OPINION ARTICLE

Factors targeting MED12 to drive tumorigenesis? [version 1; peer review: 3 approved with reservations]

Jörn Bullerdiek¹,², Birgit Rommel²

¹Institute of Medical Genetics, Medical Center, University of Rostock, Rostock, D-18057, Germany
²Human Genetics, University of Bremen, Bremen, D-28359, Germany

Abstract
Mediator Subcomplex 12 (MED12) is part of the transcriptional preinitiation machinery. Mutations of its gene predominantly occur in two types of highly frequent benign tumors, uterine leiomyomas and fibroadenomas of the breast, where they apparently act as driver mutations. Nevertheless, their presence is not restricted to benign tumors having been found at considerable frequencies in uterine leiomyosarcomas, malignant phyllodes tumors, and chronic lymphocytic leukemia also. Most of the mutations are located within exon 2 of the gene but in rare cases the intron 1/exon 2 boundary or exon 1 are affected. As to their type, predominantly single nucleotide exchanges with a hotspot in one codon are found, but small deletions clustering around that hotspot also are not uncommon. According to their presumed classification as gain-of-function mutations, these latter deletions are leaving the open reading frame intact. As to the types of mutations, so far no apparent differences between the tumor entities affected have emerged. Interestingly, this pattern with small deletions clustered around the hotspot of single nucleotide exchanges resembles that seen as a result of targeted gene editing. In contrast to other driver mutations the percentage of MED12-mutation positive tumors of independent clonal origin increases with the number of tumors per patient suggesting unknown etiological factors supporting site specific mutagenesis. These factors may act by inducing simultaneous site-specific double strand breaks the erroneous repair of which may lead to corresponding mutations. As inducers of DNA damage and its repair such as foreign nucleic acids of the microbiome displaying sequence homology to the putative target site might play a role. Interestingly, a 16 base pair homology of the hotspot to a putative terminator base-paired hairpin sequence of a Staphylococcus aureus tRNA gene cluster has been noted which might form R-loop like structures with its target sequence thus inducing said changes.

Keywords
Mediator subcomplex 12 (MED12), mutations, uterine leiomyomas, fibroadenomas of the breast, chronic lymphocytic leukemia (CLL), multiple tumors, DNA-RNA hybrids, Staphylococcus aureus

Open Peer Review

Review

Reviewer Status
Invited Reviewers
1 2 3
REVISED
version 2
published 13 Dec 2018

version 1
published 22 Mar 2018

1 Takeshi Kurita, Ohio State University, Columbus, USA
2 Eric Glasgow, Georgetown University, Washington, USA
3 Jose M. Teixeira, Michigan State University, Grand Rapids, USA

Any reports and responses or comments on the article can be found at the end of the article.
Introduction

Surprisingly, the most common mutation in human tumors does not affect one of the famous suspects in the field (for review see (Vogelstein et al., 2013)) but the much less well-known gene encoding Mediator Subcomplex 12 (MED12). MED12 is part of the transcriptional preinitiation complex CDK8 (Elmlund et al., 2006) and encoded by a gene that maps to the X-chromosome at Xq13.1. One obvious reason why its mutations so far have gained much less interest than those of other genes frequently mutated in human tumors is that they affect, to a large extent, benign tumors. Moreover, within malignant tumors, they are virtually absent from most of the predominant epithelial neoplasms like cancers of colon, breast, and lung (Kandoth et al., 2013) whereas they only have been found at considerable frequencies in some malignant tumors of non-epithelial origin.

As in the benign tumors, however, mutations of that gene occur as apparent driver mutations in a predominant subset of uterine leiomyomas (Mäkinen et al., 2011; Markowski et al., 2012; McGuire et al., 2012), constituting the by far most frequent human symptomatic tumors of all. Likewise, these mutations are also found in a large subset of fibroadenomas of the breast (Lim et al., 2014; Piscuoglio et al., 2015; Yoshida et al., 2015), another frequent benign tumor which occurs predominantly in young and middle-aged women. Interestingly, they are not restricted to benign tumors but also frequently seen in their malignant counterparts, i.e. uterine leiomyosarcomas (Kämpjärvi et al., 2012; Markowski et al., 2013a; Pérot et al., 2012) and malignant phyllodes tumors (Nagasawa et al., 2015; Piscuoglio et al., 2015; Yoshida et al., 2015). Also, their presence in malignant tumors suggests that, albeit as a very rare event, certain additional mutations can trigger malignant transformation within formerly benign tumors harboring MED12 mutations. Apart from solid tumors, MED12 mutations recently were also detected in a significant percentage of roughly 5–9% of chronic lymphocytic leukemias (CLL) (Guèze et al., 2015; Kämpjärvi et al., 2015; Wu et al., 2017a). Furthermore, the same type of MED12 mutations was found in two canine vaginal leiomyomas (Markowski et al., 2013a).

A closer look at the MED12 mutations does not reveal apparent differences between the type of mutations when comparing the different tumor entities. Predominantly, the mutations are clustered in the 5′ region of exon 2 of the gene with only a few mutations affecting the intron 1-exon 2 boundary or, much rarer, exon 1 or the exon 1-intron 1 boundary. Most of them are single base exchanges clearly clustered at two nucleotides of codon 44 where, albeit with different frequencies, guanins are found to be replaced by either A, C, or T. Besides these single base replacements, deletions and, more rarely, indels, usually affecting exon 2 or the intron 1/exon 2 boundary are found which always leave the reading frame intact indicating gain-of-function mutations. MED12 maps to the X-chromosome and, as revealed by cDNA sequencing, the mutations are apparently restricted to the active X-chromosome (Mäkinen et al., 2011; Markowski et al., 2012; McGuire et al., 2012).

We feel that for several reasons this highly frequent type of mutation might point to an unusual mechanism of mutagenesis underlying the development of the corresponding tumors. These reasons will be discussed herein and a hypothesis based on target-specific mutagenesis will be presented. Starting with a short introduction of the main tumors affected by MED12 mutations we will then address the molecular pathogenesis of uterine leiomyomas along with its clinical correlations followed by an in depth analysis of the pattern of MED12 mutations. Finally, we will present a hypothesis why these data indicate a so far unknown etiological mechanism favoring these particular highly frequent somatic alterations of the genome.

Introducing three tumor entities displaying a unique type of MED12 mutations

Besides a few very rare tumors, three main tumor entities are often affected by mutations of MED12. First, these three entities, i.e. uterine leiomyomas (fibroids), fibroadenomas of the breast, and chronic lymphocytic leukemias, will be introduced.

Uterine leiomyomas - the most frequent symptomatic human tumors

Uterine leiomyomas (UL) are benign smooth muscle tumors of myometrial origin with an apparently very low tendency to undergo malignant transformation. Depending on their location it can be distinguished between submucosal, intramural, and subserosal UL. Roughly 40–70% of women in their reproductive age will develop UL with a well-documented higher prevalence among women of African and African-American origin. In this group, the UL develop on average at younger ages (Laughlin et al., 2010). Leiomyomas are often of large size and, as a matter of debate for more than 100 years, (Figure 1), multiple nodules occur at almost the same frequency or even more frequently than single nodules.

Nevertheless, the majority of patients with UL are without symptoms. The remaining 20–30% of the patients suffer from symptoms like in particular heavy menstrual bleeding, increased menstrual periods, reduced fertility, and pelvic pressure and pain. During menopause, UL cease growing and even shrink. Nevertheless, despite their benign nature UL are the most symptomatic human tumors of all and repeatedly have been reported even in mummies with the oldest of them dating back to the Middle Neolithic age (3,200–2,500 BC) (Fornaciari & Giuffra, 2012). Surgery (hysterectomy or myomectomy) is a common type of treatment, but a variety of other non-surgical methodologies to treat UL are available (for review see: (Williams, 2017)).

While the etiology of these frequent tumors remains unclear (McWilliams & Chennathukuzhi, 2017), a bit more is known about pathogenetic factors. As to basic principles of their molecular pathogenesis, UL behave like the vast majority of other benign and malignant tumors. This includes a monoclonal origin which is triggered by so-called driver mutations. In the case of multiple UL almost every single nodule has been found to be of independent clonal origin (Masah et al., 1994). Accordingly, different nodules are usually characterized by different driver mutations in these cases.

Moreover, genetic subtypes, apparently belonging to different groups of driver mutations, exist. While MED12 mutations,
as a group, constitute the most predominant genetic subtype (Mäkinen et al., 2011), another frequent subgroup of uterine leiomyomas carries rearrangements of the gene encoding the architectural transcription factor High Mobility AT-hook 2 (HMGA2), as a rule reflected by cytogenetically visible chromosomal translocations (Schoenmakers et al., 1995). Of note, the driver mutations of both subgroups occur in a mutually exclusive manner (Markowski et al., 2012). Besides these subgroups, other, more rare but also independent genetic subgroups of UL such as one characterized by either germline (hereditary leiomyomatosis and renal cell cancer (HLRCC), OMIM 605839) (Bayley et al., 2008) or somatic loss-of-function mutations of Fumarate Dehydrogenase (FH) seem to exist (Kämpjärvi et al., 2016).

Fibroadenomas of the breast - Frequent benign tumors of adolescent and young women

Fibroadenomas of the breast are common benign tumors histologically composed of both stromal and epithelial components preferentially occurring in adolescent and young women. Their general incidence may be in the range of 10% in the corresponding age groups. Multiple tumors are not rare with some 10–15% of the patients having more than one FA.

Their name fibroadenomas (FA) and their classification as fibroepithelial tumors suggest a biphasic nature of these neoplasms. Nevertheless, as to their pathogenesis this classification seems to be misleading at least in the majority of cases because mutations are restricted to the stromal component of the tumors (Mishima et al., 2015). Between 50% and 60% of the FA harbor MED12 mutations (Lim et al., 2014; Mishima et al., 2015; Pfarr et al., 2015) Histologically, the occurrence of MED12 mutations correlates highly significant with the so-called intracanalicular growth pattern (Mishima et al., 2015; Pfarr et al., 2015). Interestingly, MED12 mutations have also been found in considerable percentages of other breast tumors of presumed stromal origin Phyllodes tumors and malignant Phyllodes tumors. In these tumors, the types of MED12 mutations are not obviously discernible from those observed in UL and FA.

Chronic lymphocytic leukemias - most frequent leukemia of the adults

In Western countries, chronic lymphocytic leukemia (CLL) is the most common type of leukemia in adults. The American Cancer Society expects an estimated number of about 20,940 new cases of CLL in the United States in 2018 with about 4,510 CLL-related deaths. Overall, CLL accounts for about one-quarter of the new cases of leukemia. The average person’s lifetime risk of getting CLL is about one in 175 (0.57%). The average age at the time of diagnosis is around 70 years, with men slightly more often affected than women. Frequently, the disease is detected in patients not yet showing any severe symptoms and in its early stages patients often undergo a ‘watch and wait’ period prior to starting therapy.

As to its pathogenesis CLL is a monoclonal leukemia of B-cell origin with a number of subsets that can be distinguished based on their genetic characterization, apparently pointing to different driver mutations. The genetic alterations also allow the stratification of groups which may require different therapeutic approaches. As a valid predictive parameter, deletions and mutations of the TP53 gene are associated with a worse prognosis than other types of genetic changes and influence therapeutic decisions. Recently, mutations of MED12 have been added to the list of potential driver mutations in CLL. Across the sub-types, they constitute a relatively rare genetic group affecting roughly 5–9% of CLL patients (Guièze et al., 2015; Kämpjärvi et al., 2015; Wu et al., 2017a). Kämpjärvi et al. (Kämpjärvi et al., 2015) have presented evidence that MED12 mutations may represent a marker of worse prognosis.
A closer look at the molecular pathogenesis of uterine leiomyomas

According to the high prevalence of uterine leiomyomas, MED12 mutations are by far the best investigated in this tumor type. Thus, we have now characterized the subset UL affected and described how they can be distinguished from other types of UL.

Leiomyomas with MED12 mutation constitute their own genetic subtype which is also characterized by a distinct clinical and histopathological appearance

There is ample evidence that HMG2 rearrangements and MED12 mutations occur mutually exclusively in UL and thus constitute independent driver mutations. Similarly, somatic MED12 mutations and biallelic Fumarate Hydratase (FH) inactivation occur in mutually exclusive manner in both HLRCC syndrome-associated and sporadic uterine leiomyomas suggesting that the latter constitutes a third small group with an independent molecular pathogenesis. Accordingly, each of these genetic alterations alone as a driver mutation seems to be sufficient to induce the development of an UL without requiring any further mutations.

As to rearrangements of HMG2 and mutations of MED12, the transcriptome of tumors of both groups clearly differs with MED12 mutation-positive and HMG2-overexpressing samples clustering in distinct branches. Both mutations allow the two major genetic subtypes of UL to be distinguished, and the question arises whether or not the genetic subtypes are also reflected by a different clinical behavior and histopathology.

Tumors carrying HMG2 rearrangements are usually solitary and, on average, of larger size than those with MED12 mutations, also usually presenting as single tumors whereas the latter are smaller and often co-occur with other clonally independent nodules of the same genetic type.

The percentage of MED12-mutated tumors is positively correlated with the total number of tumors per patient

To explain the high frequency of MED12 mutated tumors among multiple leiomyomas, Heinonen et al. speculated that the multiplicity of MED12-mutation-positive leiomyomas may derive from genetic predisposition and/or environmental factors rendering the myometrium susceptible to selection for MED12 mutations. However, as outlined later herein, the association of multiple tumors with MED12 mutations may be a key to the etiology of this type of UL.

Furthermore, MED12-mutated UL are also significantly associated with a subserous location compared to UL lacking this mutation. As to histopathological features, a recent study by Wu et al. revealed that approximately 90% of the cells in HMG2-rearranged UL were smooth muscle cells showing an overexpression of the protein, while MED12-mutated UL had a similar number of smooth muscle cells and other cells, i.e. mostly tumor-associated fibroblasts. The latter fibroblasts were lacking MED12-mutations (Wu et al., 2017b) and thus apparently can be classified as bystander cells. This fits with an earlier observation that in cell cultures of leiomyomas with MED12-mutations a rapid disappearance of mutated cells was seen that became replaced by UL wild-type cells, thus challenging the results of a variety of in vitro experiments on the biology of UL (Bloch et al., 2017; Markowski et al., 2014b).

Besides UL, the occurrence of MED12 mutations has been well-documented in malignant uterine smooth muscle tumors (leiomyosarcomas) and smooth muscle tumors of uncertain malignant potential (STUMP), too. Hence, an origin of these tumors from pre-existing UL has been suggested. In contrast, similar cases with HMG2 rearrangements have not been reported yet.

The occurrence of MED12 mutations has been well-documented in malignant uterine smooth muscle tumors (leiomyosarcomas) and smooth muscle tumors of uncertain malignant potential (STUMP), too (Holzmann et al., 2015; Péro et al., 2012). Hence, an origin of these tumors from pre-existing UL has been suggested. In contrast, similar cases with HMG2 rearrangements have not been reported yet.

The percentage of MED12-mutated tumors is positively correlated with the total number of tumors per patient

To gain further insight into the biology of MED12-mutated UL, Heinonen et al. have undertaken a systematic attempt to check all feasible distinct tumors with a size of 1 cm or larger in diameter from hysterectomy uteri for MED12 mutations. In their study, 599 out of 763 leiomyomas carried MED12 mutations (79%). Next, the data provided by the study of Heinonen et al. (Heinonen et al., 2017) have been used to analyze the number of MED12-mutation positive UL per patient. While it was shown before that in the majority of patients having surgery MED12-mutated tumors do not make their appearance as single nodules but instead are accompanied by other yet clonally independent tumors of this same genetic type (cf. Figure 2), we were interested to see how the number of these tumors per patient is distributed in this series. A slow decrease of the number of

Figure 2. Single and multiple MED12-mutation positive uterine leiomyomas. Proportion of women carrying a single MED12-mutated UL (52/176) vs. those with more than one such tumor (124/176) (A) and proportion of MED12-mutated UL appearing as single tumor (52/599) vs. those accompanied by at least one other MED12-mutated UL (547/599) (B). Data according to Heinonen et al. (2017).
\textit{MED12}-mutation positive tumors is noted (Figure 3). Nevertheless, from these Figures it is not possible to draw conclusions on the overall frequency and distribution of \textit{MED12}-positive UL in the population because the results are biased by their restriction to symptomatic patients who had undergone surgical treatment. Nevertheless, they correspond more or less to the tumor numbers in general as seen from numerous other studies and thus the question arises if the \textit{MED12}-positive tumors can be distinguished from the remaining UL as suggested by previous estimations (cf. Figure 2). Thus, we have next investigated if, in the case of multiple tumors, the percentage of \textit{MED12}-mutation positive UL remains constant independent of the total number of tumors per patient. Surprisingly, it was noted that with a growing number of tumors, the percentage of \textit{MED12}-mutated tumors clearly increased. Among solitary leiomyomas, only less than 40\% of the tumors carried \textit{MED12} mutations but their frequency was approaching nearly 100\% if twelve or more tumors were present per patient (Figure 4). Thus, in contrast to other mutations, those of \textit{MED12} seem to become more likely with an increase of the number of tumors.

Along with previous data this distribution confirms that, as for its pathogenesis, the occurrence of multiple leiomyomas near exclusively can be attributed to just one genetic mechanism, i.e. \textit{MED12} mutations. This contradicts a statement by Mehine \textit{et al.} that a shared clonal origin as a common feature of leiomyomas not carrying \textit{MED12} mutations offers one explanation for the common occurrence of multiple concurrent lesions (Mehine \textit{et al.}, 2015). Instead, a multitude of fibroids mainly appears to be a problem almost exclusively restricted to \textit{MED12}-mutation positive tumors and we thus decided to analyze and compare the different \textit{MED12} mutations in more detail.

A closer look at the patterns of \textit{MED12} mutations seen in various benign and malignant tumors

\textit{MED12} mutations occur in uterine smooth muscle tumors, fibroepithelial tumors of the breast and in chronic lymphocytic leukemia. As a rule, only subsets display these genetic abnormalities which range from the predominant alterations in UL to those detected only in a small percentage of cases in CLL. So far, there is no evidence that the tumors affected by \textit{MED12} mutations in the three tumor entities described above differ with respect to their types of mutations. It has been speculated that unknown factors favor the occurrence of these particular mutations. To get further ideas on factors favoring them, the patterns of mutations have now been analyzed in more detail and compared between the different tumor entities paying particular attention to the small deletions occurring in exon 2 or at the intron1-exon 2 boundary, respectively.

In most cases single nucleotide exchanges are found with a clear predominance of those affecting nucleotides 130 and 131 belonging to codon 44. Less frequently other codons are mutated. Besides single nucleotide exchanges, deletions of small segments of the gene with varying sizes as well as indels affecting exon 2 or the intron1-exon 2 boundary are seen in some cases. As a rule, however, the transcript though affected by the deletions remains in frame. As to these latter genomic alterations accounting for roughly 15\% of \textit{MED12} mutations, we have analyzed the positions of the deleted bases from a variety of papers analyzing UL.

\textbf{Figure 3. Number of \textit{MED12}-mutation positive uterine leiomyomas per patient.} Abscissa: number of UL/patient, ordinate: number of patients in the corresponding category. For this diagram data on \textit{MED12} alterations published by Heinonen \textit{et al.} (Heinonen \textit{et al.}, 2017) have been used.
fibroepithelial tumors, and CLL. Adding the number of deleted bases per each position reveals an almost symmetric distribution that is clustered around the hotspot of single nucleotide exchanges (Figure 5). In a previous study by our group, the beginning of the MED12 deletions observed in uterine smooth muscle tumors was mostly located within exon 2 but in rare cases also upstream of the splice site within intron 1. Their size mainly ranged between 3 and 36 bp with a clear predominance of 15 and 18 bp (Markowski et al., 2013b). Of note, an analysis of the data provided by Heinonen et al. for UL revealed that as very rare exceptions even larger deletions as well as those residing in exon 1 can occur.

As to size and position of these deletions, there is also no obvious difference between UL, fibroepithelial tumors of the breast, and CLL (Figure 5). Overall, this pattern of small deletions of various sizes clustered around the hotspot of single nucleotide exchanges resembles the results of genome editing based on targeted double-stranded breaks as for example those resulting from the usage of the CRISPR/Cas9 system (e.g. cf. Paquet et al., 2016).

If these mutations indeed arise by certain types of repair of site-specific DNA changes, one might expect that many other mutations occur in the target region of MED12. Of these, only the “active ones”, i.e. those leaving the open reading frame intact and driving tumorigenesis, will lead to a clonal proliferation of their target cell giving rise to an UL whereas cells with other mutations of the hotspot region will remain quiescent or even become apoptotic (Figure 6). Therefore, future studies aimed at the detection of these “non-driving” mutations in single cells, especially from patients suffering from a multitude of UL, may be a reasonable attempt. However, from the pattern of nucleotide exchanges and deletions, a commonly affected sequence can be depicted that may be related to the etiology of UL.

**Hypothesis and opinion**

MED12 mutations constitute highly frequent driver mutations in uterine leiomyomas and fibroadenomas, i.e. two tumor entities that occur almost exclusively in middle-aged and young women, respectively. In uterine leiomyomas, they even represent the by far most frequent genetic subtype with a clearly preferential occurrence in the case of multiple tumors. This is in sharp contrast to the other main genetic subtype of UL characterized by rearrangements of HMGA2 usually making its appearance in solitary nodules not accompanied by other tumors of the same genetic subtype.

It seems difficult to explain these findings just by independent random mutations followed by their selection. Nevertheless, additional factors favoring this multitude of tumors carrying the same type of mutations despite their independent clonal origin have remained enigmatic. After myomectomy, such factors may also account for the risk of recurrences that clearly increases with the number of UL that had been removed (Doridot et al., 2001; Fauconnier et al., 2000). Heinonen et al. (2017) have speculated that either genetic predisposition or environmental factors rendering the myometrium susceptible to selection for MED12 mutations may contribute to the multiplicity of MED12-mutation positive tumors. To describe the development of multiple tumors, those two explanations are well-compatible with a model of clonally unrelated nodules that occur successively and are endowed with a different growth rate as depicted in Figure 7A. Another alternative explaining the
Nevertheless, a variety of studies indicate that most deletions in uterine smooth muscle tumors (blue), fibroepithelial tumors of the breast (green), and chronic lymphocytic leukemia (yellow). Deletions are plotted across all deleted base positions. Minor preferred sites of single nucleotide exchanges within exon 2 are indicated by dashed red arrows. For this diagram data on MED12 deletions from the following articles have been used: (Guézé et al., 2015; Kämpjärvi et al., 2015; Lim et al., 2014; Markowski et al., 2013b; Mishima et al., 2015; Nagasawa et al., 2015; Ng et al., 2015; Pfarr et al., 2015; Yoshida et al., 2015) only those deletions beginning and ending in the displayed region have been considered.

Figure 5. Patterns of MED12 deletions in uterine leiomyomas, fibroepithelial breast tumors, and chronic lymphocytic leukemia. Left to right: Ideogram of the X-chromosome (commons.wikimedia.org), exon-intron structure of MED12 (NCBI map viewer), and plot depicting frequency of deletions at each position around the preferred site of single nucleotide exchanges (red solid arrows) seen in uterine smooth muscle tumors (blue), fibroepithelial tumors of the breast (green), and chronic lymphocytic leukemia (yellow). Deletions are plotted across all deleted base positions. Minor preferred sites of single nucleotide exchanges within exon 2 are indicated by dashed red arrows. For this diagram data on MED12 deletions from the following articles have been used: (Guézé et al., 2015; Kämpjärvi et al., 2015; Lim et al., 2014; Markowski et al., 2013b; Mishima et al., 2015; Nagasawa et al., 2015; Ng et al., 2015; Pfarr et al., 2015; Yoshida et al., 2015) only those deletions beginning and ending in the displayed region have been considered.

Figure 6. MED12 mutations in uterine smooth muscle tissue. Model illustrating the occurrence and selection of MED12 mutations during the course of leiomyoma development. The scheme suggests that of a larger number of MED12-mutations only those associated with a gain of function act as driver mutations giving rise to UL.

As possible factors certain types of bacteria e.g. those involved in reproductive tract infections (RTIs) have long been suggested to be a cause of UL (e.g. (Witherspoon & Butler, 1934)). In the United States both reportable RTIs (i.e. chlamydia and gonorrhea) and fibroids disproportionately burden African American women which lead to the conclusion that the growth of fibroids might be triggered by inflammatory infections associated with the RTIs. Nevertheless, when exploring the relationship between self-reported RTIs and fibroid size, number, and total volume Moore et al. did not find strong associations (Moore et al., 2015). In a recent contribution by the same group women seropositive for genital Chlamydia trachomatis were even found to be less likely to have fibroids (Moore et al., 2018). In line with these findings, in the study by Heinonen et al. neither a history of pelvic inflammatory disease (PID) nor of Chlamydia infection was found to be significantly associated with the MED12-type UL while PID turned out to be significantly associated with the occurrence of MED12-wild type UL (p 0.0024) (Heinonen et al., 2017).

Akin to bacteria, viruses have also been suggested to be involved in the development of UL. For example, EBV is known as a factor associated with the development of extra uterine smooth muscle tumors in HIV and post-transplant patients (see e.g. (Miettinen, 2014; Purgina et al., 2011; Ramdial et al., 2011)). However, so far no association between EBV and uterine leiomyomas has been demonstrated. As to another virus of the Herpes group, a recent study failed to reveal a significant association between HSV-2 seropositivity and the presence of fibroids (Moore et al., 2016) and in general no convincing evidence for viruses involved in the pathogenesis of UL has been presented.

While infectious diseases as etiological agents of UL repeatedly have been assumed as such the question arises as to how they could act by site-specific targeting the hotspot region of mutations residing within exon 2 of MED12 and which infectious agents are possible candidates. We will herein present the opinion that the interaction between the human DNA and foreign nucleic acids derived from the infection plays a causal role. As an unorthodox hypothesis stimulating further discussion, we would like to advance the hypothesis that the MED12 mutations result from cleavage of R-loop structures. By definition, R-loops are derived from double stranded DNA where one strand forms a stable DNA-RNA hybrid helix whereas the former associated DNA strand remains...

Figure 6. MED12 mutations in uterine smooth muscle tissue. Model illustrating the occurrence and selection of MED12 mutations during the course of leiomyoma development. The scheme suggests that of a larger number of MED12-mutations only those associated with a gain of function act as driver mutations giving rise to UL.
single-stranded. R-loops with an "exposed" stretch of single-stranded DNA can give rise to instability and DNA double-strand breaks (Aguilera & Gómez-González, 2017; Freudenreich, 2018; Su & Freudenreich, 2017). While the hybrid helix is usually composed of DNA with endogenous RNAs it seems possible that such helices can be formed with foreign RNA as well. To this end, it has been hypothesized that circulating exogenous RNA sequences after their uptake may influence the function of cells through miRNA-like mechanisms (Wang et al., 2012) suggesting direct influences of these sequences of the microbiome on its host’s cells. To search for possible sequence homologies we have depicted a target region 5'TGTAAAACAAGGTTTCAATAAC3' covering 10 nucleotides upstream and downstream each of the two c. 130 and c. 131 (GG), respectively (cf. Figure 6). Of the resulting list with at least 15 identical nucleotides, a variety of human pathogens have been identified as e.g. Bacteroides fragilis (nt 2-20), Klebsiella pneumoniae (nt 2-20), Escherichia coli (nt 3-20), Vibrio vulnificus (nt 6-22), Staphylococcus aureus (nt 6-22, and nt 1-16, respectively), Staphylococcus argenteus (nt 6-22), and Clostridium botulinum (nt 3-19). When searching for abundantly expressed sequences an interesting candidate emerged. A 16-base pair sequence identical to the sense strand of the sequence of a Staphylococcus aureus 27-tRNA gene cluster immediately 3’ to an rRNA operon (Green & Vold, 1993) was noted. The homology covers a palindromic sequence which may act as a terminator of transcription (Green & Vold, 1993) and may also lead to the formation of a hairpin structure stabilizing the RNA molecule (Figure 8). Are these molecules likely to exist as circulating RNAs? Staphylococcus aureus belongs to the phylum of Firmicutes in general constituting the third most abundant sequence population in human plasma with a significant number of the reads mapping to various bacterial ribosomal RNAs and tRNAs (Wang et al., 2012). More specifically, evidence has been presented that RNA species from Staphylococcus are commonly present in blood (Leung & Wu, 2015).

In summary, deduced from the types and patterns of MED12 mutations in human tumors we have presented evidence supporting the idea that nucleic acid sequences of the human microbiome may interact with the common hotspot of MED12 mutations. To stimulate further discussion a possible interaction of a sequence of Staphylococcus aureus with this hotspot has been considered in more detail as depicted in Figure 8 and Figure 9. As to the initial stage of leiomyoma development, the clearance of R-loops resulting from a hybrid helix between a human target cell and bacterial RNA may simultaneously give rise to multiple clonally independent cells of origin. Additional factors such as the site of origin,

![Figure 7. Different models of time and clonality of uterine leiomyoma development.](image-url)
Figure 8. Sequence adjacent to a 27-tRNA gene cluster of Staphylococcus aureus similar to the MED12 hotspot. Upper part: Homology of the human MED12 hotspot region (red arrow) with the sense strand of the sequence of the Staphylococcus aureus 27-tRNA gene cluster immediately 3' to a rRNA operon (Green & Vold, 1993). Blue dashed arrows indicate a palindromic sequence. Sequence from Staphylococcus aureus strain CFSAN007847 chromosome, complete genome; GenBank: CP017684.1; GenBank: FASTA; NCBI Blast. Lower part: bacterial tRNA and rRNA genes of the operon adjacent to the site of homology are shown in blue.

Figure 9. Model depicting the interaction between bacterial RNA with the MED12 mutational hotspot. As an example for the putative sequence of the human microbiome inducing site specific mutations this scheme refers to Figure 8.
angiogenetic support, and the type of MED12 mutation, may endow the resulting monoclonal lesions with a different growth potential.

Data availability
All data underlying the results are available as part of the article and no additional source data are required.

Competing interests
No competing interests were disclosed.

Grant information
The author declared that no grants were involved in supporting this work.

References


Jose M. Teixeira

Department of Obstetrics, Gynecology and Reproductive Biology, College of Human Medicine, Michigan State University, Grand Rapids, MI, USA

The opinion article by Bullerdiek and Rommel provides an interesting hypothesis that is partly supported by the unusual role of MED12 in the origin of uterine leiomyoma. MED12 is mutated in these benign tumors at a hotspot location at the 5' end of exon 2. How these mutations arise in the myometrium and how the mutant MED12 drives leiomyoma development are not known.

I agree with many of the previous reviewers' comments and will limit my comments to a few items so as not to be too repetitious. For example, I agree that there is no evidence that the MED12 mutations constitute a gain-of-function in the protein. Indeed, if MED12 mutation alone is sufficient for tumorigenesis, it would have been observed in a variety of other tumor types. In fact, no mechanisms for tumorigenesis unique to leiomyomas have been reported. Thus, how MED12 mutation leads to leiomyoma development is not known at this point. This should be changed in the text.

I also agree that the argument for foreign RNA from bacterial infection causing MED12 mutation, although quite novel and certainly interesting, is weak. In addition to the reasons already discussed, ascending infections, such as the reproductive tract infections described by the authors, are likely to rarely involve the myometrium. In the case of S. aureus, it is even more unlikely because myometrial infection with that organism is likely associated with bacteremia, which is even more rare and cannot account for the prevalence of MED12 mutant fibroids in women. Also, although leiomyomas are normally found in reproductive age women, that doesn't necessarily mean that sexually active women are more prone to the disease because of possible sexually transmission of pathogenic bacteria. Since parity is associated with a lower leiomyoma burden, the argument for reproductive tract infections being the culprit is not supported. Parity is also associated with increased risk of postpartum iatrogenic infection, but again parity is associated with decreased risk for leiomyomas. These caveats should be included in the text.

There is something special about the very common MED12 mutations in that hotspot and high prevalence of uterine leiomyomas (and the not-so-common fibroadenomas) in reproductive age women. It is possible that a hormonally-regulated factor expressed only in myometrium (and breast stroma) could be interacting with wt MED12 on the site where the hotspot mutations alter the protein structure, and mutation disrupts...
that interaction. How those MED12 mutations occur and why the tumors/fibroadenomas develop only in those tissues is a mystery that needs to be resolved in order to develop therapies targeting the mechanisms involved.

**Is the topic of the opinion article discussed accurately in the context of the current literature?**
Partly

**Are all factual statements correct and adequately supported by citations?**
Partly

**Are arguments sufficiently supported by evidence from the published literature?**
Partly

**Are the conclusions drawn balanced and justified on the basis of the presented arguments?**
No

**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.
-The gene name! The gene name is mediator complex subunit 12 (MED12), not Mediator Subcomplex 12 as stated in the article.

-In the Abstract it is stated “(MED12)…presumed classification as gain-of-function mutations…”, and in the Introduction “…deletions and, more rarely, indels, usually affecting exon 2 or the intron 1/exon 2 boundary are found which always leave the reading frame intact indicating gain-of-function mutations”. The term gain-of-function should be removed. The function of MED12 is unknown and there is no evidence suggesting that these mutations are gain-of-function.

Major issues:

- In Figure 2B, it is unclear how many patients in total were analyzed for tumor counts or whether the multiple MED12-mutation positive UL were concentrated in a small number of patients or spread evenly across patients analyzed.
  - Additionally, the labels appear to be switched in the key of Figure 2.

- A major problem is that the authors claim that the occurrence of multiple ULs can be exclusively attributed to MED12 mutation. However, although the authors provide ample quantitative evidence of multiple ULs resulting from MED12 mutation within the same patient, the authors show no data regarding the occurrence of multiple ULs with other genetic mechanisms (i.e. HMGA2 arrangements). In order to make the claim that multiple UL is a feature unique to MED12-mutated tumors, the authors should provide data on non-MED12-mutated multiple UL for the sake of comparison. If the number of non-MED12-mutated multiple UL is zero, this should be made clear within the text of the article.

- In reference to Figure 5, it is unclear whether the CLL mutations (yellow) cluster around the mutational hotspot as do the UL and fibroadenoma mutations, although the text of the article claims that the CLL mutations also cluster around the hotspot.

- The authors point out that the evidence for a connection between infection and UL is dubious thus far, yet they go on to propose a mechanism for the role of infection in UL pathogenesis. To make a convincing argument, the authors must clarify why they still consider a pathogenic mechanism worth considering.
  - The authors should provide more detail about the pathogens with tRNA sequences matching the mutational hotspot. What is known about the involvement of these pathogens in pelvic infections? Does the prevalence of pelvic infections involving these pathogens approximately match the prevalence of UL? Is there any existing evidence for association between these pathogens and UL?

Minor issues:

- Second paragraph of the introduction: “As in the benign tumors, however, mutations of that gene occur as apparent driver mutations in a predominant subset of uterine leiomyomas (Makinen et al., 2011; Markowski et al., 2012; McGuire et al., 2012), constituting the by far most frequent human symptomatic tumors of all.” Do the authors mean “As in the malignant tumors…” Also, “that gene” is vague. Replace with MED12.

- Second paragraph introduction: “Furthermore, the same type of MED12 mutations was found in two canine vaginal leiomyomas (Markowski et al., 2013a)” can be changed to “Furthermore, the same type of MED12 mutation was found in two canine vaginal leiomyomas (Markowski et al., 2013a)”

- Third paragraph introduction: “Predominantly, the mutations are clustered in the 5′ region of exon 2 of the
gene with only a few mutations affecting the intron 1-exon 2 boundary or, much rarer, exon 1 or the exon 1-intron 1 boundary” can be changed to: “Predominantly, the mutations are clustered in the 5’ region of exon 2 of the gene with only a few mutations affecting the intron 1-exon 2 boundary or, more rarely, exon 1 or the exon 1-intron 1 boundary”

-The heading “Introducing three tumor entities displaying a unique type of MED12 mutations” should be changed to “Introducing three tumor entities displaying MED12 mutations”

-Under subheading: Uterine Leiomyomas – the most frequent symptomatic human tumors, “Depending on their location it can be distinguished between submucosal, intramural, and subserosal UL” can be changed to “Depending on the location, the difference between submucosal, intramural, and subserosal UL can be distinguished”

-Under the heading “A closer look at the molecular pathogenesis of uterine leiomyomas,” the sentence, “According to the high prevalence of uterine leiomyomas mutations are by far best investigated in this tumor type” can be changed to “MED12 mutations are by far best investigated uterine leiomyoma tumor type.”

-Second paragraph under the subheading: “Leiomyomas with MED12 mutation constitute their own genetic subtype which is also characterized by a distinct clinical and histopathological appearance,” the sentence “Accordingly, both mutations allow the two major genetic subtypes of UL to be distinguished, and the question arises whether or not the genetic subtypes are also reflected by a different clinical behavior and histopathology” can be changed to “Accordingly, these mutations allow the two major genetic subtypes of UL to be distinguished, and the question arises whether or not the genetic subtypes are also have different clinical behaviors and histopathologies”

-At the end of the first paragraph under the heading: “A closer look at the patterns of MED12 mutations seen in various benign and malignant tumors” the word respectively is unnecessary.

-Under “Hypothesis and Opinion” the sentence “Nevertheless, additional factors favoring this multitude of tumors carrying the same type of mutations despite their independent clonal origin have remained enigmatic” should be changed to “Nevertheless, additional factors favoring this multitude of tumors with independent clonal origin carrying the same type of mutation have remained enigmatic.”

-Also, “In addition to these both models possibly accounting for other genetic subtypes, as a third alternative factors as in particular infectious agents warrant consideration” should be changed to “In addition to these models, the potential roles of infectious agents warrant consideration”

-Also, “As to another virus of the Herpes group, a recent study failed to reveal a significant association
between HSV-2 seropositivity and the presence of fibroids (Moore *et al.*, 2016) and in general not convincing evidence for viruses involved in the pathogenesis of UL has been presented" should be changed to “As to another virus of the Herpes group, a recent study failed to reveal a significant association between HSV-2 seropositivity and the presence of fibroids (Moore *et al.*, 2016) and in general no convincing evidence for involvement of viruses in the pathogenesis of UL has been presented”

To summarize, the authors build a fairly convincing argument for the need to more closely study the mechanism of MED12 mutation in tumorigenesis. However, as detailed above, there are multiple areas in which additional evidence and clarification is needed to support the authors’ claims. Additionally, correction of several grammatical errors and awkward phrasing throughout the manuscript would greatly improve readability.

**Is the topic of the opinion article discussed accurately in the context of the current literature?**
Partly

**Are all factual statements correct and adequately supported by citations?**
Partly

**Are arguments sufficiently supported by evidence from the published literature?**
Partly

**Are the conclusions drawn balanced and justified on the basis of the presented arguments?**
Partly

**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 30 April 2018

https://doi.org/10.5256/f1000research.15482.r33313

© 2018 Kurita T. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Takeshi Kurita
Department of Cancer Biology and Genetics, Comprehensive Cancer Center, Ohio State University, Columbus, OH, USA

This Opinion Article proposes a provocative idea that transcripts of *Staphylococcus aureus* contribute to an exceptionally high incidence of MED12 exon 2 mutations among human tumors. After providing an excellent summation of the current literature on demographics and characteristics of MED12 mutations in human neoplasms, the authors propose an intriguing concept: the involvement of mutagenic nucleic acids of bacterial origin in human tumors, based on the discovery of a Staphylococcus genome sequence of which the transcript may form an R-loop with human MED12 gene on the mutation
hotspot. This novel concept merits further discussion among researchers of diverse biomedical fields.

There are several issues that I recommend the authors to address.

I. The importance of the idea would increase if the following issues are resolved.

- The methods of database search for homologous sequences and the statistical significance of the finding should be described. In other words, it is unclear how frequently such homologous sequences to the human genome appear in the genome of pathogenic microorganisms, and what proportion of such homologous sequences can form an R-loop with human genome if they are transcribed.
- The palindromic sequence is located at 30 bp from the 3’ end of tRNA-Leu on the reverse strand in *Staphylococcus aureus* genome. Since it is not a typical template for transcription, inclusion of references that suggest transcriptional activity in such genomic regions of bacteria would help further discussion by increasing the feasibility of the proposed model.
- It would also be helpful to include a discussion on the possible mechanisms through which naked transcripts of low-copy number in the circulation could possibly reach the target locus in the genome of myometrial cells. For instance, internalization of shRNA for RNAi is achieved by viral transduction or chemical/physical transfection of shRNA expression vectors.
- The authors may speculate the mechanisms that unfold the minimum free energy structure of short hairpin transcripts and facilitate the hybridization to the genomic DNA. For example, shRNA is processed into siRNA, and then siRNA forms a complex with cellular proteins to elicit RNAi effect.

II. There are issues in data presentation

- Regarding the title of Figure 2: “Single and multiple MED12-mutation positive uterine leiomyomas”, the adjectives “single and multiple” could modify “MED12-mutation”. Adding a hyphen (*MED12*-mutation-positive) and using of “solitary” instead of “single” would improve the readability.
- Error: The labels for single and multiple UL groups in Figure 2 are switched.
- The graphs in Figures 2, 3 and 4 are the same data in different presentation format. The Figures 3 and 4 seem to be redundant.

III. There are other factors that also likely contribute to the high prevalence of patients with multiple MED12 mutant ULs.

1. **A single hit on the active MED12 allele is sufficient for the pathogenesis of ULs.** Human cancer cells usually carry multiple putative driver mutations even at the earliest stage. In contrast, most MED12 mutant ULs do not carry additional mutations, suggesting a MED12 mutation is sufficient to drive UL pathogenesis. Since the MED12 is on X-chromosome, a single hit on the active allele of MED12 has dominant effect. Hence, the development of UL through MED12 mutations should occur at a significantly higher rate compared to the neoplasms that require multiple genetic lesions for pathogenesis.

2. **Diversity of pathogenic MED12 mutations increases the prevalence of MED12-mutant ULs.** Generally, driver missense mutations of human neoplasms are very specific. For instance, nearly all adult-type granulosa cell tumors carry *FOXL2* c.402C>G (C134W) missense mutation. While single base replacements in the hotspot triplet bases can also result in conversion of C134 to F, S, Y, R and G, only C134W is pathogenic. In contrast, a variety of MED12 mutations, ≥ 10 missense and > 30 indel mutations, are associated with ULs. For instance, missense mutations that convert MED12 G44 to D, S, V, R, C and A are all pathogenic. Hence, even if the mutation rate per nucleotide is equal throughout the human genome, the incidence of pathogenic MED12 mutations should be many times higher than other pathogenic mutations.

3. **UL is a hormone dependent tumor.** Another key factor that this Opinion Article does not address is the systemic hormonal environment. The pathogenesis of UL depends on estrogen and progesterone. Since ULs are counted only when they grow to a grossly recognizable size, patients
with endocrine profiles favorable to the growth of ULs should have a higher number of tumors even if the incidence of pathogenic mutations in myometrial cells is equal among all women. Discussion of these factors is not essential, as the model proposed in this article would work independently. Nevertheless, these additions would help balance the discussion.

Is the topic of the opinion article discussed accurately in the context of the current literature?  
Yes

Are all factual statements correct and adequately supported by citations?  
Yes

Are arguments sufficiently supported by evidence from the published literature?  
Partly

Are the conclusions drawn balanced and justified on the basis of the presented arguments?  
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Biology of Uterine Leiomyoma

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.