely elusive, and several hypotheses have attempted to explain their cause and nature. Interestingly, several of the risk factors for leiomyomas can be linked to ovarian steroid hormones (estrogen and progesterone). Indeed, ovarian steroid hormones are clearly essential, since leiomyomas are rare before puberty and typically regress after menopause.[111] Epidemiological, clinical, and experimental data have established that leiomyoma growth is dependent on both estrogen and progesterone.[143] However, the reason behind this dependency is still unknown, and both of these hormones are considered tumor promoters rather than tumor initiators. Some researchers have hypothesized that leiomyomas mimic a pregnant myometrial cell type that is hypersensitive to ovarian steroid hormones.[144] Indeed, leiomyomas and parturient myometrium share numerous characteristics, including expression of pregnancy-related genes and increased production of extracellular matrix components.[144]

Although the hypothesis that leiomyomas are linked to inflammation dates back to 1932,[145] more recent studies have postulated that uterine injury could trigger the formation of leiomyomas through abnormal activation of growth factors involved in wound healing and inflammation.[146] Indeed, leiomyomas and keloid scars are both more common among African Americans, and both of these lesions display excessive extracellular matrix deposition.[147]

Other studies have proposed that viral infections may trigger the transformation of leiomyomas because transgenic mice expressing simian virus 40 (SV40) large T antigen – under the control of a smooth muscle promoter – develop uterine smooth muscle tumors that are estradiol dependent.[148] Furthermore, transgenic mice expressing polyomavirus large T (PyLT) antigen – under the control of the mouse mammary tumor virus promoter – develop uterine leiomyomas as well.[149]

Although leiomyomas were not historically considered a genetic disease like cancer, the prevailing theory is that their development is triggered by genetic mutations. The neoplastic transformation of a single normal myometrial cell into a fully developed leiomyoma is likely to be a multistep process involving a complex interplay between environmental factors, ovarian steroid hormones, local growth factors, and genetic mutations.

3.6 Genetics

Epidemiological, cytogenetic, and molecular studies support a central role for genetics in the predisposition, etiology, and pathogenesis of leiomyomas.[115, 117, 150] However, the identity of external or internal factors that initiate this process remains unknown. Uterine leiomyomas have been associated with a variety of genetic changes, suggesting that leiomyomas are not a single-gene disease.

3.6.1 Genetic predisposition

Familial aggregation, twin, and racial prevalence studies provided the initial evidence for an important role of genetic factors in leiomyoma development. Familial aggregation of leiomyomas was already reported in 1938, showing that leiomyomas are four times as frequent in first-degree relatives of affected probands as among first-degree relatives of unaffected probands.[113] Leiomyomas are also twice as common in women with two or more affected family members.[115] Twin studies have revealed heritability estimates of 0.26 in Finns,[117] 0.69 in the British,[116] and 0.79 in Russians.[114] Compared with dizygotic twins, monozygotic twins are twice as likely to be concordant for hysterectomy and more likely to be hospitalized for leiomyomas.[117]

Leiomyomas are three to nine times as prevalent in African-American women as in Caucasian women.[91] Furthermore, African-American women tend to develop tumors at a younger age, experience more severe symptoms, develop larger and a greater number of tumors, and are more likely to undergo hysterectomy for leiomyomas.[90, 91, 122] In addition, leiomyomas in black women tend to have higher growth rates and respond differently to medical treatment.[90] Since no environmental risk factor has yet explained this racial disparity, inherited susceptibility likely plays a major part. In support of this hypothesis, a recent admixture-mapping study of African-Americans found a positive correlation between percentage of African ancestry and leiomyoma risk.[151]

3.6.2 Genetic disorders associated with leiomyomas

Hereditary leiomyomatosis and renal cell cancer (HLRCC) is a rare autosomal dominant disorder characterized by a susceptibility to uterine and cutaneous leiomyomas, as well as to renal cell cancer in a subset of families.[152] Recent studies have proposed a slightly higher risk for pheochromocytomas and paragangliomas as well.[153, 154] This disorder was first described in 1973 as Reed's Syndrome,[155] and has since been known under several other names, including multiple cutaneous leiomyomas, and multiple cutaneous and uterine leiomyomatosis, reflecting differences in their clinical presentations.[152] Uterine leiomyomas associated with this syndrome have a higher prevalence, earlier age of onset, and require treatment more often than those that are sporadic.[152] Moreover, they typically display distinct histological features, including a large nucleus with a prominent eosinophilic nucleolus surrounded by a clear halo.[152] The cutaneous leiomyomas (piloleiomyomas) are multiple, and arise from the tiny muscles attached to hair follicles.[152] The associated renal cell cancer is highly aggressive and usually of papillary type II histology.[152]

Uterine leiomyomas are also a less common manifestation of Cowden syndrome[156] and Schwannomatosis[157] (Table 4). Furthermore, extrauterine leiomyomas have been associated with Alport syndrome,[158] Gorlin syndrome,[159] and Tuberous sclerosis (Table 4).[160] The presentation of uterine leiomyomas in some of these syndromes may be coincidental, since sporadic leiomyomas are so very common.

Table 4. Human genetic disorders associated with uterine or extrauterine leiomyomas

Genetic disorder	OMIM#	Tumor spectrum	Causative gene(s)
HLRCC	150800	Uterine leiomyoma, piloleiomyoma, renal cell cancer	*FH*
Cowden syndrome	158350	Uterine leiomyoma, multiple hamartoma	*PTEN*
Schwannomatosis	162091	Uterine leiomyoma, cutaneous schwannoma, meningioma	*SMARCB1*
Alport syndrome	308940	Esophageal leiomyomatosis, vulvar leiomyoma	*COL4A5*, *COL4A6*
Gorlin syndrome	109400	Ovarian leiomyoma, basal cell carcinoma	*PTCH1*
Tuberous sclerosis	606690	Lymphangioleiomyomatosis, angiomyolipoma, rhabdomyoma	*TSC1*, *TSC2*

OMIM: Online Mendelian Inheritance in Man

3.6.2.1 Inactivation of *FH*

In 2002, family-based linkage analysis revealed that the HLRCC syndrome is caused by heterozygous germline mutations in the *fumarate hydratase* (*FH*) gene,[161] all leading to an absent, truncated, or nonfunctional protein. Indeed, *FH* represents a classical tumor suppressor gene, since the cutaneous, uterine, and renal tumors all display somatic loss of the wild-type allele, conforming to Knudson's two-hit model of tumorigenesis.[152] Subsequent studies have revealed that only 1% of sporadic leiomyomas display somatic loss of both *FH* alleles.[162, 163] Because only 20% of HLRCC families display renal cell cancer,[152] additional genetic or environmental factors are likely to be necessary for their development.

Several hypotheses have aimed at explaining the mechanism of *FH* mutations in tumorigenesis. *FH* encodes for fumarase, a citric acid cycle enzyme that catalyzes the conversion of fumarate to L-malate in the mitochondrial matrix.[164] The citric acid cycle is a central metabolic pathway that finalizes the oxidative degradation of carbohydrates, fatty acids, and proteins into carbon dioxide and chemical energy.[164] Loss of fumarase results in reduced citric acid cycle function and accumulation of its substrate fumarate.[164] The majority of functional studies have assumed that fumarate acts as an oncometabolite that drives tumorigenesis, owing to its apparent function as a competitive inhibitor of a class of enzymes known as α-ketoglutarate-dependent dioxygenases.[165] Indeed, fumarate can inhibit prolyl 4-hydroxylases that normally degrade hypoxia-inducible factor 1-alpha (HIF-1-α).[166] Accumulation of this transcription factor results in activation of target genes involved in angiogenesis, glycolysis, cell survival, and metastasis.[167] This phenomenon has been termed "pseudohypoxia",[166] since HIF-1-α is normally active under low oxygen conditions. Fumarate can also inhibit histone demethylases and the ten-eleven-translocation family of 5-methylcytosine hydroxylases,[165] both of which are involved in epigenetic regulation. Consequently, renal cell cancers and paragangliomas with *FH* mutations display genome-wide hypermethylator patterns known as the hypermethylation phenotype or the CpG island methylator phenotype.[168, 169]

Mutations in genes encoding for other citric acid cycle enzymes are also involved in tumorigenesis.[164] Indeed, germline mutations in subunits of succinate dehydrogenase (SDH) predispose to paragangliomas, whereas somatic mutations in subunits of isocitrate dehydrogenase (IDH) occur in glioma and acute myeloid leukemia.[164] Loss-of-function mutations in the *SDHx* genes (*SDHA*, *SDHB*, *SDHC*, or *SDHD*) result in accumulation of their substrate succinate, which can also act as a competitive inhibitor of α-ketoglutarate-dependent dioxygenases.[165] Conversely, gain-of-function mutations in the *IDHx* genes (*IDH1* or *IDH2*) result in the conversion of abnormally high levels of 2-hydroxyglutarate

from α-ketoglutarate,[164] which in turn acts as another competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Indeed, both paragangliomas with *SDHx* mutations, and gliomas with *IDHx* mutations exhibit pseudohypoxia and the hypermethylation phenotype.[168, 170, 171] What remains to be seen is whether leiomyomas that harbor *FH* mutations also display hypermethylation patterns, and whether some leiomyomas are driven by *SDHx* or *IDHx* mutations.

Recent studies have shown that intracellular fumarate can activate the transcription factor Nuclear factor erythroid 2-related factor 2 (NFE2L2, also known as NRF2) through post-translational modification of its negative regulator kelch like ECH associated protein 1 (KEAP1).[172] Indeed, fumarate is able to succinate KEAP1, a chemical process in which fumarate spontaneously reacts with cysteine sulfhydryl groups to form a stable chemical modification known as S-(2-succino)-cysteine (2SC).[172] NRF2 is a master regulator of the antioxidant response pathway, and its activation can drive tumorigenesis.[173] Finally, fumarate may play a role in stimulating the repair of DNA double-strand breaks,[174] suggesting that loss of *FH* could trigger an increase in free radical generation, leading to genomic instability.

3.6.3 Common low-penetrance risk variants

A few population-based studies have recently investigated, on a genome-wide scale, the role of common genetic variants in susceptibility to leiomyomas (Table 5). The first GWAS of leiomyomas identified three genome-wide significant loci on chromosomes 10q24.33, 22q13.1, and 11p15.5 among Japanese women.[175] A subsequent GWAS was unable to replicate these findings among European Americans, but instead found a novel significant locus on chromosome 17q25.3.[176] The same study included a genome-wide linkage scan that identified two significant linkage peaks on chromosomes 10p11.21 and 3p21.31, and five suggestive peaks on chromosomes 2q37.1, 5p13.3, 11p15.5, 12q14.1, and 17q25.3. Expression of the candidate gene *fatty acid synthase* (*FASN*) – located within the 17q25.3 region – was shown by immunohistochemistry to be elevated by 3-fold in leiomyomas when compared to the corresponding myometrial tissue.

Two genome-wide admixture-mapping studies have recently aimed at identifying leiomyoma risk variants that differ between African and European Americans (Table 5). The first such study found only suggestive associations on chromosomes 2q33.3, 4p16.1, and 10q26.3.[151] The second study was unable to replicate these suggestive associations, but instead, after controlling for covariates, found a significant locus on chromosome 1q42.2.[177] Models with no stratification revealed several significant associations, of which chromosome 2q32.2 displayed the highest score.[177]

No susceptibility locus has repeatedly emerged from these four genome-wide scans, suggesting that leiomyoma risk is dependent on multiple variants with small effects. Furthermore, the variants from these studies explain only a small proportion of the estimated heritability, suggesting that additional susceptibility loci remain to be uncovered.

Several follow-up studies have attempted to replicate the findings from these four genome-wide scans. One study replicated variants within 22q13.1 and 11p15.5 in European Americans.[178] Furthermore, the variants within 22q13.1 were associated with increased leiomyoma size, whereas the variants within 11p15.5 were associated with the intramural leiomyoma subtype. A follow-up study, one using variants (both suggestive and significant) derived from the first admixture mapping study and the genome-wide linkage scan, reported several significant risk associations within genes that encodes for components of the extracellular matrix.[179] A second study by the same group identified a variant within 4p16.1 to be significantly associated with increased leiomyoma size.[180] A third study by the same group, one evaluating variants (both suggestive and significant) derived from the two GWAS, replicated the variants within 22q13.1 as significantly associated with increased leiomyoma size among European Americans.[181] The same study reported a variant within 2q22.1 to be associated with increased tumor size, but not with increased leiomyoma risk.

Table 5. Common variant susceptibility loci for uterine leiomyomas

	Chromosome	Risk variant/locus	Candidate gene(s)	Study (Reference)
Significant associations	10q24.33	rs7913069	*SLK*, *OBFC1*	Cha, P.C. et al. 2011 (175)
	22q13.1	rs12484776	*TNRC6B*, *ADSL*	Cha, P.C. et al. 2011 (175)
	11p15.5	rs2280543	*ODF3*, *BET1L*, *RIC8A*, *IRT3*	Cha, P.C. et al. 2011 (175)
	17q25.3	rs4247357	*FASN*, *CCDC57*, *SLC16A3*	Eggert, S.L. et al. 2012 (176)
	1q42.2	rs7546784	*PCNXL2*	Zhang, K. et al. 2015 (177)
	2q32.2	rs256552	*PMS1*	Zhang, K. et al. 2015 (177)
	10	9,632,527–72,985,946	*COL13A1*	Eggert, S.L. et al. 2012 (176)
	3	169,614–76,307,730	-	Eggert, S.L. et al. 2012 (176)
Suggestive associations	4p16.1	rs9715724	*SORCS2*	Wise, L.A. et al. 2012 (151)
	10q26.3	rs7100028	-	Wise, L.A. et al. 2012 (151)
	2q33.3	rs7573626	-	Wise, L.A. et al. 2012 (151)
	2	179,605,032–239,157,621	*COL6A3*	Eggert, S.L. et al. 2012 (176)
	5	2,956,307–7,419,893	-	Eggert, S.L. et al. 2012 (176)
	11	278,505–7,282,771	-	Eggert, S.L. et al. 2012 (176)
	12	38,583,007–76,110,787	*HMGA2*	Eggert, S.L. et al. 2012 (176)
	17	64,478,082–78,138,144	*FASN*, *CCDC57*, *SLC16A3*	Eggert, S.L. et al. 2012 (176)

3.6.4 Somatic chromosomal abnormalities

The first description of cytogenetic abnormalities in leiomyomas dates back to the late 1980s.[182, 183] Since then, numerous cytogenetic studies have established that approximately 40% of leiomyomas harbor cytogenetically detectable chromosomal abnormalities, some of which are leiomyoma-specific and non-random (Table 6).[150]

Chromosomally abnormal leiomyomas tend to be larger, more cellular, and have a higher mitotic index,[150] indicating that chromosomal abnormalities play fundamental roles in leiomyoma pathobiology. Furthermore, chromosomal abnormalities are generally less common among submucosal leiomyomas,[147] and the type of abnormality may affect leiomyoma size, anatomic location, or histopathology.[150] In contrast to leiomyomas, malignant leiomyosarcomas have highly complex karyotypes.[184]

Rearrangements of 12q15 and deletions of 7q22 represent the two most common type of chromosomal abnormalities in leiomyomas.[150] Other less-common abnormalities include rearrangements of 6p21, 14q24, and 10q22.[150] These cytogenetic observations have led to early attempts to categorize leiomyomas into molecularly distinct subtypes. It is still somewhat unclear whether these abnormalities represent primary or secondary changes, since some of these are present only within a subpopulation of tumor cells. This leaves approximately 60% of leiomyomas without any cytogenetically detectable chromosomal abnormalities, suggesting that submicroscopic mutations underlie the remaining tumors.

Table 6. Overview of regions recurrently affected by chromosomal abnormalities in leiomyomas

Chromosome	Type of abnormality	Putative target gene(s)
12q15	Translocation/Inversion/Trisomy	*HMGA2*
6p21	Translocation	*HMGA1*
14q24	Translocation	*RAD51B*
7q22	Deletion/Inversion/Translocation	*CUX1*, *PCOLCE*
10q22	Translocation/Deletion	*KAT6B*
1q43	Deletion	*FH*
1p36	Translocation/Deletion	*AJAP1*, *NPHP4*
3q	Deletion	-
22q	Deletion/Monosomy	-
19q	Deletion	-

3.6.4.1 Chromosome 12q15 rearrangements and *HMGA2*

Cytogenetic studies have revealed that approximately 10% of leiomyomas harbor rearrangements of chromosome 12q15,[185] typically in the form of a balanced translocation or a paracentric inversion. These rearrangements appear to be initiating or early events, since they can be seen as a sole chromosomal abnormality, or as the only consistent abnormality among several subclones.[186] Rearrangements of 12q15 are also recurrent in several other benign mesenchymal tumors, including breast fibroadenomas, endometrial polyps, lipomas, angiomyxomas, pulmonary chondroid hamartomas, salivary gland adenomas, and prolactinomas.[150] These rearrangements appear to be rare in extrauterine leiomyomas, because only a single vulvar leiomyoma[187] and a single paratesticular leiomyoma[188] have emerged as harboring this rearrangement. However, 12q15 rearrangements do occur in leiomyoma variants displaying unusual growth patterns, including intravenous leiomyomatosis[189] and disseminated peritoneal leiomyomatosis[190]. Considered together, these observations indicate that a driver gene lies within this region.

High mobility group AT-hook 2 (*HMGA2*) is nowadays widely accepted as the target gene of mesenchymal tumors carrying 12q15 rearrangements.[191] In fact, *HMGA2* was the very first human driver gene discovered to be mutated in benign tumors.[192] Positional cloning revealed that the breakpoints in leiomyomas are typically located upstream of *HMGA2*, leading to its increased expression.[191] Furthermore, these rearrangements preferentially target 14q24 as a balanced translocation partner.[191] Surprisingly, the other mesenchymal tumors usually target other regions and display breakpoints that are intragenic of *HMGA2*, resulting in fusion transcripts containing the first three exons of *HMGA2* and ectopic sequences of other genes.[191] Only a few leiomyomas have been found to harbor such fusions; these have involved *RAD51B*, *COX6C*, *ALDH2*, *CCNB1IP1*, or RTVL-H as fusion

partners of *HMGA2*.[185] However, most researchers agree that the fusion partners are not directly involved in tumorigenesis. Indeed, every fusion transcript lack the 3'UTR of *HMGA2* that contains multiple binding sites for the let-7 family of microRNAs, a family that negatively regulates the activity of *HMGA2*.[193] Furthermore, transgenic mice expressing either truncated or wild type forms of *HMGA2* develop benign tumors, although their tumor spectrum is slightly different: the former develop mainly lipomas, whereas the latter develop prolactinomas as well.[194] Of note, increased expression of *HMGA2* has been associated with an aggressive behavior in a variety of malignant neoplasias.[191] However, this overexpression is rarely accompanied by 12q15 rearrangements, with the notable exception of some hematological malignancies.[195-197]

HMGA2 encodes for a nuclear transcription factor characterized by three DNA-binding domains known as "AT-hooks".[194] HMGA2 can indirectly regulate the expression of genes by binding to the minor groove of AT-rich DNA sequences, thereby inducing changes in chromatin structure.[194] *HMGA2* is highly expressed during embryonic development and is generally silenced in differentiated adult tissues.[194] Some adult stem cells continue to express *HMGA2*, however.[198] Leiomyomas with *HMGA2* rearrangements tend to be larger than chromosomally normal leiomyomas,[199] suggesting that *HMGA2* stimulates cell growth. Indeed, *Hmga2*-null mice are born with reduced body size,[200] and a few GWAS have identified variants within the *HMGA2* locus to be associated with human height.[201] Furthermore, a specific TC dinucleotide repeat (TC227) located within the 5'UTR of *HMGA2* has been associated with small stature and an increased expression of *HMGA2* in leiomyomas.[202] In addition, the Guinness World Book record holder for the tallest teenager harbors a germline inversion involving the *HMGA2* locus.[203, 204]

Several different hypotheses have attempted to explain the oncogenic action of HMGA2 in human neoplasms, most of which are related to its ability to regulate gene expression. While the exact role of HMGA2 in leiomyomas remains unclear, HMGA2 has the ability to induce E2F1 activity, enhance expression of G1-S cell-cycle checkpoint genes, enhance AP1 activity, regulate myoblast proliferation, activate the mechanistic target of rapamycin (mTOR) signaling pathway, inactivate p53-induced apoptosis, impair DNA-repair, enhance expression of inflammatory proteins, and modulate epithelial-to-mesenchymal transition.[191]

3.6.4.2 Chromosome 6p21 rearrangements and *HMGA1*

Approximately 3% of leiomyomas harbor rearrangements of chromosome 6p21.[185] These rearrangements result in elevated levels of their target gene: *high mobility group AT-hook 1* (*HMGA1*).[205] Since *HMGA1* and *HMGA2* (*HMGAx*) are two closely related and highly evolutionarily conserved genes,[191] it is reasonable to assume that these genes have overlapping functions in tumorigenesis. *HMGA1* rearrangements appear to be common in benign metastasizing leiomyomas,[206] and in several other mesenchymal tumors as well, including pulmonary chondroid hamartomas, lipomas, and endometrial polyps.[150] Unlike *HMGA2* rearrangements, 6p21 rearrangements are never intragenic of *HMGA1*. Rearrangements of *HMGA1* may also, albeit rarely, target 14q24 in leiomyomas.[205]

3.6.4.3 Chromosome 14q24 rearrangements and *RAD51B*

Approximately 9% of leiomyomas harbor rearrangements of chromosome 14q24.[185] The recurrent involvement of 14q24 in leiomyomas and other mesenchymal tumors suggests that a driver gene is located within this region, as well. Positional cloning of the 14q24 region revealed that the breakpoints are randomly scattered within *DNA repair protein RAD51 homolog 2* (*RAD51B*).[207] Whereas overexpression of *HMGAx* is likely to be the primary pathogenic outcome, disruption of *RAD51B* may play a secondary role. Indeed, RAD51B is important in DNA double-strand break repair by homologous recombination, and haploinsufficiency of *RAD51B* causes chromosomal instability in human cell lines.[208] Both ionizing and ultraviolet radiation can induce the expression of *RAD51B*, causing a delay in the G1 phase of the cell cycle or apoptosis.[209]

Rare germline mutations in *RAD51B* and its closely related genes *RAD51C* and *RAD51D* are associated with familial breast and ovarian cancers, whereas common variants within the seventh and tenth intron of *RAD51B* are low risk factors for breast cancer.[210] Furthermore, a recent study of chemoresistant ovarian cancers reported recurrent chromosomal rearrangements that disrupted *RAD51B*.[211] In addition, germline translocations affecting the *RAD51B* locus predispose to familial thymomas.[212] It is still unclear whether *RAD51B* contributes to leiomyoma development only by providing *HMGAx* with an effective regulatory element, or whether loss of *RAD51B* itself is also selected for.

3.6.4.4 Chromosome 7q22 abnormalities

Approximately 7% of leiomyomas harbor deletions affecting the q-arm of chromosome 7.[150] Deletions of 7q are also recurrent in lipomas, endometrial polyps, and some hematological malignancies.[150] The frequent involvement of 7q deletions in leiomyomas indicates the presence of a tumor suppressor gene within this region. Defining the minimally deleted region and finding a putative tumor suppressor gene has been challenging due to inconsistent deletion maps and the gene-dense nature of the deleted regions.[213-215] The key pathogenic region has, however, been pinpointed to chromosome band 7q22.[150] Indeed, leiomyomas may occasionally harbor inversions or translocations affecting this specific region.[150] The foremost candidate is currently *Cut like homeobox 1* (*CUX1*), but no second-hit mutations have been found in the gene.[216]

Deletions of 7q may co-exist with *HMGAx* rearrangements, and are sometimes only subclonally present,[217] indicating that they are involved in tumor progression rather than tumor initiation. Some studies have proposed that 7q deletions could have a negative effect on tumor growth, since leiomyomas with 7q deletions in a mosaic state tend to be smaller than those leiomyomas with a normal karyotype.[217] In contrast to *HMGA2* rearrangements, 7q deletions are highly unstable and frequently disappear in cell cultures, suggesting that 7q deletions are not essential for the neoplastic transformation of leiomyomas.[217] Although 7q deletions are frequent in leiomyomas, all evidence points to a secondary or a passenger role for these abnormalities.

3.6.4.5 Chromosome 10q22 rearrangements and *KAT6B*

Approximately 2% of leiomyomas harbor rearrangements of chromosome 10q22,[185] and typically involve chromosome 17q21 as a translocation partner.[218, 219] Positional cloning of four leiomyomas harboring this specific translocation revealed recurrent breakpoints within the third intron of *lysine acetyltransferase 6B* (*KAT6B*), located at 10q22.[219] A recent study reported a retroperitoneal leiomyoma to harbor this specific translocation as well.[220] Using RNA sequencing, they identified a fusion gene involving the first three exons of *KAT6B* and the last five exons of *KANSL1*. *KAT6B* has also been reported to form fusion transcripts with *CREBBP* (16p13) in acute myeloid leukemia,[221] suggesting that *KAT6B* is the primary driver gene rather than *KANSL1*.

KAT6B encodes for a histone acetyltransferase that is a component of the MOZ/MORF protein complex, which play a role in epigenetic regulation.[221] Of note, 10q22 rearrangements may also occur in leiomyosarcomas, and leiomyomas with 10q22 translocations tend to be histopathologically cellular,[219, 222] suggesting that this genetic defect may lead to higher malignant potential.

3.6.4.6 Rare chromosomal abnormalities and candidate genes

A variety of chromosomal abnormalities occurs at a much lower frequency in leiomyomas (Table 6). The majority of these tend to be subclonal changes that co-occur with the previously discussed chromosomal abnormalities,[150] suggesting that they are involved in tumor progression rather than in tumor initiation. Although rare in conventional leiomyomas, deletions of 1p, 22q, and 19q frequently co-occur with each other and are one of the most common chromosomal abnormalities seen in benign metastasizing leiomyomas, intravenous leiomyomatosis, and STUMPs.[206, 223, 224] Moreover, cellular leiomyomas frequently harbor 1p deletions, and these tumors appear to display gene expression patterns similar to those seen in leiomyosarcomas.[225]

A small subset of leiomyomas display a balanced translocation involving chromosomes 1p36 and 2p24 as their only chromosomal abnormality,[185] suggesting that this represents another rare molecular pathway to leiomyoma formation. Positional cloning studies have revealed that the breakpoints on chromosome 1p36 are located between *adherens junctions associated protein 1* (*AJAP1*) and *nephrocystin 4* (*NPHP4*).[226] Of note, one study reported a retroperitoneal leiomyoma to harbor a balanced translocation that resulted in an in-frame fusion gene involving *EWSR1* (9q33) and *PBX3* (22q12),[227] suggesting that these genes may also be involved in the development of uterine leiomyomas.

3.7 Clonality

Clonality studies have revealed an identical pattern of X-chromosome inactivation among several subclones of a leiomyoma,[228-230] suggesting that each leiomyoma originates as a monoclonal tumor. X-inactivation studies have also revealed a random pattern of X-chromosome inactivation among multiple tumors from the same uterus, suggesting that each leiomyoma arises independently.[228-230] Furthermore, each tumor typically displays a unique set of chromosomal abnormalities.[150] Although most researchers agree that the majority of leiomyomas arise as monoclonal tumors through discrete initiating triggers, a few studies have encountered identical cytogenetic changes among multiple leiomyomas from the same patient,[231-234] indicating that some uterine lesions are clonally related. Furthermore, pulmonary leiomyomas and concurrent uterine leiomyomas show an identical pattern of X-chromosome inactivation,[235] suggesting that benign metastasizing leiomyoma arises from disseminated uterine lesions. Likewise, patients with disseminated peritoneal leiomyomatosis tend to have multiple concurrent peritoneal nodules that display an identical pattern of X-chromosome inactivation.[236] One patient with intravenous leiomyomatosis showed an identical pattern of X-chromosome inactivation among three retroperitoneal leiomyomas and one intracaval leiomyoma.[237] Surprisingly, patients with diffuse uterine leiomyomatosis display a random pattern of X-chromosome inactivation among different uterine sites,[97] suggesting that this condition is caused by multiple independent lesions fusing with each other and blending imperceptibly. Of note, criticism regarding the use of X-inactivation assays for determining clonality has recently grown.[238, 239]

Leiomyomas are composed of several phenotypically dissimilar cell types, including smooth muscle cells, vascular smooth muscle cells, and fibroblasts.[240] Some debate still exists as to whether leiomyomas arise from a mature smooth muscle cell (myocyte) or from a myometrial stem cell. A growing body of evidence indicates that each leiomyoma originates from a transformed somatic stem cell of the myometrium that differentiates into all the various cell types of a mature leiomyoma.[241] Indeed, recent studies show that approximately 1% of the cells of a leiomyoma display stem cell-like properties, known as the leiomyoma side population.[241]

3.8 Animal models

Animal models provide invaluable tools for investigating human genetic diseases. Currently, no suitable animal model exists that closely replicates the molecular pathogenesis of human uterine leiomyomas, but several animal models displaying leiomyoma-like lesions have provided insight into the pathobiology and treatment of these tumors (Table 7).

The most widely investigated animal model is the Eker rat. These rats, first described by Reidar Eker in 1954, spontaneously develop uterine leiomyomas and renal cell cancer at a high frequency.[242] Eker rats carry a heterozygous germline mutation in the *tuberous sclerosis 2* gene (*Tsc2*), resulting in the growth of both renal and uterine lesions.[242] These tumors display LOH at the *Tsc2* locus, resulting in biallelic loss of the gene.[242] In fact, the

Eker rat was among the very first animal models that confirmed Knudson's two-hit model of tumorigenesis.[243] These uterine tumors display increased expression of *Hmga2*, suggesting that *Tsc2* may regulate the activity of *Hmga2*.[244] Germline mutations in the human *TSC2* leads to the tuberous sclerosis syndrome, characterized by a susceptibility to a variety of neoplasms, including lymphangioleiomyomatosis and occasionally renal cell cancer.[245] These patients show no increased risk for uterine leiomyomas, however.

Several genetically engineered animal models develop leiomyoma-like lesions (Table 7). Of note, although transgenic mice overexpressing *HMGA2* do not develop leiomyomas, these mice do develop other mesenchymal tumors that could be useful in studying the mechanism of *HMGA2* in uterine leiomyomas,[194, 246] since the pathogenic mechanisms are likely to overlap.

Leiomyomas have also been studied by means of xenotransplantation of human-derived leiomyoma cells into severe combined immunodeficiency (SCID) mice (Table 7). Moreover, xenotransplantation of human myometrial cells that have been transfected with truncated forms of *HMGA2* develop leiomyoma-like tissues similar to those of human uterine leiomyomas.[247]

Spontaneous development of uterine leiomyomas is also common among several other animal species, some of which have been used to study this disease (Table 7). Female German shepherd dogs with a germline loss-of-function mutation in the *Birt-Hogg-Dubé* (*Bhd*) gene develop a variety of benign and malignant neoplasms, including uterine leiomyomas and renal tumors.[248] Germline mutations in the human homologue *FLCN* cause the Birt-Hogg-Dubé syndrome, characterized by development of skin, lung, and kidney tumors.[249] These patients show no increased risk for uterine leiomyomas, however. Approximately 8% of guinea pigs develop leiomyomas by the age of four,[250] indicating that these could serve as a feasible animal model of this disease.

Table 7. Overview of animal models in study of uterine leiomyomas

Animal model	Method	Requirements	Reference
Rat (Eker)	Selective breeding	Germline *Tsc2* mutation	(242)
Mouse	Genetic engineering	Conditional knockout of *Tsc2*	(251)
Mouse	Genetic engineering	Conditional expression of human *PRLHR*	(252)
Mouse	Genetic engineering	Conditional overexpression of β-catenin	(253)
Mouse	Genetic engineering	Conditional expression of SV40 large T antigen	(148)
Mouse	Genetic engineering	Conditional expression of PyLT antigen	(149)
Mouse (SCID)	Xenotransplantation	Adenoviral transduction with *COX2* & *VEGFA*	(254)
Mouse (SCID)	Xenotransplantation	Estrogen supplementation	(255)
Mouse (SCID)	Xenotransplantation	Ovariectomy, estrogen and progesterone supplementation, renal capsule	(256)
Mouse (SCID)	Xenotransplantation	Estrogen and progesterone supplementation	(257)
Mouse (SCID)	Xenotransplantation	Lentiviral transduction with SV40 T antigens, ovariectomy, hormone supplementation, matrigel supplementation with EGF, bFGF, & insulin	(258)
Mouse (SCID)	Xenotransplantation	Transfection of truncated *HMGA2* in myometrial cells	(247)
Rat (Wistar)	Xenotransplantation	Immunosuppressant drugs (mycophenolic acid), estrogen supplementation	(259)
Mouse (CD-1)	Drugs	Prenatal exposure to diethylstilbestrol	(142)
Guinea pig	Spontaneous	-	(250)
German Shepherd	Spontaneous	Germline *Bhd* mutation	(248)
Pot-bellied pig	Spontaneous	-	(260)
Baltic gray seal	Spontaneous	-	(261)

3.9 Signaling pathways

Whereas genetic defects are likely to initiate leiomyoma formation, it is unclear how these defects mediate neoplastic transformation. Our understanding of the molecular mechanisms underlying leiomyoma development has been somewhat limited by the lack of suitable *in vitro* cell cultures. Cultured leiomyoma cells tend to grow more poorly and senesce more quickly than the normal myometrial cells do.[262] Microarray studies have, however, provided insights into the molecular signatures of leiomyomas, revealing hundreds of differentially expressed genes between leiomyomas and normal myometrial tissue.[203, 214, 263-268] Although some of these genes are likely to act as effectors or promoters of leiomyoma growth, which of these are critical for tumorigenesis remains unclear.

Although leiomyomas are clearly ovarian steroid-dependent tumors, the molecular mechanisms by which these hormones promote leiomyoma growth remain undetermined. Their mitogenic effects may be mediated through local production of growth factors, cytokines, and chemokines.[269] However, evidence for increased levels of estrogen or progesterone receptors in leiomyomas are inconsistent,[143] and genes related to estrogen or progesterone signaling are not among the most significantly enriched genes in leiomyomas.[263] Furthermore, the presence or absence of steroid hormones does not seem to alter significantly the global gene expression pattern of leiomyomas.[203]

Several growth factors and growth-factor receptors are differentially expressed in leiomyomas.[269] Growth factors generally act over short distances and typically promote growth by binding to cell-surface receptors that activate signaling transduction cascades. Growth factors shown to influence leiomyoma growth include insulin-like growth factor (IGF), transforming growth factor alpha (TGF-α), transforming growth factor beta (TGF-β), heparin binding epidermal growth factor (HB-EGF), acidic fibroblast growth factor (aFGF), basic fibroblast growth factor (bFGF), platelet-derived growth factor (PDGF), and vascular endothelial growth factor (VEGF).[270] These appear to activate signaling pathways such as PI3K/AKT/mTOR, EGFR/MAPK, or TGF-β/Smad.[270]

The PI3K/AKT/mTOR pathway is of particular interest, because a recent study showed that the inhibition of RE1 silencing transcription factor (REST) induces upregulation of prolactin-releasing peptide receptor (PRLHR), leading to increased cell proliferation through activation of this pathway.[252] Using a transgenic mouse model, they demonstrated that conditional overexpression of human *PRLHR* results in the development of a uterus with phenotypic features similar to that of human leiomyomas, including myometrial hyperplasia and excessive extracellular matrix deposition.[252] A microarray study showed that leiomyomas from humans and Eker rats share dysregulation of genes that are part of the mTOR pathway.[267] Furthermore, *Tsc2* is a well-known negative regulator of the mTOR pathway. Progestins (synthetic progesterones) promote leiomyoma cell survival through activation of the AKT pathway,[271] whereas inactivation of AKT (also known as Protein kinase B) induces cellular senescence.[272] Phosphatidylinositol-3 kinase (PI3K) and mammalian target of rapamycin (mTOR) appear to be required for estradiol promotion of G1 cell-cycle progression and leiomyoma cell proliferation.[273]

The Wnt/β-catenin pathway is another pathway considered important in development of leiomyomas. β-catenin is, under normal conditions, regulated by the tumor suppressor gene *APC* and its loss results in stabilization of β-catenin, leading to nuclear localization of β-catenin and activation of Wnt target genes.[274] A mouse model that constitutively expresses β-catenin in the uterine mesenchyme develops tumor nodules similar to those of human uterine leiomyomas.[253] These lesions show high levels of mTOR and its downstream target phospho-S6-kinase.[253] Conversely, selective removal of β-catenin from mesenchymal tissue results in reduced uterine size and replacement of myometrial cells by adipocytes.[275] A recent study showed that cells surrounding the leiomyoma side population activate the Wnt/β-catenin in a paracrine manner.[276] Another recent study showed that inhibition of the canonical Wnt pathway attenuates cell growth in primary leiomyoma cell cultures.[277]

Microarray studies have consistently shown that genes related to extracellular matrix are predominantly differentially expressed in leiomyomas, suggesting their involvement in the formation of a myofibroblast phenotype.[266] Treatment of myometrial cells with TGF-β promotes the expression of genes encoding for extracellular matrix components, and reduces the expression of genes that degrade extracellular matrix.[278]

Only a few studies have accounted for the genetic background of leiomyomas when examining for differentially expressed genes and dysregulated pathways. Microarray studies have revealed that genetic defects can have profound consequences on the global gene expression profiles, since leiomyomas with *FH*, *HMGA2*, or 1p abnormalities show distinct gene expression patterns.[203, 264, 265] One study reported the SRF-FOS-JUNB pathway to be downregulated in leiomyomas that are deficient for *FH*,[279] whereas another study highlighted a significant dysregulation of genes related to G1/S checkpoint regulation among leiomyomas that harbored an *HMGA2* translocation.[203]

AIMS OF THE STUDY

1) To characterize the spectrum of somatic substitutions and microindels that drive uterine leiomyoma development by exome sequencing

2) To characterize the spectrum of chromosomal abnormalities that drive uterine leiomyoma development by whole-genome sequencing and SNP arrays

Discoveries from these studies clarified the subsequent aims:

3) To study the frequency and genomic architecture of clonally related leiomyomas by genome-wide methods

4) To explore genome-wide transcriptional differences and similarities between leiomyomas harboring distinct genetic drivers

MATERIALS AND METHODS

1. Study material and ethical issues (I-IV)

The study material consisted of fresh-frozen uterine leiomyoma and corresponding myometrial specimens from five separate collections and from two HLRCC patients (Table 8). All tissue specimens were collected during hysterectomies in Finland between 2001 and 2015. One of the collections (Series M) consisted of tissue specimens from anonymous patients with the approval of the director of the health care unit. The other specimens were obtained from patients who signed an informed consent upon entering the study. One of these (Series My5000) consisted of a prospective set of leiomyomas that was carefully collected in order to include all leiomyomas at least five mm in diameter. We performed a thorough investigation on these patients' medical records as well.

The Ministry of Social affairs and Health in Finland, and the Ethics Committee for gynecology and obstetrics, pediatrics and psychiatry of the Hospital Distinct of Helsinki and Uusimaa, Finland approved all of the studies.

Table 8. Overview of sample series utilized in the studies

Series ID	Leiomyomas	Patients	Hospital	Time frame	Informed consent
M	211	74	Helsinki University Central Hospital	2001-2002	No
B7 (HLRCC)	6	1	Helsinki University Central Hospital	2002	Yes
My	160	64	Helsinki University Central Hospital	2003-2008	Yes
N7 (HRLCC)	6	1	Turku University Hospital	2007	Yes
My1000	55	12	Helsinki University Central Hospital	2012	Yes
My5000	125	18	Helsinki University Central Hospital	2013	Yes
My6000	544	184	Helsinki University Central Hospital	2013-2015	Yes

2. Histopathological evaluation (I-IV)

Sections of fresh-frozen leiomyoma tissues were stained with hematoxylin and eosin according to standard procedures. A pathologist evaluated the tumor slides for number of mitotic figures per 10 high-power fields, degree of cellularity and nuclear atypia, and presence of tumor cell necrosis. As expected, all tumors showed a very high tumor percentage (>90%).

3. DNA, RNA, and cDNA preparation (I-IV)

DNA was extracted with the FastDNA Kit (MP Biomedicals). RNA was extracted with TRIzol Reagent (Invitrogen) or Tri Reagent RT (Molecular Research Center), and purified by the RNeasy MinElute clean up kit (Qiagen). The concentration and the purity of the extracted RNA were measured by the Agilent 2100 Bioanalyzer (Agilent). Extracted RNA was converted to cDNA according to standard procedures.

4. Genome-wide methods (I-IV)

Several high-throughput methods were utilized to study leiomyomas on a genome-wide scale (Table 9). The corresponding myometrium samples were included in all experiments.

Table 9. Overview of the number of tumors and the type of genome-wide tools used in each study

	Whole-exome sequencing		Whole-genome sequencing		Expression microarray		SNP array		RNA sequencing	
Study	**Tumors**	**Patients**	**Tumors**	**Patients**	**Tumors**	**Patients**	**Tumors**	**Patients**	**Tumors**	**Patients**
I	18	17	-	-	10	10	-	-	-	-
II	-	-	38	30	38	30	-	-	-	-
III	-	-	5	2	4	1	69	11	-	-
VI	-	-	63	45	94	60	50	36	2	2

4.1 Whole-exome sequencing (I)

Genomic DNA libraries were prepared using NEBNext DNA Sample Prep Reagent Set 1 Kit (New England Biolabs). Whole-exome capture was carried out using Agilent SureSelect Human All Exon Kit (Agilent). The libraries were paired-end sequenced using the Genome Analyzer II (Illumina) at the Institute of Molecular Medicine Finland (FIMM). The length of the reads was 80 to 82 nucleotides. The exome capture kit targeted approximately 38 megabase pairs (Mbps) of coding DNA.

Sequencing-read quality control was assessed using FastQC.[280] One of the tumor samples was re-sequenced due to its poor quality. Preprocessing and data analysis was carried out using the NextGENe v2.1 software (Softgenetics). The adapter sequences were trimmed with an in-house-developed script, and the sequencing reads were aligned against the hg19 reference genome. All tumor calls were filtered against the respective myometrium samples, against an in-house control set of 156 exomes sequenced with the same platform, and against polymorphisms reported in The Single Nucleotide Polymorphism database (dbSNP) build 132 and Ensembl release 59. The following criteria served to call a somatic mutation: a coverage of at least 4, an allelic fraction of at least 0.2, and a NextGENe mutation score of at least 6. We excluded silent mutations, but included mutations located at a maximum of three base pairs upstream or downstream of exon boundaries.

4.2. Whole-genome sequencing (II-IV)

Paired-end whole-genome sequencing was performed with either Complete Genomics (CG) or Illumina massively parallel sequencing technology. One tumor-normal pair (MY64 m1) was sequenced with both platforms. Genomic DNA libraries were prepared and sequenced according to their respective sequencing service protocols. The Illumina samples had a median sequencing coverage of at least 30, and consisted of reads that were 100 bps in length. The CG samples had a sequencing coverage of at least 40.

4.2.1 Detection of somatic substitutions and microindels (II-IV)

No data pre-processing was necessary with the CG samples, since their service provided us with substitution and microindel calls. Illumina-aligned (ELAND) samples were realigned with GATK IndelRealigner, duplicate reads were filtered out with Picard Tools, and substitutions and microindels were detected with VarScan 2.[281] Variations detected in the tumors were filtered against the corresponding myometrium samples, against 93 Finnish individuals from the 1000 Genomes Project (phase 1 release v2), against 157 in-house exome or whole-genome controls, and against polymorphism reported in the dbSNP Build 132. The following criteria served to call a somatic mutation: a coverage of at least 6, and an allelic fraction of at least 0.2. The minimum quality score (phred) was 29 for Illumina and 94 for CG samples. Somatic mutations located within protein-coding genes were annotated with Ensembl version 69.

Some modifications were made to the whole-genome-sequencing pipeline in Studies III and IV. All Illumina whole-genome sequencing samples (including all of the Illumina samples from Study II) were aligned against the 1000 Genomes phase 2 human reference sequence assembly hs37d5 using BWA (v0.6.2) with parameters -n 0.06 (maximum edit distance) and -q 5 (read trimming).[282] Microindels were detected by GATK SomaticIndelDetector with default parameters. Substitutions were called by MuTect 1.1.4 with default parameters.[283] Variations detected in the tumors were filtered against germline variants present in a whole-genome sequencing dataset of 139 Finnish controls. The following criteria served to call a somatic mutation: a coverage of at least 4, and a quality score of at least 20. Somatic mutations located within protein-coding genes were annotated with Ensembl version 71.

In Study III, four clonally related leiomyomas were processed further in order to gain more informative data regarding any genetic similarities or differences between the four related tumors. Here we also considered somatic substitutions located in non-coding regions, and therefore maximized specificity by exploring only those regions of high mappability (1000 Genomes Phase 1 Strict Accessibility Mask). We randomly selected 10 myometrium whole-genomes as additional controls; a somatic call was filtered out if any of these controls contained a single read harboring exactly the same substitution. We also explored whether any of the remaining mutations were present in any of the other tumors but not called by MuTect.

4.2.2 Detection of somatic copy-number alterations (II-IV)

Raw copy-number ratios were generated from Illumina tumor-normal pairs with VarScan 2,[281] and segmented by DNAcopy with a minimum of 5 markers (min.width=5).[284] For each sample, the segments were subtracted against the average of all their segments. Segments (including CG-provided segments) less than 0.3 were defined as a double deletion, 0.3 to 0.8 as a deletion, 1.2 to 1.7 as a duplication, and greater than 1.7 as a triplication. Segments at a maximum of 2 kbps apart were merged, segments located within chromosomal gaps (UCSC Gap track table) were excluded, and segments less than 100 kbps were filtered out.

Study IV included some modifications: raw copy-number ratios were generated, $\log_2$ transformed, corrected for GC-waves, smoothed, and segmented with Partek Genomic Suite™ v. 6.5 (Partek Incorporated). Segments were called with the following parameters: a minimum of 500 markers, a signal-to-noise of 0.5, and a segmentation p-value of 0.0001. Segments with a $\log_2$-ratio less than –0.2 were assessed as deletions and greater than 0.2 as amplifications. Segments located within noisy regions (such as telomeres) were assessed as technical artifacts and manually filtered out.

4.2.3 Detection of chromosomal rearrangements (II-IV)

Chromosomal rearrangements were detected from Illumina-aligned data with BreakDancer 1.2.[285] Tumor samples were run with the following parameters: a mapping quality of at least 65, and at least 4 supporting read pairs. Myometrium samples were run with the following parameters: a mapping quality of at least 1, at least 1 supporting read pair, and an output score of at least 1. All tumor calls (including the CG-provided rearrangements) were filtered against the myometrium calls; a tumor call was filtered out if both of its breakpoints were present in any of the myometrium samples within 1000 bp window. Calls reported as deletions or amplifications were similarity filtered against the Database of Genomic Variants.[286] Tumor calls were also filtered out if any one of the two breakpoints were located within a segmental duplication[287] or a region of exceptionally high sequencing depth[288]. Tumor calls were manually inspected with the Integrative Genomics Viewer, and filtered out if it resided within a repetitive or poorly aligned region, or if a similar rearrangement were observable in the paired myometrium sample but not called by BreakDancer.

4.2.3.1 Detection of interconnected chromosomal rearrangements

In order to detect interconnected rearrangements likely to have occurred simultaneously, we designed a novel computational algorithm that considered the proximity and orientation of rearrangement calls. An event graph – consisting of nodes and edges – was constructed for each sample. Each double-strand break creates two nodes, representing the left (head) and the right (tail) end of a double-strand break. If the two nodes are at most 1000 bps apart, and if one of the nodes is in the head orientation, whereas the other is in the tail orientation then the two nodes are connected by an edge. If a deletion is present between two nodes, then the distance between the nodes is shortened by the length of the deletion.

The next step was to identify components representing complex chromosomal rearrangements. Components corresponds to the total amount of nodes that are linked to each other by edges. Components that comprised four or more nodules were statistically assessed with a permutation test. Components representing a balanced translocation were manually excluded. The test statistic was the total number of breakpoint ends involved in each component. The breakpoint data were permuted by randomly repositioning each breakpoint end within the original chromosome. After 100,000 permutations, an empirical p-value was computed as p = x/n, where x is the number of cases, n is the number of all cases, and the random test statistic is higher than the observed value.

4.3. Gene expression microarrays (I-IV)

The gene expression samples in Study I were analyzed using Affymetrix GeneChip Human Genome U133 Plus 2.0 arrays (Affymetrix), whereas the samples in Studies II and IV were analyzed using Affymetrix GeneChip Human Exon 1.0 ST (Affymetrix). Hybridization and quality control were performed at the Biomedicum Functional Genomics Unit (FuGU) according to the instructions provided by Affymetrix.

All of the gene expression data analyses were performed with Partek Genomic Suite™ v. 6.5. In Study I, the gene expression data were normalized by the GC-RMA method (Gene Chips Robust Multichip Analysis), and the probes were mapped using Brainarray Custom CDF files (HGU133Plus2_Hs_ENSG,v.14.1.0). Unsupervised hierarchical clustering analysis (Euclidian distance, average linkage) used the 10% most variable genes, defined by the coefficient of variation. Differential expression analysis was carried out by a paired t-test. False discovery rate control (FDR, Benjamini and Hochberg method) served to correct for multiple testing.[289] Genes with an FDR (*q*-value) less than 0.05 and a fold change (FC) of 1.5 were considered significant. Pathway enrichment analysis was carried out with Webgestalt using the Kyoto encyclopedia of genes and genomes (KEGG) pathways.[290] Pathways with a *q*-value less than 0.05 were considered significant.

In Studies II to IV, the gene expression data were quantile-normalized by the RMA method (Robust Multichip Average), and adjusted for probe sequence and for GC-content using Partek. Probes were mapped with Brainarray Custom CDF files (HuEx10stv2_Hs_ENSG, Version 16). Technical replicates were averaged, and Partek's Batch effect removal algorithm removed technical noise originating from the five different batches. Unsupervised hierarchical clustering analysis (Cosine correlation) used the 1% most variable genes (n = 372), defined by the coefficient of variation.

A paired t-test served to identify genes differentially expressed in all leiomyomas as compared to the corresponding myometrium samples. A one-way analysis of variance (ANOVA) test identified genes differentially expressed between leiomyomas of different genetic subtypes and the myometrium samples. Multiple test correction was carried out with FDR. Genes with a *q*-value less than 0.05 and a FC of 2 were considered significant. Pathway enrichment analysis was carried out using QIAGEN'S Ingenuity Pathway Analysis

software. To identify the most uniquely expressed genes for four different leiomyoma subtypes, a two-way ANOVA test (factors: individual and mutation status) was constructed that compared each subtype against the rest of the leiomyoma and myometrium samples. These genes have a subtype-specific expression level that differs from levels in both the other leiomyoma and myometrium samples. Exon-level analysis of gene expression data were performed with default parameters using Affymetrix-provided annotations (v35, GRCh37/hg19). To identify genes most significantly downregulated by chromosome 7q22, 22q, and 1p deletions, three separate two-way ANOVA tests compared leiomyomas harboring the abnormality in question against those myometrium and leiomyoma samples that lacked the abnormality.

4.4 SNP arrays (III-IV)

All samples were analyzed using Illumina HumanOmni2.5-8 BeadChips (Illumina) arrays. Sample preparation, hybridization, and quality control were performed at the Estonian Genome Center, University of Tartu or at Illumina.

The SNP arrays were pre-processed with Genome Studio, and the data analysis performed with Partek Genomic Suite™ v. 6.5. Raw copy-number ratios were constructed, $\log_2$ transformed, and corrected for GC waves. Partek's segmentation algorithm was performed by two different sets of parameters. In the first set, we used stricter parameters: a minimum of 50 markers, a signal-to-noise of 0.5, and a segmentation p-value of 0.0001. Segments with a $\log_2$-ratio less than –0.2 were defined as deletions and greater than 0.2 as amplifications. The second set was utilized to detect large subclonal SCNAs. Here we used the following parameters: a minimum of 1000 markers, a signal-to-noise of 0.2, and a segmentation p-value of 0.0001. Additional segments (not detectable with the first set of parameters) with a $\log_2$-ratio less than –0.05 were defined as subclonal deletions and greater than 0.05 as subclonal amplifications. Regions of CN-LOH were detected from genotype calls using Partek's LOH algorithm with parameters: max probability of 0.99, genomic decay of 0, and genotype error of 0.001. A segment was defined as a CN-LOH event if LOH was detectable but no change in copy-number. All regions of SCNAs and CN-LOH were visually inspected (by $\log_2$-ratio and b-allele frequency plots), and segments assessed as technical artifact were manually filtered out. Only regions larger than 20 kbps were reported, and regions less than 200 kbp apart were merged.

4.5 RNA sequencing (IV)

RNA sequencing libraries were prepared from rRNA-depleted (RiboMinus Transcriptome Isolation Kit, Life Technologies) samples with Illumina TruSeq RNA Sample Preparation kit A (Illumina) according to the manufacturer's instructions. Quality control was performed with FastQC v0.10,[280] and the data were aligned against the human reference genome hg19 by TopHat v1.4.1.[291]

5. Sanger sequencing validation (I-IV)

Sanger sequencing served to validate genes recurrently affected by somatic substitutions or microindels that were detectable by exome or whole-genome sequencing. To confirm that the detected mutations were somatic, we performed Sanger sequencing with the corresponding myometrial tissues as well. Chromosomal rearrangements of interest with low-confidence calls were also subject to Sanger sequencing validation. Oligonucleotide primers were designed with Primer3 using hg19 as the reference genome.[292]

DNA fragments were amplified with the AmpliTaq Gold DNA Polymerase (Applied Biosystems) and purified with the ExoSAP-IT PCR Purification Kit (USB Corporation). The sequencing was performed with the BigDye Terminator v.3.1 Kit (Applied Biosystems) on an ABI3730 Automatic DNA Sequencer (Applied Biosystems) at FIMM, or an ABI3100 Capillary Sequence Analyzer (Applied Biosystems) at the Sequencing Core Facility at the Haartman Institute. DNA sequences were analyzed with the Mutation Surveyor (Softgenetics). Sequencing graphs were also manually inspected for any mutations not automatically detected by the software.

6. *In silico* prediction tools (I)

Alternative Splice Site Predictor (ASSP)[293] and NetGene2[294] served in predicting whether intronic mutations affected splicing. Multiple sequence alignment analysis was performed with Muscle[295] for the following species: human (Q93074), mouse (A2AGH6), zebrafish (Q2QCI8), fruit fly (Q9VW47) and baker's yeast (P25648). Intronic sequence alignment was performed with Multiz,[296] and secondary structure prediction was performed with PSIpred.[297]

RESULTS

1. Exome sequencing reveals high frequency of *MED12* mutations (I)

To characterize the spectrum of somatic substitutions and microindels that underlie leiomyoma etiology, we exome sequenced 18 uterine leiomyomas and corresponding myometrial tissues from 17 Finnish patients. This revealed ten leiomyomas harboring a mutation within exon 2 of *mediator complex subunit 12* (*MED12*). *MED12* was the only recurrently mutated gene, and all of the mutations were verified with Sanger sequencing. Sanger sequencing of 207 additional leiomyomas revealed *MED12* exon 2 mutations in 159 out of 225 (70.7%) leiomyomas from 80 patients (Figure 5). The majority of mutations were missense, and 110 of them affected codon 44 – leading to all six possible amino acid changes – followed by 11 affecting codon 36, and 3 affecting codon 43. Ten leiomyomas harbored an intronic T to A mutation located eight base pairs (bps) upstream of exon 2. This change was predicted to create a premature splice acceptor site. cDNA sequencing confirmed that this mutation resulted in a transcript that included the last six bases of intron 1. Several microindels – ranging from 3 to 43 bps – were detected in 25 leiomyomas, all of which were predicted to result in an in-frame transcript. *In silico* prediction tools revealed these mutations to affect an evolutionarily conserved region of the gene.

To examine whether other parts of *MED12* were mutated as well, we sequenced all 45 exons in 10 leiomyomas harboring an exon 2 *MED12* mutation, and in 20 leiomyomas lacking a *MED12* mutation, revealing no additional mutations. Since *MED12* is located on the X chromosome, one of the two alleles is epigenetically silenced. cDNA sequencing of 16 leiomyomas with a *MED12* mutation revealed that the mutant allele was predominantly expressed in all cases. We also examined whether these mutations correlate with clinical features, revealing an inverse association between *MED12* mutations and tumor size (two-sided Wilcoxon rank sum test, $p = 0.015$).

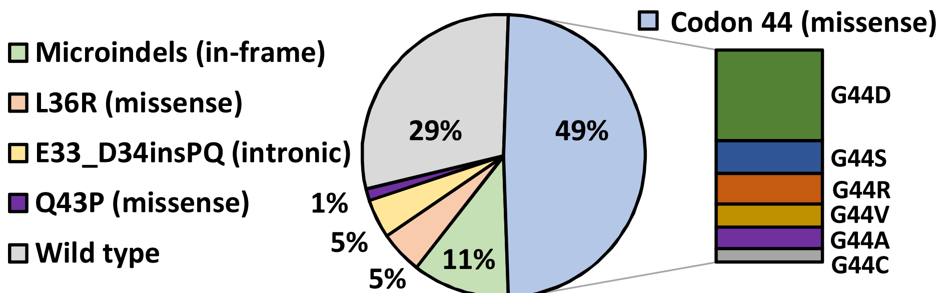


Figure 5. The spectrum of *MED12* mutations among 225 uterine leiomyomas.

2. Clonally related leiomyomas are relatively common (II-III)

In Study II, we explored the genomic landscape of 38 uterine leiomyomas from 30 patients by whole-genome sequencing. This sample set was selected to include 16 leiomyomas with a *MED12* mutation, 4 with biallelic loss of *FH*, and 18 that lacked both of these driver defects. Four of the patients had two or more leiomyomas that were whole-genome sequenced.

An unexpected finding was that two tumor pairs from two patients (MY18 & M44) displayed several identical chromosomal abnormalities. Based on the number of discordant read-pairs of chromosomal rearrangements, and the average log_2-ratio of deletions, we inferred that some chromosomal abnormalities were clonally present in MY18 m3 and M44 m2, whereas the same abnormalities were only subclonally present in MY18 m2 and M44 m1. None of the abnormalities detected in MY18 m2 and MY18 m3 were present in the third tumor (MY18 m1) from the same patient. Among the 30 patients included in the study, four (MY18, M29, M38, & M68) had additional tumors displaying either the same or no *MED12* mutation with respect to any of the other tumors that were whole-genome sequenced from the same patient. These tumors were analyzed further using Sanger sequencing for point mutations found in the whole genomes. This revealed three additional tumors (MY18 m4, MY18 m5, & MY18 m6) to be clonally related to MY18 m2 and MY18 m3, and one tumor (M38 m1) to be clonally related to M38 m5. We found altogether nine clonally related leiomyomas, each of which lacked any mutations in *MED12* or *FH*.

Using SNP arrays and *MED12* Sanger sequencing in Study III, we investigated the frequency of clonally related leiomyomas in a prospective set of 103 leiomyomas collected from 14 consecutive patients who underwent hysterectomy due to symptomatic lesions. One of these patients (MY5006) was diagnosed as an HLRCC patient, since she harbored a germline mutation in *FH* (c.583A>C, p.M195V). Leiomyomas with a different *MED12* mutation were assumed to be independent, and only tumors with an identical *MED12* mutation status were analyzed further with SNP arrays. A total of 69 tumors from 11 patients were examined for shared and unique SCNAs and CN-LOH. We detected four of these patients as having clonally related leiomyomas, including the HLRCC patient.

In 103 leiomyomas from 14 patients, we detected 13 (13%) clonally related leiomyomas in 4 (29%) patients. Each clonally related leiomyoma lacked a *MED12* mutation (Fisher's exact test, $p = 0.006$). Out of the 42 leiomyomas with a *MED12* mutation, only 19 displayed SCNAs, none of which was identical between tumors from the same patient. Since the remaining 23 tumors harbored no detectable SCNAs, these tumors could not be investigated for a clonal relationship. A careful review of all the patients' medical records revealed prior myomectomy in two of the four patients (MY5005 & MY5006) with clonally related tumors. In contrast, prior myomectomy was confirmed in only one of the ten patient (MY5013) that lacked clonally related leiomyomas. None of the clonally related leiomyomas displayed unusual histopathology.

2.1 Intratumor genetic heterogeneity and DEPDC5 mutations (III)

To gain insight into the genomic architecture of clonally related leiomyomas, we performed whole-genome sequencing with four such lesions (MY21 m1-m4) from one patient. All four tumors were confirmed as clonally related, since they harbored identical chromosomal abnormalities on chromosomes 3, 12, 14, and 22. Furthermore, all four tumors shared 295 identical somatic substitutions. We interpreted tumor m4 to be the primary tumor consisting of several genetically distinct subpopulations (Figure 6). Using SNP array data, we were also able to infer intratumor genetic heterogeneity in two other sets of clonally related leiomyomas from two other patients (MY5014 & MY5006). See Figure 11B-C for the interpretation of the tumors from patient MY5006.

We unexpectedly detected a different truncating *DEP domain containing 5* (*DEPDC5*) mutation in each of the four clonally related leiomyomas. Furthermore, all four tumors displayed an identical deletion on chromosome 22, resulting in a biallelic loss of *DEPDC5*. We found no other gene to be recurrently mutated among these four tumors. All four tumors displayed, however, an identical translocation between *HMGA2* and *RAD51B*, resulting in *HMGA2* overexpression. We searched for *DEPDC5* mutations in an in-house whole-genome sequencing dataset of 65 leiomyomas, revealing one additional leiomyoma (M26 m1) to harbor a truncating *DEPDC5* mutation. This sample also displayed a deletion of the whole chromosome 22, and an *HMGA2-RAD51B* translocation. All five *DEPDC5* mutations were validated with Sanger sequencing (Figure 10).

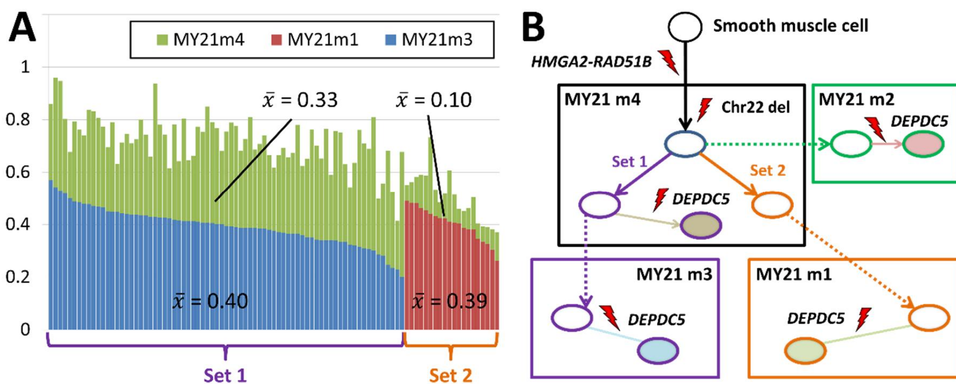


Figure 6. Genomic architecture of four clonally related leiomyomas from patient MY21. A) Two sets of substitutions (Set 1 & Set 2) were present only in two out of the four clonally related tumors from patient MY21. Set 1 (n = 75) was present in m3 and m4, whereas Set 2 (n = 20) was present in m1 and m4. Since both of these sets were present in tumor m4 and at a much lower average allelic fraction, we interpreted this lesion to consist of several genetically distinct subpopulations. **B**) A likely explanation for this pattern is that tumors m3 and m1 (and probably m2) arose from a different subpopulation that originated from tumor m4. Additionally, each tumor gradually acquired a unique set of mutations, including a separate inactivating mutation in *DEPDC5*.

3. Whole-genome sequencing reveals complex chromosomal rearrangements (II)

Whole-genome sequencing of 38 uterine leiomyomas revealed a subset of leiomyomas as displaying highly complex chromosomal rearrangements (CCRs) resembling chromothripsis. Compatible with chromothripsis, these rearrangements were locally clustered, involved one to four chromosomes, and displayed alternating copy-number states between breakpoints (Figure 7A). A CCR was defined as a series of rearrangements interconnected by a minimum of three double-strand breaks involving six DNA ends. Using an in-house-developed computational algorithm to identify such cases among the 38 leiomyomas, we identified 17 as harboring one or more CCRs. CCRs were significantly more common among leiomyomas that lacked mutations in *MED12* or *FH* mutations (two-tailed Fisher's exact test, p <0.001). Indeed, none of the four *FH*-deficient leiomyomas, and only three of the 16 leiomyomas with a *MED12* mutation displayed a CCR.

Interestingly, we detected a few leiomyomas as harboring multiple separate CCRs. For example, we detected two spatially separate CCRs in two clonally related tumors (MY18 m2 & MY18 m3): one involving chromosomes 1, 2 and 20, the other involving chromosomes 12 and 14. Although a third CCR – involving chromosome 5 – was also detected in both tumors, this CCR was only subclonally present in MY18 m2, suggesting that it occurred as a temporally separate event. Another sample (MY23 m4) displayed two separate CCRs, one of which involved chromosome 2 and one copy of chromosome 7, whereas the other involved chromosome 5 and the other copy of chromosome 7.

Furthermore, some CCRs resulted in driver changes with a selective value, such as rearrangements between *HMGA2* and *RAD51B* (Figure 7B). Of note, most CCRs did not exhibit the high number of breakpoints typical for chromothripsis in the context of cancer. Indeed, an arbitrary threshold of at least 20 intrachromosomal breakpoints has often defined chromothripsis. Therefore, only five (MY10 m3, MY23 m4, MY46 m1, MY47 m1, & MY64 m1) of the 17 leiomyomas with a CCR reflected typical examples of chromothripsis. Sanger sequencing validation was performed with all of the rearrangements detected in three selected leiomyomas (MY47 m1, MY64 m1, & MY18 m3), each of which harbored a high number of breakpoints and one or more CCRs.

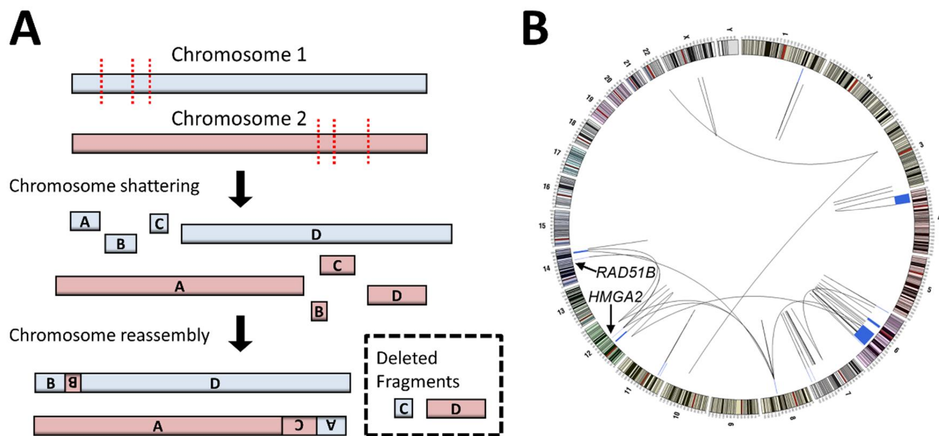


Figure 7. Complex chromosomal rearrangements (CCRs) resembling chromothripsis in a subset of leiomyomas. A) Chromothripsis is a phenomenon in which one or a few chromosomes are locally shattered and randomly reassembled in a single event. This process often leads to highly complex chromosomes and loss of DNA fragments. **B**) The circos plot above illustrates chromosomal rearrangements (lines) and deleted regions (blue rectangles) detected in a leiomyoma (MY64 m1). A chromothripsis event in this leiomyoma involved chromosomes 6, 8, 12, and 14. This event generated a driver rearrangement involving *HMGA2* and *RAD51B*. **B** is from *The New England Journal of Medicine*, **369**, Mehine et al., Characterization of uterine leiomyomas by whole-genome sequencing (2013), reprinted with the permission of Massachusetts Medical Society.

4. Chromosomal abnormalities create driver changes (II-IV)

As expected based on our exome-sequencing study, we identified no other genes besides *MED12* and *FH* to be recurrently mutated on nucleotide-level among the 38 leiomyomas whole-genome sequenced in Study II. We therefore searched for chromosomal rearrangements that could act as driver changes. In line with cytogenetic studies, the *HMGA2* locus was the region most commonly rearranged (n = 9). All of these rearrangements resulted in overexpression of *HMGA2*, confirming *HMGA2* as the target gene. The *RAD51B* locus was the second most rearranged region (n = 8), owing to the fact that the majority of *HMGA2* rearrangements involved this locus as a translocation partner. Using deletion and breakpoint mapping of the *RAD51B* locus, we concluded that a putative enhancer for *HMGA2* must be located within chr14:68,217,257–68,760,115. Only two samples (MY30 m1 & M38 m5) displayed an *HMGA2* rearrangement that targeted another region, both of which showed the weakest upregulation of *HMGA2*. One of these (MY30 m1) displayed additional breakpoints within the 3'UTR of *HMGA2*, providing an alternative mechanism for *HMGA2* overexpression. We identified one leiomyoma (M17 m1) as harboring a deletion of the whole chromosome 14 and a localized deletion within *RAD51B*, resulting in biallelic loss of *RAD51B*. This sample lacked the involvement of *HMGA2*.

A chromosomal region containing *collagen type IV alpha 5 chain* (*COL4A5*) and *collagen type IV alpha 6 chain* (*COL4A6*) was the third most commonly rearranged region. Three samples displayed a characteristic chromosomal deletion – previously reported in esophageal leiomyomas – that resulted in removal of the 5' ends of both *COL4A5* and *COL4A6*, which are located head-to-head on opposite strands (Figure 8). We found *MED12*, *FH*, *HMGA2*, and *COL4A5-COL4A6* abnormalities to be mutually exclusive, suggesting that these represent different molecular subtypes of leiomyomas. Leiomyomas lacking all of these four abnormalities are henceforth termed quadruple-negative leiomyomas.

Since none of the other recurrent rearrangements resulted in any noticeable driver changes, we explored for non-recurrent rearrangements around known driver genes. We identified and validated one quadruple-negative leiomyoma (MY22 m1) to harbor a rearrangement between *HMGA1* and *cardiac mesoderm enhancer-associated non-coding RNA* (*CARMN*, previously known as *MIR143HG*), resulting in overexpression of *HMGA1*. In one sample (MY47 m1), we detected and validated a rearrangement located upstream of *cyclin D1* (*CCND1*), resulting in the highest expression of this gene among all 38 leiomyomas analyzed. In another sample (MY10 m3), we detected and validated a rearrangement that simultaneously disrupted one copy of the tumor suppressor genes *TP53* and *NF1*.

In addition to the whole-genome sequencing datasets included in Studies II and III, Study IV involved 51 additional leiomyomas analyzed by whole-genome sequencing or SNP arrays or both. A subset of these samples were selected for the study since they were known (based on previous work) to lack a *MED12* mutation, to display biallelic loss of *FH*, or to harbor LOH at chromosome 22q. All of these samples were screened for rearrangements and deletions located within or close to *HMGA2*, *HMGA1*, or *COL4A5-COL4A6*.

We detected 18 of these to harbor an *HMGA2* rearrangement or *HMGA2* overexpression, or both, one of which (MY51 m3) harbored breakpoints within the 3'UTR of *HMGA2* and lacked the involvement of *RAD51B*. Two samples (M26 m1 & M31 m1) harbored an *HMGA2* rearrangement in which the breakpoints were located downstream of *RAD51B*. Using SNP arrays, we identified one quadruple-negative leiomyoma (MY6010 m3) and one leiomyoma of the *MED12* subtype (MY5008 m3) as harboring a deletion located upstream of *HMGA1*. We also detected one additional leiomyoma (M21 m1) as harboring a simple *COL4A5-COL4A6* deletion (Figure 8). Of note, one leiomyoma (M30 m1) of the *HMGA2* subtype exhibited a rearrangement within the *COL4A5* locus that did not result in the characteristic *COL4A5-COL4A6* deletion.

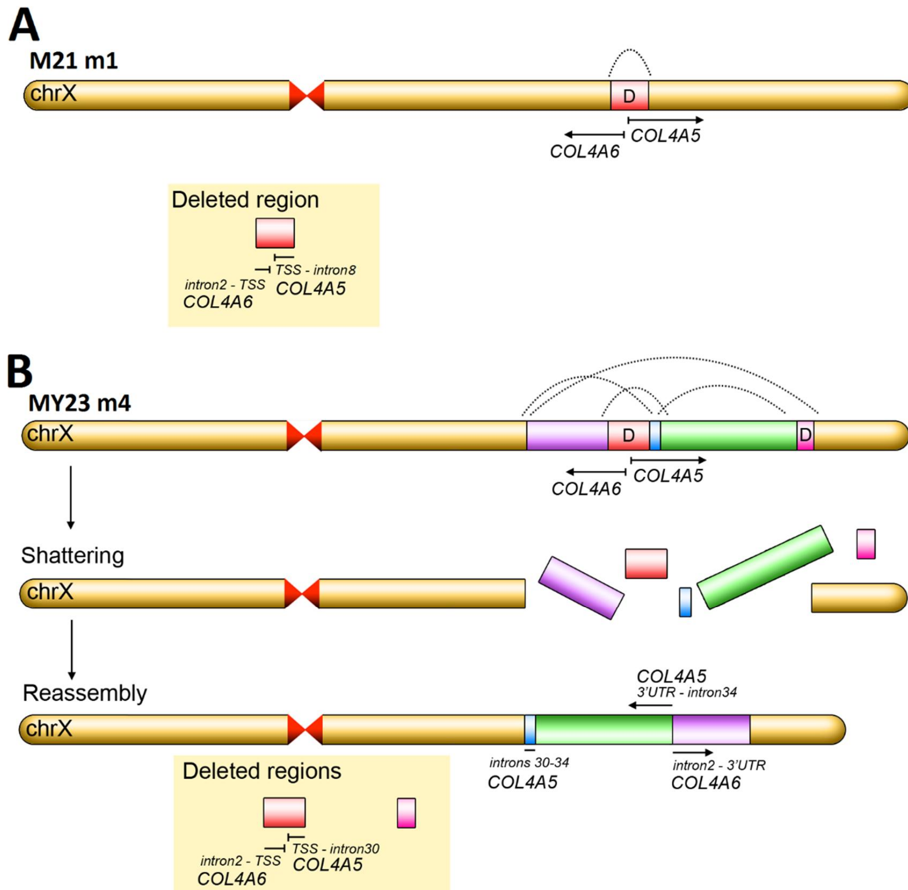


Figure 8. *COL4A5-COL4A6* abnormalities in two leiomyomas. A) We detected one leiomyoma (M21 m1) to harbor a simple deletion that resulted in the removal of the 5'ends of both *COL4A5* and *COL4A6*, located head-to-head on chromosome Xq22.3. **B**) Another sample (MY23 m4) displayed a similar abnormality that was generated by a chromothripsis-like event. Dashed lines represent rearrangements detected by whole-genome sequencing.

In addition to these recurrent abnormalities, we detected one leiomyoma (MY16 m1) of the *MED12* subtype to harbor a sole balanced translocation, t(6, 8)(q13;q12). The breakpoints were located ~2.3 kbp downstream of *PLAG1 zinc finger* (*PLAG1*) and ~21.9 kbp downstream of *collagen type XII alpha 1 chain* (*COL12A1*). Furthermore, two quadruple-negative leiomyomas (M51 m1 & MY5007 m2) harbored a large-scale amplification spanning *PLAG1*. We also identified one quadruple-negative leiomyoma (M18 m1) as harboring a rearrangement with breakpoints in intron 1 of *insulin like growth factor binding protein 5* (*IGFBP5*) and intron 10 of *platelet derived growth factor receptor beta* (*PDGFRB*). cDNA sequencing confirmed that this rearrangement resulted in an in-frame fusion joining exon 1 of *IGFBP5* to exon 11 *PDGFRB*.

4.1 Commonly deleted regions on chromosomes 7, 22 and 1 (II-IV)

In Studies II to IV, we explored commonly deleted regions on chromosome 7q. SCNA analysis revealed 14 leiomyomas harboring a deletion that spanned 7q22 (Figure 9A). We identified *CUX1* as the most commonly affected gene within this region. Sanger sequencing validation revealed one sample (MY23 m4) to harbor two separate rearrangements, each disrupting one copy of *CUX1*. We detected and validated another sample (M32 m8) to harbor a subclonal translocation between chromosomes 7 and 22, one that disrupted *CUX1*. Another leiomyoma (MY1 m1) displayed a 7q deletion and a second-hit 5 bps deletion in *zinc finger HIT-type containing 1* (*ZNHIT1*; Figure 9B). Two samples displayed targeted chromosomal rearrangements affecting *cullin 1* (*CUL1*) and *enhancer of zeste 2 polycomb repressive complex 2 subunit* (*EZH2*), resulting in significant downregulation of *CUL1* (FC = –1.7). Deletions of 7q were the SCNA most commonly detected in Study III, and these were present only in leiomyomas that harbored a *MED12* mutation (Fisher's exact test, p = 0.006). Four of these deletions were barely detectable (–0.2> log$_2$-ratio >–0.1), suggesting that they are secondary events relative to *MED12* mutations. *CUX1* was located in the minimally deleted region, and at the breakpoint end of one of the deletions (MY5013 m5).

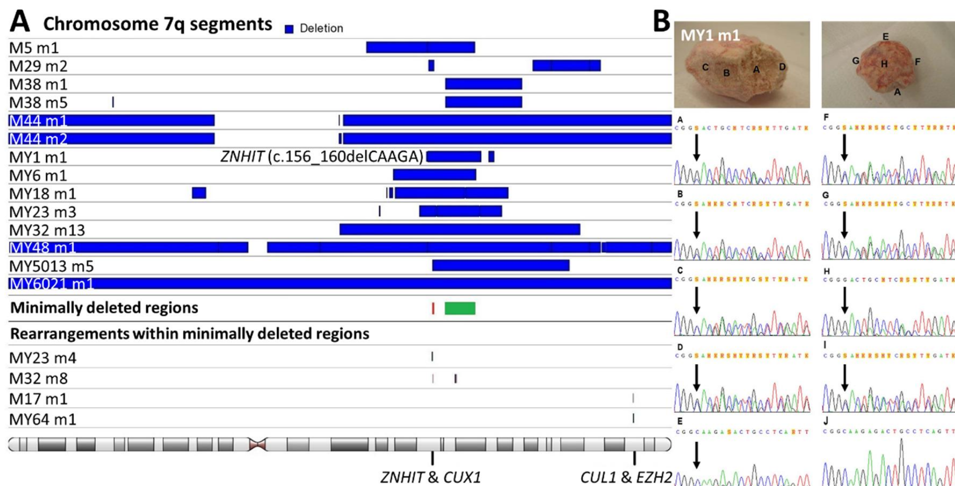


Figure 9. Chromosomal abnormalities and nucleotide-level mutations detected on chromosome 7q. **A**) We identified chr7:104,849,448–111,900,000 (green segment) as the minimally deleted region on 7q and chr7:101,732,303–102,100,000 (red) as another commonly deleted region. Two samples (MY23 m4 & M32 m8) displayed rearrangements within this region, both of which had rearrangements that disrupted *CUX1*, making it the gene most commonly affected on 7q. Furthermore, one of these (MY23 m4) displayed two separate rearrangements, each disrupting one copy of *CUX1*. Two samples (M17 m1 & MY64 m1) displayed targeted rearrangements disrupting *CUL1* and *EZH2*, located on 7q32. One sample (MY1 m1) harbored a 7q deletion and a five bp deletion in *ZNHIT* (c.156_160delCAAGA), resulting in biallelic loss of the gene. **B**) Since the five bp deletion appeared to be subclonal, we examined whether it was present in other parts of the tumor as well. We performed Sanger sequencing with nine different parts of the tumor, marked A-I. Parts H and I are on opposite sides of the tumor, and J represents the corresponding myometrium sample. The sequencing chromatograms show two distinct sequences in all nine parts, with variable ratios for the wild-type and mutant sequence. The start of the mutant sequence (lacking the reference nucleotides CAAGA) is marked with an arrow in each chromatogram.

In Study IV, we also searched for commonly deleted regions of recurrent 22q and 1p deletions. We detected 20 leiomyomas harboring a 22q deletion and one leiomyoma (MY23 m4) harboring chromosomal rearrangements within the most commonly deleted region (Figure 10). One of the rearrangements was located ~14 kbp upstream of the putative target gene *DEPDC5*. Interestingly, we detected another commonly deleted region on 22q. One sample (M9 m3) displayed a rearrangement within this region plus a deletion of the whole chromosome 22, resulting in biallelic loss of *SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1* (*SMARCB1*). We found 18 leiomyomas harboring a 1p deletion and only one protein-coding gene was located within the minimally deleted region: *NPHP4* (Figure 11A).

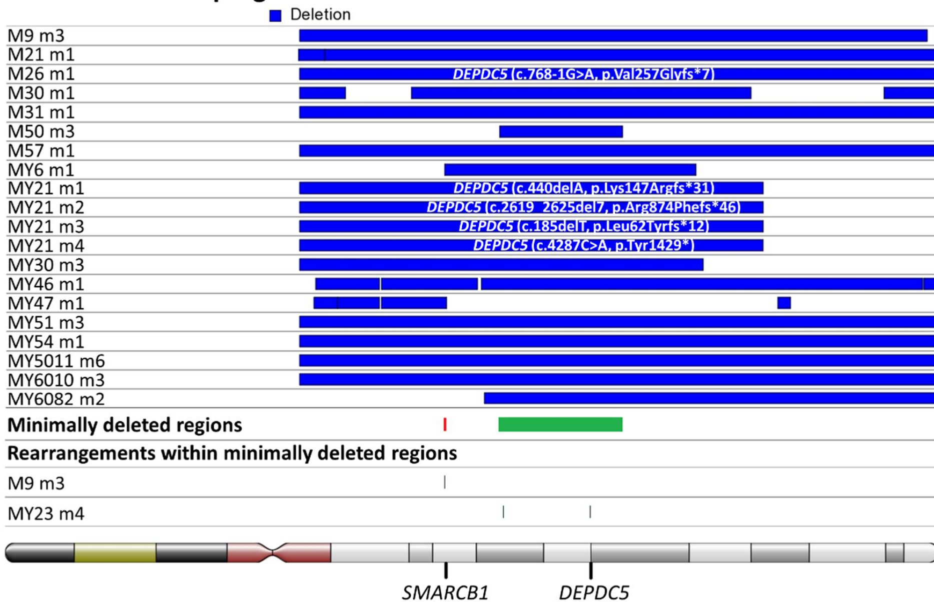


Figure 10. Chromosomal abnormalities and nucleotide-level mutations detected on chromosome 22q. One patient harbored four clonally related leiomyomas (MY21 m1-m4), each of which displayed an identical deletion on chromosome 22 and a different truncating mutation in *DEPDC5*, providing a unique second-hit mutation in each of the four tumors. Further examination revealed one additional tumor (M26 m1) to display biallelic loss of *DEPDC5*, through a truncating *DEPDC5* mutation and a deletion of the whole chromosome 22. In total, we detected 20 leiomyomas harboring a 22q deletion, and the minimally deleted region chr22:27,111,559–33,871,686 contained *DEPDC5*. We identified one additional leiomyoma (MY23 m4) to harbor chromosomal rearrangements within this region, one of which was located ~14 kbp upstream of *DEPDC5*. We identified chr22:24,087,031-24,200,000 as another commonly deleted region, and one leiomyoma (M9 m3) displayed a rearrangement in within this region in conjunction with a deletion of the whole chromosome 22, resulting in a biallelic loss of the well-known tumor suppressor gene, *SMARCB1*.

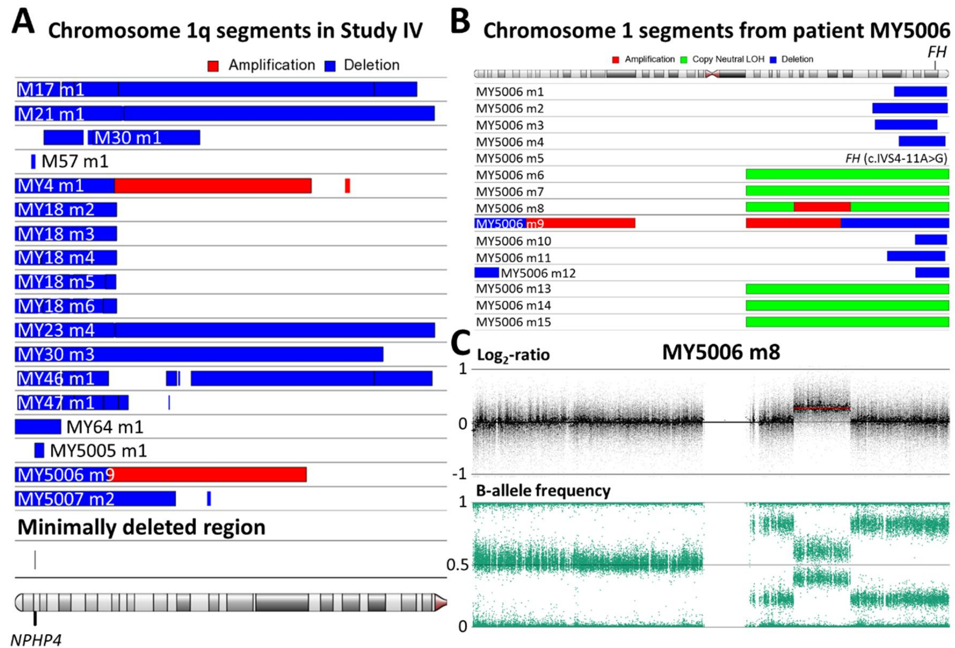


Figure 11. Deletions and amplifications affecting chromosome 1. A) In Study IV, we identified chr1:5,753,010–5,953,574 as the minimally deleted region on chromosome 1p. This region contained only one protein-coding gene: *NPHP4* **B**) In Study III, we identified a chromosome 1q deletion in 14 of 15 leiomyomas from one patient (MY5006). This patient was diagnosed with the HLRCC syndrome, because she was found to harbor a germline mutation in *FH* (c.583A>C, p.M195V). The remaining tumor (m5) displayed an inactivating point mutation in *FH* (IVS4-11A>G; predicted damaging). Six of these tumors displayed an identical CN-LOH event on chromosome 1q (green segments). **C**) One of the tumors (m8) displayed an additional amplification on chromosome 1. This amplification cannot have occurred prior to the CN-LOH event, since it was not present in the other five clonally related tumors. Nor can the amplification be a subsequent event, since the b-allele frequencies within this amplified region are heterozygous. Thus, the amplification must be present in a genetically distinct subpopulation not harboring the CN-LOH event.

5. Gene expression profiling reveals distinct expression profiles (I-II, IV)

In Study I, we performed gene expression profiling using 10 leiomyomas, 8 of which harbored a *MED12* mutation. This initial data indicated that leiomyomas with a *MED12* mutation display a global gene expression pattern differing from that of other leiomyomas. In Studies II and IV, we confirmed this observation by examining 94 leiomyomas by transcriptional profiling. Furthermore, we also found leiomyomas to cluster according to the mutation status of *HMGA2*, *FH*, and *COL4A5-COL4A6* (Figure 12).

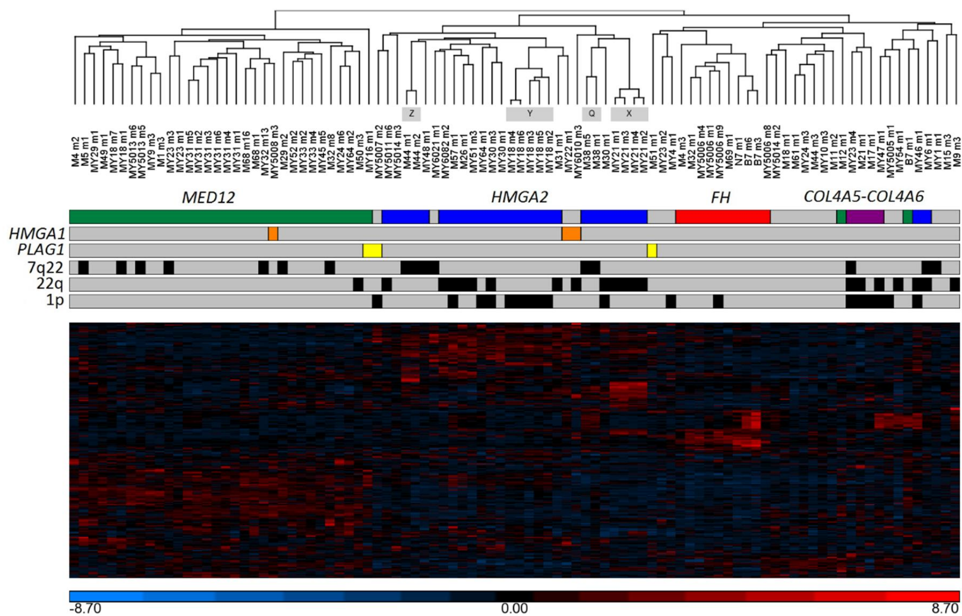


Figure 12. Unsupervised hierarchical clustering analysis of 94 leiomyomas from 60 patients. Hierarchical clustering analysis showed that most leiomyomas clustered together according to the mutation status of *MED12* (green), *HMGA2* (blue), *FH* (red), and *COL4A5-COL4A6* (purple). The remaining 19 quadruple-negative leiomyomas showed transcriptional heterogeneity and grouped into several different branches. Four of these clustered with leiomyomas of the *HMGA2* subtype, two of which harbored a genetic *HMGA1* abnormality (orange). One leiomyoma (MY5008 m3) harbored an *HMGA1* abnormality and a *MED12* mutation, consequently clustering with leiomyomas of the *MED12* subtype. Another leiomyoma (MY16 m1) harbored a *PLAG1* translocation and a *MED12* mutation, consequently clustering with leiomyomas of the *MED12* subtype. Two quadruple-negative leiomyomas (MY5007 m2 & M51 m1; yellow) displayed large-scale amplifications that spanned *PLAG1*, one of which (MY5007 m2) clustered with leiomyomas of the *HMGA2* subtype. Chromosomal abnormalities affecting 7q22, 22q, and 1p had no major impact on the clustering of samples. The sample set also included four sets of clonally related leiomyomas, each of which clustered within its own set (X, Y, Z, & Q).

5.1 Pathway enrichment analysis using differentially expressed genes (I, IV)

In Study I, we identified 924 genes (q <0.05; |FC| >1.5) to be differentially expressed between the eight leiomyomas with a *MED12* mutation and eight corresponding myometrium samples. Enrichment analysis using KEGG pathways revealed three significantly altered pathways: focal adhesion, ECM-receptor interaction, and the Wnt signaling pathway.

In Study IV, we compared 94 leiomyomas versus the corresponding 60 myometrium samples, revealing 135 genes to be significantly differentially expressed (q <0.05; |FC| >2). We found *zinc finger matrin-type 3* (*ZMAT3*) as the most significant gene. Indeed, this gene was upregulated in all leiomyomas, regardless of subtype (Figure 13A). We also compared leiomyomas of different genetic subtypes against all 60 myometrium samples. This revealed 258, 265, 296, and 198 genes to be significantly differentially expressed (q <0.05; |FC| >2) in leiomyomas of the *MED12*, *HMGA2*, *FH* and *COL4A5-COL4A6* subtypes, respectively. We did not perform a similar expression analysis with the quadruple-negative leiomyomas, since they showed high transcriptional heterogeneity in the hierarchical clustering analysis.

We performed a separate pathway analysis with each set of differentially expressed genes (Table 10). No pathway reached statistical significance for leiomyomas of the *COL4A5-COL4A6* subtype. We detected the Wnt/β-catenin pathway to be one of the most significantly dysregulated pathways in the complete set of leiomyomas (Table 10). Furthermore, the pathway analysis predicted this pathway as inhibited. Additionally, the Wnt antagonist *secreted frizzled related protein 1* (*SFRP1*) was significantly upregulated (FC = 2.1) in leiomyomas of the *MED12* subtype, whereas the Wnt antagonist *WNT inhibitory factor 1* (*WIF1*) was significantly upregulated (FC = 4.7) in leiomyomas of the *HMGA2* subtype.

We identified prolactin signaling as another significantly dysregulated pathway (Table 10). In addition, we identified *prolactin* (*PRL*) itself as one of the most highly upregulated (FC = 3) genes in the complete set of leiomyomas. This gene was particularly highly upregulated in leiomyomas of the *HMGA2* (FC = 7.7) and *COL4A5-COL4A6* (FC = 9.9) subtypes, but also significantly upregulated (FC = 2.6) in leiomyomas of the *MED12* subtype. We also identified *prolactin releasing hormone receptor* (*PRLHR*) as one of the most highly upregulated (FC = 2.9) genes in the complete set of leiomyomas. This gene was particularly highly upregulated (FC = 8.8) in leiomyomas of the *MED12* subtype, but also significantly upregulated (FC = 2.6) in leiomyomas of the *HMGA2* subtype.

Table 10. The most significantly enriched pathways according to Ingenuity Pathway Analysis

Ingenuity Canonical Pathway	All leiomyomas		*MED12* subtype		*HMGA2* subtype		*FH* subtype	
	q-value	z-score	*q*-value	z-score	*q*-value	z-score	*q*-value	z-score
Inhibition of matrix metalloproteases	8.3E-04	-	2.2E-04	-	1.8E-01	-	2.0E-01	-
Colorectal cancer metastasis signaling	1.4E-03	–0.3	1.1E-06	–0.5	3.1E-02	–0.3	3.5E-01	–1.6
RAR activation	1.5E-03	-	1.2E-02	-	8.3E-02	-	1.9E-01	-
Glucocorticoid receptor signaling	2.4E-03	-	1.1E-02	-	1.9E-02	-	6.3E-01	-
Wnt/β-catenin signaling	2.7E-03	–0.5	6.9E-04	–0.7	4.2E-02	–0.5	1.6E-01	–0.8
Endothelin-1 signaling	2.7E-03	–1.9	5.4E-02	–1.6	1.6E-02	–1.7	3.5E-01	–0.5
Prolactin signaling	2.7E-03	0.5	3.4E-03	0.8	1.6E-02	1.6	7.7E-01	-
Granulocyte adhesion and diapedesis	2.8E-03	-	9.6E-03	-	2.8E-01	-	7.4E-01	-
Agranulocyte adhesion and diapedesis	3.8E-03	-	4.6E-03	-	2.9E-01	-	5.8E-01	-
Bladder cancer signaling	4.3E-03	-	6.9E-03	-	4.2E-02	-	2.4E-01	-
IGF-1 signaling	6.5E-03	-	3.2E-02	-	5.4E-02	-	4.8E-01	-
Role of macrophages, fibroblasts, and endothelial cells in rheumatoid arthritis	7.4E-03	-	5.3E-05	-	5.4E-02	-	3.5E-01	-
HIF1α signaling	7.4E-03	-	3.4E-03	-	2.4E-01	-	5.1E-01	-
Growth hormone signaling	1.0E-02	-	1.2E-02	1.0	3.1E-02	1.0	3.5E-01	-
PI3K signaling in B lymphocytes	1.3E-02	–2.2	2.5E-03	–2.5	4.2E-02	–2.5	3.5E-01	–2.0
Neuropathic pain signaling in dorsal horn neurons	2.8E-02	0.0	8.1E-02	0.0	1.6E-02	0.4	2.7E-01	0.0
IL-8 signaling	3.6E-02	–0.5	3.2E-02	–1.1	3.1E-02	0.0	-	-
Hepatic fibrosis / hepatic stellate cell activation	3.6E-02	-	9.8E-06	-	3.1E-02	-	1.7E-01	-
D-myo-inositol-5-phosphate metabolism	5.3E-02	-	2.9E-01	-	3.1E-02	-	6.5E-01	-
Axonal guidance signaling	6.3E-02	-	2.5E-03	-	1.5E-01	-	2.4E-01	-
Glioma signaling	7.6E-02	-	1.7E-01	-	3.1E-02	1.6	-	-
Complement system	7.9E-02	-	4.6E-02	-	3.1E-02	-	1.9E-01	-
NRF2-mediated oxidative stress response	8.1E-02	-	1.2E-01	-	7.2E-02	-	2.3E-02	1.3
Glioblastoma multiforme signaling	1.4E-01	-	7.8E-02	1.3	3.1E-02	0.0	7.7E-01	-
Role of osteoblasts, osteoclasts, and chondrocytes in rheumatoid arthritis	2.3E-01	-	3.4E-03	-	5.1E-01	-	3.5E-01	-
Pentose phosphate pathway	-	-	-	-	-	-	3.9E-02	-

A positive z-score indicates a predicted activation, and a negative z-score indicates a predicted inactivation of the enriched pathway. A z-score could not be calculated for all enriched pathways.

5.2 Uniquely expressed genes in leiomyomas of different subtypes (II, IV)

To identify the most uniquely expressed protein-coding genes for each leiomyoma subtype, we compared leiomyomas of each subtype against all of the other leiomyoma and myometrium samples. These genes represent candidate biomarkers that could potentially serve to distinguish the different subtypes. The 20 most (q <0.05; |FC| >2) uniquely expressed genes are presented in Table 11.

Table 11. Top 20 most significantly differentially expressed genes in leiomyomas of the *MED12*, *HMGA2*, *FH*, and *COL4A5-COL4A6* subtypes

***MED12* subtype**			***HMGA2* subtype**			***FH* subtype**			***COL4A5-COL4A6* subtype**		
Gene	***q*-value**	**FC**	**Gene**	***q*-value**	**FC**	**Gene**	***q*-value**	**FC**	**Gene**	***q*-value**	**FC**
RAD51B	6.4E-22	3.8	*HMGA2*	5.0E-33	10.3	*AKR1B10*	4.1E-42	27.1	*IRS4*	3.4E-08	10.5
PLP1	3.5E-20	3.2	*IGF2BP2*	6.0E-28	4.4	*TKT*	6.7E-35	4.4	*NSG1*	8.8E-08	2.2
GARNL3	2.4E-19	2.3	*CCND2*	7.9E-18	2.5	*PDK1*	2.8E-24	3.6	*MXRA8*	4.9E-05	–2.5
KIAA1199	2.8E-18	5.7	*IL11RA*	7.7E-17	2.7	*SLC7A11*	4.8E-24	7.2	*FBLN1*	4.9E-05	–3.8
LAMP5	3.0E-18	5.1	*C19orf38*	1.3E-15	3.0	*G6PD*	9.9E-22	3.9	*PCSK2*	2.1E-04	3.3
MMP11	6.7E-18	5.5	*PLAG1*	3.1E-15	8.2	*PIR*	1.7E-21	3.2	*DPYD*	5.7E-04	–2.7
ADAM12	8.7E-17	8.8	*GRPR*	1.2E-13	8.3	*GCLM*	4.1E-21	3.7	*SPATA6*	7.2E-04	–2.0
POPDC2	9.7E-17	3.2	*PAPPA2*	7.4E-13	7.1	*SRXN1*	4.6E-18	2.4	*CTNNA3*	7.7E-04	2.5
CPA3	2.8E-15	–5.0	*PLA2R1*	7.4E-13	–4.3	*ENTPD7*	1.1E-17	4.1	*TMEM55A*	6.9E-03	2.1
THSD4	4.7E-15	2.5	*TBX3*	3.1E-12	–2.4	*TNFRSF21*	3.1E-16	10.3	*PCDHB8*	9.3E-03	2.4
CACNA1C	5.6E-15	2.1	*CBLN4*	3.7E-12	3.1	*SLC6A6*	8.7E-15	4.8	*SCG2*	1.4E-02	8.7
MMP16	8.0E-15	4.0	*GPR20*	1.6E-11	2.7	*NQO1*	6.4E-13	7.3	*SLAIN1*	1.6E-02	–2.1
CNTROB	1.6E-14	2.2	*GPR22*	4.6E-11	4.1	*BNIP3*	9.4E-13	3.0	*PLAGL1*	1.8E-02	–2.5
NHSL2	1.6E-14	2.0	*QPRT*	5.5E-11	2.0	*RNF128*	1.2E-12	2.4	*PARM1*	1.9E-02	–3.0
KCNAB3	1.9E-14	3.1	*PAWR*	8.7E-11	–2.7	*MGAT5*	2.5E-12	2.5	*LIX1*	2.0E-02	2.4
UNC5D	6.0E-14	2.8	*MB21D2*	1.1E-10	2.3	*PGD*	2.7E-11	3.0	*RHOB*	2.0E-02	–2.0
HPGDS	9.1E-14	–2.4	*CCND1*	2.5E-10	3.6	*FAM46C*	2.7E-11	4.4	*TGFBR3*	2.3E-02	–2.0
PCP4	1.2E-13	3.3	*WIF1*	3.3E-10	5.0	*AEBP1*	4.2E-11	–3.9	*HIST1H4H*	3.1E-02	2.1
WBSCR17	1.4E-13	2.2	*EGFR*	4.2E-10	–2.2	*SESN3*	2.4E-10	4.0	*COL4A5*	3.6E-02	–3.7
RUNDC1	1.4E-13	2.2	*AVPR1A*	4.7E-10	–4.3	*ABCC3*	5.6E-10	2.1	*PCDHB2*	3.7E-02	4.4

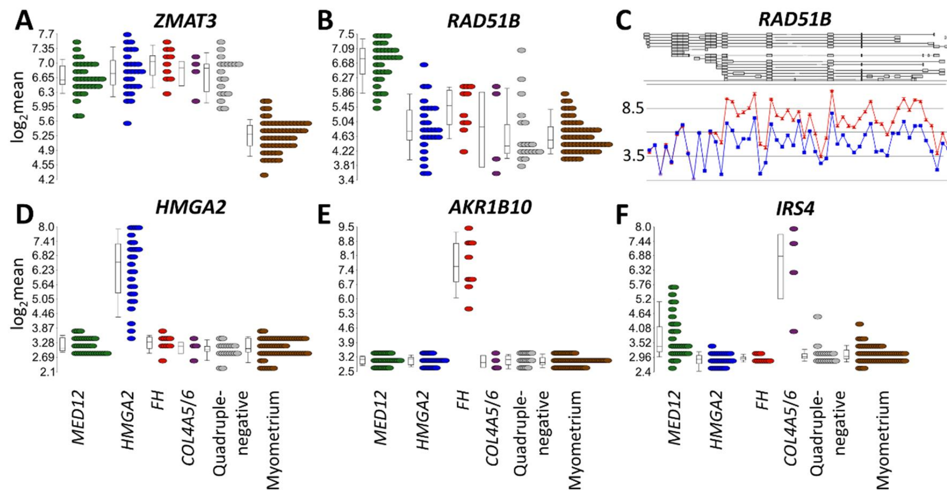


Figure 13. The most uniquely expressed gene in leiomyomas of different subtypes. A) *ZMAT3* was the most significantly differentially expressed gene in leiomyomas versus the myometrium tissue (brown). **B**) *RAD51B* was the most uniquely expressed gene in leiomyomas of the *MED12* subtype (green). **C**) Exon-level analysis revealed that the overexpression of *RAD51B* in *MED12* mutant leiomyomas (red) versus the myometrium tissue (blue) originated predominantly from a non-coding transcript (ENST00000492236). **D**) *HMGA2* was the most uniquely expressed gene in leiomyomas of the *HMGA2* (blue) subtype. **E**) *AKR1B10* was the most uniquely expressed gene in leiomyomas of the *FH* (red) subtype. **F**) *IRS4* was the most uniquely expressed gene in leiomyomas of the *COL4A5-COL4A6* (purple) subtype.

In Study II, we identified *RAD51B* as the most uniquely expressed gene (FC = 3.8) in leiomyomas of the *MED12* subtype, an observation that we confirmed in Study IV (Table 11; Figure 13B). However, exon-level analysis revealed that this upregulation corresponded to a non-coding transcript of *RAD51B* (ENST00000492236; Figure 13C). We validated this observation by RNA sequencing of two tumor-normal pairs (MY18 m1 & MY23 m1). Of note, we detected a lower expression of this non-coding transcript in their corresponding myometrium. Among the 20 most uniquely expressed genes in leiomyomas of the *MED12* subtype, three encode for matrix metalloproteases (*ADAM12*, *MMP16*, & *MMP11*). The pathway enrichment analysis confirmed a significant dysregulation of matrix metalloproteases (Table 10). Of note, leiomyomas of the *MED12* subtype displayed a significant upregulation (FC = 3.1) of *insulin like growth factor 2* (*IGF2*).

We identified *HMGA2* as the most uniquely expressed gene (FC = 10.2) in leiomyoma of *HMGA2* subtype (Table 11; Figure 13D). The second most significant gene (FC = 4.3) was *insulin like growth factor 2 mRNA binding protein 2* (*IGF2BP2*), a gene known to be directly regulated by HMGA2.[298] The proto-oncogene *PLAG1* was the sixth most uniquely expressed gene (FC = 8.2), suggesting that *HMGA2* may regulate this gene as well. The three leiomyomas with a genetic *HMGA1* abnormality also exhibited a significant upregulation (FC = 5.6) of *PLAG1*. In fact, among the 34 leiomyomas harboring a *MED12* mutation, only two showed upregulation of *PLAG1* (FC >2), one of which harbored an *HMGA1* abnormality (MY5008 m3; FC = 2.5) and the other a *PLAG1* translocation (MY16 m1; FC = 17.5). Further investigation showed upregulation of *PLAG1* in two quadruple-negative leiomyomas (MY5007 m2; FC = 11.1 & M51 m1; FC = 7.5), both of which harbored a large-scale amplification on chromosome 8 that spanned *PLAG1*. All of the three samples with a genetic *PLAG1* abnormality showed expression patterns similar to those seen in leiomyomas with an *HMGA2* or *HMGA1* abnormality. In fact, one of these samples (MY5007 m2) clustered among leiomyomas of the *HMGA2* subtype. Studies have shown that *PLAG1* directly regulates the expression of *IGF2*.[299-301] We detected a significant upregulation of *IGF2* in leiomyomas of *HMGA2* (FC = 3) and *MED12* (FC = 3.1) subtypes. Differential expression analysis also revealed a significant upregulation (FC = 4.3) of *IGF2* in the three leiomyomas with a genetic *PLAG1* abnormality.

We identified *aldo-keto reductase family 1 member B10* (*AKR1B10*) to be the most uniquely expressed gene (FC = 26.7) in leiomyomas of the *FH* subtype (Table 11; Figure 13E). *AKR1B10* was upregulated in all *FH*-deficient leiomyomas, and in none of the other leiomyoma nor myometrium samples. NRF2 mediated oxidative stress response was the most significantly dysregulated pathway (Table 10). Compatible with this observation, among the 20 most uniquely expressed genes, eight are known targets of the transcription factor NFR2 (*AKR1B10*, *TKT*, *PIR*, *SLC7A11*, *NQO1*, *SRXN1*, *SLC6A6*, & *GCLM*).[302-305] The pentose phosphate pathway was the only other significant pathway. Indeed, among the 20 most uniquely expressed genes, three encode for key enzymes of this pathway (*TKT*, *PGD*, & *G6PD*). None of the other leiomyoma subtypes displayed a significant dysregulation of these two pathways (Table 10). Of note, while leiomyomas of the *FH* subtype displayed a significant upregulation (FC = 4.8) of *PLAG1*, we found a significant downregulation (FC = –2.5) of *IGF2* in these lesions.

Although *COL4A5* and *COL4A6* are both affected by *COL4A5-COL4A6* deletions, only *COL4A5* displayed a statistically significant downregulation (FC= –3.7) when compared to the myometrium samples. We identified *insulin receptor substrate 4* (*IRS4*) to be the most uniquely expressed gene (FC=10.5) in leiomyomas of the *COL4A5-COL4A6* subtype (Table 11; Figure 13F). Our statistical analysis also revealed a unique downregulation (FC = –2.5) of the putative tumor suppressor *PLAG1 like zinc finger 1* (*PLAGL1*), which is structurally similar to *PLAG1* (Table 11).

5.3 Downregulated genes by chromosome 7q22, 22q, and 1p deletions (IV)

Assuming that 7q22, 22q, and 1p deletions, each targets a driver gene that results in reduced gene dosage, we searched for downregulated genes within these deletions by performing a separate differential expression analysis that compared leiomyomas with each abnormality in question against leiomyoma and myometrium samples lacking the abnormality. The 20 most significantly downregulated protein-coding genes (q <0.05) located within commonly deleted regions are presented in Table 12.

Table 12. Top 10 genes downregulated genes by 7q22, 22q, and 1q deletions

7q22 deletion				22q deletion				1p deletion			
Gene	*q*-value	FC	#	Gene	*q*-value	FC	#	Gene	*q*-value	FC	#
LMTK2	1.9E-04	–1.3	8	*FBXO7*	1.1E-12	–1.4	19	*UBE4B*	6.6E-11	–1.5	16
COPS6	7.5E-04	–1.3	9	*MTMR3*	7.6E-12	–1.4	19	*EXOSC10*	5.7E-08	–1.3	16
CUX1	7.9E-04	–1.5	13	*DEPDC5*	5.6E-08	–1.3	19	*DNAJC16*	5.7E-08	–1.2	15
MLL5	2.0E-03	–1.4	11	*RNF185*	1.6E-07	–1.5	19	*GNB1*	8.1E-08	–1.3	15
TNPO3	2.0E-03	–1.3	8	*EIF3D*	2.0E-07	–1.3	18	*PRDM2*	9.9E-08	–1.3	15
ZNF800	4.6E-03	–1.2	7	*DUSP18*	4.0E-07	–1.4	19	*FAM54B*	1.4E-07	–1.4	15
PNPLA8	1.9E-02	–1.4	13	*TTC28*	7.5E-07	–1.4	19	*VPS13D*	1.4E-07	–1.5	16
ZNF394	2.0E-02	–1.3	9	*EP300*	8.3E-07	–1.3	15	*RERE*	2.5E-07	–1.5	17
CADPS2	2.0E-02	–2	8	*MAPK1*	1.2E-06	–1.3	16	*KIF1B*	6.0E-07	–1.5	16
PMPCB	2.1E-02	–1.4	11	*MKL1*	1.3E-06	–1.4	16	*CLSTN1*	9.8E-07	–1.5	16

= total number of deletions that spanned the gene among a total of 94 leiomyomas.

DISCUSSION

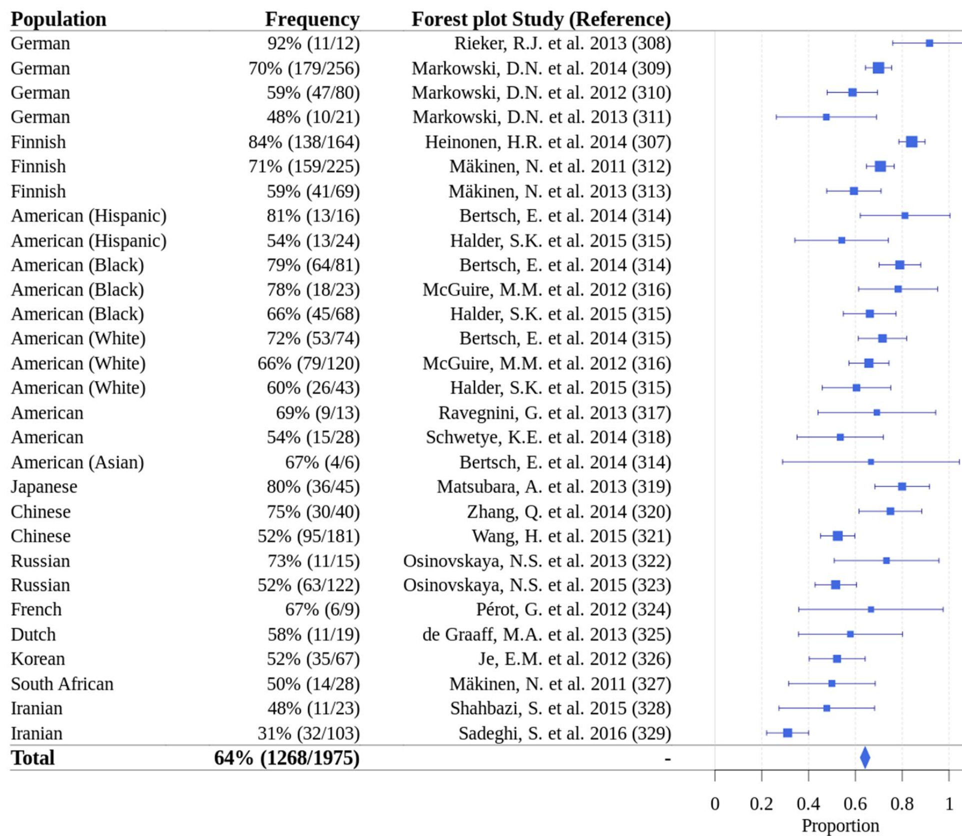
Uterine leiomyomas rank among the most common symptomatic human neoplasms, but considering their major public health impact, their etiology is severely understudied. The current classification of leiomyomas is largely based on histological features, with a molecular classification sorely lacking. Decades of work utilizing classical cytogenetic techniques have demonstrated that leiomyomas are a genetic disease characterized by chromosomal abnormalities that are non-random and leiomyoma-specific. Approximately 60% lack chromosome-level abnormalities, suggesting that submicroscopic mutations underlie the remaining tumors. Massively parallel sequencing technologies have made it possible to examine tumor genomes on a previously unprecedented scale. Utilizing genome-wide methods in this thesis work, we have made several breakthrough findings that provide invaluable insights into the molecular basis of these very common tumors.

1. *MED12*: a key driver in leiomyomas

Exome sequencing of 18 uterine leiomyomas, followed by screening of an additional 207, revealed localized mutations in the *MED12* gene in 71% of these lesions. *MED12* consists of 45 exons, and all of the mutations detected were located in exon 2 or in the preceding intron-exon boundary. We found this region to be highly evolutionary conserved; none of the mutations was present in the corresponding myometrial tissue; all were predicted to result in an in-frame protein product, and no tumor showed more than one *MED12* mutation. Considered together, these mutational patterns strongly indicate that mutated *MED12* acts an oncogene, and that these hotspot mutations represent gain-of-function mutations. Furthermore, the affected tumors expressed mainly the mutated form of *MED12*, demonstrating that the mutations reside on the active X chromosome. Exome sequencing revealed no other recurrently mutated genes, highlighting *MED12* as a key driver in the majority of leiomyomas.

Since our initial discovery, a number of subsequent studies have reported similar frequencies of *MED12* mutations among leiomyomas from various different populations, at an overall frequency of 64% (Table 13). Small sample sizes and a tendency towards selecting larger tumors may have underestimated the true frequency of *MED12* mutations. Furthermore, we recently discovered that mutations may infrequently (2%) reside at the end of exon 1 as well,[306] and the majority of studies have not accounted for this region in their sequencing analyses. Indeed, in a prospectively collected, well-documented, and unselected series of sporadic leiomyomas, we found the mutation frequency to be closer to 84%.[307]

Table 13. Frequency of *MED12* hotspot mutations in uterine leiomyomas from various populations



Population	Frequency	Forest plot Study (Reference)
German	92% (11/12)	Rieker, R.J. et al. 2013 (308)
German	70% (179/256)	Markowski, D.N. et al. 2014 (309)
German	59% (47/80)	Markowski, D.N. et al. 2012 (310)
German	48% (10/21)	Markowski, D.N. et al. 2013 (311)
Finnish	84% (138/164)	Heinonen, H.R. et al. 2014 (307)
Finnish	71% (159/225)	Mäkinen, N. et al. 2011 (312)
Finnish	59% (41/69)	Mäkinen, N. et al. 2013 (313)
American (Hispanic)	81% (13/16)	Bertsch, E. et al. 2014 (314)
American (Hispanic)	54% (13/24)	Halder, S.K. et al. 2015 (315)
American (Black)	79% (64/81)	Bertsch, E. et al. 2014 (314)
American (Black)	78% (18/23)	McGuire, M.M. et al. 2012 (316)
American (Black)	66% (45/68)	Halder, S.K. et al. 2015 (315)
American (White)	72% (53/74)	Bertsch, E. et al. 2014 (315)
American (White)	66% (79/120)	McGuire, M.M. et al. 2012 (316)
American (White)	60% (26/43)	Halder, S.K. et al. 2015 (315)
American	69% (9/13)	Ravegnini, G. et al. 2013 (317)
American	54% (15/28)	Schwetye, K.E. et al. 2014 (318)
American (Asian)	67% (4/6)	Bertsch, E. et al. 2014 (314)
Japanese	80% (36/45)	Matsubara, A. et al. 2013 (319)
Chinese	75% (30/40)	Zhang, Q. et al. 2014 (320)
Chinese	52% (95/181)	Wang, H. et al. 2015 (321)
Russian	73% (11/15)	Osinovskaya, N.S. et al. 2013 (322)
Russian	52% (63/122)	Osinovskaya, N.S. et al. 2015 (323)
French	67% (6/9)	Pérot, G. et al. 2012 (324)
Dutch	58% (11/19)	de Graaff, M.A. et al. 2013 (325)
Korean	52% (35/67)	Je, E.M. et al. 2012 (326)
South African	50% (14/28)	Mäkinen, N. et al. 2011 (327)
Iranian	48% (11/23)	Shahbazi, S. et al. 2015 (328)
Iranian	31% (32/103)	Sadeghi, S. et al. 2016 (329)
Total	**64% (1268/1975)**	-

* Confidence intervals calculated by the asymptotic (Wald) method.

1.1 MED12 hotspot mutations in leiomyoma variants, leiomyosarcomas, and extrauterine leiomyomas

Multiple subsequent studies have screened for *MED12* mutations in leiomyoma variants, leiomyosarcomas, and extrauterine leiomyomas, demonstrating that *MED12* mutations occur at a significantly lower frequency in histological variants of leiomyomas (Table 14). Indeed, cellular leiomyomas and leiomyomas with bizarre nuclei show a particularly low frequency (Table 14). Their lower prevalence indicates that other factors than *MED12* mutations are responsible for their unusual histology.

MED12 mutations also occur, albeit at a lower frequency, in STUMPs and uterine leiomyosarcomas, suggesting that these mutations are selectively advantageous in malignant tumors as well. Alternatively, some leiomyomas harboring a *MED12* mutation may have undergone malignant transformation. The role of *MED12* in malignancy remains obscure, since the majority of leiomyosarcomas lack protein expression of MED12, including those with a *MED12* mutation.[311, 314, 324] Furthermore, one study reported a leiomyosarcoma to harbor a frame-shift mutation in *MED12*.[314] Although *MED12* mutations occur in metastatic lesions as well,[317] one recent study described a patient in which the primary tumor displayed a *MED12* mutation, whereas its two metastases did not.[330] Considered together, these observations suggest that *MED12* mutations occur in the early stages of leiomyosarcoma formation and that as the tumors progress towards malignancy, their expression becomes unnecessary or is lost. It is tempting to speculate that *MED12* mutations have a negative effect on the malignant potential of leiomyomas, and that their common occurrence may explain why leiomyomas rarely, if ever, progress. It would also be reasonable to assume that other genetic defects are responsible for the aggressive behavior of leiomyosarcomas. Indeed, exome sequencing recently revealed that uterine leiomyosarcomas frequently harbor *TP53* and *ATRX, chromatin remodeler* (*ATRX*) mutations, some of which co-occurred with *MED12* mutations.[331] Furthermore, one study examined for *MED12* mutations in four leiomyosarcoma cell lines, but found none.[317]

Other studies have screened for *MED12* mutations in smooth muscle tumors of various anatomical sites (Table 14). These have revealed *MED12* mutations in a total of 22 extrauterine leiomyomas and in only 1 extrauterine leiomyosarcoma (Table 14). Some researchers have hypothesized that *MED12* mutations occur only in tumors that originate from tissues developing from the Müllerian duct.[311] Although a few leiomyomas of the kidney, ovary, and cervix have been reported to display a *MED12* mutation, these mutations are significantly more common in leiomyomas at retroperitoneal and pelvic sites,[317, 318] supporting the view that some of these actually originate from pedunculated leiomyomas that have become detached from the uterus.[101] Of note, *MED12* mutations are not restricted to humans, as similar type of mutations occur in canine vaginal leiomyomas.[311]

Table 14. Frequency of *MED12* hotspot mutations in leiomyoma variants, extrauterine leiomyomas, and leiomyosarcomas

Phenotype	Frequency	Population	Study (Reference)
Mitotically active leiomyoma	86% (6/7)	Chinese	Zhang, Q. et al. 2014 (320)
	38% (10/26)	Finnish	Mäkinen, N. et al. 2013 (313)
Total	**48% (16/33)**	-	-
Leiomyoma with bizarre nuclei	25% (1/4)	Japanese	Matsubara, A. et al. 2013 (319)
	17% (3/18)	Finnish	Mäkinen, N. et al. 2013 (313)
	10% (4/42)	Chinese	Zhang, Q. et al. 2014 (320)
	0% (0/5)	French	Pérot, G. et al. 2012 (324)
Total	**12% (8/69)**	-	-
Cellular leiomyoma	33% (2/6)	Japanese	Matsubara, A. et al. 2013 (319)
	14% (3/22)	Chinese	Zhang, Q. et al. 2014 (320)
	9% (6/67)	Finnish	Mäkinen, N. et al. 2013 (313)
Total	**12% (11/95)**	-	-
Lipoleiomyoma	**0% (0/4)**	Japanese	Matsubara, A. et al. 2013 (319)
Intravenous leiomyomatosis	**0% (0/9)**	American	Buza, N.et al. 2014 (224)
Extrauterine leiomyoma*	100% (7/7)	German	Rieker, R.J. et al. 2013 (308)
	50% (1/2)	German	Markowski, D.N. et al. 2013 (311)
	17% (10/58)	American	Schwetye, K.E. et al. 2014 (318)
	16% (3/19)	American	Ravegnini, G. et al. 2013 (317)
	5% (1/20)	Dutch	de Graaff, M.A. et al. 2013 (325)
	0% (0/42)	Finnish	Kämpjarvi, K. et al. 2012 (332)
	0% (0/51)	Japanese	Matsubara, A. et al. 2013 (319)
Total	**11% (22/199)**	-	-
STUMP	11% (2/18)	Chinese	Zhang, Q. et al. 2014 (320)
	11% (1/9)	French	Pérot, G. et al. 2012 (324)
	8% (1/12)	American	Schwetye, K.E. et al. 2014 (318)
Total	**10% (4/39)**	-	-
Uterine leiomyosarcomas**	30% (6/20)	American	Schwetye, K.E. et al. 2014 (318)
	30% (3/13)	American	Ravegnini, G. et al. 2013 (317)
	21% (4/19)	Finnish	Mäkinen, N. et al. 2016 (331)
	20% (2/10)	French	Pérot, G. et al. 2012 (324)
	17% (2/12)	Japanese	Matsubara, A. et al. 2013 (319)
	14% (1/7)	Dutch	de Graaff, M.A. et al. 2013 (325)
	11% (4/38)	Chinese	Zhang, Q. et al. 2014 (320)
	9% (3/32)	American	Bertsch, E. et al. 2014 (314)
	7% (3/41)	Finnish	Kämpjarvi, K. et al. 2012 (332)
	7% (1/14)	German	Markowski, D.N. et al. 2013 (311)
	4% (1/24)	Chinese	Wang, H. et al. 2015 (321)
	0% (0/5)	Korean	Je, E.M. et al. 2012 (326)
Total	**13% (30/235)**	-	-
Extrauterine leiomyosarcoma	4% (1/25)	American	Schwetye, K.E. et al. 2014 (318)
	0% (0/23)	Dutch	de Graaff, M.A. et al. 2013 (325)
	0% (0/10)	American	Ravegnini, G. et al. 2013 (317)
	0% (0/20)	Japanese	Matsubara, A. et al. 2013 (319)
	0% (0/20)	German	Markowski, D.N. et al. 2013 (311)
	0% (0/38)	French	Pérot, G. et al. 2012 (324)
Total	**1% (1/136)**	-	-

* Included are leiomyoma variants displaying unusual growth pattern of possible uterine origin.
** Included are extrauterine metastases of known uterine origin.

1.2 MED12 hotspot mutations in other human neoplasms

Several screening studies have shown that *MED12* mutations are rare or completely absent in the majority of human neoplasms (Table 15). Furthermore, none of the major large-scale cancer-sequencing efforts, such as The Cancer Genome Atlas (TCGA) project or the International Cancer Genome Consortium (ICGC), has highlighted recurrent *MED12* mutations. Exceptions include chronic lymphocytic leukemia and colorectal cancer, both of which exhibit *MED12* mutations at a significantly lower frequency than in leiomyomas.[333, 334] Of note, the mutational spectrum of *MED12* mutations in chronic lymphocytic leukemia differs from that of leiomyomas, since the former display an unexpectedly high proportion of exon 1 mutations.[333]

A recent exome-sequencing study described a high frequency of *MED12* mutations in fibroadenomas of the breast.[335] Similarly as seen in uterine leiomyomas, *MED12* was the only recurrently mutated gene, and the tumors' mutational spectrum was nearly identical,[335] suggesting that both of these tumors types share the same etiology. Furthermore, both are benign, ovarian steroid-dependent, and arise from smooth muscle cells. Fibroadenomas are biphasic tumors composed of epithelial and stromal components, and the mutations were exclusively present in the stromal cells.[335] Several follow-up studies have validated the occurrence of *MED12* mutations in fibroadenomas, revealing a similar frequency as in leiomyomas (Table 15). Of note, *HMGA2* rearrangements occur in fibroadenomas as well, suggesting that *MED12* and *HMGA2* have overlapping functions in tumorigenesis.[150]

Several subsequent studies have revealed a similar frequency of *MED12* mutations in phyllodes tumors of the breast (Table 15). These are rare fibroepithelial tumors that share morphological features with fibroadenomas,[336] suggesting a shared origin. However, phyllodes tumors have a variable clinical behavior, and are classified as benign, borderline, or malignant. *MED12* mutations occur at a noticeable lower frequency in the malignant cases.[337] In contrast to the benign and borderline cases, malignant lesions harbor additional mutations in classical driver genes such as *TP53*, *RB1*, and *NF1*.[338] Furthermore, *TERT* promoter mutations tend to co-occur with *MED12* mutations in phyllodes tumors, regardless of tumor grade.[339] Interestingly, a recent study showed that leiomyosarcomas harbor *ATRX* or *DAXX* mutations, resulting in alternative lengthening of telomeres.[331] It is tempting to speculate that that leiomyomas and fibroadenomas may progress to a more malignant state as they acquire defects in telomerase activity. Although *MED12* mutations are absent in most human neoplasms, it remains to be seen whether some rarer neoplasms could carry *MED12* mutations.

Table 15. Frequency of *MED12* hotspot mutations in other human neoplasms

Neoplasm	Frequency	Population	Study (Reference)
Fibroadenomas of the breast	85.7% (18/21)	Singaporean	Tan, J. et al. 2015 (336)
	66.7% (6/9)	Japanese	Nagasawa, S. et al. 2015 (340)
	65.4% (17/26)	American	Piscuoglio, S. et al. 2015 (341)
	62.1% (36/58)	Japanese	Yoshida, M. et al. 2015 (342)
	61.0% (61/100)	American	Piscuoglio, S. et al. 2016 (343)
	59.2% (58/98)	Singaporean	Lim, W.K. et al. 2014 (335)
	59.1% (13/22)	German	Pfarr, N. et al. 2015 (344)
	50.0% (9/18)	Taiwanese	Lien, H. et al. 2016 (345)
	46.6% (27/58)	Japanese	Mishima, C. et al. 2015 (346)
Total	**59.8% (245/410)**	-	-
Phyllodes tumors of the breast	80.4% (37/46)	Japanese	Yoshida, M. et al. 2015 (342)
	74.1% (20/27)	Japanese	Mishima, C. et al. 2015 (346)
	73.6% (39/53)	Taiwanese	Lien, H. et al. 2016 (347)
	70.9% (56/79)	Singaporean	Tan, J. et al. 2015 (336)
	66.7% (10/15)	American	Cani, A.K. et al. 2015 (338)
	63.8% (30/47)	American	Piscuoglio, S. et al. 2015 (341)
	62.5% (70/112)	Singaporean	Ng, C.C. et al. 2015 (348)
	60.5% (46/76)	American	Piscuoglio, S. et al. 2016 (343)
	46.0% (81/176)	Korean	Yoon, N. et al. 2016 (337)
	45.5% (5/11)	Japanese	Nagasawa, S. et al. 2015 (340)
	56.3% (9/16)	German	Pfarr, N. et al. 2015 (344)
Total	**61.2% (403/658)**	-	-
Chronic lymphocytic leukemia	5.2% (37/709)	Finnish	Kämpjarvi, K. et al. 2015 (333)
	8.8% (10/114)	French	Guièze, R. et al. 2015 (349)
	3.4% (3/88)	American	Wang, L. et al. 2011 (350)
	1.9% (3/160)	American	Landau, D. et al. 2013 (351)
	1.9% (2/105)	Spanish	Quesada, V. et al. 2012 (352)
	1.8% (3/168)	French	Damm, F. et al. 2014 (353)
Total	**4.3% (58/1344)**	-	-
Colorectal cancer	0.5% (2/392)	Finnish	Kämpjarvi, K. et al. 2012 (332)
	0.4% (1/224)	American	The Cancer Genome Atlas (334)
	0.3% (1/389)	Korean	Je, E.M. et al. 2012 (326)
Total	**0.4% (4/1005)**	-	-

1.3 Other pathogenic MED12 mutations

Exome sequencing of prostate cancer, adrenocortical carcinomas, and renal cell cancer have recently revealed recurrent mutations in other parts of *MED12* (Table 16). *MED12* mutations in prostate cancer are highly localized, affecting the leucine- and serine-rich (LS) domain (Figure 14A).[354] The mutations detected in adrenocortical carcinoma and renal cell cancer are more scattered within the *MED12* gene and predicted as damaging (Figure 14A). Furthermore, the mutations in adrenocortical carcinomas tend to affect the Med12-PQL domain, which is an interaction site for β-catenin (Figure 14B).[355] Germline mutations – typically affecting the LS domain – cause the hereditary syndromes Opitz-Kaveggia (also known as FG syndrome), Lujan-Fryns, and Ohdo syndrome.[356] These syndromes have overlapping phenotypes, including mental retardation and dysmorphic features.[356] None of these syndromes has been associated with increased tumor risk, however.

Table 16. Other pathogenic *MED12* mutations

Neoplasm/Syndrome	Frequency	Location	Study/OMIM#
Prostate cancer	5.4% (6/111)	Exon 26	Barbieri, C.E. et al. 2012 (354)
	2.3% (4/171)	Exon 26	TCGA, ICGC (357)
	1.3% (1/80)	Exon 26	Kämpjarvi, K. et al. 2016 (358)
	0% (0/226)	Exon 26	Stoehr, R. et al. 2013 (359)
Total	**1.9% (11/588)**	Exon 26	-
Adrenocortical carcinoma	**4.9% (6/122)**	Exons 34-35, 38-39 & 41	Assie, G. et al. 2014 (355)
Renal cell cancer	**4.5% (3/67)**	Exons 19, 24 & 35	Arai, E. et al. 2014 (360)
Opitz-Kaveggia syndrome	-	Exon 21	305450 (356)
Lujan-Fryns syndrome	-	Exon 22	309520 (356)
Ohdo syndrome	-	Exons 22 & 37	300895 (356)

1.4 Possible mechanisms of tumorigenesis

MED12 encodes for a subunit of the multiprotein complex known as Mediator, which serves as a bridge between transcription factors and the RNA polymerase II transcriptional machinery.[361] Mediator facilitates the transcription of almost all eukaryotic genes, and functions as a context-dependent positive or negative regulator of transcription.[361] The human core Mediator, consisting of 26 subunits, is divided into three domains known as the head, middle, and tail (Figure 14C). MED12 forms, together with MED13, Cyclin-C, and CDK8 or CDK19, a fourth submodule known as the CDK8 module.[361] MED12 connects Cyclin-C-CDK8/19 to the core Mediator and serves as an essential regulator of the kinase activity of the CDK8 module.[361] Mediator exists in two major forms, one of which contains the CDK8 module, whereas the other do not.[361] The large form, “L mediator”, acts as a transcriptional repressor, whereas the smaller form, “S mediator”, stimulates basal transcription.[361] Of note, *cyclin dependent kinase 8* (*CDK8*) acts an oncogene that is amplified in over half of colorectal cancers.[362]

We recently demonstrated that one consequence of *MED12* mutations is weakened interaction between MED12 and Cyclin-C-CDK8/19, resulting in decreased kinase activity of the CDK8 module.[363] In a subsequent study, we screened leiomyomas for mutations in other components of the CDK8 module, but found none.[364] The prostate cancer-linked mutations did not show the same effect, but rather a weakened interaction between *MED12* and other components of the Mediator.[358] The manner by which these consequences contribute to tumorigenesis remains unclear.

MED12 participates normally in various signaling pathways, including p53, Sonic hedgehog, mTOR, estrogen, TGF-β, and Wnt/β-catenin.[356] Indeed, MED12 can act as a transcriptional activator of Wnt target genes through its interaction with β-catenin,[356] suggesting that *MED12* mutations may act through the Wnt/β-catenin signaling pathway. We initially observed a significant dysregulation of genes related to Wnt signaling in leiomyomas harboring a *MED12* mutation. Several subsequent studies have examined for a possible role of activated Wnt signaling in leiomyomas, yielding somewhat contradicting results.[268, 276, 277, 310, 324] The Wnt/β-catenin pathway appears to be activated in the leiomyoma side population,[276] and its inhibition attenuates leiomyoma cell growth.[277] On the other hand, the majority of leiomyomas, including those with a *MED12* mutation, show β-catenin localization into the membrane, indicative of low transactivation activity.[324] Furthermore, our pathway analyses revealed a predicted inhibition of the Wnt/β-catenin pathway in leiomyomas. Compatible with this prediction, we detected the Wnt antagonist *SFRP1* to be specifically upregulated in leiomyomas of the *MED12* subtype and the Wnt antagonist *WIF1* to be specifically upregulated in leiomyomas of the *HMGA2* subtype.[365] It should be noted that although *SFRP1* is a well-known inhibitor of Wnt, one study postulated that *SFRP1* might demonstrate anti-apoptotic effects in leiomyomas.[366] The role of *MED12* mutations in Wnt signaling remains obscure, and at least three different types of Wnt signaling pathways exist.

One study showed that knock-down of *MED12* expression in the human leiomyoma cells result in reduced levels of extracellular matrix genes, including fibronectin and collagen type 1.[367] Furthermore, integrated expression profiling of leiomyomas and fibroadenomas carrying *MED12* mutations has revealed a shared dysregulation of genes related to extracellular matrix organization.[335] These observations suggest that *MED12* play an important role in regulating genes involved in extracellular remodeling. Compatible with this hypothesis, we found "Inhibition of matrix metalloproteases" to be the most significantly altered pathway in leiomyomas of the *MED12* subtype. Matrix metalloproteases are key proteins involved in extracellular matrix remodeling.[368] Several of these matrix metalloproteases, including *ADAM metallopeptidase domain 12* (*ADAM12*), were uniquely upregulated in leiomyomas of the *MED12* subtype. A recent study showed that knockdown of the long non-coding RNA *Intergenic 10* results in downregulation of its neighboring gene *ADAM12* and inhibition of leiomyoma cell proliferation.[369] Another study showed that inhibition of MED12 results in resistance to a wide variety of cancer drugs through activation of TGF-β receptor signaling via elevated TGF-βR2 protein expression.[370] It is well known that leiomyomas display elevated levels of TGF-β receptors,[269] suggesting that *MED12* mutations could play a role in modulating this pathway.

A recent study showed that leiomyoma cell cultures carrying a *MED12* mutation can barely survive the first passages,[371] making research into the molecular consequences of *MED12* mutations difficult. Conversely, another study showed that conditional expression of the most common *MED12* mutation (c.131G>A) in a mouse model results in the development of uterine lesions.[372] Not only does this animal model confirm a driver role for *MED12* in tumorigenesis, but also provides researchers with an accurate tool to study the mechanism of *MED12* mutations in leiomyomas; interestingly, these murine tumors also show chromosomal instability, including highly complex chromosomal abnormalities.[372]

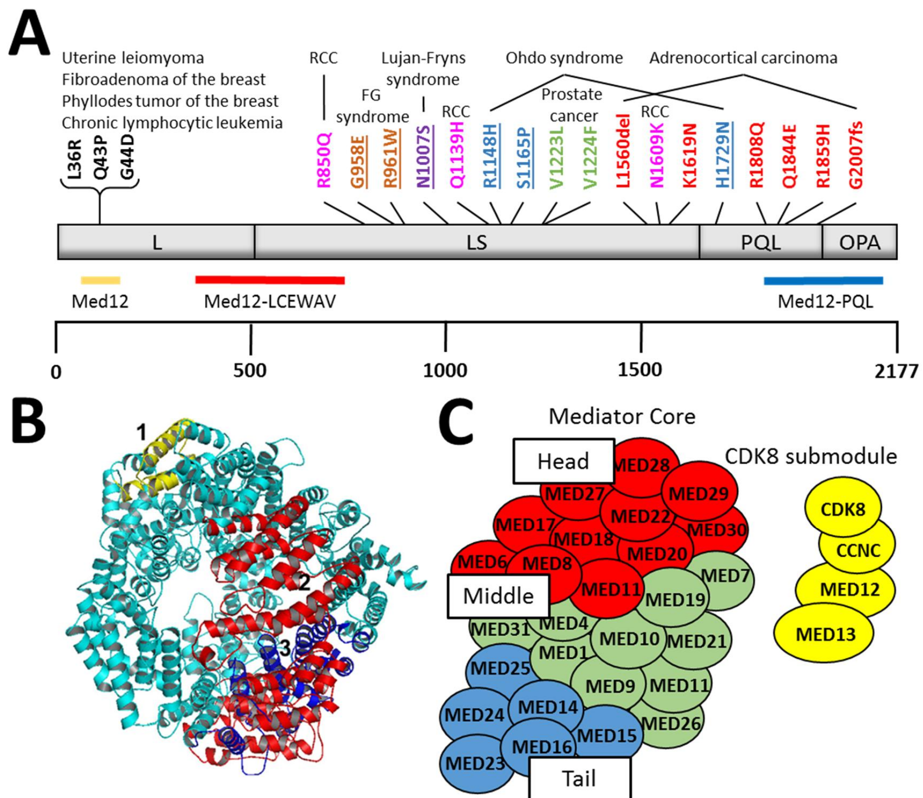


Figure 14. Schematic of the human MED12 protein and the Mediator complex. A) The MED12 protein, consisting of 2177 amino acids, is divided into four domains based on its amino-acid content: L, the leucine-rich domain; LS, the leucine- and serine-rich domain; PQL, the proline-, glutamine- and leucine-rich domain; and OPA, opposite paired domain. The protein contains also three conserved domains: Med12 (yellow), Med12-LCEWAV (red), and Med12-PQL (blue). Leiomyomas, breast epithelial tumors, and chronic lymphocytic leukemia harbor highly localized mutations affecting the early regions of MED12, whereas other pathogenic mutations are located towards the middle or end of MED12. Germline mutations are underlined. **B**) Predicted 3D structure of the MED12 protein highlighting the three conserved domains. **C**) The human core Mediator, consisting of 26 subunits, has a head, middle, and tail module. MED12 is part of a fourth submodule, known as the CDK8 kinase module. **B** is from *Journal of Cellular Biochemistry*, **117**, Banaganapalli et al. A Computational Protein Phenotype Prediction Approach to Analyze the Deleterious Mutations of Human MED12 Gene (2016), reprinted with the permission of John Wiley and Sons.

2. Clonality and intratumor genetic heterogeneity

Patients affected by uterine leiomyomas tend to have multiple separate nodules,[89] each of which is assumed to arise independently, as each shows a random pattern of X-inactivation. Furthermore, each leiomyoma tends to harbor a unique set of mutations, including distinct cytogenetic abnormalities and a different *MED12* mutation. A few studies have, however, reported separate leiomyomas with cytogenetic abnormalities that appear to be identical,[231-234] suggesting that not all leiomyomas arise independently. Using whole-genome sequencing, we were able to detect numerous rearrangements clearly identical among separate lesions from the same patient, signifying that some are in fact clonally related.

In order to investigate the frequency of clonally related leiomyomas, we examined for identical chromosomal abnormalities among 103 leiomyomas from 14 consecutive patients by use of SNP arrays. We detected four (29%) patients with 13 (13%) clonally related leiomyomas, each of which lacked any *MED12* mutations. Since the majority of leiomyomas with a *MED12* mutation displayed no chromosomal abnormalities, we cannot exclude the possibility of some being clonally related, as well. While one cannot extrapolate the exact frequency of this phenomenon from these studies, due to their small sample sizes, these preliminary observations suggest that clonally related leiomyomas are relatively common among leiomyomas that lack a *MED12* mutation. This phenomenon may occasionally explain the occurrence multiple leiomyomas within the same uterus.

2.1 Possible mechanisms of tumor spread

The very hallmark of cancers as compared to benign tumors is the ability to spread and form secondary tumors. Leiomyomas are often viewed as a textbook example of a benign tumor that does not undergo malignant transformation. However, it is now known that uterine leiomyomas can spread due to uterine surgery, demonstrating that some leiomyoma cells have the ability to attach and grow at distant locations. Indeed, pedunculated leiomyomas may become detached from the uterus and attach to peritoneal surfaces. In Study III, we investigated the surgical history of patients with clonally related leiomyomas and found prior myomectomy in only two out of five patients with clonally related leiomyomas.

Since no mutations have proven to directly cause metastasis, some researchers have argued that metastases develop when disseminated cells reach distant sites with a favorable microenvironment.[49] We, however, found an inverse relationship between the presence of *MED12* mutations and clonally related leiomyomas, suggesting that the genetic background does influence the ability to form secondary tumors. Compatible with this notion, leiomyomas of the *HMGA2* subtype are significantly larger and show a higher degree of chromosomal instability.[199] Indeed, all of the tumors assessed as clonally related by whole-genome sequencing corresponded to leiomyomas of the *HMGA2* subtype, and all of the tumors assessed as clonally related by SNP arrays were negative for *MED12* mutations. However, one set of clonally related leiomyomas consisted of six leiomyomas of the *FH* subtype, signifying that the phenomenon is not limited to leiomyomas with *HMGA2* defects.

Although the putative mechanisms by which these tumors spread are largely speculative, the ability for uterine leiomyomas to spread is not completely surprising considering the existence of leiomyoma variants displaying quasi-malignant behaviors, including benign metastasizing leiomyomas, intravenous leiomyomatosis, and disseminated peritoneal leiomyomatosis.[96] Interestingly, all of these can harbor *HMGA2* rearrangements,[189, 190] supporting the notion that leiomyomas of the *HMGA2* subtype are more prone to dissemination. On the other hand, *MED12* mutations have also been reported in disseminated peritoneal leiomyomatosis, including five distinct nodules from the same patient, each of which displayed a rare and identical *MED12* mutation, suggesting a common clonal origin.[308]

Another recent study identified *MED12* mutations in the myometrium directly adjacent to a discrete and well-circumscribed leiomyoma. These authors speculated that neoplastic cells may infiltrate the normal myometrium or that *MED12* mutations occur early in the tumorigenesis, creating a “field effect” in the normal myometrium.[318]

2.2 Branched tumor evolution

Leiomyomas are composed of smooth muscle cells, vascular smooth muscle cells, stem cell-like side population cells, and fibroblasts that presumably work together to stimulate cell proliferation.[240] Although phenotypically different, recent evidence has indicated that all of these cells arise from a single transformed progenitor stem cell.[241]

By studying the genetic architecture of clonally related leiomyomas, we were able to infer that some leiomyomas are composed of genetically distinct subpopulations. Indeed, whole-genome sequencing of four clonally related leiomyomas revealed a pattern of mutations best explained as a spread from a primary tumor. Furthermore, the mutational pattern within this primary tumor is compatible with a branching model of tumor evolution (Figure 15).[373, 374] Such sequence-level heterogeneity has emerged in several cancer types, but not, to our knowledge, in any benign tumors before.[375]

Using SNP arrays, we were able to confirm two additional cases of leiomyomas with intratumor genetic heterogeneity. Furthermore, we also detected a mutation in *ZNHIT* to be consistently subclonally present in nine different parts of a leiomyoma, supporting the fact that leiomyomas are composed of a mixture of genetically distinct cell populations. Whether these genetic subpopulations reflect phenotypically distinct cells that cooperate to stimulate growth remains to be seen.

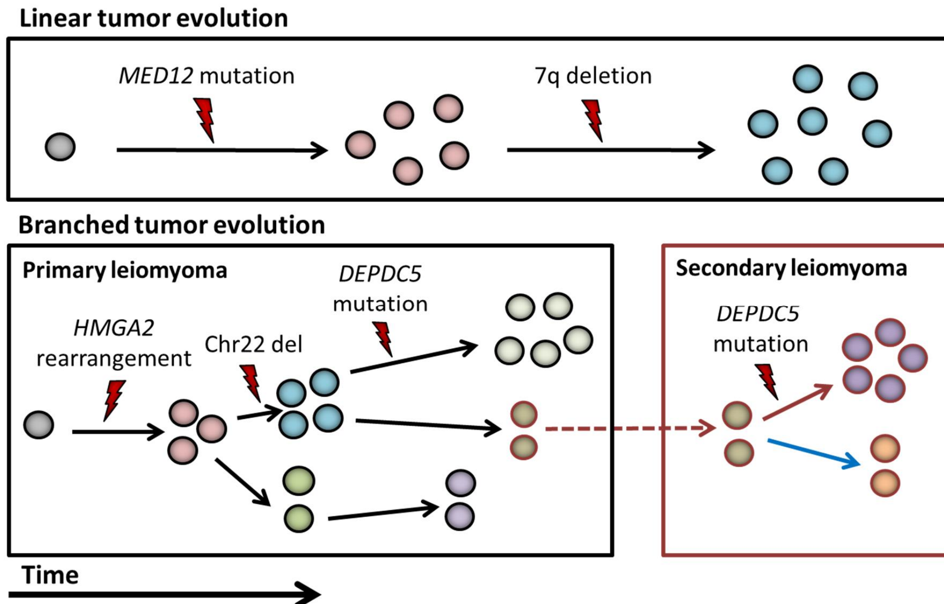


Figure 15. Linear versus branched tumor evolution. Leiomyomas are monoclonal tumors that grow through clonal expansion. The intratumor genetic heterogeneity observed in leiomyomas suggests a branching model of tumor growth, in which cells divide into multiple genetically distinct populations. *MED12* mutations and *HMGA2* rearrangements are the two most common initiating events in leiomyomas, since they appear to be clonally present. In contrast, 7q abnormalities and *DEPDC5* mutations co-occur with other genetic defects and appear to be secondary events present only in a subpopulation of cells.

2.3 DEPDC5 as a secondary driver gene

Whole-genome sequencing of four clonally related leiomyomas revealed an identical deletion on chromosome 22 and a different truncating *DEPDC5* mutation in each tumor. In the context of hereditary cancer syndromes, classical tumor suppressor genes typically follow a pattern in which the first allele is lost through a germline mutation, whereas the second is lost through a larger chromosomal deletion.[376] While the order of events are largely unknown in sporadic cases, here we show – perhaps unexpectedly – that a shared clonal deletion provided the first-hit, but an independent truncating mutation provided the second-hit. Although biallelic loss of *DEPDC5* was secondary in relation to the initiating *HMGA2* translocation, *DEPDC5* mutations were highly selectively advantageous in this patient's tumors. Indeed, we found no other gene to be recurrently mutated among these four tumors. To confirm *DEPDC5* as a recurrently mutated gene, we examined an in-house whole-genome sequencing dataset of 65 leiomyomas for additional *DEPDC5* mutations, revealing one additional case. This tumor displayed a whole chromosome 22 deletion, confirming *DEPDC5* as a tumor suppressor gene. This makes *DEPDC5* the only other gene, since the discovery of *MED12*, to be recurrently mutated on a nucleotide level. This case also provides indisputable evidence for the existence of clonally related leiomyoma nodules.

Differential expression analysis with leiomyomas harboring a 22q deletion revealed *DEPDC5* as one of the most significantly downregulated genes by deletions affecting chromosome 22, supporting the view that *DEPDC5* is target gene from these large-scale deletions. However, the majority of these leiomyomas lacked a second-hit in *DEPDC5*, suggesting that loss of one allele is sufficient to facilitate tumor growth and lead to selection.

Based on functional evidence, *DEPDC5* is a subunit of a tumor suppressor complex that negatively regulates mTORC1,[377] supporting a role for mTOR signaling in leiomyomas. *DEPDC5* is a suspected driver gene in glioblastomas and ovarian cancer, since both of these tumor types occasionally harbor truncating mutations or homozygous deletions affecting the gene.[377] SNPs within the *DEPDC5* locus have also been associated with progression to hepatocellular carcinoma among chronic hepatitis C virus carriers,[378] supporting the view that this gene play a role in tumorigenesis. On the other hand, truncating germline mutations in *DEPDC5* cause hereditary focal epilepsy and brain malformations.[379] Although these germline mutations have not been associated with increased leiomyoma risk, epilepsy is the most common symptom of tuberous sclerosis – caused by mutations in *TSC2* – a gene that, when mutated in rats, cause uterine leiomyomas and renal cell cancer.[242]

Although rare in leiomyomas,[185] deletions of 22q are one the most common abnormalities seen in intravenous leiomyomatosis[224] and benign metastasizing leiomyomas[206], both of which are characterized by quasi-malignant behavior. Deletions of 22q are also common in STUMPs[223] and leiomyosarcomas[380]. Furthermore, leiomyosarcomas may display distinct areas of leiomyoma-like or leiomyosarcoma-like histology, and one study reported that 22q deletions typically occur only in the leiomyosarcoma-like areas.[100] In short, the loss of *DEPDC5* may increase leiomyomas' malignant potential.

3. Chromothripsis in leiomyomas

Since our exome-sequencing effort revealed no other recurrently mutated genes besides *MED12*, the next step was to explore the genomic landscape of leiomyomas by whole-genome sequencing. We unexpectedly discovered a subset of leiomyomas to display highly complex chromosomal rearrangements interconnected by adjacent breakpoints, suggesting that they had occurred simultaneously. Although a few cytogenetic studies have described some leiomyomas to harbor complex chromosomal rearrangements,[150, 381] the notion that these have occurred simultaneously cannot not be inferred without high-resolution data provided by massively parallel sequencing technology. These one-off events in leiomyomas tend to be more similar to chromothripsis than to chromoplexy or chromoanasynthesis, since the events we observed frequently harbored alternating copy-number states and lacked chromosomal amplifications.

Chromothripsis was initially estimated to be a rare phenomenon present only in some 2% of all cancers.[39] Furthermore, chromothripsis has subsequently been associated with *TP53* mutations and poor prognosis,[382] neither of which is true for leiomyomas. Leiomyomas

appear to display "milder" forms of chromothripsis with significantly fewer breakpoints than the malignant tumors do. Furthermore, a few leiomyomas displayed multiple independent events of chromothripsis, an observation rarely reported in the malignant cases. However, the number of breakpoints (or events) should not be the defining feature of chromothripsis, but rather the underlying mechanism and its associated hallmark features.

Chromothripsis in its most intense form would most likely lead to apoptosis in the presence of a normally functioning cell-cycle-checkpoint system. Alternatively, these events may create genetic changes offering a selective advantage. Indeed, we discovered chromothripsis to be a recurrent generator of genetic changes with a selective value, such as the classical rearrangement between *HMGA2* and *RAD51B*. It is tempting to speculate that chromothripsis is not rare in leiomyoma precursor cells, since the vast majority of chromothripsis events are not expected to generate such specific changes. Our observations demonstrate, to our knowledge for the first time, that chromothripsis is not limited merely to malignant tumors nor to neoplasms that frequently harbor *TP53* mutations.

4. Chromosomal abnormalities as drivers

As expected based on previous cytogenetic studies,[150] whole-genome sequencing revealed *HMGA2* as the most commonly rearranged gene. While most of these rearrangements targeted 14q24, we identified two leiomyomas to harbor rearrangements that disrupted the 3′UTR sequence for the microRNA repressor let-7b, thus providing an alternative mechanism for upregulation of *HMGA2*. None of the *HMGA1* rearrangements targeted 14q24. Instead, one of them involved the *CARMN* locus, located on chromosome 5q32. *CARMN* encodes for a long non-coding RNA that is an important regulator of cardiac smooth muscle cell differentiation and homeostasis.[383]

4.1 RAD51B as a putative driver gene

The role of *RAD51B* in leiomyomas has remained a subject of debate. Since *RAD51B* is the preferential translocation partner of *HMGA2*, an effective and tissue-specific regulatory element may be located within or close to the *RAD51B* locus. *HMGA2* (and *HMGA1*) rearrangements occasionally target other regions instead, particularly in other mesenchymal tumors.[150] Interestingly, we found a weaker level of *HMGA2* overexpression in leiomyomas lacking the involvement of *RAD51B*, supporting the view that this locus contains a preferable region for *HMGA2*'s activation. We found two leiomyomas to harbor an *HMGA2* rearrangement in which the breakpoints were located downstream of *RAD51B*, suggesting that the disruption of *RAD51B* is not necessarily selectively advantageous. On the other hand, we identified biallelic loss of *RAD51B* in a leiomyoma that lacked the involvement of *HMGA2*. One study reported an *in vitro* growth advantage in a leiomyoma subpopulation harboring biallelic loss of *RAD51B* compared with a population where only one copy was disrupted through an *HMGA2* rearrangement.[384]

RAD51B encodes for a double-strand break repair enzyme, and loss of *RAD51B* may play a secondary role in inducing genomic instability. Indeed, haploinsufficiency of *RAD51B* results in hypersensitivity to DNA-damaging agents, in a reduction in sister chromatid exchange, in impaired RAD51 focus formation, and an increase in chromosomal abnormalities.[208] Compatible with these observations, we found a significantly higher number of breakpoints in leiomyomas that lacked a *MED12* mutation. In sum, *RAD51B* seems to provide the *HMGAx* genes with an effective regulatory element and at the same time provide a tumorigenic “double hit” by disrupting one copy of *RAD51B*.

We unexpectedly detected *RAD51B* to be the most uniquely expressed gene in leiomyomas of the *MED12* subtype. Exon-level examination revealed that the upregulation of *RAD51B* corresponded to a (long) non-coding transcript of the gene. Interestingly, a previous study reported a unique alternative last exon of *RAD51B* to be exclusively expressed in the uterus,[207] suggesting that disruption of this unique isoform could be selectively advantageous. Since *RAD51B* is specifically targeted in leiomyomas, and because of the existence of uterine-specific *RAD51B* isoforms, this gene may have a unique role in driving tumors that arise from the uterus. It is therefore tempting to speculate that the overexpression of this non-coding *RAD51B* transcript actually contributes to tumorigenesis. Alternatively, overexpression of *RAD51B* may just be a passenger consequence, since this non-coding transcript was also detected in the myometrium – but at lower levels. Functional studies will provide a better understanding of the role of *RAD51B* in tumorigenesis.

Of note, the *HMGA2* and the *RAD51B* loci have both been highlighted as frequent integration hotspots of HPV.[385, 386] Since HPV integration can cause genomic instability, it is tempting to speculate that HPV integration may be involved in forming these rearrangements. *RAD51B* is also located within a known fragile site: regions frequently targeted by viruses.[387]

4.2 CUX1 as a haploinsufficient tumor suppressor gene

Deletions affecting the q-arm of chromosome 7 make up the second most common chromosomal abnormality seen in leiomyomas, and the putative target genes have remained largely elusive.[150] *CUX1* has been highlighted as the strongest candidate, since two leiomyomas recently emerged as harboring inversions disrupting one copy of *CUX1*.[388] With whole-genome sequencing, we discovered that 7q abnormalities are often highly complex, consisting of inversions, translocations, and deletions at various different sites along the chromosome arm. Not only did we discover *CUX1* as the single most commonly deleted gene on 7q, but also found that biallelic loss of *CUX1* may occur in leiomyomas.

Differential expression analysis revealed *CUX1* to be one of the most significantly downregulated genes by 7q deletions. Loss of one allele appears to be sufficient for selection, since, of the 14 leiomyomas with a 7q deletion, only one displayed biallelic loss of *CUX1*. Compatible with these observations, *CUX1* is a haploinsufficient tumor suppressor gene in acute myeloid leukemia.[389] Surprisingly, we and others have been unable

to identify any nucleotide-level mutations affecting *CUX1*,[390] suggesting that it is easier to inactivate *CUX1* through a variety of large-scale chromosomal abnormalities. *CUX1* plays an important role in the ATM/ATR DNA-damage response pathway,[391] suggesting that loss of this gene may induce genomic instability. In sum, it is safe to assume that loss of *CUX1* drives leiomyoma development, although the exact mechanism awaits elucidation.

4.3 IRS4 as a driver of COL4A5-COL4A6 deletions

We discovered a subset of leiomyomas to harbor characteristic deletions that simultaneously affected *COL4A5* and *COL4A6*, located head-to-head on chromosome Xq22.3. Similar deletions have previously been discovered in the germline of Alport syndrome patients displaying diffuse leiomyomatosis (ATS-DL) of the esophageal, tracheobronchial, and genitourinary tract. Such deletions can also occur as somatic events in esophageal leiomyomas,[392] confirming that they play a role in stimulating smooth muscle overgrowth.

A variety of germline loss-of-function mutations in *COL4A2*, *COL4A3*, and *COL4A5* causes Alport syndrome,[393] but only a specific type of deletion that affects *COL4A5* and *COL4A6* leads to diffuse leiomyomatosis as well, suggesting that the minimally deleted region contains important driver sequences. A recent study described diffuse leiomyomatosis in a patient harboring a germline deletion that did not span *COL4A6*,[158] suggesting that the critical sequences are located within *COL4A5*. Moreover, diffuse leiomyomatosis is not a symptom if the patient harbors a germline deletion that extends past intron 3 of *COL4A6*,[394] suggesting that other important sequences are located beyond this region. Compatible with these observations, the somatic deletions we detected did not extend past intron 3 of *COL4A6*.

We identified *IRS4* – a gene located downstream of *COL4A5* – to be the most significantly upregulated gene in leiomyomas harboring these deletions, suggesting that *IRS4* is the actual driver from these deletions rather than *COL4A5* or *COL4A6*. One logical explanation for these patterns is that the deleted region contains an insulator that normally blocks the interaction between the promoter of *IRS4* and a distant enhancer located in proximity to intron 3 of *COL4A6* (Figure 16). *IRS4* encodes for the insulin receptor substrate 4, which can stimulate cell proliferation by enhancing the function of IGF-1.[395] Interestingly, subungual exostosis and pediatric T-cell acute lymphoblastic leukemia have been shown to harbor recurrent translocations that result in *IRS4* upregulation.[396-398] Furthermore, the *COL12A1* locus is a recurrent translocation partner in subungual exostosis,[397] suggesting that this region also contains an enhancer that can upregulate *IRS4*.

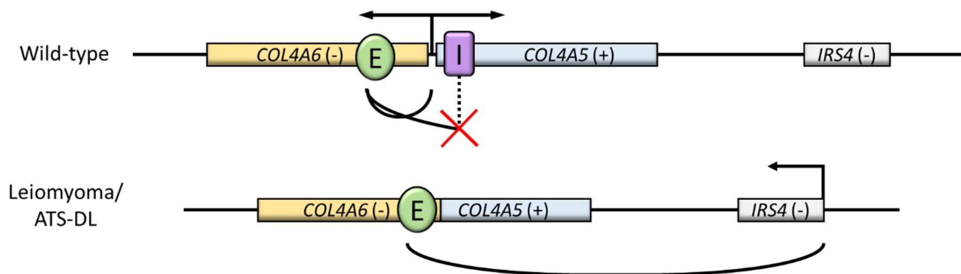


Figure 16. The enhancer-insulator model of *IRS4* activation in uterine leiomyomas and ATS-DL patients with a *COL4A5-COL4A6* deletion. A deletion that typically affects the 5' end of both *COL4A5* and *COL4A6* occurs in a subset of uterine leiomyomas and ATS-DL patients. According to the enhancer-insulator model, the loss of a putative insulator (I) – located within *COL4A5* – results in activation of *IRS4* though a distal enhancer (E) – located beyond intron 3 of *COL4A6*.

4.4 PLAG1 as a putative driver gene

We identified one leiomyoma of the *MED12* subtype as displaying a translocation between *PLAG1* and *COL12A1*, resulting in high expression of *PLAG1*. Interestingly, the *COL12A1* locus is also targeted by *IRS4* translocations in subungual exostosis,[396, 397] supporting the view that the *COL12A1* locus contains a selectively advantageous regulatory element. We also found upregulation of *PLAG1* in two quadruple-negative leiomyomas that harbored a large-scale amplification of the gene. *PLAG1* encodes for a zinc finger transcription factor that is frequently upregulated by translocations in several benign mesenchymal tumors.[399] Furthermore, *PLAG1* and *HMGA2* translocations are mutually exclusive in pleomorphic adenomas of the salivary gland.[400] Interestingly, we detected *PLAG1* to be distinctly upregulated in leiomyomas harboring an *HMGA2* or *HMGA1* abnormality, suggesting that these genes are upstream regulators of *PLAG1*. This hypothesis is supported by the fact that *PLAG1* translocations can also target *RAD51B* in lipoblastomas.[401] Furthermore, *HMGA2* and *PLAG1* expression levels correlate in thyroid tumors, leiomyomas, and experimental models.[402] Of note, we detected the putative tumors suppressor *PLAGL1* to be among the most uniquely downregulated genes in leiomyomas of the *COL4A5-COL4A6* subtype. *PLAGL1* is structurally similar, but functionally different to *PLAG1*.[403]

4.5 Other candidate driver genes

The 7q deletions we detected in leiomyomas were often large, complex, and a few did not span *CUX1*, suggesting that other genes are simultaneously targeted. According to the "cancer gene island model", regions recurrently affected by hemizygous deletions are likely to harbor multiple target genes that maximize proliferative fitness through cumulative haploinsufficiencies.[404] We identified one sample as displaying a second-hit mutation in *ZNHIT*, and two other samples as displaying highly localized rearrangements affecting *CUL1*, suggesting that they are also targets of 7q deletions. ZNHIT1 is involved in cyclin dependent kinase 6 (CDK6)-driven cell-cycle arrest at the G1 phase,[405] whereas CUL-1 forms the major structural scaffold part of the SCF-complex and play a role in ubiquitin-

dependent degradation of numerous cell-cycle regulators including CCND1.[406] Interestingly, we identified one leiomyoma as displaying a *CCND1* rearrangement, resulting in the highest expression of the gene among the 38 leiomyomas analyzed. Compatible with our observations, others have reported a significant upregulation of genes related to cell-cycle progression in leiomyomas of the *HMGA2* subtype, including *CCND1*, *CCND2*, *CCND3*, and *CDK6*.[203]

We also identified another commonly deleted region on 22q that contained the known tumor suppressor *SMARCB1*. Furthermore, one leiomyoma harbored an additional chromosomal rearrangement within this region, resulting in biallelic loss of *SMARCB1*. *SMARCB1* is of special interest, since this gene was recently associated with an increased risk for leiomyomas though a germline *SMARCB1* mutation.[157]

Deletions of 1p represent another frequently occurring event in leiomyomas that has been associated with distinct histopathological features and possible malignant progression.[225, 265] Chromosome 1p deletions are common in several cancers and the target genes remain elusive in these as well. *CHD5*, *CAMTA1*, *KIF1B*, *CASZ*1, and *miR-34a* have been highlighted as the foremost candidates in some cancers,[407] and among these, we identified *kinesin family member 1B* (*KIF1B*) as one of the most significantly downregulated gene in leiomyomas with 1p deletions. However, deletion mapping revealed *NPHP4* to be the most commonly deleted gene on chromosome 1p. *NPHP4* has previously been highlighted as a putative target gene in leiomyomas, owing to recurrent translocation breakpoints located upstream of the *NPHP4* locus.[226] We detected no significant dysregulation of *NPHP4* in leiomyomas with 1p deletions, however.

We also identified one quadruple-negative leiomyoma to harbor a fusion gene involving *IGFBP5* and *PDGFRB* (Figure 17). Fusions involving the same 3'end of *PDGFRB* are frequent in myeloid and lymphoid neoplasms and are associated with hypereosinophilia.[408] Such fusions encode for constitutively activated tyrosine kinases; patients with such defects respond well to low doses of the tyrosine-kinase inhibitor imatinib.[408] While identified in only one quadruple-negative leiomyomas thus far, leiomyomas with such fusions may represent a rare leiomyomas subtype. Although fusions involving *PDGFRB* are known to drive hematopoietic cancers, the disruption of *IGFBP5* may further enhance leiomyoma development.

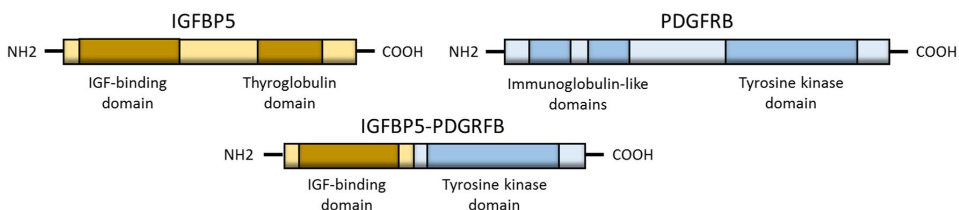


Figure 17. IGFBP5-PDGFRB fusion protein as a putative driver. We detected a quadruple-negative leiomyoma to harbor a chromosomal rearrangement resulting in the formation of an in-frame fusion gene consisting of the first exon of *IGFBP5* and the last five exons of *PDGFRB*. The putative protein product includes the IGF-binding domain of IGFBP5 and the tyrosine-kinase domain of PDGFRB. This fusion protein is likely to cause continuous activation of the tyrosine-kinase domain.

5. Molecular classification of uterine leiomyomas

Uterine leiomyomas are still widely regarded as a single entity, although evidence of heterogeneity exists at several different levels including symptoms, histopathology, and therapeutic requirements. Our high-throughput sequencing studies have emphasized heterogeneity on the genetic level as well, suggesting that molecularly distinct subtypes of leiomyomas exist (Figure 18). Oncogenic *MED12* mutations clearly represent the most common genetic defect in leiomyoma, with *HMGA2* rearrangements as the second most common. Interestingly, these two defects are mutually exclusive,[314] suggesting that they represent two distinct subtypes of leiomyomas. Biallelic loss of *FH* is present only in 1% of sporadic leiomyomas, and although the majority of HLRCC patients' tumors show loss of the second *FH* allele, some may occasionally harbor a *MED12* mutation. Interestingly, such lesions do not show loss of the second *FH* allele,[409] indicating that *FH* and *MED12* mutations are also mutually exclusive. We have also identified a few leiomyomas with *COL4A5-COL4A6* deletions, all of which have been mutually exclusive with the other three drivers. More samples are required for validation, however. These four genetic defects appear to account for over 95% of all leiomyomas, leaving only a fraction of leiomyomas with no clearly identifiable initiating event (Figure 18).

In contrast to these four mutually exclusive driver genes, deletions of 7q22, 22q, and 1p frequently co-occur with other genetic changes, suggesting that they are related to tumor progression rather than tumor initiation. Furthermore, we identified several examples where these changes were subclonal. Surprisingly, others and we have shown that *HMGA1* rearrangements can co-occur with *MED12* mutations as a secondary abnormality relative to the initiating *MED12* mutation.[310] However, we also identified some quadruple-negative leiomyomas with an *HMGA1* abnormality to display expression patterns that are similar to those seen in leiomyomas of the *HMGA2* subtype, suggesting that *HMGA1* rearrangements may also occur as initiating events.

One interesting candidate driver gene in leiomyomas is *KAT6B*, which is rearranged in leiomyomas at approximately 2%.[185] This specific translocation may exist as a sole cytogenetic abnormality, suggesting that it represents another molecularly distinct pathway to leiomyoma formation. Indeed, a recent study screened for *MED12* mutations in a retroperitoneal leiomyoma harboring a *KAT6B* rearrangement, but found none.[220] Interestingly, germline mutations in either *KAT6B* or *MED12* result in Ohdo syndrome,[410] suggesting that these two genes are functionally related.

Our gene expression results support the existence of various leiomyoma subtypes, since the mutation status of *MED12*, *HMGA2*, *FH*, and *COL4A5-COL4A6* clearly affected global gene expression patterns. These distinct patterns suggest that different genetic defects act through separate pathways. Leiomyomas with *HMGA1* or *HMGA2* abnormalities displayed similar global expression signatures, supporting that these structurally and evolutionarily related transcription factors function similarly in tumorigenesis. In contrast, deletions of 7q22, 22q, and 1p exerted no major influence on the clustering, supporting their role in

tumor progression. The majority of quadruple-negative leiomyomas clustered into several unique branches, indicating the presence of a few rare and possibly novel subtypes.

These different mutations are also associated with distinct phenotypes. Indeed, leiomyomas of the *MED12* subtype tends to be smaller,[307] and leiomyomas of the *HMGA2* subtype tends to be larger.[199] Furthermore, *HMGA2* subtype tumors tend to develop as solitary lesions, whereas *MED12* subtype tumors frequently develop as multiple independent lesions.[309] We also found leiomyomas harboring *MED12* or *FH* mutations to be more chromosomally stable. Unlike cultured leiomyoma cells with *MED12* or 7q abnormalities, cells with *HMGA2* rearrangements are able to survive *in vitro* for numerous passages.[371]

Some researchers have hypothesized that only a subset of leiomyomas with variant histology or rare genetic defects have the potential for malignant progression. Indeed, uterine leiomyomas with bizarre nuclei was recently shown to harbor *TP53* and *FH* mutations.[320] Another recent study proposed that leiomyomas with bizarre nuclei should be further divided into two distinct subtypes.[411] Indeed, one of these subtypes showed a higher rate of *MED12* mutations and immunoreactivity for *HMGA2*,[411] suggesting that the other subtype corresponds to leiomyomas with *FH* mutations. Other studies have shown that leiomyomas with bizarre nuclei and cellular leiomyomas resemble leiomyosarcomas more than they do leiomyomas, since both of these can harbor *TP53* mutations, and both tend to cluster with leiomyosarcomas in gene expression studies.[320] Considering that *KAT6B* rearrangements tend to occur in cellular leiomyomas and that they have been described in a few leiomyosarcomas,[219, 222] these defects may increase the malignant potential of leiomyomas. Gene expression profiling has also shown that leiomyomas with 1p36 deletions tend to cluster with leiomyosarcomas,[225] suggesting that 1p36 deletions are related to malignant progression as well.

Of note, many of the genes involved in leiomyoma formation have been implicated as drivers of some hematological malignancies. Indeed, chronic lymphocytic leukemia is one of the few cancer types that harbor *MED12* mutations,[333] whereas *HMGA2*, *KAT6B* and *PDGFRB* rearrangements occur in acute myeloid leukemia as well.[195-197, 221, 408] In addition, *CUX1* is frequently inactivated in acute myeloid leukemia, where it acts as a haploinsufficient tumor suppressor gene.[389] A very recent study showed that MED12 is essential to the growth of hematopoietic stem cells in the bone marrow.[412] Transgenic mice overexpressing human *HMGA2* under control of the VH promoter/Eμ enhancer develop a lymphoproliferative disease that resembles human T-cell acute lymphoblastic leukemia.[246] Chronic lymphocytic leukemia is also the very first cancer type in which chromothripsis was discovered.[39]

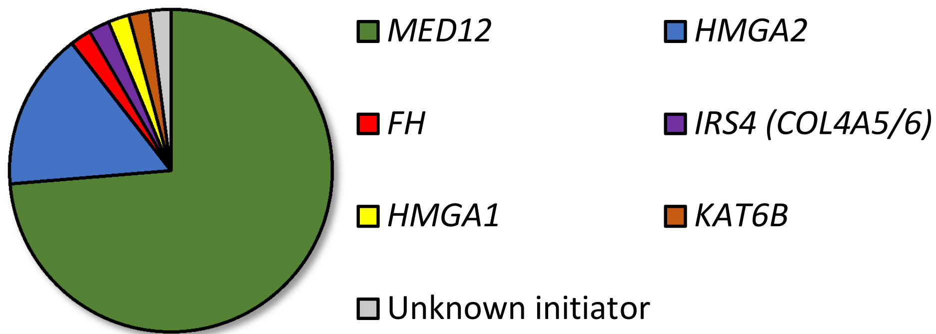


Figure 18. Initiating mutations in leiomyomas. Oncogenic *MED12* mutations are clearly the most common initiating mutation in leiomyomas. Rearrangements resulting in upregulation of *HMGA2* are another common initiating event. Rearrangements of *HMGA1* may represent a rare initiating event, although it can occur as secondary driver as well. Biallelic loss of *FH* can also initiate tumorigenesis, and the first-hit can be inherited. Upregulation of *IRS4* resulting from deletions affecting the *COL4A5* and *COL4A6* locus appear to be a rare initiating event. Rearrangements resulting in fusion genes involving *KAT6B* are likely to represent another rare initiating event. In only a small fraction of leiomyomas is the initiating event unknown.

6. Putative biomarkers and dysregulated pathways

Our gene expression analyses revealed several genes to be uniquely expressed in leiomyomas of different subtypes. The most significant of these represent putative biomarkers with the potential to classify leiomyomas in future studies. Although succination of proteins by fumarate has previously been used as a valuable immunohistochemical biomarker for *FH*-deficiency,[413] we identified *AKR1B10* to be uniquely expressed in leiomyomas of the *FH* subtype, suggesting that it may serve as an alternative biomarker. We also detected specific driver mutations to be the main determinants of expression changes in leiomyomas. In addition to the previously described changes in Wnt/β-catenin signaling, we identified subtype-specific expression changes in other key tumorigenic pathways, including prolactin, IGF-1, and NRF2 signaling (Figure 20).

6.1 Activation of Prolactin signaling

We detected a significant activation of prolactin signaling in leiomyomas, particularly in leiomyomas of the *MED12* and *HMGA2* subtypes. Indeed, *PRL* and *PRLHR* were among the most upregulated genes in leiomyomas. *PRL* encodes for the growth hormone prolactin, and we observed a particularly high expression of *PRL* in the *HMGA2* subgroup. Interestingly, transgenic mice overexpressing *HMGA2* develop pituitary adenomas secreting prolactin.[414] In contrast, leiomyomas of the *MED12* subtype displayed a

particularly high *PRLHR* expression. *PRLHR* encodes for a prolactin-releasing peptide receptor that binds to the prolactin-releasing peptide, which then stimulates the release of prolactin.[415] Increased expression of *PRLHR* recently emerged as stimulating the proliferation of cultured human leiomyoma cells, and transgenic mice overexpressing *PRLHR* develop myometrial hyperplasia with excessive extracellular matrix deposition.[252] Prolactin may act as an autocrine or paracrine growth factor in several tumor types, including leiomyomas.[416]

6.2 Dysregulation of IGF-1 signaling

In most leiomyoma subtypes, we identified a significant dysregulation of IGF-1 signaling. Furthermore, statistical analysis revealed *IGF2BP2* as the second most uniquely expressed gene in leiomyomas of the *HMGA2* subtype. *HMGA2* directly regulates the expression of *IGF2BP2* by binding to an AT-rich regulatory region within its first intron.[298] *IGF2BP2* encodes for the insulin-like growth factor 2 mRNA binding protein 2 that plays a role in promoting *IGF2* mRNA translation.[417] Interestingly, *PLAG1* regulates the expression of *IGF2* by binding to its P3 promoter.[299-301] We detected upregulation of *IGF2* in the majority of leiomyomas of the *HMGA2* subtype; *IGF2* encodes for the insulin-like growth factor 2, which binds to the IGF-1 receptor and thereby promotes growth. *HMGA2*, *PLAG1*, *IGF2BP2*, and *IGF2* are all silenced in adult tissues,[298, 418, 419] supporting the view that they are functionally related.

Statistical analysis revealed *ADAM12* to be among the most uniquely upregulated genes in leiomyomas of the *MED12* subtype, and *pappalysin 2* (*PAPPA2*) to be among the most uniquely upregulated genes in leiomyomas of the *HMGA2* subtypes. Both of these genes are involved in placental development and encode for proteases of IGFBP-5.[420, 421]. Insulin-like growth factor binding proteins (IGFBPs) inhibit the activity of the IGF type 1 receptor by binding to IGFs. IGFBP-5 inhibits IGF-1-induced proliferation and migration of smooth muscle cells.[422] The role of IGFBP-5 in leiomyomas remains unclear, because IGFBP-5 has been described to both stimulate as well as inhibit cancer development.[423]

6.3 Activation of the NRF2 pathway in leiomyomas of the FH subtype

In leiomyomas of the *FH* subtype, we found the NRF2 pathway to be the most significantly dysregulated pathway, supporting the theory that high levels of fumarate leads to an activation of the oncogenic transcription factor NRF2.[172] The pentose phosphate pathway was the only other significant pathway, supporting that NRF2 play a role in redirecting glucose and glutamine into anabolic pathways.[305] Conversely, the HIF-1-α signaling pathway was not significantly altered, challenging the pseudo-hypoxia hypothesis.[166] Furthermore, a few studies have hypothesized that loss of *FH* could lead to tumorigenesis though DNA damage caused by an increased production of free radicals.[174] However, we found a low number of chromosomal abnormalities and nucleotide-level mutations in leiomyomas of the *FH* subtype.

Activation of the NRF2 signaling pathway has recently emerged as a common oncogenic feature of many cancers.[424] Indeed, the NRF2 pathway can be activated through gain-of-function mutations in *NFE2L2*, or loss-of-function mutations in either *cullin 3* (*CUL3*) or *KEAP1* (Figure 19).[425] These mutations result in loss of the interaction between KEAP1 and NRF2,[425] leading to stabilization of NRF2. A recent comprehensive study on the molecular characteristics of renal cell cancers revealed that renal tumors of papillary type 2 histology frequently display activation of the NRF2 pathway though mutations in *NFE2L2*, *CUL3*, *KEAP1*, or *FH*. Since the hypermethylation phenotype was detected only in the renal tumors that harbored *FH* mutations,[169] *NRF2* activation – rather than hypermethylation – is likely the primary mechanism by which these gene defects drive tumorigenesis. If NRF2 activation is indeed the underlying mechanism, it would be reasonable to assume that a subset of leiomyomas may harbor mutations in other genes of the KEAP1-NRF2-CUL3 apparatus. Conversely, if hypermethylation is the driving force, a subset of leiomyomas could harbor mutations in other genes that result in the hypermethylation phenotype, such as in *IDHx* or *SDHx*. Novel therapeutic strategies have recently emerged as targeting the NRF2 pathway,[426] suggesting that NRF2 inhibitors could be useful to combat leiomyomas that are driven by *FH* mutations.

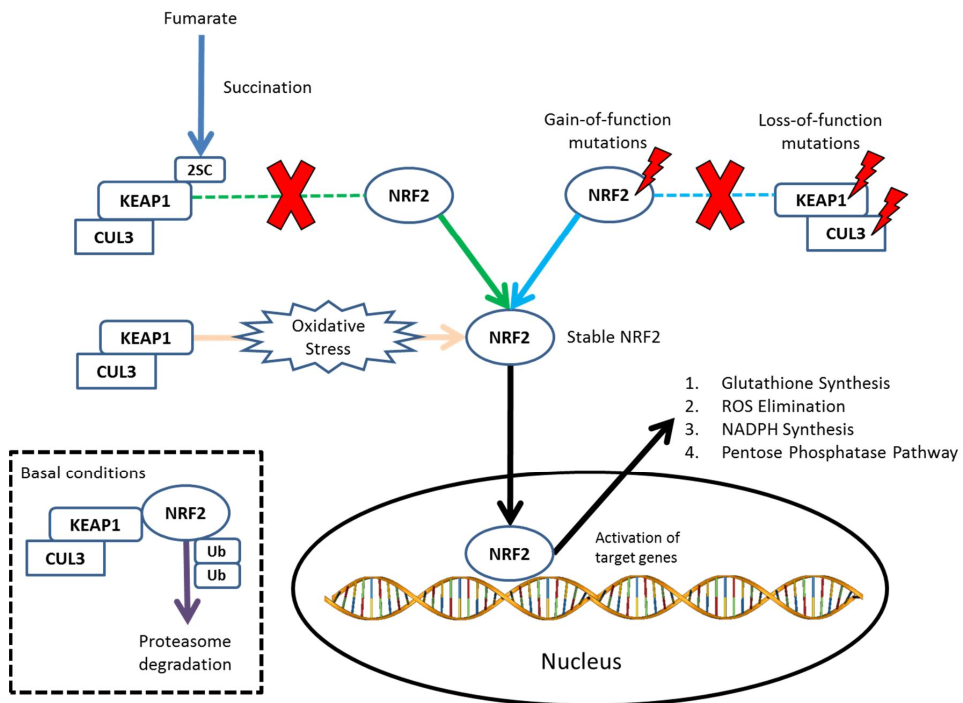


Figure 19. Different routes to NRF2 activation. NRF2 is, under basal conditions, degraded in the cytosol by a KEAP1 homodimer that facilitates the ubiquitination and proteasomal degradation of NRF2. Under high levels of oxidative stress, NRF2 is relocated into the nucleus, where it activates the expression of genes involved in the antioxidant pathway. The interaction between NRF2 and KEAP1 can be lost through gain-of-function mutations in *NRF2* or through loss-of-function mutations in *KEAP1* or *CUL3*. Recent studies have proposed that KEAP1 becomes succinated by high levels of fumarate, leading to loss of this interaction. Abnormally high levels of NRF2 activation are likely to stimulate tumor growth.

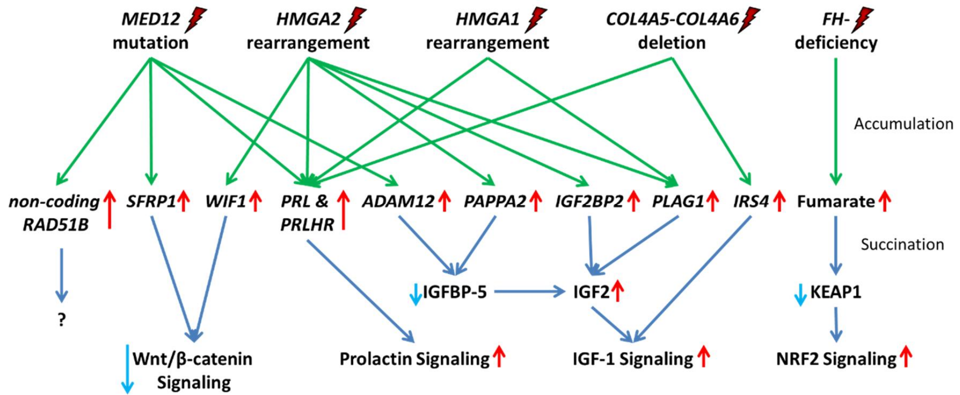


Figure 20. Schematic of highlighted driver genes and pathways in leiomyoma development and growth. Leiomyomas display subtype-specific differences in key driver pathways including Wnt/β-catenin, prolactin, IGF-1, and NRF2 signaling. Leiomyomas of the *MED12* subtype exhibit high expression of a non-coding transcript of *RAD51B*; *RAD51B* is also the most common translocation partner of *HMGA2* in leiomyomas.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The findings presented in this thesis work represent an important contribution to our understanding of uterine leiomyoma etiology. Although these studies have provided answers to several longstanding questions, they have also raised several new questions: why are *MED12* mutations so common? Is one mutation sufficient to generate a clinically relevant leiomyoma? What is the mechanism by which leiomyomas spread? Are some rare leiomyoma subtypes prone to malignant progression? Are ATS-DL patients at increased risk for uterine leiomyomas as well? Do molecular differences in leiomyomas explain variations frequently seen in clinical outcomes or responses to treatments?

Our massively parallel sequencing efforts have revealed not only several novel driver genes in leiomyomas, but also revealed that chromothripsis-like events are not limited to malignant tumors. Furthermore, these technologies have made it possible to demonstrate that some separate leiomyomas are indeed of common clonal origin. While the most important genetic drivers of this disease appear to be discovered, other less common genetic defects related to tumor initiation or progression remain to be clarified. This may be particularly true for leiomyoma variants displaying rare histopathology or unusual growth patterns. Further research into leiomyoma variants displaying some features of malignancy may also improve our understanding of the early stages of malignant transformation. The genetic factors that predispose to leiomyomas demand further research as well, because the currently discovered variants explain only a small fraction of the entire estimated heritability.

Discoveries made through genome-wide methods in these studies have now laid a basis for the molecular classification of leiomyomas. Further research will determine whether the candidate biomarkers presented here will provide guidance to researchers and health professionals. The ability to stratify each lesion into a clinically relevant subclass will pave the way for improved management through personalized treatments. While these studies highlight the importance of leiomyoma stratification, they also show that several genes and pathways are commonly dysregulated in leiomyomas, regardless of subtype. Targeting these shared abnormalities would be an ideal step in the development of treatments for this disease. A comprehensive understanding of leiomyoma etiology cannot be reached through genetics only, however. Leiomyoma research has historically been limited by lack of suitable cell cultures and animal models. The recent emergence of an animal model accurately resembling human leiomyomas may finally prove to be an invaluable tool for investigating targeted therapies and obtaining a comprehensive understanding of this very common disease.

ACKNOWLEDGEMENTS

This thesis work was carried out at the Department of Medical and Clinical Genetics and Research Programs Unit, Faculty of Medicine, University of Helsinki during 2011-2016. I warmly thank the former and the present heads of the departments for their excellent research facilities at my disposal. I acknowledge the Doctoral Programme in Biomedicine (DPBM) for their high-quality education and financial support. This thesis was also financially supported by personal grants from the Biomedicum Helsinki Foundation, the Orion Research Foundation, the Emil Aaltonen Foundation, the Cancer Society of Finland, the Ida Montin Foundation, and the Maud Kuistila Memorial Foundation, to whom I am sincerely grateful. I also thank the Institute for Molecular Medicine Finland, the Biomedicum Functional Genomics Unit, the Estonian Genome Center, and the CSC-IT Center for Science for their excellent services.

My deepest gratitude goes to my supervisor Lauri Aaltonen for his continuous support and faith in me during these years. I admire Lauri's scientific curiosity, open-mindedness, and big ambitions, which have inspired me to grow as a researcher. My heartfelt thanks goes to my other supervisors, Esa Pitkänen and Pia Vahteristo, for guiding me throughout this journey. I thank Esa "Mr. p-value" for his kindness and his expertise in statistics and bioinformatics. I thank Pia for her encouragement and for introducing me to the mysterious world of leiomyomas. I am also grateful to my thesis advisory members, Ville Hietakangas and Matti Nykter, for their continuous support.

I thank Merja Heinäniemi and Ville Hietakangas for their critical review of this thesis and their valuable comments and suggestions. I owe special thanks to Carol Norris for her excellent author-reading and for her kind hospitality.

I sincerely thank the clinicians: Nanna Sarvilinna, Jari Sjöberg, Ralf Bützow, Elina Virolainen, Annukka Pasanen, Tom Böhling, and Oskari Heikinheimo for providing us with invaluable sample materials and their excellent clinical expertise. I warmly thank all other collaborators and co-authors without whom this work would have been impossible: Massimiliano Gentile, Anna Vähärautio, Virpi Launonen, Taru Claypole, Anna Kuosmanen, Jian Yan, Martin Enge, Minna Taipale, and Jussi Taipale.

It has been a pleasure to work with all the present and former members of the Aaltonen group. Special thanks goes to the myoma gang: Netta, Hanna, Kati, Jaana, and Simona. I thank Netta for making sure that everything runs smoothly; I would have been in trouble without you. I thank Hanna for having the patience to teach me wet-lab work and for her friendship. I would also like to thank Kati for her endless positivity and mental support.

I extend my sincere thanks to Sini, Sirpa, Alison, Inga-Lill, Iina, Alison, Jiri, Javier, Lauri S., Heikki R., and Marjo for their excellent technical assistance. I thank Sini for creating an atmosphere that is less “nerdy” to work in. I would also like to thank all the postdocs and senior scientists for their valuable guidance: Heli L., Alex, Sari, Silva, Linda F., Outi, Kimmo, Niko, Rainer, and Auli. I thank Heli for reassuring me that life does not have to be boring after you turn 30. I would like to wish my fellow graduate students good luck with their work: Johanna, Ulrika, and Tomas.

I thank the “original incubators” and the “Menorca gang” for making the Aaltonen group feel like a community of friends. I thank Riku for his contagious enthusiasm towards science and for reminding me to put life into perspective; Mervi “the original hipster” for her honesty; Iikki for her spontaneity and endless energy; Yilong for all the fun (especially on our inter-rail trip), and Eevi for challenging me in our endless “scientific” discussions and for joining me at the Lady Gaga concert. Heikki M., I henceforth pronounce you the future king of floorball. Big thanks go to my “movie-night” trio of friends: Tatiana, Sofie, and Linda B. I am grateful for all the help in the preparations for my party!

I am thankful for my entourage outside of the research world: Aamos, Macce, Jonathan, Bisse, and Panu. Halo Rulez!!

I am indebted to all of the patients who participated in the studies.

Finally, I would like to thank my parents, Pirkko and Eduard, and my siblings, Mooses and Merike, for their lifetime of love. Missan, happy 20th birthday in advance!

Helsinki, September 2016

Miika Mehine

REFERENCES

1. Sleeman, J. & Steeg, P. S. Cancer metastasis as a therapeutic target *European journal of cancer* **46:** 1177-1180 (2010)

2. Stewart, E. A. Uterine fibroids *Lancet* **357:** 293-298 (2001)

3. Colao, A. & Savastano, S. Medical treatment of prolactinomas *Nature Reviews Endocrinology* **7:** 267-278 (2011)

4. Morson, B. C. Precancerous Lesions of the Colon and Rectum *JAMA: The Journal of the American Medical Association* **179:** 316-321 (1962)

5. Vogelstein, B. & Kinzler, K. W. The multistep nature of cancer *Trends in Genetics* **9:** 138-141 (1993)

6. de Magalhães, J. P. How ageing processes influence cancer *Nature Reviews Cancer* **13:** 357-365 (2013)

7. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation *Cell* **144:** 646-674 (2011)

8. Narayanan, D. L., Saladi, R. N. & Fox, J. L. Ultraviolet radiation and skin cancer *International journal of dermatology* **49:** 978-986 (2010)

9. Hecht, S. S. Tobacco smoke carcinogens and lung cancer *Journal of the National Cancer Institute* **91:** 1194-1210 (1999)

10. WHO | The top 10 causes of death (2016)

11. Ligibel, J. A., Alfano, C. M., Courneya, K. S., Demark-Wahnefried, W., Burger, R. A., Chlebowski, R. T., Fabian, C. J., Gucalp, A., Hershman, D. L., Hudson, M. M., Jones, L. W., Kakarala, M., Ness, K. K., Merrill, J. K., Wollins, D. S. & Hudis, C. A. American Society of Clinical Oncology Position Statement on Obesity and Cancer *Journal of Clinical Oncology* **32:** 3568-3574 (2014)

12. Polk, D. B. & Peek, R. M. Helicobacter pylori: gastric cancer and beyond *Nature Reviews Cancer* **10:** 403-414 (2010)

13. Moody, C. A. & Laimins, L. A. Human papillomavirus oncoproteins: pathways to transformation *Nature Reviews Cancer* **10:** 550-560 (2010)

14. Nagy, R., Sweet, K. & Eng, C. Highly penetrant hereditary cancer syndromes *Oncogene* **23:** 6445-6470 (2004)

15. Tomasetti, C. & Vogelstein, B. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions *Science* **347:** 78-81 (2015)

16. Pitot, H. C. The molecular biology of carcinogenesis *Cancer* **72:** 962-970 (1993)

17. Greaves, M. & Maley, C. C. Clonal evolution in cancer *Nature* **481:** 306-313 (2012)

18. Kreso, A. & Dick, J. E. Evolution of the cancer stem cell model *Cell stem cell* **14:** 275-291 (2014)

19. Whiteside, T. L. The tumor microenvironment and its role in promoting tumor growth *Oncogene* **27:** 5904-5912 (2008)

20. Eng, C., Leone, G., Orloff, M. S. & Ostrowski, M. C. Genomic Alterations in Tumor Stroma *Cancer Research* **69:** 6759-6764 (2009)

21. Karki, R., Pandya, D., Elston, R. C. & Ferlini, C. Defining "mutation" and "polymorphism" in the era of personal genomics *BMC medical genomics* **8:** 37-015-0115-z (2015)

22. 1000 Genomes Project Consortium, Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A. & Abecasis, G. R. A global reference for human genetic variation *Nature* **526:** 68-74 (2015)

23. Baralle, D. & Baralle, M. Splicing in action: assessing disease causing sequence changes *Journal of medical genetics* **42:** 737-748 (2005)

24. Madan, K. What Is a Complex Chromosome Rearrangement? *American Journal of Medical Genetics Part A* **161:** 1181-1184 (2013)

25. Geigl, J. B., Obenauf, A. C., Schwarzbraun, T. & Speicher, M. R. Defining ‘chromosomal instability’ *Trends in Genetics* **24:** 64-69 (2008)

26. Vijg, J. & Suh, Y. Genome instability and aging *Annual Review of Physiology* **75:** 645-668 (2013)

27. Govindan, R., Ding, L., Griffith, M., Subramanian, J., Dees, N. D., Kanchi, K. L., Maher, C. A., Fulton, R., Fulton, L., Wallis, J., Chen, K., Walker, J., McDonald, S., Bose, R., Ornitz, D., Xiong, D., You, M., Dooling, D. J., Watson, M., Mardis, E. R. & Wilson, R. K. Genomic landscape of non-small cell lung cancer in smokers and never-smokers *Cell* **150:** 1121-1134 (2012)

28. Barnard, S., Bouffler, S. & Rothkamm, K. The shape of the radiation dose response for DNA double-strand break induction and repair *Genome Integrity* **4:** 1 (2013)

29. McGranahan, N., Burrell, R. A., Endesfelder, D., Novelli, M. R. & Swanton, C. Cancer chromosomal instability: therapeutic and diagnostic challenges *EMBO reports* **13:** 528-538 (2012)

30. Chen, J. M., Cooper, D. N., Ferec, C., Kehrer-Sawatzki, H. & Patrinos, G. P. Genomic rearrangements in inherited disease and cancer *Seminars in cancer biology* **20:** 222-233 (2010)

31. Moldovan, G. & D'Andrea, A. D. How the Fanconi Anemia Pathway Guards the Genome *Annual Review of Genetics* **43:** 223-249 (2009)

32. Bunting, S. F. & Nussenzweig, A. End-joining, translocations and cancer *Nature Reviews Cancer* **13:** 443-454 (2013)

33. Bastians, H. Causes of Chromosomal Instability *Recent Results in Cancer Research* **200:** 95-113 (2015)

34. Grady, W. M. Genomic instability and colon cancer *Cancer metastasis reviews* **23:** 11-27 (2004)

35. Roberts, S. A. & Gordenin, D. A. Hypermutation in human cancer genomes: footprints and mechanisms *Nature Reviews Cancer* **14:** 786-800 (2014)

36. Shendure, J. & Ji, H. Next-generation DNA sequencing *Nature biotechnology* **26:** 1135-1145 (2008)

37. Helleday, T., Eshtad, S. & Nik-Zainal, S. Mechanisms underlying mutational signatures in human cancers *Nature Reviews Genetics* **15:** 585-598 (2014)

38. Maciejowski, J., Li, Y., Bosco, N., Campbell, P. J. & de Lange, T. Chromothripsis and Kataegis Induced by Telomere Crisis *Cell* **163:** 1641-1654 (2015)

39. Stephens, P. J., Greenman, C. D., Fu, B., Yang, F., Bignell, G. R., Mudie, L. J., Pleasance, E. D., Lau, K. W., Beare, D., Stebbings, L. A., McLaren, S., Lin, M. L., McBride, D. J., Varela, I., Nik-Zainal, S., Leroy, C., Jia, M., Menzies, A., Butler, A. P., Teague, J. W., Quail, M. A., Burton, J., Swerdlow, H., Carter, N. P., Morsberger, L. A., Iacobuzio-Donahue, C., Follows, G. A., Green, A. R., Flanagan, A. M., Stratton, M. R., Futreal, P. A. & Campbell, P. J. Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development *Cell* **144:** 27-40 (2011)

40. Kinsella, M., Patel, A. & Bafna, V. The elusive evidence for chromothripsis *Nucleic acids research* **42:** 8231-8242 (2014)

41. Korbel, J. O. & Campbell, P. J. Criteria for inference of chromothripsis in cancer genomes *Cell* **152:** 1226-1236 (2013)

42. Kloosterman, W. P., Koster, J. & Molenaar, J. J. Prevalence and clinical implications of chromothripsis in cancer genomes *Current opinion in oncology* **26:** 64-72 (2014)

43. Kloosterman, W. P., Tavakoli-Yaraki, M., van Roosmalen, M. J., van Binsbergen, E., Renkens, I., Duran, K., Ballarati, L., Vergult, S., Giardino, D., Hansson, K., Ruivenkamp, C. A., Jager, M., van Haeringen, A., Ippel, E. F., Haaf, T., Passarge, E., Hochstenbach, R., Menten, B., Larizza, L., Guryev, V., Poot, M. & Cuppen, E. Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms *Cell reports* **1:** 648-655 (2012)

44. McDermott, D. H., Gao, J. L., Liu, Q., Siwicki, M., Martens, C., Jacobs, P., Velez, D., Yim, E., Bryke, C. R., Hsu, N., Dai, Z., Marquesen, M. M., Stregevsky, E., Kwatemaa, N., Theobald, N., Long Priel, D. A., Pittaluga, S., Raffeld, M. A., Calvo, K. R., Maric, I., Desmond, R., Holmes, K. L., Kuhns, D. B., Balabanian, K., Bachelerie, F., Porcella, S. F., Malech, H. L. & Murphy, P. M. Chromothriptic cure of WHIM syndrome *Cell* **160:** 686-699 (2015)

45. Shen, M. M. Chromoplexy: a new category of complex rearrangements in the cancer genome *Cancer cell* **23:** 567-569 (2013)

46. Plaisancie, J., Kleinfinger, P., Cances, C., Bazin, A., Julia, S., Trost, D., Lohmann, L. & Vigouroux, A. Constitutional chromoanasynthesis: description of a rare chromosomal event in a patient *European journal of medical genetics* **57:** 567-570 (2014)

47. Holland, A. J. & Cleveland, D. W. Chromoanagenesis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements *Nature medicine* **18:** 1630-1638 (2012)

48. Zhang, C. Z., Spektor, A., Cornils, H., Francis, J. M., Jackson, E. K., Liu, S., Meyerson, M. & Pellman, D. Chromothripsis from DNA damage in micronuclei *Nature* **522:** 179-184 (2015)

49. Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A.,Jr & Kinzler, K. W. Cancer genome landscapes *Science* **339:** 1546-1558 (2013)

50. Croce, C. M. Oncogenes and cancer *The New England journal of medicine* **358:** 502-511 (2008)

51. Cox, A. D. & Der, C. J. Ras history: The saga continues *Small GTPases* **1:** 2-27 (2010)

52. Huang, F. W., Bielski, C. M., Rinne, M. L., Hahn, W. C., Sellers, W. R., Stegmeier, F., Garraway, L. A. & Kryukov, G. V. TERT promoter mutations and monoallelic activation of TERT in cancer *Oncogenesis* **4:** e176 (2015)

53. Wiestner, A., Tehrani, M., Chiorazzi, M., Wright, G., Gibellini, F., Nakayama, K., Liu, H., Rosenwald, A., Muller-Hermelink, H. K., Ott, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Vose, J., Armitage, J. O., Gascoyne, R. D., Connors, J. M., Campo, E., Montserrat, E., Bosch, F., Smeland, E. B., Kvaloy, S., Holte, H., Delabie, J., Fisher, R. I., Grogan, T. M., Miller, T. P., Wilson, W. H., Jaffe, E. S. & Staudt, L. M. Point mutations and genomic deletions in CCND1 create stable truncated cyclin D1 mRNAs that are associated with increased proliferation rate and shorter survival *Blood* **109:** 4599-4606 (2007)

54. Uren, A. G., Kool, J., Berns, A. & van Lohuizen, M. Retroviral insertional mutagenesis: past, present and future *Oncogene* **24:** 7656-7672 (2005)

55. Lane, D. P. Cancer. p53, guardian of the genome *Nature* **358:** 15-16 (1992)

56. Kinzler, K. W. & Vogelstein, B. Cancer-susceptibility genes. Gatekeepers and caretakers *Nature* **386:** 761, 763 (1997)

57. Kinzler, K. W. & Vogelstein, B. Landscaping the Cancer Terrain *Science* **280:** 1036-1037 (1998)

58. Payne, S. R. & Kemp, C. J. Tumor suppressor genetics *Carcinogenesis* **26:** 2031-2045 (2005)

59. O'Keefe, C., McDevitt, M. A. & Maciejewski, J. P. Copy neutral loss of heterozygosity: a novel chromosomal lesion in myeloid malignancies *Blood* **115:** 2731-2739 (2010)

60. Sehgal, R., Sheahan, K., O'Connell, P. R., Hanly, A. M., Martin, S. T. & Winter, D. C. Lynch syndrome: an updated review *Genes* **5:** 497-507 (2014)

61. Scheffer, H., te Meerman, G. J., Kruize, Y. C., van den Berg, A. H., Penninga, D. P., Tan, K. E., der Kinderen, D. J. & Buys, C. H. Linkage analysis of families with hereditary retinoblastoma: nonpenetrance of mutation, revealed by combined use of markers within and flanking the RB1 gene *American Journal of Human Genetics* **45:** 252-260 (1989)

62. Kittur, S. D., Bagdon, M. M., Lubs, M. L., Phillips, J. A.,3rd, Murray, J. C., Slaugenhaupt, S. A., Chakravarti, A. & Adler, W. H. Linkage analysis of neurofibromatosis type I, using chromosome 17 DNA markers *American Journal of Human Genetics* **44:** 48-50 (1989)

63. Schork, N. J., Murray, S. S., Frazer, K. A. & Topol, E. J. Common vs. rare allele hypotheses for complex diseases *Current opinion in genetics & development* **19:** 212-219 (2009)

64. McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. & Hirschhorn, J. N. Genome-wide association studies for complex traits: consensus, uncertainty and challenges *Nature Reviews Genetics* **9:** 356-369 (2008)

65. Sur, I. K., Hallikas, O., Vaharautio, A., Yan, J., Turunen, M., Enge, M., Taipale, M., Karhu, A., Aaltonen, L. A. & Taipale, J. Mice lacking a Myc enhancer that includes human SNP rs6983267 are resistant to intestinal tumors *Science* **338:** 1360-1363 (2012)

66. Winkler, C. A., Nelson, G. W. & Smith, M. W. Admixture mapping comes of age *Annual review of genomics and human genetics* **11:** 65-89 (2010)

67. Boland, M. J., Nazor, K. L. & Loring, J. F. Epigenetic regulation of pluripotency and differentiation *Circulation research* **115:** 311-324 (2014)

68. Deans, C. & Maggert, K. A. What do you mean, "epigenetic"? *Genetics* **199:** 887-896 (2015)

69. Kanherkar, R. R., Bhatia-Dey, N. & Csoka, A. B. Epigenetics across the human lifespan *Frontiers in cell and developmental biology* **2:** 49 (2014)

70. Heard, E. & Martienssen, R. A. Transgenerational epigenetic inheritance: myths and mechanisms *Cell* **157:** 95-109 (2014)

71. Morgan, H. D., Santos, F., Green, K., Dean, W. & Reik, W. Epigenetic reprogramming in mammals *Human molecular genetics* **14:** R47-R58 (2005)

72. Razin, A. & Kantor, B. DNA methylation in epigenetic control of gene expression *Progress in molecular and subcellular biology* **38:** 151-167 (2005)

73. Sharma, S., Kelly, T. K. & Jones, P. A. Epigenetics in cancer *Carcinogenesis* **31:** 27-36 (2010)

74. Struhl, K. Is DNA methylation of tumour suppressor genes epigenetic? *eLife* **3:** e02475 (2014)

75. Bilitewski, U. DNA microarrays: an introduction to the technology *Methods in molecular biology* **509:** 1-14 (2009)

76. Moorthie, S., Mattocks, C. J. & Wright, C. F. Review of massively parallel DNA sequencing technologies *The HUGO journal* **5:** 1-12 (2011)

77. Collins, F. S., Morgan, M. & Patrinos, A. The Human Genome Project: Lessons from Large-Scale Biology *Science* **300:** 286 (2003)

78. Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome *Nature* **458:** 719-724 (2009)

79. Nielsen, R., Paul, J. S., Albrechtsen, A. & Song, Y. S. Genotype and SNP calling from next-generation sequencing data *Nature Reviews Genetics* **12:** 443-451 (2011)

80. Li, H. & Homer, N. A survey of sequence alignment algorithms for next-generation sequencing *Briefings in Bioinformatics* **11:** 473-483 (2010)

81. Roberts, N. D., Kortschak, R. D., Parker, W. T., Schreiber, A. W., Branford, S., Scott, H. S., Glonek, G. & Adelson, D. L. A comparative analysis of algorithms for somatic SNV detection in cancer *Bioinformatics* **29:** 2223-2230 (2013)

82. Hasan, M. S., Wu, X. & Zhang, L. Performance evaluation of indel calling tools using real short-read data *Human Genomics* **9:** 1-14 (2015)

83. Liu, B., Conroy, J. M., Morrison, C. D., Odunsi, A. O., Qin, M., Wei, L., Trump, D. L., Johnson, C. S., Liu, S. & Wang, J. Structural variation discovery in the cancer genome using next generation sequencing: Computational solutions and perspectives *Oncotarget* **6:** 5477-5489 (2015)

84. Liu, B., Morrison, C. D., Johnson, C. S., Trump, D. L., Qin, M., Conroy, J. C., Wang, J. & Liu, S. Computational methods for detecting copy number variations in cancer genome using next generation sequencing: principles and challenges *Oncotarget* **4:** 1868-1881 (2013)

85. Griffith, M., Miller, C., Griffith, O., Krysiak, K., Skidmore, Z., Ramu, A., Walker, J., Dang, H., Trani, L., Larson, D., Demeter, R., Wendl, M., McMichael, J., Austin, R., Magrini, V., McGrath, S., Ly, A., Kulkarni, S., Cordes, M., Fronick, C., Fulton, R., Maher, C., Ding, L., Klco, J., Mardis, E., Ley, T. & Wilson, R. Optimizing Cancer Genome Sequencing and Analysis *Cell Systems* **1:** 210-223 (2015)

86. Bozini, N. & Baracat, E. C. The history of myomectomy at the Medical School of University of Sao Paulo *Clinics* **62:** 209-210 (2007)

87. Tinelli, A., Malvasi, A., Rahimi, S., Negro, R., Cavallotti, C., Vergara, D., Vittori, G. & Mettler, L. Myoma pseudocapsule: a distinct endocrino-anatomical entity in gynecological surgery *Gynecological endocrinology : the official journal of the International Society of Gynecological Endocrinology* **25:** 661-667 (2009)

88. Malik, M., Norian, J., McCarthy-Keith, D., Britten, J. & Catherino, W. H. Why leiomyomas are called fibroids: the central role of extracellular matrix in symptomatic women *Seminars in reproductive medicine* **28:** 169-179 (2010)

89. Cramer, S. F. & Patel, A. The frequency of uterine leiomyomas *American Journal of Clinical Pathology* **94:** 435-438 (1990)

90. Peddada, S. D., Laughlin, S. K., Miner, K., Guyon, J. P., Haneke, K., Vahdat, H. L., Semelka, R. C., Kowalik, A., Armao, D., Davis, B. & Baird, D. D. Growth of uterine leiomyomata among premenopausal black and white women *Proceedings of the National Academy of Sciences of the United States of America* **105:** 19887-19892 (2008)

91. Baird, D. D., Dunson, D. B., Hill, M. C., Cousins, D. & Schectman, J. M. High cumulative incidence of uterine leiomyoma in black and white women: ultrasound evidence *American Journal of Obstetrics and Gynecology* **188:** 100-107 (2003)

92. Sabry, M. & Al-Hendy, A. Medical treatment of uterine leiomyoma *Reproductive sciences* **19:** 339-353 (2012)

93. Whiteman, M. K., Hillis, S. D., Jamieson, D. J., Morrow, B., Podgornik, M. N., Brett, K. M. & Marchbanks, P. A. Inpatient hysterectomy surveillance in the United States, 2000-2004 *American Journal of Obstetrics and Gynecology* **198:** 34.e1-34.e7 (2008)

94. Cardozo, E. R., Clark, A. D., Banks, N. K., Henne, M. B., Stegmann, B. J. & Segars, J. H. The estimated annual cost of uterine leiomyomata in the United States *American Journal of Obstetrics and Gynecology* **206:** 211.e1-211.e9 (2012)

95. Kurman, R. J., Carcangiu, M. L., Herrington, S. & Young, R. H. WHO Classification of Tumours of Female Reproductive Organs. Fourth Edition (2014)

96. Mahmoud, M. S., Desai, K. & Nezhat, F. R. Leiomyomas beyond the uterus; benign metastasizing leiomyomatosis with paraaortic metastasizing endometriosis and intravenous leiomyomatosis: a case series and review of the literature *Archives of Gynecology and Obstetrics* **291:** 223-230 (2015)

97. Baschinsky, D. Y., Isa, A., Niemann, T. H., Prior, T. W., Lucas, J. G. & Frankel, W. L. Diffuse leiomyomatosis of the uterus: a case report with clonality analysis *Human pathology* **31:** 1430-1432 (2000)

98. Arleo, E. K., Schwartz, P. E., Hui, P. & McCarthy, S. Review of Leiomyoma Variants *American journal of roentgenology* **205:** 912-921 (2015)

99. Yanai, H., Wani, Y., Notohara, K., Takada, S. & Yoshino, T. Uterine leiomyosarcoma arising in leiomyoma: Clinicopathological study of four cases and literature review *Pathology international* **60:** 506-509 (2010)

100. Mittal, K. R., Chen, F., Wei, J. J., Rijhvani, K., Kurvathi, R., Streck, D., Dermody, J. & Toruner, G. A. Molecular and immunohistochemical evidence for the origin of uterine leiomyosarcomas from associated leiomyoma and symplastic leiomyoma-like areas *Modern pathology* **22:** 1303-1311 (2009)

101. Fasih, N., Prasad Shanbhogue, A. K., Macdonald, D. B., Fraser-Hill, M. A., Papadatos, D., Kielar, A. Z., Doherty, G. P., Walsh, C., McInnes, M. & Atri, M. Leiomyomas beyond the uterus: unusual locations, rare manifestations *Radiographics* **28:** 1931-1948 (2008)

102. Wilde, S. & Scott-Barrett, S. Radiological appearances of uterine fibroids *The Indian journal of radiology & imaging* **19:** 222-231 (2009)

103. Stewart, E. A. Uterine Fibroids *The New England journal of medicine* **372:** 1646-1655 (2015)

104. Fauconnier, A., Chapron, C., Babaki-Fard, K. & Dubuisson, J. B. Recurrence of leiomyomata after myomectomy *Human reproduction update* **6:** 595-602 (2000)

105. Siedhoff, M. T. & Kim, K. H. Morcellation and myomas: Balancing decisions around minimally invasive treatments for fibroids *Journal of surgical oncology* **112:** 769-771 (2015)

106. Shakir, F. & Hill, N. Morcelloma Formation Following Laparoscopic Myomectomy and Sub-Total Hysterectomy *Journal of Minimally Invasive Gynecology* **19:** S130 (2012)

107. Brohl, A. S., Li, L., Andikyan, V., Obican, S. G., Cioffi, A., Hao, K., Dudley, J. T., Ascher-Walsh, C., Kasarskis, A. & Maki, R. G. Age-stratified risk of unexpected uterine sarcoma following surgery for presumed benign leiomyoma *The oncologist* **20:** 433-439 (2015)

108. Parker, W. H., Kaunitz, A. M., Pritts, E. A., Olive, D. L., Chalas, E., Clarke-Pearson, D. L., Berek, J. S. & Leiomyoma Morcellation Review Group. U.S. Food and Drug Administration's Guidance Regarding Morcellation of Leiomyomas: Well-Intentioned, But Is It Harmful for Women? *Obstetrics and gynecology* **127:** 18-22 (2016)

109. Bortoletto, P., Einerson, B. D., Miller, E. S. & Milad, M. P. Cost-Effectiveness Analysis of Morcellation Hysterectomy for Myomas *Journal of minimally invasive gynecology* **22:** 820-826 (2015)

110. Eltoukhi, H. M., Modi, M. N., Weston, M., Armstrong, A. Y. & Stewart, E. A. The health disparities of uterine fibroid tumors for African American women: a public health issue *American Journal of Obstetrics and Gynecology* **210:** 194-199 (2014)

111. Cramer, S. F., Marchetti, C., Freedman, J. & Padela, A. Relationship of myoma cell size and menopausal status in small uterine leiomyomas *Archives of Pathology & Laboratory Medicine* **124:** 1448-1453 (2000)

112. Sommer, E. M., Balkwill, A., Reeves, G., Green, J., Beral, D. V., Coffey, K. & Million Women Study Collaborators. Effects of obesity and hormone therapy on surgically-confirmed fibroids in postmenopausal women *European journal of epidemiology* **30:** 493-499 (2015)

113. Winkler, V. & Hoffmann, W. Regarding the question of inheritance of uterine myoma. *Deutsche-Medizinische Wochenschrift* **68:** 235-257 (1938)

114. Kurbanova, M. K., Koroleva, A. G. & Sergeev, A. S. Genetic-epidemiologic analysis of uterine myoma: assessment of repeated risk *Genetika* **25:** 1896-1898 (1989)

115. Vikhlyaeva, E. M., Khodzhaeva, Z. S. & Fantschenko, N. D. Familial predisposition to uterine leiomyomas *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* **51:** 127-131 (1995)

116. Snieder, H., MacGregor, A. J. & Spector, T. D. Genes control the cessation of a woman's reproductive life: a twin study of hysterectomy and age at menopause *The Journal of clinical endocrinology and metabolism* **83:** 1875-1880 (1998)

117. Luoto, R., Kaprio, J., Rutanen, E. M., Taipale, P., Perola, M. & Koskenvuo, M. Heritability and risk factors of uterine fibroids--the Finnish Twin Cohort study *Maturitas* **37:** 15-26 (2000)

118. Baird, D. D. & Dunson, D. B. Why is parity protective for uterine fibroids? *Epidemiology* **14:** 247-250 (2003)

119. Velez Edwards, D. R., Baird, D. D. & Hartmann, K. E. Association of age at menarche with increasing number of fibroids in a cohort of women who underwent standardized ultrasound assessment *American Journal of Epidemiology* **178:** 426-433 (2013)

120. Chiaffarino, F., Parazzini, F., La Vecchia, C., Chatenoud, L., Di Cintio, E. & Marsico, S. Diet and uterine myomas *Obstetrics and gynecology* **94:** 395-398 (1999)

121. Shen, Y., Wu, Y., Lu, Q. & Ren, M. Vegetarian diet and reduced uterine fibroids risk: A case-control study in Nanjing, China *The journal of obstetrics and gynaecology research* **42:** 87-94 (2016)

122. Kjerulff, K. H., Langenberg, P., Seidman, J. D., Stolley, P. D. & Guzinski, G. M. Uterine leiomyomas. Racial differences in severity, symptoms and age at diagnosis *The Journal of reproductive medicine* **41:** 483-490 (1996)

123. Wise, L. A., Palmer, J. R., Ruiz-Narvaez, E., Reich, D. E. & Rosenberg, L. Is the observed association between dairy intake and fibroids in African Americans explained by genetic ancestry? *American Journal of Epidemiology* **178:** 1114-1119 (2013)

124. Sheiner, E., Bashiri, A., Levy, A., Hershkovitz, R., Katz, M. & Mazor, M. Obstetric characteristics and perinatal outcome of pregnancies with uterine leiomyomas *The Journal of reproductive medicine* **49:** 182-186 (2004)

125. Zhang, D., Al-Hendy, M., Richard-Davis, G., Montgomery-Rice, V., Rajaratnam, V. & Al-Hendy, A. Antiproliferative and proapoptotic effects of epigallocatechin gallate on human leiomyoma cells *Fertility and sterility* **94:** 1887-1893 (2010)

126. Harmon, Q. E. & Baird, D. D. Use of depot medroxyprogesterone acetate and prevalent leiomyoma in young African American women *Human reproduction* **30:** 1499-1504 (2015)

127. Ott, J., Kurz, C., Braun, R., Promberger, R., Seemann, R., Vytiska-Binstorfer, E. & Walch, K. Overt hypothyroidism is associated with the presence of uterine leiomyoma: a retrospective analysis *European journal of obstetrics, gynecology, and reproductive biology* **177:** 19-22 (2014)

128. Ross, R. K., Pike, M. C., Vessey, M. P., Bull, D., Yeates, D. & Casagrande, J. T. Risk factors for uterine fibroids: reduced risk associated with oral contraceptives *British medical journal* **293:** 359-362 (1986)

129. Luoto, R., Rutanen, E. M. & Auvinen, A. Fibroids and hypertension. A cross-sectional study of women undergoing hysterectomy *The Journal of reproductive medicine* **46:** 359-364 (2001)

130. Baird, D. D., Dunson, D. B., Hill, M. C., Cousins, D. & Schectman, J. M. Association of physical activity with development of uterine leiomyoma *American Journal of Epidemiology* **165:** 157-163 (2007)

131. Wise, L. A., Palmer, J. R., Stewart, E. A. & Rosenberg, L. Polycystic ovary syndrome and risk of uterine leiomyomata *Fertility and sterility* **87:** 1108-1115 (2007)

132. Brakta, S., Diamond, J. S., Al-Hendy, A., Diamond, M. P. & Halder, S. K. Role of vitamin D in uterine fibroid biology *Fertility and sterility* **104:** 698-706 (2015)

133. Qin, J., Yang, T., Kong, F. & Zhou, Q. Oral contraceptive use and uterine leiomyoma risk: a meta-analysis based on cohort and case-control studies *Archives of Gynecology and Obstetrics* **288:** 139-148 (2013)

134. Shikora, S. A., Niloff, J. M., Bistrian, B. R., Forse, R. A. & Blackburn, G. L. Relationship between obesity and uterine leiomyomata *Nutrition* **7:** 251-255 (1991)

135. Moore, K. R., Cole, S. R., Dittmer, D. P., Schoenbach, V. J., Smith, J. S. & Baird, D. D. Self-Reported Reproductive Tract Infections and Ultrasound Diagnosed Uterine Fibroids in African-American Women *Journal of women's health* **24:** 489-495 (2015)

136. Faerstein, E., Szklo, M. & Rosenshein, N. B. Risk factors for uterine leiomyoma: a practice-based case-control study. II. Atherogenic risk factors and potential sources of uterine irritation *American Journal of Epidemiology* **153:** 11-19 (2001)

137. Vines, A. I., Ta, M. & Esserman, D. A. The association between self-reported major life events and the presence of uterine fibroids *Women's health issues* **20:** 294-298 (2010)

138. Baird, D. D., Travlos, G., Wilson, R., Dunson, D. B., Hill, M. C., D'Aloisio, A. A., London, S. J. & Schectman, J. M. Uterine leiomyomata in relation to insulin-like growth factor-I, insulin, and diabetes *Epidemiology* **20:** 604-610 (2009)

139. Wise, L. A., Palmer, J. R., Harlow, B. L., Spiegelman, D., Stewart, E. A., Adams-Campbell, L. L. & Rosenberg, L. Risk of uterine leiomyomata in relation to tobacco, alcohol and caffeine consumption in the Black Women's Health Study *Human Reproduction* **19:** 1746-1754 (2004)

140. Nagata, C., Nakamura, K., Oba, S., Hayashi, M., Takeda, N. & Yasuda, K. Association of intakes of fat, dietary fibre, soya isoflavones and alcohol with uterine fibroids in Japanese women *The British journal of nutrition* **101:** 1427-1431 (2009)

141. Kawamura, S., Kasagi, F., Kodama, K., Fujiwara, S., Yamada, M., Ohama, K. & Ito, K. Prevalence of Uterine Myoma Detected by Ultrasound Examination in the Atomic Bomb Survivors *Radiation research* **147:** 753-758 (1997)

142. Baird, D. D. & Newbold, R. Prenatal diethylstilbestrol (DES) exposure is associated with uterine leiomyoma development *Reproductive toxicology* **20:** 81-84 (2005)

143. Reis, F. M., Bloise, E. & Ortiga-Carvalho, T. M. Hormones and pathogenesis of uterine fibroids *Best Practice & Research Clinical Obstetrics & Gynaecology* **34:** 13-24 (2016)

144. Cesen-Cummings, K., Copland, J. A., Barrett, J. C., Walker, C. L. & Davis, B. J. Pregnancy, Parturition, and Prostaglandins: Defining Uterine Leiomyomas *Environmental health perspectives* **108:** 817-820 (2000)

145. Witherspoon, J. T. & Butler, V. W. The etiology of uterine fibroids with special reference to the frequency of their occurrence in the Negro: A hypothesis *Surgery, Gynecology, & Obstetrics* **58:** 57-61 (1934)

146. Ciavattini, A., Di Giuseppe, J., Stortoni, P., Montik, N., Giannubilo, S. R., Litta, P., Islam, M. S., Tranquilli, A. L., Reis, F. M. & Ciarmela, P. Uterine fibroids: pathogenesis and interactions with endometrium and endomyometrial junction *Obstetrics and gynecology international* **2013:** 173-184 (2013)

147. Luo, X., Pan, Q., Liu, L. & Chegini, N. Genomic and proteomic profiling II: Comparative assessment of gene expression profiles in leiomyomas, keloids, and surgically-induced scars *Reproductive Biology and Endocrinology* **5:** 35 (2007)

148. Romagnolo, B., Molina, T., Leroy, G., Blin, C., Porteux, A., Thomasset, M., Vandewalle, A., Kahn, A. & Perret, C. Estradiol-dependent uterine leiomyomas in transgenic mice *The Journal of clinical investigation* **98:** 777-784 (1996)

149. Webster, M. A., Martin-Soudant, N., Nepveu, A., Cardiff, R. D. & Muller, W. J. The induction of uterine leiomyomas and mammary tumors in transgenic mice expressing polyomavirus (PyV) large T (LT) antigen is associated with the ability of PyV LT antigen to form specific complexes with retinoblastoma and CUTL1 family members *Oncogene* **16:** 1963-1972 (1998)

150. Sandberg, A. A. Updates on the cytogenetics and molecular genetics of bone and soft tissue tumors: leiomyoma *Cancer genetics and cytogenetics* **158:** 1-26 (2005)

151. Wise, L. A., Ruiz-Narvaez, E. A., Palmer, J. R., Cozier, Y. C., Tandon, A., Patterson, N., Radin, R. G., Rosenberg, L. & Reich, D. African ancestry and genetic risk for uterine leiomyomata *American Journal of Epidemiology* **176:** 1159-1168 (2012)

152. Lehtonen, H. J. Hereditary leiomyomatosis and renal cell cancer: update on clinical and molecular characteristics *Familial Cancer* **10:** 397-411 (2011)

153. Castro-Vega, L. J., Buffet, A., De Cubas, A. A., Cascon, A., Menara, M., Khalifa, E., Amar, L., Azriel, S., Bourdeau, I., Chabre, O., Curras-Freixes, M., Franco-Vidal, V., Guillaud-Bataille, M., Simian, C., Morin, A., Leton, R., Gomez-Grana, A., Pollard, P. J., Rustin, P., Robledo, M., Favier, J. & Gimenez-Roqueplo, A. P. Germline mutations in FH confer predisposition to malignant pheochromocytomas and paragangliomas *Human molecular genetics* **23:** 2440-2446 (2014)

154. Clark, G. R., Sciacovelli, M., Gaude, E., Walsh, D. M., Kirby, G., Simpson, M. A., Trembath, R. C., Berg, J. N., Woodward, E. R., Kinning, E., Morrison, P. J., Frezza, C. & Maher, E. R. Germline FH mutations presenting with pheochromocytoma *The Journal of clinical endocrinology and metabolism* **99:** E2046-50 (2014)

155. Reed, W. B., Walker, R. & Horowitz, R. Cutaneous leiomyomata with uterine leiomyomata *Acta Dermato-Venereologica* **53:** 409-416 (1973)

156. Waite, K. A. & Eng, C. Protean PTEN: form and function *American Journal of Human Genetics* **70:** 829-844 (2002)

157. Hulsebos, T. J., Kenter, S., Siebers-Renelt, U., Hans, V., Wesseling, P. & Flucke, U. SMARCB1 involvement in the development of leiomyoma in a patient with schwannomatosis *The American Journal of Surgical Pathology* **38:** 421-425 (2014)

158. Sa, M. J., Fieremans, N., de Brouwer, A. P., Sousa, R., e Costa, F. T., Brito, M. J., Carvalho, F., Rodrigues, M., de Sousa, F. T., Felgueiras, J., Neves, F., Carvalho, A., Ramos, U., Vizcaino, J. R., Alves, S., Carvalho, F., Froyen, G. & Oliveira, J. P. Deletion of the 5'exons of COL4A6 is not needed for the development of diffuse leiomyomatosis in patients with Alport syndrome *Journal of medical genetics* **50:** 745-753 (2013)

159. Garavelli, L., Piemontese, M. R., Cavazza, A., Rosato, S., Wischmeijer, A., Gelmini, C., Albertini, E., Albertini, G., Forzano, F., Franchi, F., Carella, M., Zelante, L. & Superti-Furga, A. Multiple tumor types including leiomyoma and Wilms tumor in a patient with Gorlin syndrome due to 9q22.3 microdeletion encompassing the PTCH1 and FANC-C loci *American Journal of Medical Genetics Part A* **161:** 2894-2901 (2013)

160. Hancock, E., Tomkins, S., Sampson, J. & Osborne, J. Lymphangioleiomyomatosis and tuberous sclerosis *Respiratory medicine* **96:** 7-13 (2002)

161. Tomlinson, I. P., Alam, N. A., Rowan, A. J., Barclay, E., Jaeger, E. E., Kelsell, D., Leigh, I., Gorman, P., Lamlum, H., Rahman, S., Roylance, R. R., Olpin, S., Bevan, S., Barker, K., Hearle, N., Houlston, R. S., Kiuru, M., Lehtonen, R., Karhu, A., Vilkki, S., Laiho, P., Eklund, C., Vierimaa, O., Aittomaki, K., Hietala, M., Sistonen, P., Paetau, A., Salovaara, R., Herva, R., Launonen, V., Aaltonen, L. A. & Multiple Leiomyoma Consortium. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer *Nature genetics* **30:** 406-410 (2002)

162. Lehtonen, R., Kiuru, M., Vanharanta, S., Sjoberg, J., Aaltonen, L. M., Aittomaki, K., Arola, J., Butzow, R., Eng, C., Husgafvel-Pursiainen, K., Isola, J., Jarvinen, H., Koivisto, P., Mecklin, J. P., Peltomaki, P., Salovaara, R., Wasenius, V. M., Karhu, A., Launonen, V., Nupponen, N. N. & Aaltonen, L. A. Biallelic inactivation of fumarate hydratase (FH) occurs in nonsyndromic uterine leiomyomas but is rare in other tumors *The American journal of pathology* **164:** 17-22 (2004)

163. Harrison, W. J., Andrici, J., Maclean, F., Madadi-Ghahan, R., Farzin, M., Sioson, L., Toon, C. W., Clarkson, A., Watson, N., Pickett, J., Field, M., Crook, A., Tucker, K., Goodwin, A., Anderson, L., Srinivasan, B., Grossmann, P., Martinek, P., Ondic, O., Hes, O., Trpkov,

K., Clifton-Bligh, R. J., Dwight, T. & Gill, A. J. Fumarate Hydratase-deficient Uterine Leiomyomas Occur in Both the Syndromic and Sporadic Settings *The American Journal of Surgical Pathology* **40:** 599-607 (2016)

164. Raimundo, N., Baysal, B. E. & Shadel, G. S. Revisiting the TCA cycle: signaling to tumor formation *Trends in molecular medicine* **17:** 641-649 (2011)

165. Xiao, M., Yang, H., Xu, W., Ma, S., Lin, H., Zhu, H., Liu, L., Liu, Y., Yang, C., Xu, Y., Zhao, S., Ye, D., Xiong, Y. & Guan, K. L. Inhibition of alpha-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors *Genes & development* **26:** 1326-1338 (2012)

166. Shanmugasundaram, K., Nayak, B., Shim, E. H., Livi, C. B., Block, K. & Sudarshan, S. The oncometabolite fumarate promotes pseudohypoxia through noncanonical activation of NF-kappaB signaling *The Journal of biological chemistry* **289:** 24691-24699 (2014)

167. Powis, G. & Kirkpatrick, L. Hypoxia inducible factor-1alpha as a cancer drug target *Molecular cancer therapeutics* **3:** 647-654 (2004)

168. Letouzé, E., Martinelli, C., Loriot, C., Burnichon, N., Abermil, N., Ottolenghi, C., Janin, M., Menara, M., Nguyen, A., Benit, P., Buffet, A., Marcaillou, C., Bertherat, J., Amar, L., Rustin, P., De Reyniès, A., Gimenez-Roqueplo, A. & Favier, J. SDH Mutations Establish a Hypermethylator Phenotype in Paraganglioma *Cancer Cell* **23:** 739-752 (2013)

169. Cancer Genome Atlas Research Network, Linehan, W. M., Spellman, P. T., Ricketts, C. J., Creighton, C. J., Fei, S. S., Davis, C., Wheeler, D. A., Murray, B. A., Schmidt, L., Vocke, C. D., Peto, M., Al Mamun, A. A., Shinbrot, E., Sethi, A., Brooks, S., Rathmell, W. K., Brooks, A. N., Hoadley, K. A., Robertson, A. G., Brooks, D., Bowlby, R., Sadeghi, S., Shen, H., Weisenberger, D. J., Bootwalla, M., Baylin, S. B., Laird, P. W., Cherniack, A. D., Saksena, G., Haake, S., Li, J., Liang, H., Lu, Y., Mills, G. B., Akbani, R., Leiserson, M. D., Raphael, B. J., Anur, P., Bottaro, D., Albiges, L., Barnabas, N., Choueiri, T. K., Czerniak, B., Godwin, A. K., Hakimi, A. A., Ho, T. H., Hsieh, J., Ittmann, M., Kim, W. Y., Krishnan, B., Merino, M. J., Mills Shaw, K. R., Reuter, V. E., Reznik, E., Shelley, C. S., Shuch, B., Signoretti, S., Srinivasan, R., Tamboli, P., Thomas, G., Tickoo, S., Burnett, K., Crain, D., Gardner, J., Lau, K., Mallery, D., Morris, S., Paulauskis, J. D., Penny, R. J., Shelton, C., Shelton, W. T., Sherman, M., Thompson, E., Yena, P., Avedon, M. T., Bowen, J., Gastier-Foster, J. M., Gerken, M., Leraas, K. M., Lichtenberg, T. M., Ramirez, N. C., Santos, T., Wise, L., Zmuda, E., Demchok, J. A., Felau, I., Hutter, C. M., Sheth, M., Sofia, H. J., Tarnuzzer, R., Wang, Z., Yang, L., Zenklusen, J. C., Zhang, J., Ayala, B., Baboud, J., Chudamani, S., Liu, J., Lolla, L., Naresh, R., Pihl, T., Sun, Q., Wan, Y., Wu, Y., Ally, A., Balasundaram, M., Balu, S., Beroukhim, R., Bodenheimer, T., Buhay, C., Butterfield, Y. S., Carlsen, R., Carter, S. L., Chao, H., Chuah, E., Clarke, A., Covington, K. R., Dahdouli, M., Dewal, N., Dhalla, N., Doddapaneni, H. V., Drummond, J. A., Gabriel, S. B., Gibbs, R. A., Guin, R., Hale, W., Hawes, A., Hayes, D. N., Holt, R. A., Hoyle, A. P., Jefferys, S. R., Jones, S. J., Jones, C. D., Kalra, D., Kovar, C., Lewis, L., Li, J., Ma, Y., Marra, M. A., Mayo, M., Meng, S., Meyerson, M., Mieczkowski, P. A., Moore, R. A., Morton, D., Mose, L. E., Mungall, A. J., Muzny, D., Parker, J. S., Perou, C. M., Roach, J., Schein, J. E., Schumacher, S. E., Shi, Y., Simons, J. V., Sipahimalani, P., Skelly, T., Soloway, M. G., Sougnez, C., Tam, A., Tan, D., Thiessen, N., Veluvolu, U., Wang, M., Wilkerson, M. D., Wong, T., Wu, J., Xi, L., Zhou, J., Bedford, J., Chen, F., Fu, Y., Gerstein, M., Haussler, D., Kasaian, K., Lai, P., Ling, S., Radenbaugh, A., Van Den Berg, D., Weinstein, J. N., Zhu, J., Albert, M., Alexopoulou, I., Andersen, J. J., Auman, J. T., Bartlett, J., Bastacky, S., Bergsten, J., Blute, M. L., Boice, L., Bollag, R. J., Boyd, J., Castle, E., Chen, Y. B., Cheville, J. C., Curley, E., Davies, B., DeVolk, A., Dhir, R., Dike, L., Eckman, J., Engel, J., Harr, J., Hrebinko, R., Huang, M., Huelsenbeck-Dill, L., Iacocca, M., Jacobs, B., Lobis, M., Maranchie, J. K., McMeekin, S., Myers, J., Nelson, J., Parfitt, J., Parwani, A., Petrelli, N., Rabeno, B., Roy, S., Salner, A. L., Slaton, J., Stanton, M., Thompson, R. H., Thorne, L., Tucker, K., Weinberger, P. M., Winemiller, C., Zach, L. A. & Zuna, R. Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma *The New England journal of medicine* **374:** 135-145 (2016)

170. Zhao, S., Lin, Y., Xu, W., Jiang, W., Zha, Z., Wang, P., Yu, W., Li, Z., Gong, L., Peng, Y., Ding, J., Lei, Q., Guan, K. L. & Xiong, Y. Glioma-derived mutations in IDH1 dominantly inhibit IDH1 catalytic activity and induce HIF-1alpha *Science* **324:** 261-265 (2009)

171. Turcan, S., Rohle, D., Goenka, A., Walsh, L. A., Fang, F., Yilmaz, E., Campos, C., Fabius, A. W., Lu, C., Ward, P. S., Thompson, C. B., Kaufman, A., Guryanova, O., Levine, R., Heguy, A., Viale, A., Morris, L. G., Huse, J. T., Mellinghoff, I. K. & Chan, T. A. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype *Nature* **483:** 479-483 (2012)

172. Kinch, L., Grishin, N. V. & Brugarolas, J. Succination of Keap1 and activation of Nrf2-dependent antioxidant pathways in FH-deficient papillary renal cell carcinoma type 2 *Cancer cell* **20:** 418-420 (2011)

173. DeNicola, G. M., Karreth, F. A., Humpton, T. J., Gopinathan, A., Wei, C., Frese, K., Mangal, D., Yu, K. H., Yeo, C. J., Calhoun, E. S., Scrimieri, F., Winter, J. M., Hruban, R. H., Iacobuzio-Donahue, C., Kern, S. E., Blair, I. A. & Tuveson, D. A. Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis *Nature* **475:** 106-109 (2011)

174. Lees-Miller, S. P. Fumarate in DNA repair *Nature cell biology* **17:** 1096-1097 (2015)

175. Cha, P. C., Takahashi, A., Hosono, N., Low, S. K., Kamatani, N., Kubo, M. & Nakamura, Y. A genome-wide association study identifies three loci associated with susceptibility to uterine fibroids *Nature genetics* **43:** 447-450 (2011)

176. Eggert, S. L., Huyck, K. L., Somasundaram, P., Kavalla, R., Stewart, E. A., Lu, A. T., Painter, J. N., Montgomery, G. W., Medland, S. E., Nyholt, D. R., Treloar, S. A., Zondervan, K. T., Heath, A. C., Madden, P. A., Rose, L., Buring, J. E., Ridker, P. M., Chasman, D. I., Martin, N. G., Cantor, R. M. & Morton, C. C. Genome-wide linkage and association analyses implicate FASN in predisposition to Uterine Leiomyomata *American Journal of Human Genetics* **91:** 621-628 (2012)

177. Zhang, K., Wiener, H. & Aissani, B. Admixture mapping of genetic variants for uterine fibroids *Journal of human genetics* **60:** 533-538 (2015)

178. Edwards, T. L., Hartmann, K. E. & Velez Edwards, D. R. Variants in BET1L and TNRC6B associate with increasing fibroid volume and fibroid type among European Americans *Human genetics* **132:** 1361-1369 (2013)

179. Aissani, B., Zhang, K. & Wiener, H. Follow-up to genome-wide linkage and admixture mapping studies implicates components of the extracellular matrix in susceptibility to and size of uterine fibroids *Fertility and sterility* **103:** 528-534.e13 (2015)

180. Aissani, B., Zhang, K. & Wiener, H. Genetic determinants of uterine fibroid size in the multiethnic NIEHS uterine fibroid study *International journal of molecular epidemiology and genetics* **6:** 9-19 (2015)

181. Aissani, B., Zhang, K. & Wiener, H. Evaluation of GWAS candidate susceptibility loci for uterine leiomyoma in the multi-ethnic NIEHS uterine fibroid study *Frontiers in genetics* **6:** 241 (2015)

182. Heim, S., Nilbert, M., Vanni, R., Floderus, U. M., Mandahl, N., Liedgren, S., Lecca, U. & Mitelman, F. A specific translocation, t(12;14)(q14-15;q23-24), characterizes a subgroup of uterine leiomyomas *Cancer genetics and cytogenetics* **32:** 13-17 (1988)

183. Gibas, Z., Griffin, C. A. & Emanuel, B. S. Clonal chromosome rearrangements in a uterine myoma *Cancer genetics and cytogenetics* **32:** 19-24 (1988)

184. Sreekantaiah, C., Davis, J. R. & Sandberg, A. A. Chromosomal abnormalities in leiomyosarcomas *The American journal of pathology* **142:** 293-305 (1993)

185. Mitelman, F., Johansson, B. & Mertens, F. Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (2016)

186. Pandis, N., Heim, S., Bardi, G., Floderus, U. M., Willen, H., Mandahl, N. & Mitelman, F. Parallel karyotypic evolution and tumor progression in uterine leiomyoma *Genes, chromosomes & cancer* **2:** 311-317 (1990)

187. Guardiola, M. T., Dobin, S. M., Dal Cin, P. & Donner, L. R. Pericentric inversion (12)(p12q13-14) as the sole chromosomal abnormality in a leiomyoma of the vulva *Cancer genetics and cytogenetics* **199:** 21-23 (2010)

188. Gorunova, L., Bjerkehagen, B. & Heim, S. Paratesticular leiomyoma with a der(14)t(12;14)(q15;q24) *Cancer genetics* **204:** 465-468 (2011)

189. Dal Cin, P., Quade, B. J., Neskey, D. M., Kleinman, M. S., Weremowicz, S. & Morton, C. C. Intravenous leiomyomatosis is characterized by a der(14)t(12;14)(q15;q24) *Genes, chromosomes & cancer* **36:** 205-206 (2003)

190. Ordulu, Z., Cin, P. D., Chong, W. W. S., Choy, K. W., Lee, C., Muto, M. G., Quade, B. J. & Morton, C. C. Disseminated peritoneal leiomyomatosis after laparoscopic supracervical hysterectomy with characteristic molecular cytogenetic findings of uterine leiomyoma *Genes, Chromosomes and Cancer* **49:** 1152-1160 (2010)

191. Fusco, A. & Fedele, M. Roles of HMGA proteins in cancer *Nature Reviews Cancer* **7:** 899-910 (2007)

192. Van de Ven, W. J. M. Genetic basis of uterine leiomyoma: involvement of high mobility group protein genes *European Journal of Obstetrics & Gynecology and Reproductive Biology* **81:** 289-293 (1998)

193. Mayr, C., Hemann, M. T. & Bartel, D. P. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation *Science* **315:** 1576-1579 (2007)

194. Young, A. R. & Narita, M. Oncogenic HMGA2: short or small? *Genes & development* **21:** 1005-1009 (2007)

195. Pierantoni, G. M., Santulli, B., Caliendo, I., Pentimalli, F., Chiappetta, G., Zanesi, N., Santoro, M., Bulrich, F. & Fusco, A. HMGA2 locus rearrangement in a case of acute lymphoblastic leukemia *International journal of oncology* **23:** 363-367 (2003)

196. Odero, M. D., Grand, F. H., Iqbal, S., Ross, F., Roman, J. P., Vizmanos, J. L., Andrieux, J., Lai, J. L., Calasanz, M. J. & Cross, N. C. Disruption and aberrant expression of HMGA2 as a consequence of diverse chromosomal translocations in myeloid malignancies *Leukemia* **19:** 245-252 (2005)

197. Nyquist, K. B., Panagopoulos, I., Thorsen, J., Roberto, R., Wik, H. S., Tierens, A., Heim, S. & Micci, F. t(12;13)(q14;q31) leading to HMGA2 upregulation in acute myeloid leukaemia *British journal of haematology* **157:** 769-771 (2012)

198. Nishino, J., Kim, I., Chada, K. & Morrison, S. J. Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression *Cell* **135:** 227-239 (2008)

199. Hennig, Y., Deichert, U., Bonk, U., Thode, B., Bartnitzke, S. & Bullerdiek, J. Chromosomal translocations affecting 12q14-15 but not deletions of the long arm of chromosome 7 associated with a growth advantage of uterine smooth muscle cells *Molecular human reproduction* **5:** 1150-1154 (1999)

200. Federico, A., Forzati, F., Esposito, F., Arra, C., Palma, G., Barbieri, A., Palmieri, D., Fedele, M., Pierantoni, G. M., De Martino, I. & Fusco, A. Hmga1/Hmga2 double knock-out mice display a "superpygmy" phenotype *Biology open* **3:** 372-378 (2014)

201. Yang, T. L., Guo, Y., Zhang, L. S., Tian, Q., Yan, H., Guo, Y. F. & Deng, H. W. HMGA2 is confirmed to be associated with human adult height *Annals of Human Genetics* **74:** 11-16 (2010)

202. Hodge, J. C., T Cuenco, K., Huyck, K. L., Somasundaram, P., Panhuysen, C. I., Stewart, E. A. & Morton, C. C. Uterine leiomyomata and decreased height: a common HMGA2 predisposition allele *Human genetics* **125:** 257-263 (2009)

203. Hodge, J. C., Kim, T. M., Dreyfuss, J. M., Somasundaram, P., Christacos, N. C., Rousselle, M., Quade, B. J., Park, P. J., Stewart, E. A. & Morton, C. C. Expression profiling of uterine leiomyomata cytogenetic subgroups reveals distinct signatures in matched myometrium: transcriptional profilingof the t(12;14) and evidence in support of predisposing genetic heterogeneity *Human molecular genetics* **21:** 2312-2329 (2012)

204. Ligon, A. H., Moore, S. D. P., Parisi, M. A., Mealiffe, M. E., Harris, D. J., Ferguson, H. L., Quade, B. J. & Morton, C. C. Constitutional Rearrangement of the Architectural Factor HMGA2: A Novel Human Phenotype Including Overgrowth and Lipomas *The American Journal of Human Genetics* **76:** 340-348 (2005)

205. Nezhad, M. H., Drieschner, N., Helms, S., Meyer, A., Tadayyon, M., Klemke, M., Belge, G., Bartnitzke, S., Burchardt, K., Frantzen, C., Schmidt, E. H. & Bullerdiek, J. 6p21 rearrangements in uterine leiomyomas targeting HMGA1 *Cancer genetics and cytogenetics* **203:** 247-252 (2010)

206. Bowen, J. M., Cates, J. M., Kash, S., Itani, D., Gonzalez, A., Huang, D., Oliveira, A. & Bridge, J. A. Genomic imbalances in benign metastasizing leiomyoma: characterization by conventional karyotypic, fluorescence in situ hybridization, and whole genome SNP array analysis *Cancer genetics* **205:** 249-254 (2012)

207. Schoenmakers, E. F., Huysmans, C. & Van de Ven, W. J. Allelic knockout of novel splice variants of human recombination repair gene RAD51B in t(12;14) uterine leiomyomas *Cancer research* **59:** 19-23 (1999)

208. Date, O., Katsura, M., Ishida, M., Yoshihara, T., Kinomura, A., Sueda, T. & Miyagawa, K. Haploinsufficiency of RAD51B causes centrosome fragmentation and aneuploidy in human cells *Cancer research* **66:** 6018-6024 (2006)

209. Liu, Y. Gamma-radiation sensitivity and polymorphisms in RAD51L1 modulate glioma risk *Carcinogenesis* **31:** 1762-1769 (2010)

210. Pelttari, L. M., Khan, S., Vuorela, M., Kiiski, J. I., Vilske, S., Nevanlinna, V., Ranta, S., Schleutker, J., Winqvist, R., Kallioniemi, A., DÃ¶rk, T., Bogdanova, N. V., Figueroa, J., Pharoah, P. D. P., Schmidt, M. K., Dunning, A. M., GarcÃa-Closas, M., Bolla, M. K., Dennis, J., Michailidou, K., Wang, Q., Hopper, J. L., Southey, M. C., Rosenberg, E. H., Fasching, P. A., Beckmann, M. W., Peto, J., dos-Santos-Silva, I., Sawyer, E. J., Tomlinson, I., Burwinkel, B., Surowy, H., GuÃ©nel, P., Truong, T., Bojesen, S. E., Nordestgaard, B., Benitez, J., GonzÃ¡lez-Neira, A., Neuhausen, S. L., Anton-Culver, H., Brenner, H., Arndt, V., Meindl, A., Schmutzler, R. K., Brauch, H., BrÃ¼ning, T., Lindblom, A., Margolin, S., Mannermaa, A., Hartikainen, J. M., Chenevix-Trench, G., Van Dyck, L., Janssen, H., Chang-Claude, J., Rudolph, A., Radice, P., Peterlongo, P., Hallberg, E., Olson, J. E., Giles, G. G., Milne, R. L., Haiman, C. A., Schumacher, F., Simard, J., Dumont, M., Kristensen, V., Borresen-Dale, A., Zheng, W., Beeghly-Fadiel, A., Grip, M., Andrulis, I. L., Glendon, G., Devilee, P., Seynaeve, C., Hooning, M. J., CollÃ©e, M., Cox, A., Cross, S. S., Shah, M., Luben, R. N., Hamann, U., Torres, D., Jakubowska, A., Lubinski, J., Couch, F. J., Yannoukakos, D., Orr, N., Swerdlow, A., Darabi, H., Li, J., Czene, K., Hall, P., Easton, D. F., Mattson, J., Blomqvist, C., AittomÃ¤ki, K., Nevanlinna, H. & kConFab/AOCS Investigators. RAD51B in Familial Breast Cancer *PLoS ONE* **11:** e0153788 (2016)

211. Patch, A., Christie, E. L., Etemadmoghadam, D., Garsed, D. W., George, J., Fereday, S., Nones, K., Cowin, P., Alsop, K., Bailey, P. J., Kassahn, K. S., Newell, F., Quinn, M. C. J., Kazakoff, S., Quek, K., Wilhelm-Benartzi, C., Curry, E., Leong, H. S., The Australian Ovarian Cancer,Study Group, Hamilton, A., Mileshkin, L., Au-Yeung, G., Kennedy, C., Hung, J., Chiew, Y., Harnett, P., Friedlander, M., Quinn, M., Pyman, J., Cordner, S., O/'Brien, P., Leditschke, J., Young, G., Strachan, K., Waring, P., Azar, W., Mitchell, C., Traficante, N., Hendley, J., Thorne, H., Shackleton, M., Miller, D. K., Arnau, G. M., Tothill, R. W., Holloway, T. P., Semple, T., Harliwong, I., Nourse, C., Nourbakhsh, E., Manning, S., Idrisoglu, S., Bruxner, T. J. C., Christ, A. N., Poudel, B., Holmes, O., Anderson, M., Leonard, C., Lonie, A., Hall, N., Wood, S., Taylor, D. F., Xu, Q., Fink, J. L., Waddell, N., Drapkin, R., Stronach, E., Gabra, H., Brown, R., Jewell, A., Nagaraj, S. H., Markham, E., Wilson, P. J., Ellul, J., McNally, O., Doyle, M. A., Vedururu, R., Stewart, C., Lengyel, E., Pearson, J. V., Waddell, N., deFazio, A., Grimmond, S. M. & Bowtell, D. D. L. Whole-genome characterization of chemoresistant ovarian cancer *Nature* **521:** 489-494 (2015)

212. Nicodème, F., Geffroy, S., Conti, M., Delobel, B., Soenen, V., Grardel, N., Porte, H., Copin, M., Laï, J. & Andrieux, J. Familial occurrence of thymoma and autoimmune diseases with the constitutional translocation t(14;20)(q24.1;p12.3) *Genes, Chromosomes and Cancer* **44:** 154-160 (2005)

213. Sell, S. M., Tullis, C., Stracner, D., Song, C. Y. & Gewin, J. Minimal interval defined on 7q in uterine leiomyoma *Cancer genetics and cytogenetics* **157:** 67-69 (2005)

214. Vanharanta, S., Wortham, N. C., Laiho, P., Sjoberg, J., Aittomaki, K., Arola, J., Tomlinson, I. P., Karhu, A., Arango, D. & Aaltonen, L. A. 7q Deletion Mapping and Expression Profiling in Uterine Fibroids *Oncogene* **24:** 6545-6554 (2005)

215. Vanharanta, S., Wortham, N. C., Langford, C., El-Bahrawy, M., van der Spuy, Z., Sjoberg, J., Lehtonen, R., Karhu, A., Tomlinson, I. P. & Aaltonen, L. A. Definition of a minimal region of deletion of chromosome 7 in uterine leiomyomas by tiling-path microarray CGH and mutation analysis of known genes in this region *Genes, chromosomes & cancer* **46:** 451-458 (2007)

216. Moon, N. S., Rong Zeng, W., Premdas, P., Santaguida, M., Berube, G. & Nepveu, A. Expression of N-terminally truncated isoforms of CDP/CUX is increased in human uterine leiomyomas *International journal of cancer* **100:** 429-432 (2002)

217. Polly Xing, Y., Lee Powell, W. & Morton, C. C. The del(7q) subgroup in uterine leiomyomata: Genetic and biologic characteristics: Further evidence for the secondary nature of cytogenetic abnormalities in the pathobiology of uterine leiomyomata *Cancer genetics and cytogenetics* **98:** 69-74 (1997)

218. Ozisik, Y. Y., Meloni, A. M., Surti, U. & Sandberg, A. A. Involvement of 10q22 in leiomyoma *Cancer genetics and cytogenetics* **69:** 132-135 (1993)

219. Moore, S. D., Herrick, S. R., Ince, T. A., Kleinman, M. S., Dal Cin, P., Morton, C. C. & Quade, B. J. Uterine leiomyomata with t(10;17) disrupt the histone acetyltransferase MORF *Cancer research* **64:** 5570-5577 (2004)

220. Panagopoulos, I., Gorunova, L., Bjerkehagen, B. & Heim, S. Novel KAT6B-KANSL1 fusion gene identified by RNA sequencing in retroperitoneal leiomyoma with t(10;17)(q22;q21) *PloS one* **10:** e0117010 (2015)

221. Troke, P. J., Kindle, K. B., Collins, H. M. & Heery, D. M. MOZ fusion proteins in acute myeloid leukaemia *Biochemical Society symposium* **73:** 23-39 (2006)

222. Kiechle-Schwarz, M., Berger, C. S., Surti, U. & Sandberg, A. A. Rearrangement of band 10q22 in leiomyoma and leiomyosarcoma of the uterus *Cancer genetics and cytogenetics* **47:** 95-100 (1990)

223. Holzmann, C., Markowski, D. N., VON Leffern, I., Loning, T. & Bullerdiek, J. Patterns of Chromosomal Abnormalities that Can Improve Diagnosis of Uterine Smooth Muscle Tumors *Anticancer Research* **35:** 6445-6456 (2015)

224. Buza, N., Xu, F., Wu, W., Carr, R. J., Li, P. & Hui, P. Recurrent chromosomal aberrations in intravenous leiomyomatosis of the uterus: high-resolution array comparative genomic hybridization study *Human pathology* **45:** 1885-1892 (2014)

225. Hodge, J. C., Pearce, K. E., Clayton, A. C., Taran, F. A. & Stewart, E. A. Uterine cellular leiomyomata with chromosome 1p deletions represent a distinct entity *American Journal of Obstetrics and Gynecology* **210:** 572.e1-572.e7 (2014)

226. van Rijk, A., Sweers, M., Huys, E., Kersten, M., Merkx, G., van Kessel, A. G., Debiec-Rychter, M. & Schoenmakers, E. F. Characterization of a recurrent t(1;2)(p36;p24) in human uterine leiomyoma *Cancer genetics and cytogenetics* **193:** 54-62 (2009)

227. Panagopoulos, I., Gorunova, L., Bjerkehagen, B. & Heim, S. Fusion of the genes EWSR1 and PBX3 in retroperitoneal leiomyoma with t(9;22)(q33;q12) *PloS one* **10:** e0124288 (2015)

228. Townsend, D. E., Sparkes, R. S., Baluda, M. C. & McClelland, G. Unicellular histogenesis of uterine leiomyomas as determined by electrophoresis by glucose-6-phosphate dehydrogenase *American Journal of Obstetrics and Gynecology* **107:** 1168-1173 (1970)

229. Hashimoto, K., Azuma, C., Kamiura, S., Kimura, T., Nobunaga, T., Kanai, T., Sawada, M., Noguchi, S. & Saji, F. Clonal determination of uterine leiomyomas by analyzing differential inactivation of the X-chromosome-linked phosphoglycerokinase gene *Gynecologic and obstetric investigation* **40:** 204-208 (1995)

230. Wang, S., Su, Q., Zhu, S., Liu, J., Hu, L. & Li, D. Clonality of multiple uterine leiomyomas *Chinese journal of pathology* **31:** 107-111 (2002)

231. Nibert, M. & Heim, S. Uterine leiomyoma cytogenetics *Genes, chromosomes & cancer* **2:** 3-13 (1990)

232. Nilbert, M., Heim, S., Mandahl, N., Floderus, U. M., Willen, H. & Mitelman, F. Characteristic chromosome abnormalities, including rearrangements of 6p, del(7q), +12, and t(12;14), in 44 uterine leiomyomas *Human genetics* **85:** 605-611 (1990)

233. Mashal, R. D., Fejzo, M. L., Friedman, A. J., Mitchner, N., Nowak, R. A., Rein, M. S., Morton, C. C. & Sklar, J. Analysis of androgen receptor DNA reveals the independent clonal origins of uterine leiomyomata and the secondary nature of cytogenetic aberrations in the development of leiomyomata *Genes, chromosomes & cancer* **11:** 1-6 (1994)

234. Canevari, R. A., Pontes, A., Rosa, F. E., Rainho, C. A. & Rogatto, S. R. Independent clonal origin of multiple uterine leiomyomas that was determined by X chromosome inactivation and microsatellite analysis *American Journal of Obstetrics and Gynecology* **193:** 1395-1403 (2005)

235. Patton, K. T., Cheng, L., Papavero, V., Blum, M. G., Yeldandi, A. V., Adley, B. P., Luan, C., Diaz, L. K., Hui, P. & Yang, X. J. Benign metastasizing leiomyoma: clonality, telomere length and clinicopathologic analysis *Modern pathology* **19:** 130-140 (2006)

236. Quade, B. J., McLachlin, C. M., Soto-Wright, V., Zuckerman, J., Mutter, G. L. & Morton, C. C. Disseminated peritoneal leiomyomatosis. Clonality analysis by X chromosome inactivation and cytogenetics of a clinically benign smooth muscle proliferation *The American Journal of Pathology* **150:** 2153-2166 (1997)

237. Quade, B. J., Dal Cin, P., Neskey, D. M., Weremowicz, S. & Morton, C. C. Intravenous Leiomyomatosis: Molecular and Cytogenetic Analysis of a Case *Modern pathology* **15:** 351-356 (2002)

238. Swierczek, S. I., Piterkova, L., Jelinek, J., Agarwal, N., Hammoud, S., Wilson, A., Hickman, K., Parker, C. J., Cairns, B. R. & Prchal, J. T. Methylation of AR locus does not always reflect X chromosome inactivation state *Blood* **119:** e100-9 (2012)

239. Gomes, C. C. & Gomez, R. S. Interpretation of clonality and X-chromosome inactivation assays urge attention *Journal of clinical pathology* **65:** 473 (2012)

240. Holdsworth-Carson, S. J., Zaitseva, M., Girling, J. E., Vollenhoven, B. J. & Rogers, P. A. Common fibroid-associated genes are differentially expressed in phenotypically dissimilar cell populations isolated from within human fibroids and myometrium *Reproduction* **147:** 683-692 (2014)

241. Mas, A., Cervello, I., Gil-Sanchis, C. & Simón, C. Current understanding of somatic stem cells in leiomyoma formation *Fertility and sterility* **102:** 613-620 (2014)

242. Cook, J. D. & Walker, C. L. The Eker rat: establishing a genetic paradigm linking renal cell carcinoma and uterine leiomyoma *Current Molecular Medicine* **4:** 813-824 (2004)

243. Yeung, R. S. Lessons from the Eker rat model: from cage to bedside *Current Molecular Medicine* **4:** 799-806 (2004)

244. Hunter, D. S., Klotzbucher, M., Kugoh, H., Cai, S. L., Mullen, J. P., Manfioletti, G., Fuhrman, U. & Walker, C. L. Aberrant expression of HMGA2 in uterine leiomyoma associated with loss of TSC2 tumor suppressor gene function *Cancer research* **62:** 3766-3772 (2002)

245. Muzykewicz, D. A., Sharma, A., Muse, V., Numis, A. L., Rajagopal, J. & Thiele, E. A. TSC1 and TSC2 mutations in patients with lymphangioleiomyomatosis and tuberous sclerosis complex *Journal of medical genetics* **46:** 465-468 (2009)

246. Efanov, A., Zanesi, N., Coppola, V., Nuovo, G., Bolon, B., Wernicle-Jameson, D., Lagana, A., Hansjuerg, A., Pichiorri, F. & Croce, C. M. Human HMGA2 protein overexpressed in mice induces precursor T-cell lymphoblastic leukemia *Blood Cancer Journal* **4:** e227 (2014)

247. Mas, A., Cervello, I., Fernandez-Alvarez, A., Faus, A., Diaz, A., Burgues, O., Casado, M. & Simon, C. Overexpression of the truncated form of High Mobility Group A proteins (HMGA2) in human myometrial cells induces leiomyoma-like tissue formation *Molecular human reproduction* **21:** 330-338 (2015)

248. Lingaas, F., Comstock, K. E., Kirkness, E. F., Sorensen, A., Aarskaug, T., Hitte, C., Nickerson, M. L., Moe, L., Schmidt, L. S., Thomas, R., Breen, M., Galibert, F., Zbar, B. & Ostrander, E. A. A mutation in the canine BHD gene is associated with hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis in the German Shepherd dog *Human molecular genetics* **12:** 3043-3053 (2003)

249. Menko, F. H., van Steensel, M. A., Giraud, S., Friis-Hansen, L., Richard, S., Ungari, S., Nordenskjold, M., Hansen, T. V., Solly, J., Maher, E. R. & European BHD Consortium. Birt-Hogg-Dube syndrome: diagnosis and management *The Lancet Oncology* **10:** 1199-1206 (2009)

250. Porter, K. B., Tsibris, J. C., Nicosia, S. V., Murphy, J. M., O'Brien, W. F., Rao, P. S. & Spellacy, W. N. Estrogen-induced guinea pig model for uterine leiomyomas: do the ovaries protect? *Biology of reproduction* **52:** 824-832 (1995)

251. Prizant, H., Sen, A., Light, A., Cho, S. N., DeMayo, F. J., Lydon, J. P. & Hammes, S. R. Uterine-specific loss of Tsc2 leads to myometrial tumors in both the uterus and lungs *Molecular endocrinology* **27:** 1403-1414 (2013)

252. Varghese, B. V., Koohestani, F., McWilliams, M., Colvin, A., Gunewardena, S., Kinsey, W. H., Nowak, R. A., Nothnick, W. B. & Chennathukuzhi, V. M. Loss of the repressor REST in uterine fibroids promotes aberrant G protein-coupled receptor 10 expression and activates mammalian target of rapamycin pathway *Proceedings of the National Academy of Sciences of the United States of America* **110:** 2187-2192 (2013)

253. Tanwar, P. S., Lee, H. J., Zhang, L., Zukerberg, L. R., Taketo, M. M., Rueda, B. R. & Teixeira, J. M. Constitutive activation of Beta-catenin in uterine stroma and smooth muscle leads to the development of mesenchymal tumors in mice *Biology of reproduction* **81:** 545-552 (2009)

254. Hassan, M. H., Eyzaguirre, E., Arafa, H. M., Hamada, F. M., Salama, S. A. & Al-Hendy, A. Memy I: a novel murine model for uterine leiomyoma using adenovirus-enhanced human fibroid explants in severe combined immune deficiency mice *American Journal of Obstetrics and Gynecology* **199:** 156.e1-156.e8 (2008)

255. Suo, G., Sadarangani, A., Lamarca, B., Cowan, B. & Wang, J. Y. Murine xenograft model for human uterine fibroids: an in vivo imaging approach *Reproductive sciences* **16:** 827-842 (2009)

256. Ishikawa, H., Ishi, K., Serna, V. A., Kakazu, R., Bulun, S. E. & Kurita, T. Progesterone is essential for maintenance and growth of uterine leiomyoma *Endocrinology* **151:** 2433-2442 (2010)

257. Tsuiji, K., Takeda, T., Li, B., Kondo, A., Ito, M. & Yaegashi, N. Establishment of a novel xenograft model for human uterine leiomyoma in immunodeficient mice *The Tohoku journal of experimental medicine* **222:** 55-61 (2010)

258. Drosch, M., Bullerdiek, J., Zollner, T. M., Prinz, F., Koch, M. & Schmidt, N. A novel mouse model that closely mimics human uterine leiomyomas *Fertility and sterility* **99:** 927-935.e6 (2013)

259. Sousa, W. B., Garcia, J. B., Nogueira Neto, J., Furtado, P. G. & Anjos, J. A. Xenotransplantation of uterine leiomyoma in Wistar rats: a pilot study *European journal of obstetrics, gynecology, and reproductive biology* **190:** 71-75 (2015)

260. Mozzachio, K., Linder, K. & Dixon, D. Uterine smooth muscle tumors in potbellied pigs (Sus scrofa) resemble human fibroids: a potential animal model *Toxicologic pathology* **32:** 402-407 (2004)

261. Bredhult, C., Backlin, B. M., Bignert, A. & Olovsson, M. Study of the relation between the incidence of uterine leiomyomas and the concentrations of PCB and DDT in Baltic gray seals *Reproductive toxicology* **25:** 247-255 (2008)

262. Carney, S. A., Tahara, H., Swartz, C. D., Risinger, J. I., He, H., Moore, A. B., Haseman, J. K., Barrett, J. C. & Dixon, D. Immortalization of human uterine leiomyoma and myometrial cell lines after induction of telomerase activity: molecular and phenotypic characteristics *Laboratory investigation* **82:** 719-728 (2002)

263. Arslan, A. A., Gold, L. I., Mittal, K., Suen, T. C., Belitskaya-Levy, I., Tang, M. S. & Toniolo, P. Gene expression studies provide clues to the pathogenesis of uterine leiomyoma: new evidence and a systematic review *Human reproduction* **20:** 852-863 (2005)

264. Vanharanta, S., Pollard, P. J., Lehtonen, H. J., Laiho, P., Sjoberg, J., Leminen, A., Aittomaki, K., Arola, J., Kruhoffer, M., Orntoft, T. F., Tomlinson, I. P., Kiuru, M., Arango, D. & Aaltonen, L. A. Distinct expression profile in fumarate-hydratase-deficient uterine fibroids *Human molecular genetics* **15:** 97-103 (2006)

265. Christacos, N. C., Quade, B. J., Dal Cin, P. & Morton, C. C. Uterine leiomyomata with deletions of Ip represent a distinct cytogenetic subgroup associated with unusual histologic features *Genes, chromosomes & cancer* **45:** 304-312 (2006)

266. Leppert, P. C., Catherino, W. H. & Segars, J. H. A new hypothesis about the origin of uterine fibroids based on gene expression profiling with microarrays *American Journal of Obstetrics and Gynecology* **195:** 415-420 (2006)

267. Crabtree, J. S., Jelinsky, S. A., Harris, H. A., Choe, S. E., Cotreau, M. M., Kimberland, M. L., Wilson, E., Saraf, K. A., Liu, W., McCampbell, A. S., Dave, B., Broaddus, R. R., Brown, E. L., Kao, W., Skotnicki, J. S., Abou-Gharbia, M., Winneker, R. C. & Walker, C. L. Comparison of human and rat uterine leiomyomata: identification of a dysregulated mammalian target of rapamycin pathway *Cancer research* **69:** 6171-6178 (2009)

268. Blagosklonny, M. V., Zavadil, J., Ye, H., Liu, Z., Wu, J., Lee, P., Hernando, E., Soteropoulos, P., Toruner, G. A. & Wei, J. Profiling and Functional Analyses of MicroRNAs and Their Target Gene Products in Human Uterine Leiomyomas *PLoS ONE* **5:** e12362 (2010)

269. Chegini, N. Proinflammatory and profibrotic mediators: principal effectors of leiomyoma development as a fibrotic disorder *Seminars in reproductive medicine* **28:** 180-203 (2010)

270. Ciarmela, P., Islam, M. S., Reis, F. M., Gray, P. C., Bloise, E., Petraglia, F., Vale, W. & Castellucci, M. Growth factors and myometrium: biological effects in uterine fibroid and possible clinical implications *Human reproduction update* **17:** 772-790 (2011)

271. Hoekstra, A. V., Sefton, E. C., Berry, E., Lu, Z., Hardt, J., Marsh, E., Yin, P., Clardy, J., Chakravarti, D., Bulun, S. & Kim, J. J. Progestins Activate the AKT Pathway in Leiomyoma Cells and Promote Survival *The Journal of clinical endocrinology and metabolism* **94:** 1768-1774 (2009)

272. Xu, X., Lu, Z., Qiang, W., Vidimar, V., Kong, B., Kim, J. J. & Wei, J. Inactivation of AKT Induces Cellular Senescence in Uterine Leiomyoma *Endocrinology* **155:** 1510-1519 (2014)

273. Yin, X. J., Wang, G. & Khan-Dawood, F. S. Requirements of phosphatidylinositol-3 kinase and mammalian target of rapamycin for estrogen-induced proliferation in uterine leiomyoma- and myometrium-derived cell lines *American Journal of Obstetrics and Gynecology* **196:** 176.e1-176.e5 (2007)

274. Jamieson, C., Sharma, M. & Henderson, B. R. Wnt signaling from membrane to nucleus: β-catenin caught in a loop *The international journal of biochemistry & cell biology* **44:** 847-850 (2012)

275. Arango, N. A., Szotek, P. P., Manganaro, T. F., Oliva, E., Donahoe, P. K. & Teixeira, J. Conditional deletion of β-catenin in the mesenchyme of the developing mouse uterus results in a switch to adipogenesis in the myometrium *Developmental biology* **288:** 276-283 (2005)

276. Ono, M., Yin, P., Navarro, A., Moravek, M. B., Coon JS, 5., Druschitz, S. A., Serna, V. A., Qiang, W., Brooks, D. C., Malpani, S. S., Ma, J., Ercan, C. M., Mittal, N., Monsivais, D., Dyson, M. T., Yemelyanov, A., Maruyama, T., Chakravarti, D., Kim, J. J., Kurita, T., Gottardi, C. J. & Bulun, S. E. Paracrine activation of WNT/beta-catenin pathway in uterine leiomyoma stem cells promotes tumor growth *Proceedings of the National Academy of Sciences of the United States of America* **110:** 17053-17058 (2013)

277. Ono, M., Yin, P., Navarro, A., Moravek, M. B., Coon, V. J. S., Druschitz, S. A., Gottardi, C. J. & Bulun, S. E. Inhibition of canonical WNT signaling attenuates human leiomyoma cell growth *Fertility and sterility* **101:** 1441-1449 (2014)

278. Joseph, D. S., Malik, M., Nurudeen, S. & Catherino, W. H. Myometrial cells undergo fibrotic transformation under the influence of transforming growth factor β-3 *Fertility and sterility* **93:** 1500-1508 (2010)

279. Raimundo, N., Vanharanta, S., Aaltonen, L. A., Hovatta, I. & Suomalainen, A. Downregulation of SRF-FOS-JUNB pathway in fumarate hydratase deficiency and in uterine leiomyomas *Oncogene* **28:** 1261-1273 (2009)

280. Andrews, S. FastQC: a quality control tool for high throughput sequence data (2010)

281. Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., Miller, C. A., Mardis, E. R., Ding, L. & Wilson, R. K. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing *Genome research* **22:** 568-576 (2012)

282. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform *Bioinformatics* **25:** 1754-1760 (2009)

283. Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E. S. & Getz, G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples *Nature biotechnology* **31:** 213-219 (2013)

284. Venkatraman, E. S. & Olshen, A. B. A faster circular binary segmentation algorithm for the analysis of array CGH data *Bioinformatics* **23:** 657-663 (2007)

285. Chen, K., Wallis, J. W., McLellan, M. D., Larson, D. E., Kalicki, J. M., Pohl, C. S., McGrath, S. D., Wendl, M. C., Zhang, Q., Locke, D. P., Shi, X., Fulton, R. S., Ley, T. J., Wilson, R. K., Ding, L. & Mardis, E. R. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation *Nature methods* **6:** 677-681 (2009)

286. MacDonald, J. R., Ziman, R., Yuen, R. K. C., Feuk, L. & Scherer, S. W. The Database of Genomic Variants: a curated collection of structural variation in the human genome *Nucleic acids research* **42:** D986-D992 (2014)

287. Bailey, J. A., Yavor, A. M., Massa, H. F., Trask, B. J. & Eichler, E. E. Segmental duplications: organization and impact within the current human genome project assembly *Genome research* **11:** 1005-1017 (2001)

288. Pickrell, J. K., Gaffney, D. J., Gilad, Y. & Pritchard, J. K. False positive peaks in ChIP-seq and other sequencing-based functional assays caused by unannotated high copy number regions *Bioinformatics* **27:** 2144-2146 (2011)

289. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing *Journal of the Royal Statistical Society.Series B (Methodological)* **57:** 289-300 (1995)

290. Wang, J., Duncan, D., Shi, Z. & Zhang, B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013 *Nucleic acids research* **41:** W77-W83 (2013)

291. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq *Bioinformatics* **25:** 1105-1111 (2009)

292. Koressaar, T. & Remm, M. Enhancements and modifications of primer design program Primer3 *Bioinformatics* **23:** 1289-1291 (2007)

293. Wang, M. & Marin, A. Characterization and prediction of alternative splice sites *Gene* **366:** 219-227 (2006)

294. Hebsgaard, S. M., Korning, P. G., Tolstrup, N., Engelbrecht, J., Rouze, P. & Brunak, S. Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information *Nucleic acids research* **24:** 3439-3452 (1996)

295. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput *Nucleic acids research* **32:** 1792-1797 (2004)

296. Blanchette, M., Kent, W. J., Riemer, C., Elnitski, L., Smit, A. F., Roskin, K. M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E. D., Haussler, D. & Miller, W. Aligning multiple genomic sequences with the threaded blockset aligner *Genome research* **14:** 708-715 (2004)

297. Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices *Journal of Molecular Biology* **292:** 195-202 (1999)

298. Cleynen, I., Brants, J. R., Peeters, K., Deckers, R., Debiec-Rychter, M., Sciot, R., Van de Ven, W. J. & Petit, M. M. HMGA2 regulates transcription of the Imp2 gene via an intronic regulatory element in cooperation with nuclear factor-kappaB *Molecular Cancer Research* **5:** 363-372 (2007)

299. Voz, M. L., Agten, N. S., Van de Ven, W. J. & Kas, K. PLAG1, the main translocation target in pleomorphic adenoma of the salivary glands, is a positive regulator of IGF-II *Cancer research* **60:** 106-113 (2000)

300. Voz, M. L., Mathys, J., Hensen, K., Pendeville, H., Van Valckenborgh, I., Van Huffel, C., Chavez, M., Van Damme, B., De Moor, B., Moreau, Y. & Van de Ven, W. J. Microarray screening for target genes of the proto-oncogene PLAG1 *Oncogene* **23:** 179-191 (2004)

301. Akhtar, M., Holmgren, C., Gondor, A., Vesterlund, M., Kanduri, C., Larsson, C. & Ekstrom, T. J. Cell type and context-specific function of PLAG1 for IGF2 P3 promoter activity *International journal of oncology* **41:** 1959-1966 (2012)

302. Malhotra, D., Portales-Casamar, E., Singh, A., Srivastava, S., Arenillas, D., Happel, C., Shyr, C., Wakabayashi, N., Kensler, T. W., Wasserman, W. W. & Biswal, S. Global mapping of binding sites for Nrf2 identifies novel targets in cell survival response through ChIP-Seq profiling and network analysis *Nucleic acids research* **38:** 5718-5734 (2010)

303. Ooi, A., Wong, J., Petillo, D., Roossien, D., Perrier-Trudova, V., Whitten, D., Min, B., Tan, M., Zhang, Z., Yang, X., Zhou, M., Gardie, B., Molinié, V., Richard, S., Tan, P., Teh, B. & Furge, K. An Antioxidant Response Phenotype Shared between Hereditary and Sporadic Type 2 Papillary Renal Cell Carcinoma *Cancer Cell* **20:** 511-523 (2011)

304. Adam, J., Hatipoglu, E., O'Flaherty, L., Ternette, N., Sahgal, N., Lockstone, H., Baban, D., Nye, E., Stamp, G., Wolhuter, K., Stevens, M., Fischer, R., Carmeliet, P., Maxwell, P., Pugh, C., Frizzell, N., Soga, T., Kessler, B., El-Bahrawy, M., Ratcliffe, P. & Pollard, P. Renal Cyst Formation in Fh1-Deficient Mice Is Independent of the Hif/Phd Pathway: Roles for Fumarate in KEAP1 Succination and Nrf2 Signaling *Cancer Cell* **20:** 524-537 (2011)

305. Mitsuishi, Y., Taguchi, K., Kawatani, Y., Shibata, T., Nukiwa, T., Aburatani, H., Yamamoto, M. & Motohashi, H. Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming *Cancer cell* **22:** 66-79 (2012)

306. Kampjarvi, K., Park, M. J., Mehine, M., Kim, N. H., Clark, A. D., Butzow, R., Bohling, T., Bohm, J., Mecklin, J. P., Jarvinen, H., Tomlinson, I. P., van der Spuy, Z. M., Sjoberg, J., Boyer, T. G. & Vahteristo, P. Mutations in Exon 1 highlight the role of MED12 in uterine leiomyomas *Human mutation* **35:** 1136-1141 (2014)

307. Heinonen, H., Sarvilinna, N. S., Sjöberg, J., Kämpjärvi, K., Pitkänen, E., Vahteristo, P., Mäkinen, N. & Aaltonen, L. A. MED12 mutation frequency in unselected sporadic uterine leiomyomas *Fertility and sterility* **102:** 1137-1142 (2014)

308. Rieker, R. J., Agaimy, A., Moskalev, E. A., Hebele, S., Hein, A., Mehlhorn, G., Beckmann, M. W., Hartmann, A. & Haller, F. Mutation status of the mediator complex subunit 12 (MED12) in uterine leiomyomas and concurrent/metachronous multifocal peritoneal smooth muscle nodules (leiomyomatosis peritonealis disseminata) *Pathology* **45:** 388-392 (2013)

309. Markowski, D. N., Helmke, B. M., Bartnitzke, S., Loning, T. & Bullerdiek, J. Uterine fibroids: do we deal with more than one disease? *International journal of gynecological pathology* **33:** 568-572 (2014)

310. Markowski, D. N., Bartnitzke, S., Loning, T., Drieschner, N., Helmke, B. M. & Bullerdiek, J. MED12 mutations in uterine fibroids--their relationship to cytogenetic subgroups *International journal of cancer* **131:** 1528-1536 (2012)

311. Markowski, D. N., Huhle, S., Nimzyk, R., Stenman, G., Loning, T. & Bullerdiek, J. MED12 mutations occurring in benign and malignant mammalian smooth muscle tumors *Genes, chromosomes & cancer* **52:** 297-304 (2013)

312. Mäkinen, N., Mehine, M., Tolvanen, J., Kaasinen, E., Li, Y., Lehtonen, H. J., Gentile, M., Yan, J., Enge, M., Taipale, M., Aavikko, M., Katainen, R., Virolainen, E., Bohling, T., Koski, T. A., Launonen, V., Sjoberg, J., Taipale, J., Vahteristo, P. & Aaltonen, L. A. MED12, the mediator complex subunit 12 gene, is mutated at high frequency in uterine leiomyomas *Science* **334:** 252-255 (2011)

313. Mäkinen, N., Vahteristo, P., Kämpjärvi, K., Arola, J., Butzow, R. & Aaltonen, L. A. MED12 exon 2 mutations in histopathological uterine leiomyoma variants *European journal of human genetics* **21:** 1300-1303 (2013)

314. Bertsch, E., Qiang, W., Zhang, Q., Espona-Fiedler, M., Druschitz, S., Liu, Y., Mittal, K., Kong, B., Kurita, T. & Wei, J. J. MED12 and HMGA2 mutations: two independent genetic events in uterine leiomyoma and leiomyosarcoma *Modern pathology* **27:** 1144-1153 (2014)

315. Halder, S. K., Laknaur, A., Miller, J., Layman, L. C., Diamond, M. & Al-Hendy, A. Novel MED12 gene somatic mutations in women from the Southern United States with symptomatic uterine fibroids *Molecular genetics and genomics* **290:** 505-511 (2015)

316. McGuire, M. M., Yatsenko, A., Hoffner, L., Jones, M., Surti, U. & Rajkovic, A. Whole exome sequencing in a random sample of North American women with leiomyomas identifies MED12 mutations in majority of uterine leiomyomas *PloS one* **7:** e33251 (2012)

317. Ravegnini, G., Marino-Enriquez, A., Slater, J., Eilers, G., Wang, Y., Zhu, M., Nucci, M. R., George, S., Angelini, S., Raut, C. P. & Fletcher, J. A. MED12 mutations in leiomyosarcoma and extrauterine leiomyoma *Modern pathology* **26:** 743-749 (2013)

318. Schwetye, K. E., Pfeifer, J. D. & Duncavage, E. J. MED12 exon 2 mutations in uterine and extrauterine smooth muscle tumors *Human pathology* **45:** 65-70 (2014)

319. Matsubara, A., Sekine, S., Yoshida, M., Yoshida, A., Taniguchi, H., Kushima, R., Tsuda, H. & Kanai, Y. Prevalence of MED12 mutations in uterine and extrauterine smooth muscle tumours *Histopathology* **62:** 657-661 (2013)

320. Zhang, Q., Ubago, J., Li, L., Guo, H., Liu, Y., Qiang, W., Kim, J. J., Kong, B. & Wei, J. J. Molecular analyses of 6 different types of uterine smooth muscle tumors: Emphasis in atypical leiomyoma *Cancer* **120:** 3165-3177 (2014)

321. Wang, H., Ye, J., Qian, H., Zhou, R., Jiang, J. & Ye, L. High-resolution melting analysis of MED12 mutations in uterine leiomyomas in Chinese patients *Genetic testing and molecular biomarkers* **19:** 162-166 (2015)

322. Osinovskaya, N. S., Ivashchenko, T. E., Dolinskii, A. K., Sultanov, I. Y., Ghimbovchi, S., Hoffman, E., Bezhenar, V. F. & Baranov, V. S. MED12 gene mutations in women with uterine myoma *Genetika* **49:** 1426-1431 (2013)

323. Osinovskaya, N. S., Malysheva, O. V., Shved, N. Y., Ivashchenko, T. E., Sultanov, I. Y., Efimova, O. A., Yarmolinskaya, M. I., Bezhenar, V. F. & Baranov, V. S. Frequency and Spectrum of MED12 Exon 2 Mutations in Multiple Versus Solitary Uterine Leiomyomas From Russian Patients *International journal of gynecological pathology* [Epub ahead of print] (2015)

324. Pérot, G. MED12 Alterations in Both Human Benign and Malignant Uterine Soft Tissue Tumors *PloS one* **7:** e40015 (2012)

325. de Graaff, M. A., Cleton-Jansen, A. M., Szuhai, K. & Bovee, J. V. Mediator complex subunit 12 exon 2 mutation analysis in different subtypes of smooth muscle tumors confirms genetic heterogeneity *Human pathology* **44:** 1597-1604 (2013)

326. Je, E. M., Kim, M. R., Min, K. O., Yoo, N. J. & Lee, S. H. Mutational analysis of MED12 exon 2 in uterine leiomyoma and other common tumors *International journal of cancer* **131:** E1044-7 (2012)

327. Makinen, N., Heinonen, H. R., Moore, S., Tomlinson, I. P., van der Spuy, Z. M. & Aaltonen, L. A. MED12 exon 2 mutations are common in uterine leiomyomas from South African patients *Oncotarget* **2:** 966-969 (2011)

328. Shahbazi, S., Fatahi, N. & Amini-Moghaddam, S. Somatic mutational analysis of MED12 exon 2 in uterine leiomyomas of Iranian women *American Journal of Cancer Research* **5:** 2441-2446 (2015)

329. Sadeghi, S., Khorrami, M., Amin-Beidokhti, M., Abbasi, M., Kamalian, Z., Irani, S., Omrani, M., Azmoodeh, O. & Mirfakhraie, R. The study of MED12 gene mutations in uterine leiomyomas from Iranian patients *Tumour biology* **37:** 1567-1571 (2016)

330. Liegl-Atzwanger, B., Heitzer, E., Flicker, K., Muller, S., Ulz, P., Saglam, O., Tavassoli, F., Devouassoux-Shisheboran, M., Geigl, J. & Moinfar, F. Exploring chromosomal abnormalities and genetic changes in uterine smooth muscle tumors *Modern pathology* [Epub ahead of print] (2016)

331. Makinen, N., Aavikko, M., Heikkinen, T., Taipale, M., Taipale, J., Koivisto-Korander, R., Butzow, R. & Vahteristo, P. Exome Sequencing of Uterine Leiomyosarcomas Identifies Frequent Mutations in TP53, ATRX, and MED12 *PLoS genetics* **12:** e1005850 (2016)

332. Kämpjärvi, K., Mäkinen, N., Kilpivaara, O., Arola, J., Heinonen, H. R., Bohm, J., Abdel-Wahab, O., Lehtonen, H. J., Pelttari, L. M., Mehine, M., Schrewe, H., Nevanlinna, H., Levine, R. L., Hokland, P., Bohling, T., Mecklin, J. P., Butzow, R., Aaltonen, L. A. & Vahteristo, P. Somatic MED12 mutations in uterine leiomyosarcoma and colorectal cancer *British journal of cancer* **107:** 1761-1765 (2012)

333. Kampjarvi, K., Jarvinen, T. M., Heikkinen, T., Ruppert, A. S., Senter, L., Hoag, K. W., Dufva, O., Kontro, M., Rassenti, L., Hertlein, E., Kipps, T. J., Porkka, K., Byrd, J. C., de la Chapelle, A. & Vahteristo, P. Somatic MED12 mutations are associated with poor prognosis markers in chronic lymphocytic leukemia *Oncotarget* **6:** 1884-1888 (2015)

334. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer *Nature* **487:** 330-337 (2012)

335. Lim, W. K., Ong, C. K., Tan, J., Thike, A. A., Ng, C. C., Rajasegaran, V., Myint, S. S., Nagarajan, S., Nasir, N. D., McPherson, J. R., Cutcutache, I., Poore, G., Tay, S. T., Ooi, W. S., Tan, V. K., Hartman, M., Ong, K. W., Tan, B. K., Rozen, S. G., Tan, P. H., Tan, P. & Teh, B. T. Exome sequencing identifies highly recurrent MED12 somatic mutations in breast fibroadenoma *Nature genetics* **46:** 877-880 (2014)

336. Tan, J., Ong, C. K., Lim, W. K., Ng, C. C., Thike, A. A., Ng, L. M., Rajasegaran, V., Myint, S. S., Nagarajan, S., Thangaraju, S., Dey, S., Nasir, N. D., Wijaya, G. C., Lim, J. Q., Huang, D., Li, Z., Wong, B. H., Chan, J. Y., McPherson, J. R., Cutcutache, I., Poore, G., Tay, S. T., Tan, W. J., Putti, T. C., Ahmad, B. S., Iau, P., Chan, C. W., Tang, A. P., Yong, W. S., Madhukumar, P., Ho, G. H., Tan, V. K., Wong, C. Y., Hartman, M., Ong, K. W., Tan, B. K., Rozen, S. G., Tan, P., Tan, P. H. & Teh, B. T. Genomic landscapes of breast fibroepithelial tumors *Nature genetics* **47:** 1341-1345 (2015)

337. Yoon, N., Bae, G. E., Kang, S. Y., Choi, M. S., Hwang, H. W., Kim, S. W., Lee, J. E., Nam, S. J., Gong, G., Lee, H. J., Bae, Y. K., Lee, A. & Cho, E. Y. Frequency of MED12 mutations in phyllodes tumors: Inverse correlation with histologic grade *Genes, chromosomes & cancer* **55:** 495-504 (2016)

338. Cani, A. K., Hovelson, D. H., McDaniel, A. S., Sadis, S., Haller, M. J., Yadati, V., Amin, A. M., Bratley, J., Bandla, S., Williams, P. D., Rhodes, K., Liu, C. J., Quist, M. J., Rhodes, D. R., Grasso, C. S., Kleer, C. G. & Tomlins, S. A. Next-Gen Sequencing Exposes Frequent MED12 Mutations and Actionable Therapeutic Targets in Phyllodes Tumors *Molecular Cancer Research* **13:** 613-619 (2015)

339. Yoshida, M., Ogawa, R., Yoshida, H., Maeshima, A., Kanai, Y., Kinoshita, T., Hiraoka, N. & Sekine, S. TERT promoter mutations are frequent and show association with MED12 mutations in phyllodes tumors of the breast *British journal of cancer* **113:** 1244-1248 (2015)

340. Nagasawa, S., Maeda, I., Fukuda, T., Wu, W., Hayami, R., Kojima, Y., Tsugawa, K. & Ohta, T. MED12 exon 2 mutations in phyllodes tumors of the breast *Cancer Medicine* **4:** 1117-1121 (2015)

341. Piscuoglio, S., Murray, M., Fusco, N., Marchio, C., Loo, F. L., Martelotto, L. G., Schultheis, A. M., Akram, M., Weigelt, B., Brogi, E. & Reis-Filho, J. S. MED12 somatic mutations in fibroadenomas and phyllodes tumours of the breast *Histopathology* **67:** 719-729 (2015)

342. Yoshida, M., Sekine, S., Ogawa, R., Yoshida, H., Maeshima, A., Kanai, Y., Kinoshita, T. & Ochiai, A. Frequent MED12 mutations in phyllodes tumours of the breast *British journal of cancer* **112:** 1703-1708 (2015)

343. Piscuoglio, S., Ng, C. K., Murray, M., Burke, K. A., Edelweiss, M., Geyer, F. C., Macedo, G. S., Inagaki, A., Papanastasiou, A. D., Martelotto, L. G., Marchio, C., Lim, R. S., Ioris, R. A., Nahar, P. K., Bruijn, I. D., Smyth, L., Akram, M., Ross, D., Petrini, J. H., Norton, L., Solit, D. B., Baselga, J., Brogi, E., Ladanyi, M., Weigelt, B. & Reis-Filho, J. S. Massively parallel sequencing of phyllodes tumours of the breast reveals actionable mutations, and TERT promoter hotspot mutations and TERT gene amplification as likely drivers of progression *The Journal of pathology* **238:** 508-518 (2016)

344. Pfarr, N., Kriegsmann, M., Sinn, P., Klauschen, F., Endris, V., Herpel, E., Muckenhuber, A., Jesinghaus, M., Klosterhalfen, B., Penzel, R., Lennerz, J. K., Weichert, W. & Stenzinger, A. Distribution of MED12 mutations in fibroadenomas and phyllodes tumors of the breast--implications for tumor biology and pathological diagnosis *Genes, chromosomes & cancer* **54:** 444-452 (2015)

345. Lien, H., Huang, C., Yang, Y. & Jeng, Y. Mutational analysis of MED12 exon 2 in a spectrum of fibroepithelial tumours of the breast: implications for pathogenesis and histogenesis *Histopathology* **68:** 433-441 (2016)

346. Mishima, C., Kagara, N., Tanei, T., Naoi, Y., Shimoda, M., Shimomura, A., Shimazu, K., Kim, S. J. & Noguchi, S. Mutational analysis of MED12 in fibroadenomas and phyllodes tumors of the breast by means of targeted next-generation sequencing *Breast cancer research and treatment* **152:** 305-312 (2015)

347. Lien, H., Huang, C., Yang, Y. & Jeng, Y. MED12 exon 2 mutation as a highly sensitive and specific marker in distinguishing phyllodes tumours from other spindle neoplasms of the breast *APMIS* **124:** 356-364 (2016)

348. Ng, C. C., Tan, J., Ong, C. K., Lim, W. K., Rajasegaran, V., Nasir, N. D., Lim, J. C., Thike, A. A., Salahuddin, S. A., Iqbal, J., Busmanis, I., Chong, A. P., Teh, B. T. & Tan, P. H. MED12 is frequently mutated in breast phyllodes tumours: a study of 112 cases *Journal of clinical pathology* **68:** 685-691 (2015)

349. Guièze, R., Robbe, P., Clifford, R., de Guibert, S., Pereira, B., Timbs, A., Dilhuydy, M., Cabes, M., Ysebaert, L., Burns, A., Nguyen-Khac, F., Davi, F., VÃ©ronÃ¨se, L., Combes, P., Le Garff-Tavernier, M., Leblond, V., Merle-BÃ©ral, H., Alsolami, R., Hamblin, A., Mason, J., Pettitt, A., Hillmen, P., Taylor, J., Knight, S. J. L., Tournilhac, O. & Schuh, A. Presence of multiple recurrent mutations confers poor trial outcome of relapsed/refractory CLL *Blood* **126:** 2110-2117 (2015)

350. Wang, L., Lawrence, M. S., Wan, Y., Stojanov, P., Sougnez, C., Stevenson, K., Werner, L., Sivachenko, A., DeLuca, D. S., Zhang, L., Zhang, W., Vartanov, A. R., Fernandes, S. M., Goldstein, N. R., Folco, E. G., Cibulskis, K., Tesar, B., Sievers, Q. L., Shefler, E., Gabriel, S., Hacohen, N., Reed, R., Meyerson, M., Golub, T. R., Lander, E. S., Neuberg, D., Brown, J. R., Getz, G. & Wu, C. J. SF3B1 and Other Novel Cancer Genes in Chronic Lymphocytic Leukemia *The New England journal of medicine* **365:** 2497-2506 (2011)

351. Landau, D., Carter, S., Stojanov, P., McKenna, A., Stevenson, K., Lawrence, M., Sougnez, C., Stewart, C., Sivachenko, A., Wang, L., Wan, Y., Zhang, W., Shukla, S., Vartanov, A., Fernandes, S., Saksena, G., Cibulskis, K., Tesar, B., Gabriel, S., Hacohen, N., Meyerson, M., Lander, E., Neuberg, D., Brown, J., Getz, G. & Wu, C. Evolution and Impact of Subclonal Mutations in Chronic Lymphocytic Leukemia *Cell* **152:** 714-726 (2013)

352. Quesada, V., Conde, L., Villamor, N., Ordonez, G. R., Jares, P., Bassaganyas, L., Ramsay, A. J., Bea, S., Pinyol, M., Martinez-Trillos, A., Lopez-Guerra, M., Colomer, D., Navarro, A., Baumann, T., Aymerich, M., Rozman, M., Delgado, J., Gine, E., Hernandez, J. M., Gonzalez-Diaz, M., Puente, D. A., Velasco, G., Freije, J. M. P., Tubio, J. M. C., Royo, R., Gelpi, J. L., Orozco, M., Pisano, D. G., Zamora, J., Vazquez, M., Valencia, A., Himmelbauer, H., Bayes, M., Heath, S., Gut, M., Gut, I., Estivill, X., Lopez-Guillermo, A., Puente, X. S., Campo, E. & Lopez-Otin, C. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia *Nature genetics* **44:** 47-52 (2012)

353. Damm, F., Mylonas, E., Cosson, A., Yoshida, K., Della Valle, V., Mouly, E., Diop, M., Scourzic, L., Shiraishi, Y., Chiba, K., Tanaka, H., Miyano, S., Kikushige, Y., Davi, F., Lambert, J., Gautheret, D., Merle-Béral, H., Sutton, L., Dessen, P., Solary, E., Akashi, K., Vainchenker, W., Mercher, T., Droin, N., Ogawa, S., Nguyen-Khac, F. & Bernard, O. A. Acquired Initiating Mutations in Early Hematopoietic Cells of CLL Patients *Cancer Discovery* **4:** 1088-1101 (2014)

354. Barbieri, C. E., Baca, S. C., Lawrence, M. S., Demichelis, F., Blattner, M., Theurillat, J. P., White, T. A., Stojanov, P., Van Allen, E., Stransky, N., Nickerson, E., Chae, S. S., Boysen, G., Auclair, D., Onofrio, R. C., Park, K., Kitabayashi, N., MacDonald, T. Y., Sheikh, K., Vuong, T., Guiducci, C., Cibulskis, K., Sivachenko, A., Carter, S. L., Saksena, G., Voet, D., Hussain, W. M., Ramos, A. H., Winckler, W., Redman, M. C., Ardlie, K., Tewari, A. K., Mosquera, J. M., Rupp, N., Wild, P. J., Moch, H., Morrissey, C., Nelson, P. S., Kantoff, P. W., Gabriel, S. B., Golub, T. R., Meyerson, M., Lander, E. S., Getz, G., Rubin, M. A. & Garraway, L. A. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer *Nature genetics* **44:** 685-689 (2012)

355. Assie, G., Letouze, E., Fassnacht, M., Jouinot, A., Luscap, W., Barreau, O., Omeiri, H., Rodriguez, S., Perlemoine, K., Rene-Corail, F., Elarouci, N., Sbiera, S., Kroiss, M., Allolio, B., Waldmann, J., Quinkler, M., Mannelli, M., Mantero, F., Papathomas, T., De Krijger, R., Tabarin, A., Kerlan, V., Baudin, E., Tissier, F., Dousset, B., Groussin, L., Amar, L., Clauser, E., Bertagna, X., Ragazzon, B., Beuschlein, F., Libe, R., de Reynies, A. & Bertherat, J. Integrated genomic characterization of adrenocortical carcinoma *Nature genetics* **46:** 607-612 (2014)

356. Wang, H., Shen, Q., Ye, L. & Ye, J. MED12 mutations in human diseases *Protein & Cell* **4:** 643-646 (2013)

357. The International Cancer Genome Consortium. ICGC data portal 3.8.5.6.

358. Kampjarvi, K., Kim, N. H., Keskitalo, S., Clark, A. D., von Nandelstadh, P., Turunen, M., Heikkinen, T., Park, M. J., Makinen, N., Kivinummi, K., Lintula, S., Hotakainen, K., Nevanlinna, H., Hokland, P., Bohling, T., Butzow, R., Bohm, J., Mecklin, J. P., Jarvinen, H., Kontro, M., Visakorpi, T., Taipale, J., Varjosalo, M., Boyer, T. G. & Vahteristo, P. Somatic MED12 mutations in prostate cancer and uterine leiomyomas promote tumorigenesis through distinct mechanisms *The Prostate* **76:** 22-31 (2016)

359. Stoehr, R., Taubert, H., Gaisa, N. T., Smeets, D., Kneitz, B., Giedl, J., Ruemmele, P., Wieland, W. F., Rau, T. T. & Hartmann, A. Lack of evidence for frequent MED12 p.L1224F mutation in prostate tumours from Caucasian patients *The Journal of pathology* **230:** 453-456 (2013)

360. Arai, E., Sakamoto, H., Ichikawa, H., Totsuka, H., Chiku, S., Gotoh, M., Mori, T., Nakatani, T., Ohnami, S., Nakagawa, T., Fujimoto, H., Wang, L., Aburatani, H., Yoshida, T. & Kanai, Y. Multilayer-omics analysis of renal cell carcinoma, including the whole exome, methylome and transcriptome *International journal of cancer* **135:** 1330-1342 (2014)

361. Taatjes, D. J. The human Mediator complex: a versatile, genome-wide regulator of transcription *Trends in biochemical sciences* **35:** 315-322 (2010)

362. Firestein, R., Bass, A. J., Kim, S. Y., Dunn, I. F., Silver, S. J., Guney, I., Freed, E., Ligon, A. H., Vena, N., Ogino, S., Chheda, M. G., Tamayo, P., Finn, S., Shrestha, Y., Boehm, J. S., Jain, S., Bojarski, E., Mermel, C., Barretina, J., Chan, J. A., Baselga, J., Tabernero, J., Root, D. E., Fuchs, C. S., Loda, M., Shivdasani, R. A., Meyerson, M. & Hahn, W. C. CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity *Nature* **455:** 547-551 (2008)

363. Turunen, M., Spaeth, J. M., Keskitalo, S., Park, M. J., Kivioja, T., Clark, A. D., Mäkinen, N., Gao, F., Palin, K., Nurkkala, H., Vaharautio, A., Aavikko, M., Kämpjärvi, K., Vahteristo, P., Kim, C. A., Aaltonen, L. A., Varjosalo, M., Taipale, J. & Boyer, T. G. Uterine Leiomyoma-Linked MED12 Mutations Disrupt Mediator-Associated CDK Activity *Cell reports* **7:** 654-660 (2014)

364. Mäkinen, N., Heinonen, H. R., Sjoberg, J., Taipale, J., Vahteristo, P. & Aaltonen, L. A. Mutation analysis of components of the Mediator kinase module in MED12 mutation-negative uterine leiomyomas *British journal of cancer* **110:** 2246-2249 (2014)

365. Hu, J., Dong, A., Fernandez-Ruiz, V., Shan, J., Kawa, M., Martinez-Anso, E., Prieto, J. & Qian, C. Blockade of Wnt signaling inhibits angiogenesis and tumor growth in hepatocellular carcinoma *Cancer research* **69:** 6951-6959 (2009)

366. Fukuhara, K., Kariya, M., Kita, M., Shime, H., Kanamori, T., Kosaka, C., Orii, A., Fujita, J. & Fujii, S. Secreted Frizzled Related Protein 1 Is Overexpressed in Uterine Leiomyomas, Associated with a High Estrogenic Environment and Unrelated to Proliferative Activity *The Journal of Clinical Endocrinology & Metabolism* **87:** 1729-1736 (2002)

367. Halder, S. K., Diamond, M. P. & Al-Hendy, A. The mediator complex subunit 12 (MED12) modulates the expression of extracellular matrix genes in human uterine fibroid cells *Fertility and sterility* **102:** e88 (2014)

368. Stamenkovic, I. Extracellular matrix remodelling: the role of matrix metalloproteinases *The Journal of pathology* **200:** 448-464 (2003)

369. Guo, H., Zhang, X., Dong, R., Liu, X., Li, Y., Lu, S., Xu, L., Wang, Y., Wang, X., Hou, D., Wei, J. J., Shao, C. & Liu, Z. Integrated analysis of long noncoding RNAs and mRNAs reveals their potential roles in the pathogenesis of uterine leiomyomas *Oncotarget* **5:** 8625-8636 (2014)

370. Huang, S., Holzel, M., Knijnenburg, T., Schlicker, A., Roepman, P., McDermott, U., Garnett, M., Grernrum, W., Sun, C., Prahallad, A., Groenendijk, F. H., Mittempergher, L., Nijkamp, W., Neefjes, J., Salazar, R., Ten Dijke, P., Uramoto, H., Tanaka, F., Beijersbergen, R. L., Wessels, L. F. & Bernards, R. MED12 controls the response to multiple cancer drugs through regulation of TGF-beta receptor signaling *Cell* **151:** 937-950 (2012)

371. Nadine Markowski, D., Tadayyon, M., Bartnitzke, S., Belge, G., Maria Helmke, B. & Bullerdiek, J. Cell cultures in uterine leiomyomas: Rapid disappearance of cells carrying MED12 mutations *Genes, Chromosomes and Cancer* **53:** 317-323 (2014)

372. Mittal, P., Shin, Y. H., Yatsenko, S. A., Castro, C. A., Surti, U. & Rajkovic, A. Med12 gain-of-function mutation causes leiomyomas and genomic instability *The Journal of clinical investigation* **125:** 3280-3284 (2015)

373. Gerlinger, M., Rowan, A. J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., Varela, I., Phillimore, B., Begum, S., McDonald, N. Q., Butler, A., Jones, D., Raine, K., Latimer, C., Santos, C. R., Nohadani, M., Eklund, A. C., Spencer-Dene, B., Clark, G., Pickering, L., Stamp, G., Gore, M., Szallasi, Z., Downward, J., Futreal, P. A. & Swanton, C. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing *The New England journal of medicine* **366:** 883-892 (2012)

374. Swanton, C. Intratumor heterogeneity: evolution through space and time *Cancer research* **72:** 4875-4882 (2012)

375. Caldas, C. Cancer sequencing unravels clonal evolution *Nature biotechnology* **30:** 408-410 (2012)

376. Tomlinson, I. P., Roylance, R. & Houlston, R. S. Two hits revisited again *Journal of medical genetics* **38:** 81-85 (2001)

377. Bar-Peled, L., Chantranupong, L., Cherniack, A. D., Chen, W. W., Ottina, K. A., Grabiner, B. C., Spear, E. D., Carter, S. L., Meyerson, M. & Sabatini, D. M. A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1 *Science* **340:** 1100-1106 (2013)

378. Miki, D., Ochi, H., Hayes, C. N., Abe, H., Yoshima, T., Aikata, H., Ikeda, K., Kumada, H., Toyota, J., Morizono, T., Tsunoda, T., Kubo, M., Nakamura, Y., Kamatani, N. & Chayama, K. Variation in the DEPDC5 locus is associated with progression to hepatocellular carcinoma in chronic hepatitis C virus carriers *Nature genetics* **43:** 797-800 (2011)

379. Ishida, S., Picard, F., Rudolf, G., Noe, E., Achaz, G., Thomas, P., Genton, P., Mundwiller, E., Wolff, M., Marescaux, C., Miles, R., Baulac, M., Hirsch, E., Leguern, E. & Baulac, S. Mutations of DEPDC5 cause autosomal dominant focal epilepsies *Nature genetics* **45:** 552-555 (2013)

380. Riva, P., Dalpra, L., Gualandri, V., Volonte, M., Miozzo, M., Malgara, R., Conti, A. F. & Larizza, L. 19p Deletion in Recurring Leiomyosarcoma Lesions from the Same Patient *Cancer genetics and cytogenetics* **119:** 102-108 (2000)

381. Pandis, N., Bardi, G., Sfikas, K., Panayotopoulos, N., Tserkezoglou, A. & Fotiou, S. Complex chromosome rearrangements involving 12q14 in two uterine leiomyomas *Cancer genetics and cytogenetics* **49:** 51-56 (1990)

382. Rausch, T., Jones, D. T., Zapatka, M., Stutz, A. M., Zichner, T., Weischenfeldt, J., Jager, N., Remke, M., Shih, D., Northcott, P. A., Pfaff, E., Tica, J., Wang, Q., Massimi, L., Witt, H., Bender, S., Pleier, S., Cin, H., Hawkins, C., Beck, C., von Deimling, A., Hans, V., Brors, B., Eils, R., Scheurlen, W., Blake, J., Benes, V., Kulozik, A. E., Witt, O., Martin, D., Zhang, C., Porat, R., Merino, D. M., Wasserman, J., Jabado, N., Fontebasso, A., Bullinger, L., Rucker, F. G., Dohner, K., Dohner, H., Koster, J., Molenaar, J. J., Versteeg, R., Kool, M., Tabori, U., Malkin, D., Korshunov, A., Taylor, M. D., Lichter, P., Pfister, S. M. & Korbel, J. O. Genome Sequencing of Pediatric Medulloblastoma Links Catastrophic DNA Rearrangements with TP53 Mutations *Cell* **148:** 59-71 (2012)

383. Ounzain, S., Micheletti, R., Arnan, C., Plaisance, I., Cecchi, D., Schroen, B., Reverter, F., Alexanian, M., Gonzales, C., Ng, S. Y., Bussotti, G., Pezzuto, I., Notredame, C., Heymans, S., Guigó, R., Johnson, R. & Pedrazzini, T. CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis *Journal of Molecular and Cellular Cardiology* **89, Part A:** 98-112 (2015)

384. Amant, F., Debiec-Rychter, M., Schoenmakers, E. F., Hagemeijer-Hausman, A. & Vergote, I. Cumulative dosage effect of a RAD51L1/HMGA2 fusion and RAD51L1 loss in a case of pseudo-Meigs' syndrome *Genes, chromosomes & cancer* **32:** 324-329 (2001)

385. Rusan, M., Li, Y. Y. & Hammerman, P. S. Genomic landscape of human papillomavirus-associated cancers *Clinical cancer research* **21:** 2009-2019 (2015)

386. Hu, Z., Zhu, D., Wang, W., Li, W., Jia, W., Zeng, X., Ding, W., Yu, L., Wang, X., Wang, L., Shen, H., Zhang, C., Liu, H., Liu, X., Zhao, Y., Fang, X., Li, S., Chen, W., Tang, T., Fu, A., Wang, Z., Chen, G., Gao, Q., Li, S., Xi, L., Wang, C., Liao, S., Ma, X., Wu, P., Li, K., Wang, S., Zhou, J., Wang, J., Xu, X., Wang, H. & Ma, D. Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism *Nature genetics* **47:** 158-163 (2015)

387. Tang, K. W., Alaei-Mahabadi, B., Samuelsson, T., Lindh, M. & Larsson, E. The landscape of viral expression and host gene fusion and adaptation in human cancer *Nature communications* **4:** 2513 (2013)

388. Schoenmakers, E. F. P. M., Bunt, J., Hermers, L., Schepens, M., Merkx, G., Janssen, B., Kersten, M., Huys, E., Pauwels, P., Debiec-Rychter, M. & van Kessel, A. G. Identification of CUX1 as the recurrent chromosomal band 7q22 target gene in human uterine leiomyoma *Genes, Chromosomes and Cancer* **52:** 11-23 (2013)

389. McNerney, M. E., Brown, C. D., Wang, X., Bartom, E. T., Karmakar, S., Bandlamudi, C., Yu, S., Ko, J., Sandall, B. P., Stricker, T., Anastasi, J., Grossman, R. L., Cunningham, J. M., Le Beau, M. M. & White, K. P. CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia *Blood* **121:** 975-983 (2013)

390. Zeng, W. R., Scherer, S. W., Koutsilieris, M., Huizenga, J. J., Filteau, F., Tsui, L. C. & Nepveu, A. Loss of heterozygosity and reduced expression of the CUTL1 gene in uterine leiomyomas *Oncogene* **14:** 2355-2365 (1997)

391. Vadnais, C., Davoudi, S., Afshin, M., Harada, R., Dudley, R., Clermont, P., Drobetsky, E. & Nepveu, A. CUX1 transcription factor is required for optimal ATM/ATR-mediated responses to DNA damage *Nucleic acids research* **40:** 4483-4495 (2012)

392. Zhou, J., Mochizuki, T., Smeets, H., Antignac, C., Laurila, P., de Paepe, A., Tryggvason, K. & Reeders, S. T. Deletion of the paired alpha 5(IV) and alpha 6(IV) collagen genes in inherited smooth muscle tumors *Science* **261:** 1167-1169 (1993)

393. Kruegel, J., Rubel, D. & Gross, O. Alport syndrome--insights from basic and clinical research *Nature Reviews Nephrology* **9:** 170-178 (2013)

394. Uliana, V., Marcocci, E., Mucciolo, M., Meloni, I., Izzi, C., Manno, C., Bruttini, M., Mari, F., Scolari, F., Renieri, A. & Salviati, L. Alport syndrome and leiomyomatosis: the first deletion extending beyond COL4A6 intron 2 *Pediatric Nephrology* **26:** 717-724 (2011)

395. Qu, B. H., Karas, M., Koval, A. & LeRoith, D. Insulin receptor substrate-4 enhances insulin-like growth factor-I-induced cell proliferation *The Journal of biological chemistry* **274:** 31179-31184 (1999)

396. Mertens, F., Moller, E., Mandahl, N., Picci, P., Perez-Atayde, A. R., Samson, I., Sciot, R. & Debiec-Rychter, M. The t(X;6) in subungual exostosis results in transcriptional deregulation of the gene for insulin receptor substrate 4 *International journal of cancer* **128:** 487-491 (2011)

397. Storlazzi, C. T., Wozniak, A., Panagopoulos, I., Sciot, R., Mandahl, N., Mertens, F. & Debiec-Rychter, M. Rearrangement of the COL12A1 and COL4A5 genes in subungual exostosis: molecular cytogenetic delineation of the tumor-specific translocation t(X;6)(q13-14;q22) *International journal of cancer* **118:** 1972-1976 (2006)

398. Karrman, K., Kjeldsen, E., Lassen, C., Isaksson, M., Davidsson, J., Andersson, A., Hasle, H., Fioretos, T. & Johansson, B. The t(X;7)(q22;q34) in paediatric T-cell acute lymphoblastic leukaemia results in overexpression of the insulin receptor substrate 4 gene through illegitimate recombination with the T-cell receptor beta locus *British journal of haematology* **144:** 546-551 (2009)

399. Van Dyck, F., Declercq, J., Braem, C. V. & Van de Ven, W. J. PLAG1, the prototype of the PLAG gene family: versatility in tumour development (review) *International journal of oncology* **30:** 765-774 (2007)

400. Bahrami, A., Perez-Ordonez, B., Dalton, J. D. & Weinreb, I. An analysis of PLAG1 and HMGA2 rearrangements in salivary duct carcinoma and examination of the role of precursor lesions *Histopathology* **63:** 250-262 (2013)

401. Deen, M., Ebrahim, S., Schloff, D. & Mohamed, A. N. A novel PLAG1-RAD51L1 gene fusion resulting from a t(8;14)(q12;q24) in a case of lipoblastoma *Cancer genetics* **206:** 233-237 (2013)

402. Klemke, M., Muller, M. H., Wosniok, W., Markowski, D. N., Nimzyk, R., Helmke, B. M. & Bullerdiek, J. Correlated expression of HMGA2 and PLAG1 in thyroid tumors, uterine leiomyomas and experimental models *PloS one* **9:** e88126 (2014)

403. Abdollahi, A. LOT1 (ZAC1/PLAGL1) and its family members: mechanisms and functions *Journal of cellular physiology* **210:** 16-25 (2007)

404. Davoli, T., Xu, A. W., Mengwasser, K. E., Sack, L. M., Yoon, J. C., Park, P. J. & Elledge, S. J. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome *Cell* **155:** 948-962 (2013)

405. Yang, Z., Cao, Y., Zhu, X., Huang, Y., Ding, Y. & Liu, X. Znhit1 causes cell cycle arrest and down-regulates CDK6 expression *Biochemical and biophysical research communications* **386:** 146-152 (2009)

406. Yu, Z. K., Gervais, J. L. & Zhang, H. Human CUL-1 associates with the SKP1/SKP2 complex and regulates p21(CIP1/WAF1) and cyclin D proteins *Proceedings of the National Academy of Sciences of the United States of America* **95:** 11324-11329 (1998)

407. Henrich, K. O., Schwab, M. & Westermann, F. 1p36 Tumor Suppression--a Matter of Dosage? *Cancer research* **72:** 6079-6088 (2012)

408. Savage, N., George, T. I. & Gotlib, J. Myeloid neoplasms associated with eosinophilia and rearrangement of PDGFRA, PDGFRB, and FGFR1: a review *International journal of laboratory hematology* **35:** 491-500 (2013)

409. Kampjarvi, K., Makinen, N., Mehine, M., Valipakka, S., Uimari, O., Pitkanen, E., Heinonen, H., Heikkinen, T., Tolvanen, J., Ahtikoski, A., Frizzell, N., Sarvilinna, N., Sjoberg, J., Butzow, R., Aaltonen, L.,A. & Vahteristo, P. MED12 mutations and FH inactivation are mutually exclusive in uterine leiomyomas *British journal of cancer* **114:** 1405-1411 (2016)

410. Campeau, P. M., Lu, J. T., Dawson, B. C., Fokkema, I. F., Robertson, S. P., Gibbs, R. A. & Lee, B. H. The KAT6B-related disorders genitopatellar syndrome and Ohdo/SBBYS syndrome have distinct clinical features reflecting distinct molecular mechanisms *Human mutation* **33:** 1520-1525 (2012)

411. Ubago, J. M., Zhang, Q., Kim, J. J., Kong, B. & Wei, J. Two Subtypes of Atypical Leiomyoma: Clinical, Histologic, and Molecular Analysis *The American Journal of Surgical Pathology* **40:** (2016)

412. Aranda-Orgilles, B., Saldaña-Meyer, R., Wang, E., Trompouki, E., Fassl, A., Lau, S., Mullenders, J., Rocha, P., Raviram, R., Guillamot, M., Sánchez-Díaz, M., Wang, K., Kayembe, C., Zhang, N., Amoasii, L., Choudhuri, A., Skok, J., Schober, M., Reinberg, D.,

Sicinski, P., Schrewe, H., Tsirigos, A., Zon, L. & Aifantis, I. MED12 Regulates HSC-Specific Enhancers Independently of Mediator Kinase Activity to Control Hematopoiesis *Cell Stem Cell* [Epub ahead of print] (2016)

413. Bardella, C., El-Bahrawy, M., Frizzell, N., Adam, J., Ternette, N., Hatipoglu, E., Howarth, K., O'Flaherty, L., Roberts, I., Turner, G., Taylor, J., Giaslakiotis, K., Macaulay, V. M., Harris, A. L., Chandra, A., Lehtonen, H. J., Launonen, V., Aaltonen, L. A., Pugh, C. W., Mihai, R., Trudgian, D., Kessler, B., Baynes, J. W., Ratcliffe, P. J., Tomlinson, I. P. & Pollard, P. J. Aberrant succination of proteins in fumarate hydratase-deficient mice and HLRCC patients is a robust biomarker of mutation status *The Journal of pathology* **225:** 4-11 (2011)

414. Fedele, M., Battista, S., Kenyon, L., Baldassarre, G., Fidanza, V., Klein-Szanto, A. J., Parlow, A. F., Visone, R., Pierantoni, G. M., Outwater, E., Santoro, M., Croce, C. M. & Fusco, A. Overexpression of the HMGA2 gene in transgenic mice leads to the onset of pituitary adenomas *Oncogene* **21:** 3190-3198 (2002)

415. Hinuma, S., Habata, Y., Fujii, R., Kawamata, Y., Hosoya, M., Fukusumi, S., Kitada, C., Masuo, Y., Asano, T., Matsumoto, H., Sekiguchi, M., Kurokawa, T., Nishimura, O., Onda, H. & Fujino, M. A prolactin-releasing peptide in the brain *Nature* **393:** 272-276 (1998)

416. Ben-Jonathan, N., Liby, K., McFarland, M. & Zinger, M. Prolactin as an autocrine/paracrine growth factor in human cancer *Trends in Endocrinology & Metabolism* **13:** 245-250 (2002)

417. Dai, N., Rapley, J., Angel, M., Yanik, M. F., Blower, M. D. & Avruch, J. mTOR phosphorylates IMP2 to promote IGF2 mRNA translation by internal ribosomal entry *Genes & development* **25:** 1159-1172 (2011)

418. Hirning-Folz, U., Wilda, M., Rippe, V., Bullerdiek, J. & Hameister, H. The expression pattern of the Hmgic gene during development *Genes, chromosomes & cancer* **23:** 350-357 (1998)

419. Hensen, K., Braem, C., Declercq, J., Van Dyck, F., Dewerchin, M., Fiette, L., Denef, C. & Van de Ven, W. J. Targeted disruption of the murine Plag1 proto-oncogene causes growth retardation and reduced fertility *Development, growth & differentiation* **46:** 459-470 (2004)

420. Yan, X., Baxter, R. C. & Firth, S. M. Involvement of pregnancy-associated plasma protein-A2 in insulin-like growth factor (IGF) binding protein-5 proteolysis during pregnancy: a potential mechanism for increasing IGF bioavailability *The Journal of clinical endocrinology and metabolism* **95:** 1412-1420 (2010)

421. Loechel, F., Fox, J. W., Murphy, G., Albrechtsen, R. & Wewer, U. M. ADAM 12-S cleaves IGFBP-3 and IGFBP-5 and is inhibited by TIMP-3 *Biochemical and biophysical research communications* **278:** 511-515 (2000)

422. Parker, A., Rees, C., Clarke, J., Busby, W. H.,Jr & Clemmons, D. R. Binding of insulin-like growth factor (IGF)-binding protein-5 to smooth-muscle cell extracellular matrix is a major determinant of the cellular response to IGF-I *Molecular biology of the cell* **9:** 2383-2392 (1998)

423. Gullu, G., Karabulut, S. & Akkiprik, M. Functional roles and clinical values of insulin-like growth factor-binding protein-5 in different types of cancers *Chinese journal of cancer* **31:** 266-280 (2012)

424. DeNicola, G. M., Karreth, F. A., Humpton, T. J., Gopinathan, A., Wei, C., Frese, K., Mangal, D., Yu, K. H., Yeo, C. J., Calhoun, E. S., Scrimieri, F., Winter, J. M., Hruban, R. H., Iacobuzio-Donahue, C., Kern, S. E., Blair, I. A. & Tuveson, D. A. Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis *Nature* **475:** 106-109 (2011)

425. Sporn, M. B. & Liby, K. T. NRF2 and cancer: the good, the bad and the importance of context *Nature Reviews Cancer* **12:** 564-571 (2012)

426. Tong, Y., Zhang, B., Fan, Y. & Lin, N. Keap1–Nrf2 pathway: A promising target towards lung cancer prevention and therapeutics *Chronic Diseases and Translational Medicine* **1:** 175-186 (2015)