Panel Sequencing in Uterine Leiomyomas identifies Missense Mutations associated with immunohistochemical Fumarate Hydratase (FH)-Loss

Bernt Popp[1*], Ramona Erber[2*], Cornelia Kraus[1], Georgia Vasileiou[1], Juliane Hoyer[1], Stefanie Burghaus[3], Arndt Hartmann[2], Matthias W. Beckmann[3], André Reis[2#], Abbas Agaimy[2#]

Running title: **Missense Mutations in FH deficient ULs**

[1]Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

[2]Institute of Pathology, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

[3]Department of Obstetrics and Gynecology, University Hospital Erlangen, Comprehensive Cancer Center ER-EMN, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Erlangen, Germany

[*, #]These authors contributed equally to this work

Correspondence to:

André Reis, Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Schwabachanlage 10, 91054 Erlangen, Germany, E-mail: andre.reis@uk-erlangen.de, phone: +49 9131 8522319, fax: +49 9131 23232

Abbas Agaimy, Institute of Pathology, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Krankenhausstrasse 8-10, 91054 Erlangen, Germany, E-mail: abbas.agaimy@uk-erlangen.de, phone: +49 9131 8522288, fax: +49 9131 8524745

ABSTRACT

Uterine leiomyomas (ULs) constitute a considerable health burden in the general female population. The fumarate hydratase (FH) deficient subtype is found in up to 1.6% of cases. Identifying these individuals is important as a subset of cases is related to germline *FH* variants in the setting of the hereditary leiomyomatosis and renal cell carcinoma (HLRCC) syndrome which is associated with aggressive renal cell cancer.

We sequenced 13 FH-deficient ULs using the TruSight Cancer Panel. In all 13 cases, we could identify biallelic *FH* variants. In eight cases, we found a *FH* point mutation (two truncating, six missense) with evidence for loss of the second allele. Variant allele-frequencies pointed to somatic variants in all cases with a point mutation. Spatial clustering of the identified missense variants in the lyase domain indicated their importance for proper fumarase oligomerization. In five tumors, we found a biallelic deletion of *FH*. Analysis of driver mutations of other genes revealed two truncating *TP53* variants in one case. Copy number (CN) analysis found a large *RB1* deletion in three tumors and enrichment for *ALK* CN gains.

Following our recommendation, one individual presented for genetic consultation. A germline *FH* variant could not be identified. Clinical examination and family history raised the suspicion of tuberous sclerosis complex, which could subsequently be confirmed due to the splice variant c.2040A>G in *TSC1*. This case highlights that presence of FH deficiency does not exclude other concurrent genetic syndromes.

By comparing the *FH* mutation carrier frequency in public databases and the prevalence of FH deficient ULs, we conclude that most of these are likely sporadic and estimate 3.5 - 14.3% of females with an FH deficient UL to carry a germline *FH* variant. Further prospective tumor/normal sequencing studies are needed to develop a reliable screening strategy for HLRCC in women with ULs.

Keywords: uterine leiomyoma, fumarate hydratase, fumarase, FH, HLRCC, panel sequencing, TSC1

INTRODUCTION

Uterine leiomyomas (ULs; fibroids) are benign smooth muscle tumors of the uterine myometrium with an estimated lifetime risk of 70% for females(1). As about 30% of women are symptomatic and present with abdominal pain, vaginal bleeding or anemia,(2, 3) ULs represent a considerable health burden(4). These hormone dependent tumors usually do not occur before adolescence, but increase in size in the reproductive period and frequently decrease in size after menopause(2). Besides these hormonal influences other factors associated with modulating the individual risk for ULs are certain dietary habits, caffeine, alcohol, smoking, components of the metabolic syndrome (central obesity, high blood pressure, hyperlipidemia) and ethnic background(5). The observation that ULs are more common in women of African origin than in Caucasian women(6) implicates genetic factors as additional risk factors.

Indeed, genome-wide association studies (GWAS) have linked several genomic regions and biological processes like mRNA degradation(7), thyroid function(8) and fatty acid synthesis(9) with UL risk. Recently, a large study(10) identified several novel loci associated with estrogen metabolism, uterine development and genetic instability. Interestingly, this study could also associate the combined polygenic risk of all loci with the most common UL subtype, positive for somatic *MED12* mutations(10).

From a molecular pathologist view, there are currently four different and mutually exclusive groups of ULs defined by their typical driver variants(11). ULs with either *MED12* hotspot driver mutations or with genomic rearrangements involving *HMGA2* are the most common subtypes and represent up to 90%(12). The two other known subtypes, defined by typical deletions in the *COL4A5*/*COL4A6* genes or *FH* deficiency, are much rarer. The *COL4A5*/*COL4A6* deletion positive UL subtype constitutes about 2%(12). We and other estimated that FH deficient UL subtype makes up 0.4% to 1.6% of all ULs(13, 14).

While the most common *MED12* mutated subtype is associated with polygenic risk and additional external factors(10), both rare subtypes are associated with certain highly penetrant heritable mendelian syndromes. Recurrent deletions at the *COL4A5*/*COL4A6* gene locus lead to the X-linked dominantly inherited diffuse leiomyomatosis with Alport syndrome (DL-ATS; OMIM #308940). Affected individuals present with esophageal leiomyomatosis and hematuria. Germline Loss-of-Function (LOF) variants (both truncating and missense) in the fumarate hydratase (fumarase) gene *FH* cause the dominantly inherited hereditary leiomyomatosis and renal cell cancer syndrome (HLRCC; OMIM #150800), which is characterized by multiple cutaneous and uterine leiomyomata and increased risk for aggressive renal tumors. Carriers of a pathogenic variant in the *FH* gene have an increased risk for offspring with the recessively inherited fumarase deficiency (FMRD; OMIM #606812),

which results in a severe encephalopathy with developmental delay and dysmorphic features. While both monogenic heritable UL subtypes are rare, their occurrence in a syndromic setting with additional health problems in affected individuals and increased risk for relatives makes a timely diagnosis particularly important.

We recently showed that FH deficient ULs can reliably be suspected based on consistent morphological features and confirmed by FH immunohistochemistry (IHC)(13). Based on this cohort, we now analyzed 13 ULs with available tumor DNA by targeted panel sequencing to investigate their somatic variant spectrum for both small genetic (single nucleotide variants: SNVs, base insertions or deletions: indels) and copy number variants (CNVs).

MATERIALS AND METHODS

Tumor samples

We included 13 FH deficient ULs with sufficient high-quality DNA from a previously described cohort of 22 prospectively diagnosed tumors collected from routine surgical pathology (n = 10) or consultation files (n = 3) of one of the authors (A.A.)(13). The histopathological characteristics of theses ULs have been reviewed by an experienced pathologist and analyzed by immunohistochemistry for FH in all 13 cases (as described previously(13)) and Retinoblastoma-1 (RB1) protein loss in 12 cases (see Supplementary notes). Detailed cohort descriptions are provided in the Supplementary data file 1 in sheet “cohort”.

Genetic counseling and germline molecular genetic testing

After the identification of FH deficiency in one resected UL, genetic counseling and molecular genetic testing for HLRCC at our Center for Rare Diseases was recommended in the pathology report. Genetic consultation included detailed medical, family history and clinical examination by a trained geneticist. Upon informed written consent, molecular genetic testing on DNA derived from peripheral blood lymphocytes was performed by PCR and Sanger sequencing of the *FH* gene with custom primers (Supplementary notes and Table S1) together with multiplex ligation probe amplification (MLPA; Kit P198, MRC-Holland). In individual S11 tuberous sclerosis was suspected clinically and targeted sequencing using the TruSight Cancer Sequencing Panel (Illumina, Inc., San Diego, USA) was additionally performed as described previously(15).

Panel sequencing of UL tumor DNA

Genomic DNA was extracted from formalin-fixed paraffin-embedded (FFPE) tumor tissue using the QIAamp DNA FFPE Tissue kit (QIAGEN, Redwood City, USA) after tumors areas were marked by a trained pathologist and manually macro-dissected. Enrichment and library preparation were performed with the TruSight Cancer Sequencing Panel v1 (Illumina, Inc., San Diego, USA) using 500ng of genomic DNA. This commercial panel includes 94 genes associated with different cancer syndromes. Besides the higher DNA input, all procedures were performed according to the manufacturers’ instructions. Libraries were sequenced with 150 bp paired-end reads on an MiSeq system (Illumina, Inc., San Diego, USA).

After demultiplexing, quality and adapter trimming was performed on the reads using cutadapt(16) version 1.10 from within the wrapper tool Trim Galore version 0.4.1. Read

alignment to the hg19 reference genome was performed with BWA-MEM(17) version 0.7.15. After removing duplicate reads with Picard tools version 2.5.0 and local realignment of indels, tumor-only variant calling was performed on the final tumor BAM files (for alignment statistics see Supplementary file 1 sheet “panel_stats”) using MuTect2(18) from GATK version 3.7-0(19) and a panel of 100 in-house germline control samples sequenced on the same platform. For annotation of the resulting variant files, SnpEff and SnpSift(20, 21) were used with dbNSFP(22) version 2.93 and variant frequencies from the ExAC database(23) version 3.1 and COSMIC database(24) version 79 based on the files provided from the respective website. The annotated variants were filtered to have a coverage (DP) of at least 10 reads and an allele frequency (AF) of at least 10%. Variants present in the ExAC(23) database ≥ 100 times were filtered out unless they were also reported in the COSMIC database ≥ 10 times or were reported as (likely) pathogenic in ClinVar(25). Only coding and splice site variants were further analyzed. Subsequently, the resulting lists were examined using the IGV browser(26) and evaluated for their biological plausibility.

CNV calling from panel data was performed on the same BAM-files used for variant calling utilizing the CNVkit(27) version 0.8.3 with standard parameters against the same 100 germline control samples used for variant calling. The CNVkit “call” was used with a purity setting of 0.8 to convert log2 ratios into integer CN-(copy number) values. Results were visualized with the “scatter” and “heatmap” functions in CNVkit.

Estimation of germline probability

To estimate whether the *FH* SNVs/indels identified from our tumor-only sequencing approach were germline or somatic variants, we first generated a plausible range of purity estimates for uterine tumors from published reports. As we could not identify large studies estimating the purity of macro-dissected ULs from sequencing data, we used purity estimates(28) of two other uterine tumor types (uterine carcinosarcoma, uterine endometrial carcinoma) from The Cancer Genome Atlas Program (TCGA)(29). We then plotted the theoretical relationship between tumor purity (TP) and expected variant allele fraction (VAF) for germline and somatic variants assuming a somatic second-hit deleting the other allele and compared this to the variant allele frequency observed in our samples (Figure S4).

Collection and computational analyses of *FH* variants

To analyze enrichment of likely disease associated *FH* variants in domains, we downloaded all described variants from the ClinVar(25), LOVD(30, 31) and COSMIC(24) databases and

scored these with InterVar(32) according to the American College of Medical Genetics and Genomics (ACMG) 5-tier classification(33). To assess the carrier frequency, we annotated all *FH* variants from the gnomAD (https://gnomad.broadinstitute.org/) and BRAVO (https://bravo.sph.umich.edu/) databases with our curated set of (likely) pathogenic variants (see Supplementary file 2 for details).

By generating all possible missense variants for *FH* and plotting different computational scores(34-36) along the linear protein representation, we analyzed domain regions intolerant to missense variation.

All variant-sets were harmonized to a common reference with VariantValidator(37) (NM_000143.3 transcript, hg19 reference genome) and annotated with the same pipeline (Supplementary notes).

Protein structure analysis of identified *FH* missense variants

We visualized the spatial clustering of identified *FH* missense variants in 3D using the publicly available tertiary protein structure data of human fumarase (PDB-ID: 5D6B)(38) with the Pymol molecular visualization software (Version 1.8.6.0; Schrödinger LLC, New York, USA) installed through Conda (Anaconda Inc., Austin, USA).

To estimate the probability of the observed spatial clustering, we employed the online version of mutation3D(39), which uses a bootstrapping approach to estimate an empiric p-value, with the 5D6B protein structure as template. As a baseline, we compared the 3D clustering analysis of the herein identified six somatic *FH* missense variants to all (likely) pathogenic missense variants from the curated ClinVar and LOVD datasets (Figure S3).

RESULTS

Study cohort

The median age at diagnosis in the 13 females included was 37 years (y) with a range of 25y to 72y. Seven individuals were treated by hysterectomy and five by enucleation (no data for one case). Eight individuals had more than one UL nodule (no data for one case). Three individuals had a personal history of tumors/cancer of other organ systems including thyroid adenoma (S06), colorectal adenocarcinoma and endometrioid carcinoma of the uterus (S07) and breast cancer (S12). For three individuals, a family history of tumors/cancer in first degree relatives was reported to the pathologist: lung carcinoma at age 56y in the mother of S05, chronic leukemia at age 56y in the father and adenoma of the thyroid at age 30y in the mother of S06 and colorectal adenocarcinoma at age 69y in the sister and gastric carcinoma at age 84y in the father of S07.

FH loss and additionally reduced RB1 expression in IHC

All 13 ULs, selected for panel sequencing in this study, showed a complete loss of FH staining by IHC as reported previously(13). RB1 IHC showed a reduced expression in all 12 (100%) analyzed FH deficient ULs with no sample showing a complete loss. For one sample (S02), no RB1 IHC was performed.

Tuberous sclerosis in individual S11

Though recommended to all 10 routine cases, only one individual presented for genetic consultation at our center.

At age 34 years, she had a symptomatic fibroid of the uterus (causing severe dysmenorrhea) which increased in size for about seven years. The posterior wall myoma was surgically removed followed by reconstruction of the uterus. Histopathological examination revealed atypical leiomyoma and IHC showed deficiency of FH expression. Perivascular epitheloid cell tumor (PEComa) was excluded by presence of histological features characteristic of FH deficient leiomyoma and by lack of melanocytic marker expression. An examination of the kidneys by means of ultrasound as well as the other gynecological examination was inconspicuous. A specialist dermatological investigation was not done, but melanocytic skin nevi were removed by the general practitioner and the histopathological examination of these was unremarkable. At the age of 18 years, she had a seizure disorder which had been treated with lamotrigine and resulted in a seizure free interval for three years. A brain MRI

conducted because of dizziness at age 33 years raised the suspicion of multiple sclerosis-like unclassified lesions. Examination of the skin revealed multiple small papules on the forehead, and a larger (2 x 2.5mm) similar change on the scalp. She had one large 8cm hypomelanotic spot on the left lower back and a smaller one on the right upper arm. Iris abnormalities or signs of nail fibromas were not noted. Family history showed that her mother had a double-walled uterus, her maternal aunt had only one kidney, her maternal grandmother had colorectal cancer at age of 88 years and this grandmother's mother died of renal cancer at about 60 years of age. Her 27-year-old half-brother also had epilepsy. Clinical examination and family history thus raised the suspicion of tuberous sclerosis complex.

Molecular genetic testing by means of Sanger sequencing and MLPA identified no pathogenic germline variant in the *FH* gene. Also, panel-based sequencing of other genes associated with increased renal cancer risk (*SDHB*, *SDHC*, *SDHD*, *SDHAF2*, *PTEN*) was unremarkable. Extended panel analysis identified the synonymous variant c.2040A>G affecting the penultimate position of exon 16 of *TSC1*. Subsequent RT-PCR and Sanger sequencing showed aberrant RNA splicing and confirmed the suspected diagnosis. Tumor panel sequencing in this case showed biallelic loss of *FH* without evidence for a second-hit in the *TSC1* gene.

Properties of identified *FH* SNV/indels

Tumor-only variant calling for small genetic variants using a somatic variant caller identified a single SNV/indel in the *FH* gene in 8/13 (61.5%) of ULs. In no sample, we identified two SNVs/indels. Two of the identified variants (25.0%) were annotated as likely gene disrupting. The variant c.457delG in individual S06 causes a frame-shift which directly introduces a termination codon (p.(Val153*)). The variant c.379-2A>G in S03 disrupts the conserved splice-acceptor of exon 4, is predicted to cause aberrant splicing (r.spl?) by different computational algorithms and has been described to reduce FH activity(40).

The remaining six variants were annotated as missense variants causing the substitution of different single amino acids. Five of the missense variants were predicted as disrupting protein function throughout all eight computational missense prediction scores. Only the variant c.1236G>A, p.(Met412Ile) identified in S09 was predicted as benign by five out of eight scores. Instead, this variant is predicted to cause aberrant splicing by four of the five splice prediction tools used (Supplementary file 1 sheet "FH_SNVindel_summary"). Indeed the c.1236G>A variant changes the last coding guanin base of exon 8 into an adenine. This base position is typically highly conserved in the consensus splice-site sequence(41). Furthermore, c.1236G>A is the only variant annotated as missense which does not affect a

conserved protein domain like the Lyase domain. This domain contains the other five missense variants and most of the (likely) pathogenic missense variants reported in databases (Figure 1A). Therefore, it seems likely that the c.1236G>A variant in fact disrupts normal splicing (r.spl?) and should be regarded as a likely gene disrupting variant (p.0?). However, further studies on RNA are needed to confirm the exact consequence of this variant. For the other five missense variants, computational splice-effect prediction scores where either unremarkable or not available, which points to the conclusion that these are indeed missense variants.

As missense variants are not expected to cause protein loss, but all 13 ULs were preselected based on complete FH-loss in IHC, the identification of true missense variants in 5/13 (38.5%) of ULs was unanticipated. When analyzing the proximity of the missense variants in the tertiary fumarase protein structure, we observed that the variants lie very close to each other (Figure 1B). We therefore performed 3D clustering analyses using mutation3D, which showed that the three variants identified in samples S01 (c.944T>C, p.(Leu315Pro)), S07 (c.824G>C, p.(Gly275Ala)) and S10 (c.817G>A, p.(Ala273Thr)) form a protein-wide significant cluster (p-value: 0.0411, empirical bootstrapping approach) within the fumarase protein (Figure S3).

While the genomic position of seven SNV/indel variants showed high read coverage allowing reliable estimation of VAF, only the variant c.824G>C, p.(Gly275Ala) in individual S07 was covered relatively low with 19 reads. All eight variants showed high VAF between 0.443 and 0.793 (median 0.766), indicating a CN-loss of the second allele as second-hit.

Somatic variant status in FH deficient ULs

As SNV/indel calling only identified one *FH*-variant in eight of the 13 ULs, we next performed CN-calling from the capture-based sequencing data using the CNVkit algorithm to search for second-hits. Assuming a tumor purity of 80%, we were able to identify a CN-loss in all 13 ULs studied. The *FH* CN-losses ranged in size between 15.9 and 42,757.4 kilo-bases (kb) with a median of 4,897.1 kb (Figure 2A; Table 1; Supplementary file 2 sheet “FH_CNVkit_summary”). In the five ULs with no SNV/indel identified, the log2 ratios and CN-values indicate biallelic deletions. Due to the relatively low resolution of the panel (when compared to chromosomal microarrays), the likely independent two deletions in theses sample have often been called with the same breakpoints. In sample S12, the CNVkit algorithm indeed called a small 15.9 kb deletion affecting exons 3 to 10 of the *FH* gene and a large 1407.7 kb deletion in the chromosomal band 1q43 affecting the whole *FH* gene (Figure 2B and C). For seven of the eight ULs with a SNV/indel variant identified, the CN-analysis

indicated the deletion of a single allele, confirming the suspicion from the observed high VAF. In UL sample S05, with the c.587A>T, p.(His196Leu) variant at a relatively low VAF of 0.443, a 15.9 kb small deletion was called. This deletion had an exceptionally low (more negative than the monoallelic deletions and less negative than the biallelic deletions) log2 value of -1.87 when compared to other samples with one deletion only. Together with the low VAF of the SNV in this sample, this observation points to the explanation of tumor-heterogeneity in this UL with multiple second-hit events occurring after an initial deletion of one *FH* allele.

The analysis of other genes covered by the panel identified the two pathogenic likely gene disrupting variants c.782+1G>T, p.0? and c.309C>A, p.(Tyr103*) in the *TP53* gene in UL sample S04 with VAFs (40.8% and 37.0%, respectively) pointing to somatic variants. No other likely somatic SNVs/indels were identified in the other 12 samples. The only other pathogenic variant identified was the c.1001C>A, p.(Pro334His) in the *XPC* gene in UL sample S07, which is a known pathogenic variant (dbSNP rs74737358) causing the recessively inherited xeroderma pigmentosum (OMIM # 278720) when inherited with a second pathogenic variant on the other allele. The high allele frequency of this variant in the general population (854/278,442 ~ 0.31% in gnomAD v2.1.1) together with its VAF of 56.0% indicated that this is very likely a germline variant.

In contrast to the relatively stable situation for somatic SNVs/indels, all ULs showed several CN-aberrations especially on chromosome 1 (Figure 2A). Unsurprisingly, the *FH*-gene showed significantly more CN-losses in the 13 ULs, but no other gene reached panel-wide significance when correcting for multiple testing. Interestingly, the *RB1* gene locus indicated a deletion ≥ 500 kb in only three FH deficient ULs (two additional when considering CNVs < 500 kb), despite the reduced expression in IHC in all 12 samples analyzed. In regards of CN-gains, we found a panel wide significance only for the *ALK*-gene locus (Figure S5).

Both the *FH* and *TERT* genes encode for moonlighting proteins (fumarase and telomerase) performing functions in the mitochondria and the cell nucleus and associated with cellular aging and tumorigenesis(42, 43). Therefore, we also investigated the telomere content (TC) and mitochondrial gene dosage (MGD) in ULs. While the 13 ULs showed significantly higher TC and MGD, when compared to germline controls (from peripheral blood lymphocytes), the difference was not significant for both analyses when compared with 15 in-house hereditary breast and ovarian cancer (HBOC)- associated tumors (Figure S7).

Pathogenic *FH* variants and carrier frequency in databases

By collecting described *FH* variants from the ClinVar and LOVD databases and classifying them according to ACMG criteria, we summarized 280 unique (likely) pathogenic variants. These included 130 protein truncating variants (46.4%), 36 variants affecting the splice regions (12.9%), 104 missense variants (37.1%) and 10 in-frame indel (3.6%) variants (Figure S2, Supplementary file 2). Truncating variants were distributed throughout the protein, while missense variants were enriched in the Lyase domain (Figure 1A middle panel, Figure S2).

By summarizing these 280 curated (likely) pathogenic *FH* variants in public databases, we estimated a carrier frequency of 1 in 2.563 (0.0390%) to 1 in 3.247 (0.0308%) individuals for the BRAVO and gnomAD databases, respectively (Table S2, Table S3 and Supplementary file 2).

DISCUSSION

Identifying individuals who carry a pathogenic variant in the *FH* gene is considered important as they have an increased lifetime risk of about 15%(44) for the aggressive renal cell carcinoma (RCC) type associated with HLRCC. A timely diagnosis can enable clinical surveillance for RCC and thus, may benefit these individuals and their families.(44) Due to the distinctive histomorphological appearance of the FH deficiency-driven neoplasms of uterus and kidney, pathologists play a central role in their initial recognition and hence identification of patients with increased risk for such hereditary tumor syndromes. Nevertheless, they are often blinded for the patients detailed medical and family history. Cutaneous leiomyomas, especially if multiple, are considered pathognomonic and FH-deficient RCC in young adults before age 50 years also raises a strong suspicion for HLRCC. Due to the high lifetime risk for ULs in women and the associated frequent need for surgical procedures as treatment, the FH deficient UL subtype is frequently encountered in the pathologist routine despite constituting only up to 1.3% of all ULs. However, there is currently no consensus approach to identify the subgroup among patients with ULs who carry a germline *FH*-variant.

Different screening approaches based on morphological or immunohistochemical methods have been recently proposed and tested(45, 46). Those recent studies highlighted the value of routine morphological assessment as strong screening tool assisted by adjunct IHC in the initial recognition of FH-related ULs. Further investigating our previously reported cohort(13), we now characterized the somatic variant status in 13 ULs with distinctive histomorphological features and confirmed FH loss by IHC. By using an established capture-based panel sequencing approach together with comprehensive bioinformatic analyses for SNVs/indels and CNVs, we could identify biallelic variants in all analyzed cases. In eight cases, we identified a SNV/indel together with a CN-loss on the second allele, while the remaining five cases showed biallelic CN-losses. The observation that no UL had two SNVs/indels, points to an initial *FH* mutation (SNV/indel or CNV) in a progenitor cell which then increased the probability for a CN-loss in descendent cells. Indeed, fumarase is known to play a role in response to double strand-breaks (DSB) by activating the non-homologous end joining (NHEJ) mechanism(47). When the NHEJ pathway is inactivated the DSB are repaired by more error prone mechanisms, like microhomology-mediated end joining, which can lead to deletions. The proportion of SNVs/indels detected in our study is comparable to previous reports(14, 45, 48, 49). However, the authors of these publications could often not identify CN-loss of the second allele since the CN-analysis and the sequencing technique used, did not allow to reliably estimate the allele frequency. Interestingly, Joseph and colleagues used a similar approach to identify a biallelic CN-loss in one FH deficient UL which had escaped

Sanger based mutation detection before(45). The detection/identification of 5/13 ULs with biallelic CN-loss and 8/13 monoallelic CN-loss in this study, suggests CN-loss as a predominant mutational mechanism which has been underestimated, but can be reliably detected using panel sequencing. The combination of CN- and VAF-analyses not only allowed us to detect both *FH* mutations in all 13 tumors, but also raised the suspicion of intratumoral heterogeneity in one case (S05), a mechanism only recently proposed to further complicate detection of the cause of FH-associated ULs(46).

Notably, our series of ULs had been preselected based on complete FH-loss in IHC. While biallelic CN-losses or the combination of a CN-loss of one allele and a truncating variant on the other allele (S03: splice-acceptor variant, S06: frameshifting variant) are expected to result in the lack of protein product, missense variants usually do not cause reduced protein dosage. Rabban and colleagues indeed reported normal FH immunoexpression in tumors of two females with pathogenic germline *FH* missense variants(46). However, we observed no difference in IHC and morphology for ULs with biallelic deletions, likely gene disrupting and missense variants in *FH* (Figure 3). The observation of six variants annotated as missense was thus unexpected, also as the majority of (likely) pathogenic mutations (59.3%) in databases are either truncating or affect the splice-sites (Figure S2) and identifying six missense variants out of 8 total variants is unlikely (Binomial p-value 0.045). A possible explanation would be that the six missense variants we identified in fact pose a different effect on the gene product by, for example, altering the mRNA splicing as has been described for other exonic variants(50). Indeed, the predicted missense variant c.1236G>A, identified in individual S09, affects the last base of exon 8 and is predicted to result in aberrant splicing. For the other five variants though, algorithms indicated normal splicing behavior, leaving another mechanism to explain loss of protein. We therefore utilized a published crystal structure to perform a three-dimensional variant clustering, which showed that the distribution of missense variants in our samples are unlikely by chance (Figure 1B; Figure S3). Hence, it seemed reasonable that the protein domain was important for proper FH expression. Indeed, a literature search showed that other missense variants (A308T and H318Y) affecting the same domain result in defective fumarase oligomerization(51) and we could show that these two variants significantly colocalize with the missense variants identified herein (Figure S3B). These results are particularly interesting as the combination of classical pathology techniques with molecular genetics and bioinformatic analyses pinpointed the specific function of certain variants in a protein domain. This indicates that larger data collections and systematic evaluation will yield novel hypothesis and insights, which can then be evaluated by functional studies. Nevertheless, our finding of a specific missense cluster causing loss of FH expression also indirectly confirms the concerned observation that other missense variants will be missed by screening methods based on FH-IHC.

Despite the successful complete characterization of the dual FH-hits in all 13 ULs, the question remained whether the identified aberrations are somatic, indicating sporadic disease, or in fact are germline and potentially associated with HLRCC. The ethics consent for this study did not allow us to test adjacent normal tissue or to recontact the patients regarding germline testing. Instead, we estimated the probability of a germline event based on the VAF in the eight ULs sequenced with an SNV/indel. Assuming the typically relatively high purity of a usually monoclonal and non-invasive tumor like ULs, the observed VAF between 0.443 and 0.793 is likely best explained by a somatic SNV/indel with a second somatic CN-loss on the other allele. The alternative hypothesis of a germline variant would require a tumor purity well below 57.7%. As we could not directly estimate tumor purity from the sequencing data due to the low mutational load, we cannot fully exclude that some of the identified eight SNVs/indels and maybe also CNVs in the remaining five cases are indeed germline variants. Our reasoning is further supported by the high expected prevalence of FH deficient ULs (UL prevalence of up to 70% (1) x proportion of FH deficient ULs (13, 14) between 0.4 to 1.6% ≈ 0.3% to 1.1%) but relatively low estimated carrier frequencies of up to 1/2.563 for (likely) pathogenic germline variants in public databases. Thus, only 1 in 7 (14.3%) to 29 (3.5%) females with an FH deficient UL is expected to carry a germline *FH* variant. In this regard, the screening strategy proposed by Rabban and colleagues, based on FH deficient UL morphology (staghorn-shape or thin/curved blood vessels, cells with perinucleolar halos or vesicular chromatin, cytoplasmic eosinophilic globules, alveolar pattern edema), is especially interesting as it allowed identification of a germline variant in 5/2.060 individuals(46) which equates to a 6.2x enrichment of germline carriers.

Detection of a *TSC1* germline mutation in a patient with FH deficient UL in this study is novel. This observation on the one hand highlights the diagnostic strength of genetic counseling combined with broad panel testing in young tumor patients, enabling the identification of rare monogenic tumor syndromes. On the other hand, the association of FH deficient ULs and this *TSC1* variant is intriguing as both fumarase and the tuberous sclerosis proteins hamartin (*TSC1*) and tuberin (*TSC2*) influence HIF1α signaling(52) and heterozygous *TSC1* knock-out mice develop uterine leiomyoma and leiomyosarcoma with subsequent loss of the second allele(53). Perivascular epithelioid cell tumors (PEComas) represent the main uterine manifestation the tuberous sclerosis complex (TSC). These neoplasms overlap strongly with uterine smooth muscle tumors but can be separated by their additional distinctive morphological and immunophenotypic features. In this individual S11, the histology of the tumor (with characteristic features of FH deficient UL; Figure 4) and lack of melanocytic marker expression precluded a diagnosis of PEComa and confirmed instead UL with FH deficiency. Also, while we could confirm aberrant splicing of the identified *TSC1* germline variant and thus its pathogenicity, the VAF in the sequenced UL from the individual showed

no significant deviation from 50% and did therefore not further confirm a causal relationship. Currently, no data is available about FH status in uterine PEComa. Notably, a recent report described renal PEComa in a patient with proven HLRCC syndrome and the PEComa was initially judged as a potential RCC on screening of the kidney(54). FH expression was retained in that PEComa indicating an alternative pathogenesis. However, taken together, the current and that previous report suggest a possibility of interaction between the *FH* and *TSC1* gene mutations in rare cases and it remains to be clarified if these two genes might on occasion represent alternate mechanisms in the pathogenesis of rare smooth muscle neoplasms.

Panel sequencing further allowed us to identify two somatic likely gene disrupting mutations in *TP53* in sample S04, which is comparable to a homozygous *TP53* deletion described in a UL sample with FH-loss in IHC and a missense mutation with loss-of heterozygosity(48). While IHC showed a reduced expression of RB1 in all samples analyzed, CN-analysis revealed a likely heterozygous loss of the *RB1* gene locus in 25% of the examined ULs. This finding is comparable to the results of Bennett and colleagues who identified homozygous *RB1*-losses in 40% of ULs with normal FH staining but not in FH deficient ULs(48). Additionally, our CN-analysis identified an enrichment of CN-gains at *ALK*-gene locus previously not reported in ULs. This observation is interesting as it could offer novel treatment options with *ALK*-inhibitors(55) but needs further confirmation. While our analysis of mitochondrial genome dosage and telomere content from panel sequencing data did not identify significant differences compared to other tumors, it further confirms the added value of next-generation based sequencing to investigate novel hypotheses from available data.

Finally, further systematic sequencing analyses like the initial studies of Mehine and colleagues(11, 56) on larger cohorts of ULs will be required to fully define the mechanism involved in the development of these benign uterine tumors causing significant morbidity in the female population and to investigate associations with heritable tumor syndromes like the HLRCC. We anticipate that this will first require unbiased tumor/normal sequencing (panels, exomes or genomes) but also RNAseq and methylation studies in a representative cohort. Future interesting fields of investigation would be individuals with multiple ULs who do not have a germline variant, where sequencing of multiple tumors from the same individual might uncover somatic mosaicism.

In conclusion, the combination of IHC screening and panel sequencing with detailed bioinformatic analyses allowed the identification of both genetic hits in all the 13 ULs studied, confirming the established Knudson hypothesis in *FH*-related tumor development and the role of *FH* as a tumor suppressor gene. This successful approach allowed us to identify a cluster of missense variants associated with immunohistochemical protein reduction, proving

that missense variants contribute to FH deficient ULs. We agree with previous concerns(45, 46) that some pathogenic missense variants in individuals with HLRCC might be missed using only FH-IHC as screening method. While screening of certain morphologic features in tumors has been shown to enrich for patients with HLRCC(46), the statistical performance of this approach can currently not be assessed reliably. Thus, a prospective tumor/normal sequencing study, which represent the current gold-standard(45), in a representative risk group is needed. Based on literature recommendations(44, 57) and our experience with heritable tumor syndromes, a possible strategy would be to offer this tumor/normal screening to every symptomatic woman below age of 40 years who has multiple ULs or one UL ≥ 10cm in diameter without prior selection based on morphological features. Further inclusion criteria could be a positive personal or family history for tumors, if this information is available to the pathologist. The use of a relatively small but standardized commercial gene panel for screening, like in this study, can reduce the associated costs and allow inter-institutional data collection and collaboration between pathologists and geneticists.

FIGURES AND TABLES

Figure 1 | Somatic SNV/indel spectrum variant properties

(A) Upper panel: Schematic representation of the FH protein, domains (ticks below x-axis numbered after NP_000134.2 and P07954) and localization of herein identified variants. Likely gene disrupting variants are presented in black and missense variants in red. TransPep, mitochondrion transit peptide; Lyase, N-terminal fumarate lyase domain; small light blue boxes in Lyase domain: substrate binding site A and B; FumC-C, C-terminal fumarase C domain. Middle panel: Density blot of (likely) pathogenic truncating (red) and missense (blue) variants reported in ClinVar(25). Lower panel: Generalized linear models of the CADD(34) scores for all possible *FH* missense variants. (B) Structural analysis of human fumarase (FH, dark grey) based on the crystal structure with PDB code 5D6B(38). Green, Lyase domain; pale-green, FumC-C domain; red, amino acid residues for the mutations L315P, H196P, G275A, M412I, A273T, H196L lying in the Lyase domain or close to it. The missense mutations in the Lyase domain affect highly conserved amino acid residues which likely disrupt the protein structure (see also Figure S02 and S03). One letter amino acid code wwas used due to space constraint.

Figure 2 | Somatic copy number aberrations

(A) Heat-plot of copy number (CN) aberrations detected in the 13 analyzed ULs using panel data. Upper panel: all chromosomes. Lower panel: zoomed in chromosome 1 containing the *FH* gene. Blue: CN loss, red: CN gain, CR: copy ratio. The position of the *FH* gene is indicated by the red stroke in the lower panel. Note the dark blue squares at the locus indicating samples with biallelic CN-loss in the tumor. In contrast to the relatively stable situation for SNVs/indels, all ULs show several larger CNVs with different recurrently affected genomic loci. When comparing the CNV-spectrum of the FH deficient ULs to 15 in-house HBOC tumors and correcting for multiple testing, only the *FH*-gene remained significant for CN-losses and the *ALK*-gene for CN-gains (see also Figure S5). Note that the 15.9 kb small deletion on one allele S05 is merely identifiable at this resolution (compare Figure S6 CR-profile of this sample at the *FH*-gene locus). (B) Exemplary CR profile for sample S12 at chromosome 1 showing different rearrangements and especially the *FH*-gene locus affected by both a larger CN-loss and a second smaller one (marked by red ellipses and black arrows). (C) Zoomed in CR profile for sample S12 at the *FH*-gene locus with VAF (variant allele frequency) plot (lower panel). Grey dots represent markers used by CNVkit(27) (target and anti-target regions) and shading the dots indicates weight within the analysis. Vertical yellow bars mark gene regions. Horizontal orange bars represent CN segmentation calls.

Figure 3 | Exemplary histopathology and FH IHC

Exemplary FH IHC staining and morphologic features for ULs with biallelic deletions, likely gene disrupting and missense variants. (A, B) S12 called with both a 15.9 kb and a 1407.7 kb deletion. (C, D) S06 with the LGF variant c.457delG, p.(Val153*) and a large 42,557.9 kb deletion. (E, F) S05 with the missense variant c.587A>T, p.(His196Leu) and a small 15.9 kb deletion. Note the similar immunohistochemical FH loss and morphology in all three samples despite the diverse mutations.

Figure 4 | Histopathology and IHC of case S11

Exemplary morphologic features and IHC staining of an UL from individual S11 with the germline with *TSC1* variant c.2040A>G. (A)-(C) S11: Histomorphology of FH deficient UL, H&E (note the hyaline globular bodies in (B) and perinucleolar halos in (C); (D) S11 with immunohistochemical FH loss (FH IHC).

ID	*Age (y)*	*FH IHC*	*Rb1 IHC*	*Nodules*	*FH variants (NM_000143.3)*	*AF FH variants*	*Other (likely) pathogenic SNV/indel*	*FH CNVs (hg19)*	*log2 CNVs*
S01	37	Loss	Reduced	Multiple	c.944T>C, p.(Leu315Pro)	137/178 (0.767)	-	chr1:g.208278362_244894272x1	-0.77
S02	43	Loss	n.d.	Multiple	c.587A>C, p.(His196Pro)	108/140 (0.771)	-	chr1:g.221246882_247101663x1	-0.85
S03	33	Loss	Reduced	5	c.379-2A>G, r.spl?, p.0?	110/141 (0.780)	-	chr1:g.237286185_242686882x1	-1.26
S04	72	Loss	Reduced	1	n.d.	n.a.	*TP53*: c.782+1G>T, p.0?, likely somatic; c.309C>A, p.(Tyr103*), likely somatic	chr1:g.237694880_243690241x0	-9.97
S05	25	Loss	Reduced	14	c.587A>T, p.(His196Leu)	74/167 (0.443)	-	chr1:g.241661127_241677014x1	-1.87
S06	35	Loss	Reduced	1	c.457delG, p.(Val153*)	153/193 (0.793)	-	chr1:g.206682236_249240121x1	-1.02
S07	70	Loss	Reduced	Multiple	c.824G>C, p.(Gly275Ala)	10/19 (0.526)	*XPC*: c.1001C>A, p.(Pro334His); ACMG 5, likely germline	chr1:g.206482721_249240121x1	-0.87
S09	36	Loss	Reduced	1	c.1236G>A, r.(spl?), p.[(Met412Ile);0?]	104/136 (0.765)	-	chr1:g.208078846_249240121x1	-1.08
S10	27	Loss	Reduced	n.a.	c.817G>A, p.(Ala273Thr)	93/154 (0.604)	-	chr1:g.237490532_241884194x1	-0.9
S11	33	Loss	Reduced	2	n.d.	n.a.	*TSC1*: c.2040A>G, p.[Pro668Serfs*5, Leu667Phefs*1]; ACMG 5, confirmed germline	chr1:g.240082627_242686882x0	-5.88
S12	54	Loss	Reduced	Multiple	n.d.	n.a.	-	chr1:g.241661127_241677014x0; chr1:g.241680481_243088226x0	-9.97; -5.35
S14	47	Loss	Reduced	1	n.d.	n.a.	-	chr1:g.241266127_244292257x0	-5.31
S16	49	Loss	Reduced	Multiple	n.d.	n.a.	-	chr1:g.238701877_242084866x0	-9.97

Table 1 | Overview of analyzed cases and identified somatic variant status

n.d. = none determined; n.a. = not applicable

SUPPLEMENTARY

Supplementary notes | Six additional figures, three additional tables and supplementary methods detailing the results and proving further information about the analyses performed.

Supplementary File S1 | Excel file containing the worksheets “summary”, "cases", "panel_stats", "FH_SNVindel_summary", "SNVindel_all", "FH_CNVkit_summary", "CNVkit_aberrations", “mitochondria” and “telomere”. The “summary” worksheet contains a detailed description of all other worksheets and the respective data columns.

Supplementary File S2 | Excel file containing the worksheets “summary”, "FH_domains", "allFH_missense", "FH_ClinVar", "FH_LOVD", "FH_COSMIC", "FH_gnomAD", and “FH_BRAVO”. The “summary” worksheet contains a detailed description of all other worksheets and the respective data columns.

ACKNOWLEDGEMENTS

The authors would like to thank Rudolph Jung, Christa Winkelmann, Andrea Eberwein and Leonora Klassen for excellent technical assistance.

COMPLIANCE WITH ETHICAL STANDARDS

The authors confirm that the study has been performed in accordance with accepted principles of ethical and professional conduct for biomedical scientific research. This study is covered by the ethical vote for retrospective translational research studies of the Ethical Committee of the Medical Faculty of the Friedrich-Alexander-Universität Erlangen-Nürnberg.

CONFLICT OF INTEREST

The authors have no conflict of interest to disclose.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website.

REFERENCES

1. Zimmermann A, Bernuit D, Gerlinger C, Schaefers M, Geppert K. Prevalence, symptoms and management of uterine fibroids: an international internet-based survey of 21,746 women. BMC Womens Health 2012;12:6.

2. Williams ARW. Uterine fibroids - what's new? F1000Res 2017;6:2109.

3. Moroni RM, Vieira CS, Ferriani RA, Reis RM, Nogueira AA, Brito LG. Presentation and treatment of uterine leiomyoma in adolescence: a systematic review. BMC Womens Health 2015;15:4.

4. Cardozo ER, Clark AD, Banks NK, Henne MB, Stegmann BJ, Segars JH. The estimated annual cost of uterine leiomyomata in the United States. Am J Obstet Gynecol 2012;206:211 e1-9.

5. Pavone D, Clemenza S, Sorbi F, Fambrini M, Petraglia F. Epidemiology and Risk Factors of Uterine Fibroids. Best Pract Res Clin Obstet Gynaecol 2018;46:3-11.

6. Marshall LM, Spiegelman D, Barbieri RL, Goldman MB, Manson JE, Colditz GA*, et al.* Variation in the incidence of uterine leiomyoma among premenopausal women by age and race. Obstet Gynecol 1997;90:967-73.

7. Cha PC, Takahashi A, Hosono N, Low SK, Kamatani N, Kubo M*, et al.* A genome-wide association study identifies three loci associated with susceptibility to uterine fibroids. Nat Genet 2011;43:447-50.

8. Hellwege JN, Jeff JM, Wise LA, Gallagher CS, Wellons M, Hartmann KE*, et al.* A multi-stage genome-wide association study of uterine fibroids in African Americans. Hum Genet 2017;136:1363-73.

9. Eggert SL, Huyck KL, Somasundaram P, Kavalla R, Stewart EA, Lu AT*, et al.* Genome-wide linkage and association analyses implicate FASN in predisposition to Uterine Leiomyomata. Am J Hum Genet 2012;91:621-8.

10. Valimaki N, Kuisma H, Pasanen A, Heikinheimo O, Sjoberg J, Butzow R*, et al.* Genetic predisposition to uterine leiomyoma is determined by loci for genitourinary development and genome stability. Elife 2018;7.

11. Mehine M, Kaasinen E, Heinonen HR, Makinen N, Kampjarvi K, Sarvilinna N*, et al.* Integrated data analysis reveals uterine leiomyoma subtypes with distinct driver pathways and biomarkers. Proc Natl Acad Sci U S A 2016;113:1315-20.

12. Mehine M, Makinen N, Heinonen HR, Aaltonen LA, Vahteristo P. Genomics of uterine leiomyomas: insights from high-throughput sequencing. Fertil Steril 2014;102:621-9.

13. Siegler L, Erber R, Burghaus S, Brodkorb T, Wachter D, Wilkinson N*, et al.* Fumarate hydratase (FH) deficiency in uterine leiomyomas: recognition by histological features versus blind immunoscreening. Virchows Arch 2018;472:789-96.

14. Miettinen M, Felisiak-Golabek A, Wasag B, Chmara M, Wang Z, Butzow R*, et al.* Fumarase-deficient Uterine Leiomyomas: An Immunohistochemical, Molecular Genetic, and Clinicopathologic Study of 86 Cases. Am J Surg Pathol 2016;40:1661-9.

15. Kraus C, Hoyer J, Vasileiou G, Wunderle M, Lux MP, Fasching PA*, et al.* Gene panel sequencing in familial breast/ovarian cancer patients identifies multiple novel mutations also in genes others than BRCA1/2. Int J Cancer 2017;140:95-102.

16. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal; Vol 17, No 1: Next Generation Sequencing Data AnalysisDO - 1014806/ej171200 2011.

17. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:13033997 2013.

18. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C*, et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nature biotechnology 2013;31:213-9.

19. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A*, et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 2010;20:1297-303.

20. Cingolani P, Patel VM, Coon M, Nguyen T. Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. Frontiers in … 2012.

21. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L*, et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012;6:80-92.

22. Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. Hum Mutat 2013;34:E2393-402.

23. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T*, et al.* Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285-91.

24. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H*, et al.* COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic acids research 2015;43:D805-11.

25. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM*, et al.* ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic acids research 2014;42:D980-5.

26. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G*, et al.* Integrative genomics viewer. Nature biotechnology 2011;29:24-6.

27. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol 2016;12:e1004873.

28. Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. Nat Commun 2015;6:8971.

29. Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA*, et al.* The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013;45:1113-20.

30. Bayley JP, Launonen V, Tomlinson IP. The FH mutation database: an online database of fumarate hydratase mutations involved in the MCUL (HLRCC) tumor syndrome and congenital fumarase deficiency. BMC Med Genet 2008;9:20.

31. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 2011;32:557-63.

32. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. Am J Hum Genet 2017;100:267-80.

33. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J*, et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-24.

34. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic acids research 2019;47:D886-D94.

35. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S*, et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet 2016;99:877-85.

36. Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN*, et al.* M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nat Genet 2016;48:1581-6.

37. Freeman PJ, Hart RK, Gretton LJ, Brookes AJ, Dalgleish R. VariantValidator: Accurate validation, mapping, and formatting of sequence variation descriptions. Hum Mutat 2018;39:61-8.

38. Pereira de Padua RA, Nonato MC. Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of recombinant human fumarase. Acta Crystallogr F Struct Biol Commun 2014;70:120-2.

39. Meyer MJ, Lapcevic R, Romero AE, Yoon M, Das J, Beltran JF*, et al.* mutation3D: Cancer Gene Prediction Through Atomic Clustering of Coding Variants in the Structural Proteome. Hum Mutat 2016;37:447-56.

40. Muller M, Ferlicot S, Guillaud-Bataille M, Le Teuff G, Genestie C, Deveaux S*, et al.* Reassessing the clinical spectrum associated with hereditary leiomyomatosis and renal cell carcinoma syndrome in French FH mutation carriers. Clin Genet 2017;92:606-15.

41. Padgett RA. New connections between splicing and human disease. Trends Genet 2012;28:147-54.

42. Monaghan RM, Whitmarsh AJ. Mitochondrial Proteins Moonlighting in the Nucleus. Trends Biochem Sci 2015;40:728-35.

43. Lionaki E, Gkikas I, Tavernarakis N. Differential Protein Distribution between the Nucleus and Mitochondria: Implications in Aging. Front Genet 2016;7:162.

44. Menko FH, Maher ER, Schmidt LS, Middelton LA, Aittomaki K, Tomlinson I*, et al.* Hereditary leiomyomatosis and renal cell cancer (HLRCC): renal cancer risk, surveillance and treatment. Fam Cancer 2014;13:637-44.

45. Joseph NM, Solomon DA, Frizzell N, Rabban JT, Zaloudek C, Garg K. Morphology and Immunohistochemistry for 2SC and FH Aid in Detection of Fumarate Hydratase Gene Aberrations in Uterine Leiomyomas From Young Patients. Am J Surg Pathol 2015;39:1529-39.

46. Rabban JT, Chan E, Mak J, Zaloudek C, Garg K. Prospective Detection of Germline Mutation of Fumarate Hydratase in Women With Uterine Smooth Muscle Tumors Using Pathology-based Screening to Trigger Genetic Counseling for Hereditary Leiomyomatosis Renal Cell Carcinoma Syndrome: A 5-Year Single Institutional Experience. Am J Surg Pathol 2019;43:639-55.

47. Leshets M, Silas YBH, Lehming N, Pines O. Fumarase: From the TCA Cycle to DNA Damage Response and Tumor Suppression. Front Mol Biosci 2018;5:68.

48. Bennett JA, Weigelt B, Chiang S, Selenica P, Chen YB, Bialik A*, et al.* Leiomyoma with bizarre nuclei: a morphological, immunohistochemical and molecular analysis of 31 cases. Mod Pathol 2017;30:1476-88.

49. Harrison WJ, Andrici J, Maclean F, Madadi-Ghahan R, Farzin M, Sioson L*, et al.* Fumarate Hydratase-deficient Uterine Leiomyomas Occur in Both the Syndromic and Sporadic Settings. Am J Surg Pathol 2016;40:599-607.

50. Anna A, Monika G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. J Appl Genet 2018;59:253-68.

51. Bulku A, Weaver TM, Berkmen MB. Biochemical Characterization of Two Clinically-Relevant Human Fumarase Variants Defective for Oligomerization. Open Biochem J 2018;12:1-15.

52. King A, Selak MA, Gottlieb E. Succinate dehydrogenase and fumarate hydratase: linking mitochondrial dysfunction and cancer. Oncogene 2006;25:4675-82.

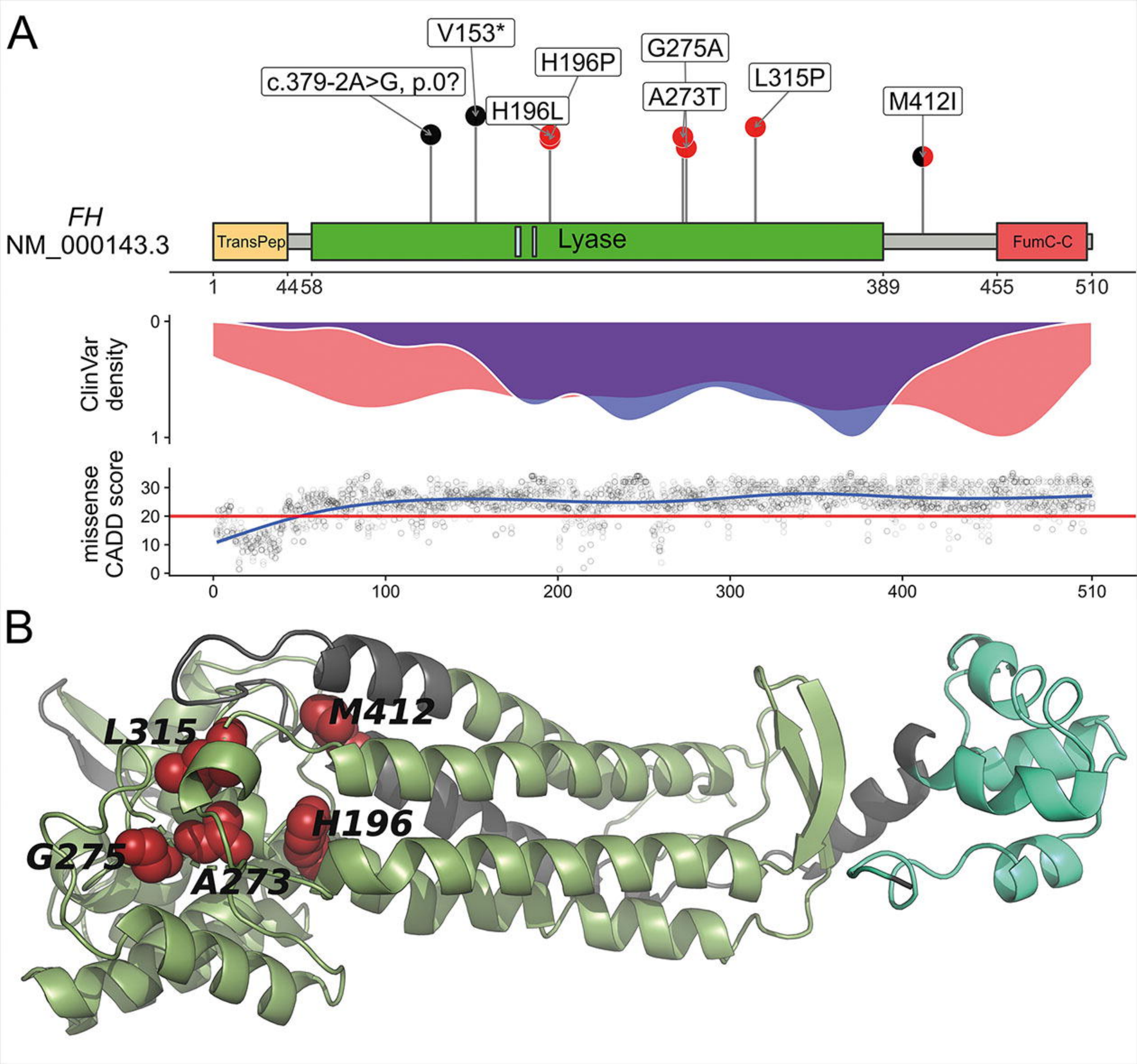
53. Kobayashi T, Minowa O, Sugitani Y, Takai S, Mitani H, Kobayashi E*, et al.* A germ-line Tsc1 mutation causes tumor development and embryonic lethality that are similar, but not identical to, those caused by Tsc2 mutation in mice. Proc Natl Acad Sci U S A 2001;98:8762-7.

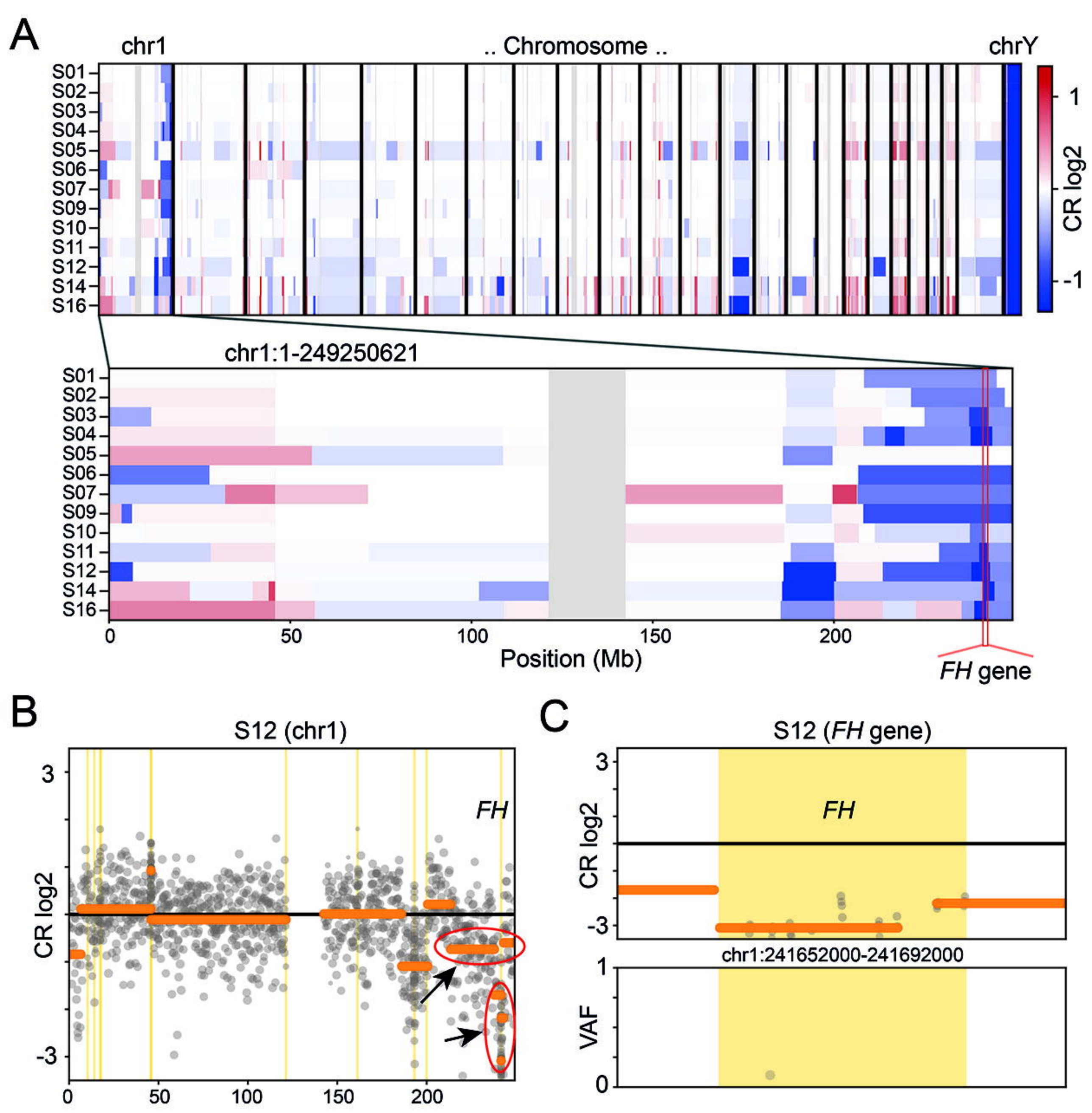
54. Tsagoudis K, Magiera-Lappann I, Haverkamp T, Drebber U, Agaimy A, Tantcheva-Poomicronr I. Hereditary leiomyomatosis and renal cell cancer syndrome: A novel mutation in the FH gene. J Dermatol 2018;45:373-5.

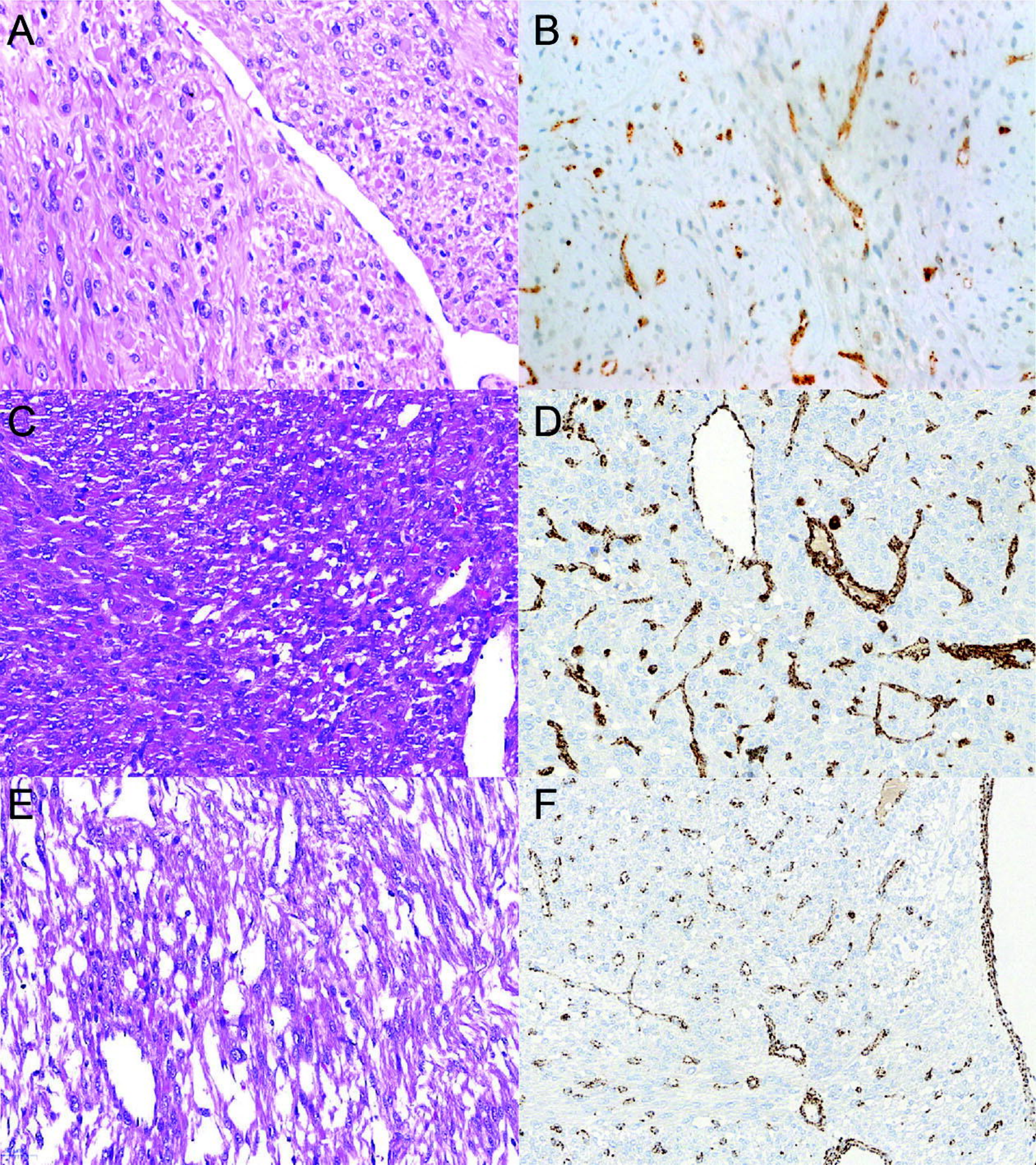
55. Holla VR, Elamin YY, Bailey AM, Johnson AM, Litzenburger BC, Khotskaya YB*, et al.* ALK: a tyrosine kinase target for cancer therapy. Cold Spring Harb Mol Case Stud 2017;3:a001115.

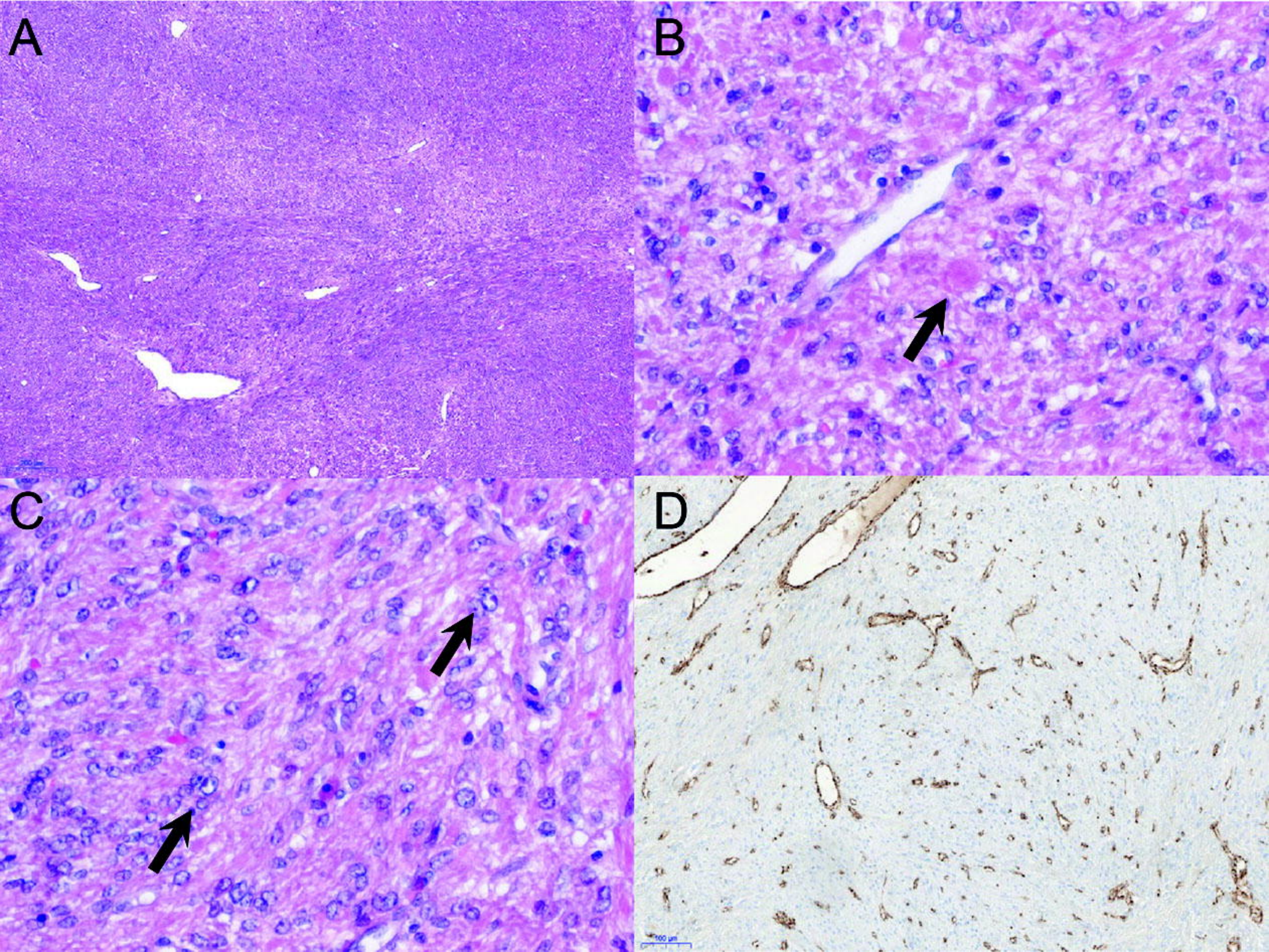
56. Mehine M, Kaasinen E, Makinen N, Katainen R, Kampjarvi K, Pitkanen E*, et al.* Characterization of uterine leiomyomas by whole-genome sequencing. N Engl J Med 2013;369:43-53.

57. Smit DL, Mensenkamp AR, Badeloe S, Breuning MH, Simon ME, van Spaendonck KY*, et al.* Hereditary leiomyomatosis and renal cell cancer in families referred for fumarate hydratase germline mutation analysis. Clin Genet 2011;79:49-59.


A
V153*
c.379-2A>G, p.0?
H196P
H196L
G275A
A273T
L315P
M412I
FH
NM_000143.3
TransPep
Lyase
FumC-C
1
44
58
389
455
510
ClinVar density
0
1
missense CADD score
30
20
10
0
0
100
200
300
400
510
B
M412
L315
H196
G275
A273


A
chr1
.. Chromosome ..
chrY
S01
S02
S03
S04
S05
S06
S07
S09
S10
S11
S12
S14
S16
1
CR log2
-1
chr1:1-249250621
0
50
100
150
200
Position (Mb)
FH gene
B
S12 (chr1)
3
FH
CR log2
-3
0
50
100
150
200
C
S12 (FH gene)
3
FH
CR log2
-3
chr1:241652000-241692000
1
VAF
0


A
B
C
D
E
F


A
B
C
D