



**HAL**  
open science

## GTn repeat microsatellite instability in uterine fibroids

Bineta Kénémé, Mbacké Sembène

► **To cite this version:**

Bineta Kénémé, Mbacké Sembène. GTn repeat microsatellite instability in uterine fibroids. *Frontiers in Genetics*, 2019, <10.3389/fgene.2019.00810>. <hal-02517494>

**HAL Id: hal-02517494**

**<https://hal.science/hal-02517494v1>**

Submitted on 24 Mar 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



# GT<sub>n</sub> Repeat Microsatellite Instability in Uterine Fibroids

Bineta Kénémé<sup>1\*</sup> and Mbacké Sembène<sup>1,2</sup>

<sup>1</sup> GenGesPop, Cheikh Anta Diop University, Animal Biology, Dakar-Fann, Senegal, <sup>2</sup> Biopass, IRD, Dakar-Bel Air, Senegal

**Background:** *Type I collagen* is a triple helix structure with two  $\alpha 1$  and one  $\alpha 2$  chains. Coordinated biosynthesis of  $\alpha 1$  and  $\alpha 2$  subunits is very important for tissue morphogenesis, growth, and repair. In contrast, abnormal deposition in response to proinflammatory cytokines is associated with organ dysfunction. In humans, *COL1A2* contains two microsatellite loci: one located at the 5'-flanking region is composed of poly CA and poly CG; the other located in the 1st intron is constituted of poly GT. Expression of *COL1A2* has been noted in gastric cancer and was positively correlated with degree of invasion and metastases. But no genetic study taking into account polymorphism of *COL1A2* in uterine fibroids has been undertaken.

**Methods:** In this study, repeated dinucleotide GT<sub>n</sub> of intron 1 *COL1A2* was highlighted in 55 patients with uterine fibroids (UF). Clinical and pathological data were obtained from patient's records, and other parameters were recorded. Mutation Surveyor version 5.0.1, DnaSP version 5.10, MEGA version 7.0.26, and Arlequin version 3.5.1.3 were used to determine genetics parameters. To estimate genetic variation according to epidemiological parameters, index of genetic differentiation (Fst) and genetic structure (AMOVA) were determined with Arlequin version.

**Results:** Based on reference microsatellite pattern (GT)<sub>14</sub>CT(GT)<sub>3</sub>CT(GT)<sub>3</sub>, 15 haplotypes were found. Among the 15 haplotypes, 12 have mutation at position **2284C > G** and 7 at position **2292C > G**. Insertions of repeated dinucleotide GT<sub>n</sub> were found on three haplotypes against eight haplotypes in which they are deletions. Intron 1 of *COL1A2* gene exhibits high genetic diversity in uterine fibroids with 35.34% polymorphic sites, 95.74% of which were parsimoniously variable and an average number of nucleotide difference of 10.442, which reflects an important genetic variability. According to epidemiological parameters, our results showed, for the first time, a genetic structuring of uterine fibroids according to ethnicity, marital status, use of contraception, diet, and physical activity, beyond confirming the involvement dinucleotide length polymorphism GT<sub>n</sub> in occurrence of uterine fibroids in Senegalese women.

**Conclusion:** Results obtained open up avenues for understanding the mechanisms involved in the racial variation in the prevalence of uterine fibroids as well as the predisposing factors.

**Keywords:** uterine fibroid, *COL1A2* polymorphism, risk factors, Senegal, microsatellite genetic marker

## OPEN ACCESS

### Edited by:

Barbara Karen Dunn,  
National Institutes of Health (NIH),  
United States

### Reviewed by:

Pawel Buczkowicz,  
Gene42, Inc., Canada  
Jun Zhong,  
National Cancer Institute (NCI),  
United States

### \*Correspondence:

Bineta Keneme  
bineta.keneme@ucad.edu.sn

### Specialty section:

This article was submitted to  
Cancer Genetics,  
a section of the journal  
Frontiers in Genetics

**Received:** 28 January 2019

**Accepted:** 02 August 2019

**Published:** 20 September 2019

### Citation:

Kénémé B and Sembène M (2019)  
GT<sub>n</sub> Repeat Microsatellite Instability  
in Uterine Fibroids.  
Front. Genet. 10:810.  
doi: 10.3389/fgene.2019.00810

## INTRODUCTION

Uterine fibroids (UF), more commonly known as myomas or uterine leiomyomas, are the most common benign tumors of female reproductive organs. They are associated with significant morbidity and therefore constitute a real public health problem. UF, which are highly variable within uterus, develop at the expense of smooth muscle and are often separated from the myometrium by a pseudocapsule associated with connective tissue condensation (Audebert, 1990). Heterogeneity of UF localization and their progression in the same patient illustrate the complex biological mechanism involved in their development. Clinically, UF are firm, stiff nodular tumours, a fact confirmed by biomechanical studies. Proteins of extracellular matrix, especially interstitial collagens, are responsible for this property of “firmness” and mechanical strength of tissue. Indeed, UF have an accumulation of altered collagens and different amounts of glycosaminoglycans and a proliferation of cells, which is by definition a fibrosis. A complete understanding of the role of extracellular matrix proteins, in particular collagen, and their effect on the growth and development of UF becomes an important issue for elucidating molecular mechanisms involved in their etiology.

Located on chromosome 7, *COL1A2* is an essential component of matrix tissue. It is predominantly produced by mesenchymal cells such as fibroblasts, osteoblasts, and smooth muscle cells. Transcription of *COL1A2* is under control of a regulatory complex that includes several DNA elements and several trans-activating factors. During the last two decades, *type I alpha chain collagen 2 (COL1A2)* has been considered as an informative model for studying principles that govern the control of extracellular matrix transcription for normal and fibrotic tissues (Kirkland, 2009; Krasny et al., 2010; Trojonowska, 2002; Yasul et al., 2004). In humans, *COL1A2* contains two microsatellite loci: one located at the 5'-flanking region of the gene is composed of poly CA and poly CG; the other located in the 1st intron is constituted of poly GT. In a study led by Akai et al. (1999), it has been shown that complete transcription of *COL1A2* gene is regulated by these repeated dinucleotides. Analysis of polymorphism in these two regions indicates that these two sequences show a variation in their repetition number, suggesting that these dinucleotides constitute microsatellites. Lei et al. (2005) hypothesized that GT<sub>n</sub> polymorphism triggers transcription of the gene, and variation in the number of repetitions can partly be responsible for the difference in transcriptional activity. In this study, we evaluate instability of repeated dinucleotide GT<sub>n</sub> in Senegalese patients with UF.

## MATERIALS AND METHODS

### Clinical Sampling

Tumor tissue samples were collected from 55 patients with UF (from Military Hospital of Ouakam and General Hospital of Grand Yoff). Clinical and pathological data were recorded including age, ethnicity, age at menarche, marital status, number of pregnancies, number of childbirth, hormonal contraception,

diet, and physical activity (Table 1). None of the patients surveyed claimed to consuming alcohol and using tobacco, which is why these factors are not included in this study.

### DNA Extraction, Amplification, and Sequencing of Intron 1 *COL1A2* Gene

Total DNA of each sample was extracted using Qiagen protocol (Qiagen Dneasy Tissue kit). After extraction, repeated dinucleotide GT<sub>n</sub> were amplified using forward 5'-TGTCT ACCACTGCATAATTTC-3 and reverse 5'-AATATGAACTCG GTAATGTGA-3' primers (Lei et al., 2005). The 35 cycle PCR for *COL1A2* intron 1 amplification was carried out using 4 μl of human genomic DNA in a 50 μl reaction mixture, which contained 0.1 μl of Taq DNA polymerase, 2.5 μl of forward and reverse primers, 1 μl of magnesium chloride, 2 μl of mix dNTPs, and 5 μl of 10X ammonium sulfate buffer. Thermal cycle conditions for amplification PCR consisted of 1st step-3 min

**TABLE 1 |** Clinical and pathological characteristics of 55 cases analyzed.

| Epidemiological factors                | Number of patients (%) |
|--|------------------------|
| <b>Age (n = 36)</b>                    |                        |
| ≤35                                    | 11 (30.55%)            |
| ]35–45]                                | 18 (50%)               |
| > 45                                   | 7 (19.45%)             |
| <b>Ethnicity (n = 39)</b>              |                        |
| Wolof                                  | 13 (33.33%)            |
| Sérère                                 | 4 (10.26%)             |
| Lébou                                  | 7 (17.95%)             |
| Bambara                                | 3 (7.69%)              |
| Diola                                  | 5 (12.82%)             |
| Alpulaar                               | 7 (17.95%)             |
| <b>Marital status (n = 31)</b>         |                        |
| Single                                 | 8 (25.80%)             |
| Married                                | 20 (64.52%)            |
| Divorced                               | 3 (9.68%)              |
| <b>Age at menarche (n = 18)</b>        |                        |
| ≤12                                    | 1 (5.56%)              |
| ]12–15]                                | 13 (72.22%)            |
| > 15                                   | 4 (22.22%)             |
| <b>Number of pregnancies (n = 31)</b>  |                        |
| 0                                      | 20 (64.51%)            |
| I                                      | 4 (12.91%)             |
| II                                     | 4 (12.91%)             |
| III                                    | 1 (3.22%)              |
| > III                                  | 2 (6.45%)              |
| <b>Number of childbirth (n = 33)</b>   |                        |
| 0                                      | 23 (69.70%)            |
| I                                      | 7 (21.21%)             |
| II                                     | 1 (3.03%)              |
| III                                    | 2 (6.06%)              |
| > III                                  | 0 (0%)                 |
| <b>Hormonal contraception (n = 23)</b> |                        |
| Yes                                    | 2 (8.69%)              |
| No                                     | 21 (91.31%)            |
| <b>Diet (n = 23)</b>                   |                        |
| Meat preference                        | 7 (30.43%)             |
| Vegetarian preference                  | 6 (26.09%)             |
| No preference                          | 10 (43.48%)            |
| <b>Physical activity (n = 23)</b>      |                        |
| Yes                                    | 5 (21.74%)             |
| No                                     | 18 (78.26%)            |

cycle of initial denaturation at a temperature of 94°C, followed by 2nd step consisting of 35 cycles each of 45 s of denaturation at 94°C, annealing at 60°C/1min and primer extension at 72°C/1 min, and 3rd step: final extension or polymerization at 72°C for 10 min. After PCR reaction, all products were electrophoresed on 1.5% agarose gel, followed by its analysis in an UVitec Gel Documentation system for imaging the gel and to determine the amplicon lengths. Sequencing reactions were performed in a thermal cycler MJ Research PTC-225 Peltier type with ABI PRISM BigDye TM Terminator Cycle kit. Each sample was sequenced using forward primer. Fluorescent fragments were purified with the BigDye Xterminator purification protocol. The samples were suspended in distilled water and subjected to electrophoresis in 3730xl ABI sequencer (Applied Biosystems).

## Molecular Analysis

To determine length polymorphism of dinucleotide GT<sub>n</sub> of intron 1 COL1A2 gene, the raw sequencing data were submitted to Mutation Surveyor software version 5.0.1 (www.softgenetics.com). This program can directly compare chromatograms with genomic DNA of reference sequence of COL1A2 (NT\_007933\_94023373). Alignment of the sequences was carried out using BioEdit software version 8.0.5 and ClustalW algorithm (Thompson et al., 1994). Sequences obtained (Hall, 1999) were thoroughly checked, cleaned, and aligned to identify homologies among sites, and also to perform other phylogenetic analysis including the determination of variability index and genetic diversity as well as the parameters of genetic differentiation. Genetic variability parameters (number of polymorphic sites, total number of haplotype, average number of nucleotide difference K) were obtained through DnaSP 5.10 software (Librado and Rozas, 2009) and MEGA 7.0.26 (Kumar et al., 2016). To estimate genetic variation according to epidemiological parameters, the factor of genetic differentiation (Fst) and the analysis of molecular variance (AMOVA) were determined with Arlequin software version 3.5.1.3 (Excoffier and Lischer, 2010). Values of P less than 0.05 are considered significant at a 5% confidence interval.

## RESULTS

### Mutations Status of Microsatellite GT<sub>n</sub>

COL1A2 was sequenced in 55 tumour tissues. Of these sequences, five were removed from the genetic analysis because of a strong polymorphism. Based on the microsatellite reference pattern in the form (GT)<sub>14</sub>CT(GT)<sub>3</sub>CT(GT)<sub>3</sub>, 15 haplotypes were found in 50 Senegalese women with UF (Table 2). These haplotypes indicate a variation in GT repetition number ranging from 13 to 25 (Figure 1).

Of the 15 haplotypes, 12 have mutation at position 2284C > G (first site that interrupts GT dinucleotide repeat) and 7 have mutation at 2292C > G (2nd site that interrupts GT dinucleotide repeat). Insertions of repeated dinucleotide GT<sub>n</sub> were found on three haplotypes (microsatellite elongation) compared to eight haplotypes in which they were deletions (microsatellite shortening). Haplotype 7 representing 20% of the haplotypes was characterized by the presence of two types of transversions 2284C > G and 2292C > G. Haplotypes 11 (14%) and 8 (12%) were respectively characterized by deletions at position 2280-2283\_DelGTGT and 2282-2283\_DelGT (Table 2). Some microsatellite length polymorphisms of GT<sub>n</sub> COL1A2 were summarized in Figure 2.

### COL1A2 Intron 1 Polymorphisms

Intron 1 of COL1A2 gene exhibits high genetic diversity in UF with 35.34% polymorphic sites, 95.74% of which were parsimoniously variable. Average number of nucleotide differences was 10.442 (Table 3). The high haplotypic diversity (Hd = 0.9984) and the low nucleotide diversity (Pi = 0.0877) showed a rapid evolution of microsatellite polymorphism in UF in Senegalese women.

### Microsatellite GT<sub>n</sub> Instability and Genetic Differentiation

Depending on epidemiological parameters studied, only repeated dinucleotide is taken into account in analysis. This allows us to highlight the role of the length polymorphism of intron 1 COL1A2. Results obtained show a variable expressivity of dinucleotide GT<sub>n</sub> in

**TABLE 2** | Length and pattern polymorphism of repeated dinucleotide GT<sub>n</sub> of intron 1 COL1A2 gene in uterine fibroids.

| Haplotype | Number (%) | Microsatellite pattern                                     | Variants                                    |
|-----------|------------|--|---|
| H1        | 5 (10%)    | (GT) <sub>14</sub> CT(GT) <sub>3</sub> CT(GT) <sub>3</sub> | Wide type                                   |
| H2        | 4 (8%)     | (GT) <sub>17</sub> CT(GT) <sub>3</sub> CT(GT) <sub>3</sub> | 2276_2281_InsGTGTGT                         |
| H3        | 2 (4%)     | GT <sub>25</sub>   | 2276_2281_InsGTGTGT; 2284C > G; 2292C > G   |
| H4        | 1 (2%)     | (GT) <sub>21</sub> CT(GT) <sub>3</sub>                     | 2276_2281_InsGTGTGT; 2284C > G              |
| H5        | 2 (4%)     | GT <sub>23</sub>   | 2281-2282_InsGT; 2284C > G; 2292C > G       |
| H6        | 2 (4%)     | (GT) <sub>18</sub> CT(GT) <sub>3</sub>                     | 2284C > G                                   |
| H7        | 10 (20%)   | GT <sub>22</sub>   | 2284C > G; 2292C > G                        |
| H8        | 6 (12%)    | (GT) <sub>21</sub>   | 2282-2283_DelGT; 2284C > G; 2292C > G       |
| H9        | 2 (4%)     | (GT) <sub>13</sub> CT(GT) <sub>3</sub> CT(GT) <sub>3</sub> | 2282-2283_DelGT                             |
| H10       | 1 (2%)     | (GT) <sub>17</sub> CT(GT) <sub>3</sub>                     | 2282-2283_DelGT; 2284C > G                  |
| H11       | 7 (14%)    | (GT) <sub>16</sub> CT(GT) <sub>3</sub>                     | 2280-2283_DelGTGT; 2284C > G                |
| H12       | 4 (8%)     | GT <sub>20</sub>   | 2280-2283_DelGTGT; 2284C > G; 2292C > G     |
| H13       | 1 (2%)     | GT <sub>19</sub>   | 2276-2281_DelGTGTGT; 2284C > G; 2292C > G   |
| H14       | 2 (4%)     | (GT) <sub>15</sub> CT(GT) <sub>3</sub>                     | 2276-2281_DelGTGTGT; 2284C > G              |
| H15       | 1 (2%)     | GT <sub>18</sub>   | 2274-2281_DelGTGTGTGT; 2284C > G; 2292C > G |

Ins, insertion; Del, deletion.



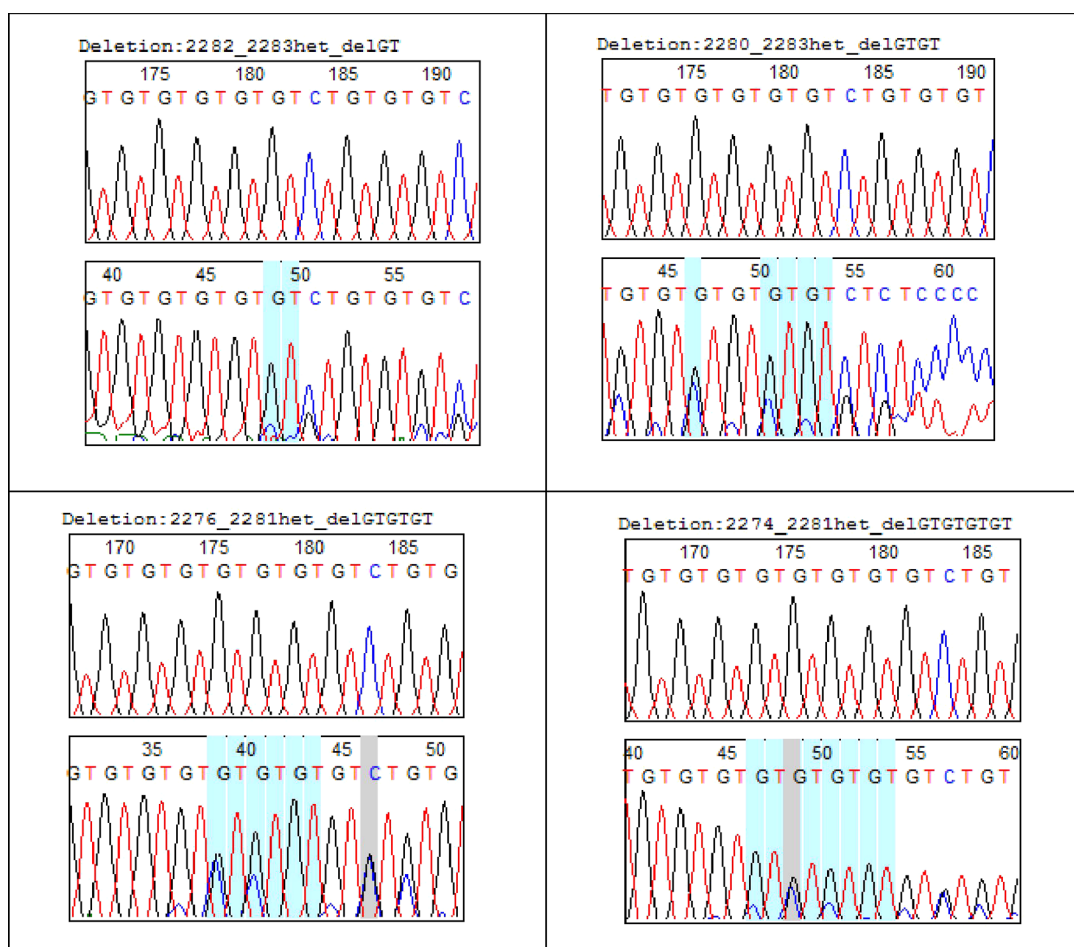


FIGURE 2 | GT<sub>n</sub> COL1A2 deletion in uterine fibroids.

TABLE 3 | Index of variability and genetic diversity of intron 1 COL1A2 in fibroid cases.

| Variability index                            |                  |            |
|--|------------------|------------|
| Parameters                                   | Number           | Percentage |
| Number of sequences                          | 50               |            |
| Number of sites                              | 133              |            |
| Monomorphic sites                            | 86               | 64.66%     |
| Polymorphic sites                            | 47               | 35.34%     |
| Singleton variable sites                     | 2                | 4.26%      |
| Parsimony informative sites                  | 45               | 95.74%     |
| Average number of nucleotide differences (k) | 10.442           |            |
| Genetic diversity index                      |                  |            |
| Pi ± variance                                | 0.0877 ± 0.00002 |            |
| Hd ± variance                                | 0.9984 ± 0.00003 |            |

Hd, haplotypic diversity; Pi, nucleotide diversity.

are modulated by risk factors such as ethnicity, marital status, hormonal contraception, diet, and physical activity. Indeed, the polymorphism of the repeated dinucleotide GT<sub>n</sub> of intron 1 COL1A2 in UF is explained to:

- 8.83% by differentiation between women of different ethnic groups;
- 10.57% by a differentiation between women according to their marital status;
- 24.88% by differentiation according to whether or not use of hormonal contraception;
- 7.94% by dietary preference; and
- 15.10% by a differentiation according to physical activity.

Since there is no multivariate analysis and the sample sizes are small for some of these variables, more research is needed to highlight these results.

## DISCUSSION

### COL1A2 Polymorphisms in Uterine Fibroids

Located on chromosome 7, COL1A2 is an essential component of the tissue matrix. It is predominantly produced by mesenchymal cells such as fibroblasts, osteoblasts, and smooth muscle cells (Rossert et al., 2000). Transcription of COL1A2 is under control

**TABLE 4 |** Degree of genetic differentiation of repeated dinucleotide GT<sub>n</sub> of COL1A2 gene in relation to the epidemiological parameters studied.

| Epidemiological parameters    | Genetic differentiation (Fst) |                    |                    |
|-------------------------------|-------------------------------|--------------------|--------------------|
|                               | Within sub-groups             | Between sub-groups |                    |
| Groups<br>Sub-groups          | Fst                           | Between sub-groups | Fst (P-value)      |
| <b>Age</b>                    |                               |                    |                    |
| ≤35                           | 0.04237                       | ≤35 & ]35–45]      | 0.01274 (0.32422)  |
| ]35–45]                       | 0.03152                       | ≤35 & >45          | 0.20818 (0.07812)  |
| > 45                          | 0.06574                       | ]35–45] & >45      | –0.02089 (0.46582) |
| <b>Ethnicity</b>              |                               |                    |                    |
| Wolof                         | 0.00180                       | Wolof & Sérère     | 0.15238 (0.17773)  |
| Sérère                        | 0.20123                       | Wolof & Lébou      | 0.06714 (0.29785)  |
| Lébou                         | 0.14805                       | Wolof & Bambara    | –0.03106 (0.99902) |
| Bambara                       | 0.30276                       | Wolof & Diola      | 0.10515 (0.19531)  |
| Diola                         | –0.00183                      | Wolof & Alpulaar   | 0.18605 (0.07324)  |
| Alpulaar                      | 0.11805                       | Sérère & Lébou     | –0.16129 (0.76562) |
|                               |                               | Sérère et Bambara  | 0.06667 (0.99902)  |
|                               |                               | Sérère & Diola     | 0.15152 (0.15918)  |
|                               |                               | Sérère & Alpulaar  | 0.12513 (0.12891)  |
|                               |                               | Lébou & Bambara    | –0.33333 (0.99902) |
|                               |                               | Lébou & Diola      | –0.09091 (0.60938) |
|                               |                               | Lébou & Alpulaar   | –0.08691 (0.72852) |
|                               |                               | Bambara & Diola    | –0.40000 (0.99902) |
|                               |                               | Bambara & Alpulaar | –0.52941 (0.99902) |
|                               |                               | Diola & Alpulaar   | 0.13754 (0.11523)  |
| <b>Marital status</b>         |                               |                    |                    |
| Single                        | 0.08964                       | Single & Married   | 0.03216 (0.24805)  |
| Married                       | 0.10771                       | Single & Divorcée  | 0.39683 (0.99902)  |
| Divorced                      | 0.23556                       | Married & Divorced | 0.38444 (0.99902)  |
| <b>Age at menarche</b>        |                               |                    |                    |
| ≤12                           | 0.20408                       | ≤12 & ]12 – 15]    | –0.45799 (0.91992) |
| ]12 – 15]                     | –0.14031                      | ≤12 & >15          | 0.00000 (0.39453)  |
| >15                           | –0.08291                      | ]12 – 15] & >15    | –0.01850 (0.48242) |
| <b>Number of pregnancies</b>  |                               |                    |                    |
| 0                             | 0.00988                       | 0 & I              | –0.06950 (0.70508) |
| I                             | 0.02840                       | 0 & II             | 0.00284 (0.38184)  |
| II                            | –0.11721                      | 0 & III            | –0.01720 (0.49707) |
| III                           | 0.34321                       | 0 & > III          | 0.42177 (0.99902)  |
| > III                         | 0.34321                       | I & II             | –0.35429 (0.80273) |
|                               |                               | I & III            | –0.10092 (0.70703) |
|                               |                               | I & > III          | –0.17647 (0.99902) |
|                               |                               | II & III           | 0.13333 (0.31641)  |
|                               |                               | II & > III         | –1.00000 (0.99902) |
|                               |                               | III & > III        | 1.00000 (0.99902)  |
| <b>Number of childbirth</b>   |                               |                    |                    |
| 0                             | –0.06871                      | 0 & I              | 0.07169 (0.18652)  |
| I                             | –0.13999                      | 0 & II             | –0.41520 (0.99902) |
| II                            | 0.26463                       | 0 & III            | –0.41520 (0.99902) |
| III                           | 0.26463                       | I & II             | –0.16049 (0.99902) |
|                               |                               | I & III            | –0.16049 (0.99902) |
|                               |                               | II & III           | 0.00000 (0.99902)  |
| <b>Hormonal contraception</b> |                               |                    |                    |
| Yes                           | 0,17376                       | Yes & No           | 0.24883 (0.07129)  |
| No                            | –0,07280                      |                    |                    |
| <b>Diet</b>                   |                               |                    |                    |
| 1 Meat preference             | 0.01042                       | 1 & 2              | 0.17374 (0.07520)  |
| 2 Vegetarian preference       | –0.01543                      | 1 & 3              | –0.11892 (0.92285) |
| 3 No preference               | 0.01406                       | 2 & 3              | 0.20431 (0.04785)  |
| <b>Physical activity</b>      |                               |                    |                    |
| Yes                           | 0.14435                       | Yes & No           | 0.15107 (0.06445)  |
| No                            | 0.07661                       |                    |                    |

**TABLE 5 |** Genetic structuring of GT<sub>n</sub> COL1A2 according to epidemiological parameters.

| Epidemiological parameters | Source of variation | Percentage of variation | Fst (P-value)      |
|----------------------------|---------------------|-------------------------|--------------------|
| Age                        | Within sub-groups   | 96.03862                | 0.03961 (0.06843)  |
|                            | Between sub-groups  | 3.96138                 |                    |
| Ethnicity                  | Within sub-groups   | 91.16560                | 0.08834 (0.03421)  |
|                            | Between sub-groups  | 8.83440                 |                    |
| Marital status             | Within sub-groups   | 89.42471                | 0.10575 (0.00782)  |
|                            | Between sub-groups  | 10.57529                |                    |
| Age at menarche            | Within sub-groups   | 108.72040               | -0.08720 (0.88368) |
|                            | Between sub-groups  | -8.72040                |                    |
| Number of pregnancies      | Within sub-groups   | 95.80665                | 0.04193 (0.21994)  |
|                            | Between sub-groups  | 4.19335                 |                    |
| Number of childbirth       | Within sub-groups   | 105.40304               | -0.05403 (0.79570) |
|                            | Between sub-groups  | -5.40304                |                    |
| Hormonal contraception     | Within sub-groups   | 75.11658                | 0.24883 (0.00098)  |
|                            | Between sub-groups  | 24.88342                |                    |
| Diet                       | Within sub-groups   | 92.05111                | 0.07949 (0.04301)  |
|                            | Between sub-groups  | 7.94889                 |                    |
| Physical activity          | Within sub-groups   | 84.89311                | 0.15107 (0.00684)  |
|                            | Between sub-groups  | 15.10689                |                    |

of a regulatory complex that includes several DNA elements and several trans-activating factors. In humans, *COL1A2* contains two microsatellite loci: one located at the 5'-flanking region of the gene is composed of poly CA and poly CG; the other located in the 1st intron is constituted of poly GT. In this study, microsatellite polymorphism GT<sub>n</sub> of intron 1 *COL1A2* was highlighted in cases of UF in Senegalese women. Based on microsatellite reference pattern that is (GT)<sub>14</sub>(CT)(GT)<sub>3</sub>(CT)(GT)<sub>3</sub>, 15 haplotypes were found. These haplotypes indicate a variation in number of GT repeats ranging from 13 to 25. Of the 15 haplotypes, 12 have the 2284C > G mutation and 7 have the 2292C > G mutation. Insertions of the dinucleotide GT<sub>n</sub> were found on three haplotypes (microsatellite elongation) compared to eight haplotypes in which they were deletions (microsatellite shortening). This suggests an altered mechanism of the role of *COL1A2* gene in UF. Indeed, in a study led by Lei et al. (2005), it has been hypothesized that intron 1 GT<sub>n</sub> polymorphism triggers transcription of gene and variation in number of repeats may be partly responsible for the difference in transcriptional activity. In addition, about 200 different chromosomal abnormalities have been described in UF including long-arm translocations of chromosome 7 occurring in about 17% of karyotypically abnormal UF (Sandberg, 2005). In contrast to normal tissues where collagen is organized into long, thin, wavy fibrils parallel to the epithelial boundary, collagen fibrils in the tumor stroma are thicker and shorter (Cho et al., 2015). In epithelial ovarian cancer, collagenous pathways perpendicular to the epithelial boundary have been observed (Adur et al., 2014).

Intron 1 of *COL1A2* gene exhibits high genetic diversity in UF with 35.34% polymorphic sites, 95.74% of which are parsimoniously variable and an average number of nucleotide differences of 10.442, which reflects an important genetic variability. This could be explained by the fact that compared to the myometrium, in UF, not only expression of collagen genes increases (Stewart et al., 1994), but also amount of mature reticulated collagen protein is increased and the more

important is modified (Leppert et al., 2004). UF are firm, stiff nodular tumors, a fact understood by all clinicians and confirmed by biomechanical studies (Rogers et al., 2008; Jayes et al., 2013). Extracellular matrix (ECM) proteins, including interstitial collagens, are responsible for this property of "firmness" and mechanical strength of tissues. ECM is a structure that has a supporting role, but on the other hand, it provides signals to cells that determines their behavior. The role of ECM and mechanotransduction as an important signaling factor in human uterus is just beginning to be appreciated. ECM is not just substance surrounding cells, but rigidity compresses cells or stretches them into signals converted into chemical changes, depending on amount of collagen, crosslinking and hydration, as well as other components of ECM. Since connective tissue integrity, architecture, and function result from specific interactions between collagen and other components of ECM, the presence of abnormal collagen chains may have a strong influence on metabolism of non-collagenic components (Tenni et al., 1988). According to study by Hauptman et al. (2018) in colorectal cancers, the results showed that 9 of 16 genes that show differential expression in carcinomas compared to adenomas are components of ECM. Among these components, two collagen type I proteins (COL1A1, COL1A2) are significantly over-regulated in cancerous tissues compared to normal tissues. Studies on cell lines suggest that type I collagen adhesion promotes intracellular signaling pathways.

## Microsatellite GT<sub>n</sub> Instability in Uterine Fibroids: Correlation With Epidemiological Parameters Ethnicity

In addition to great variability, repeated dinucleotide GT<sub>n</sub> of intron 1 *COL1A2* exhibits heterogeneity given to clinico-pathological parameters in women with UF. Heterogeneity of predisposing factors involved in UF illustrates the complex biological mechanism involved in their development. This suggests the involvement of several molecular mechanisms in occurrence of UF. Prospective studies with larger number of samples would strengthen the correlation observed in the current study. Epidemiological data have mentioned racial disparity in occurrence of UF. Ethnicity has a major influence on development and clinical severity of UF. African-American women develop UF at higher frequency and with more severe symptoms. Hispanic women have an intermediate disease profile, and Caucasian women are the least severely affected ethnic group (Velebil et al., 1995; Baird et al., 2003; Wise et al., 2012). It appears that increased incidence and severity of disease in African-American women may be due to a combination of specific genetic and environmental factors that are not independent risk factors for the disease (Commandeur et al., 2015). In this study, we took ethnicity into account, although these women are all black. UF are genetically heterogeneous in Bambara, Sérère, Lébou, and Alpelaar and more homogeneous in Diola and Wolof. In a study conducted by Thiaw (2018) on ethnic diversity of Senegalese population (unpublished data), analysis of GT<sub>n</sub> pattern polymorphism shows a genetic differentiation between Diola and Wolof compared to other

ethnic groups. This differentiation may explain in part 8.83% of genetic structure of *COL1A2* observed in UF by ethnicity.

### Marital Status

Genetic differentiation observed (10.57%) is also explained by the differentiation between women according to their marital status; greater differentiation is observed among married and divorced women compared to single. This could be explained by the difference in hormonal status in these women. Studies of Barrett et al. (2015) on ovarian steroid status in marital status showed that estradiol was higher among married women than among unmarried women ( $\beta = 0.19$ , 95% CI: 0.02–0.36) as well as progesterone ( $\beta = 0.19$ , 95% CI: 0.01–0.39). In addition, many clinical observations indicate that the development of UF is related to hormonal status (Ross et al., 1986). For example, UF do not occur in prepubertal women and are rarely seen in adolescent girls (Fields and Neinstein, 1996).

### Hormonal Contraception

Our results also indicated that 24.88% of differentiation observed in cases of UF is explained by a differentiation following hormonal contraception use. Relationship between oral contraceptives and UF has been largely elucidated. But epidemiological data between contraceptive use and UF seems controversial. Published studies show a reduction or absence of risk between oral contraceptives use combined with appearance of UF (Berisavac et al., 2009). One study has shown that oral contraception may play a role in development of UF. Others have found no association between occurrence of UF and use of contraception (Parazzini et al., 1992).

### Diet

In relation to diet, a positive correlation was noted between dietary preference and genetic expression of *COL1A2* in UF (7.94% of genetic differentiation). Genetic differentiation is more observed in patients with meat preference. Recently, Wise and Laughlin-Tommaso (2016) published results on relationship between dietary fat intake and UF risk in African-American women, confirming an increased risk associated with consumption of omega-3 fatty acids long chain. They validated hypothesis that a diet rich in fruits and vegetables reduced risk. According to studies of Chiaffarino (1999), women with UF consume beef, other red meats, and ham more frequently and have less frequent consumption of green vegetables, fruits, and fish. Multivariate rib ratios were 1.7 for beef and other red meats, 1.3 for ham, and 0.8 for fruit consumption. Limitation of this current diet study is the lack of data on total energy intake because information was collected only on the frequency of vegetable consumption compared to red meat and in interviews with patients. Further research would be interesting to evaluate the effect of fat intake on uterine fibroids biology.

### Physical Activity

There have been few studies on effect of physical activity on risk of developing UF. Nevertheless, our results showed a genetic

structuring of UF according to practice or not of sport (15.10% of genetic differentiation). Since this is a modifiable factor, more research is needed to evaluate effects of physical activity on UF biology.

## CONCLUSION

Results obtained show, for the first time, a genetic structuring of UF according to ethnicity, marital status, use of contraception, diet, and physical activity, beyond confirming the involvement of *COL1A2* gene, in particular dinucleotide length polymorphism GT<sub>n</sub> in occurrence of UF in Senegalese women. In addition to this, results obtained open up avenues for understanding the mechanisms involved in racial variation in the prevalence of UF as well as the predisposing factors. Given the admitted results, it is clear that more research is needed to determine risk factors associated with appearance and growth of UF, as they cause significant morbidity and affect quality of life. A clear overview of the epidemiology of UF has not yet been realized and future research on modifiable risk factors such as vegetarian diet, contraception, physical activity, among others could inform the prevention of myomas and provide new non-surgical approaches to treatment.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of World Medical Association's Declaration of Helsinki. The protocol was approved by the Institutional Ethics Committee on Human Research of Cheikh Anta Diop University (Reference: Protocol 0267/2017/CER/UCAD). All subjects gave written informed consent according to a standardized form.

## AUTHOR CONTRIBUTIONS

BK performed molecular analysis, organized the database, performed data analysis, and wrote the first draft of the manuscript. MS contributed to conception and design of the study, revised the manuscript, and read and approved the submitted version.

## ACKNOWLEDGMENTS

We acknowledge the African Center of Excellence for Mother and Child Health (ACE-MCH), Grant number B041715-P00041/2017 and University Cheikh Anta Diop, Dakar, UCAD (<http://www.ucad.sn>) for technical support. We are most grateful to all Senegalese women who participated in the present study. We are extremely grateful to Dr Daouda CISS Dr Sidy KA and Pr Ahmadou DEM who helped with the collection of samples. Also Pr SEMBENE the head of molecular biology platform of BIOPASS institute of Senegal for all the molecular studies done.

## REFERENCES

- Adur, J., Pelegati, V. B., and De Thomaz, A. A. (2014). Second harmonic generation microscopy as a powerful diagnostic imaging modality for human ovarian cancer. *J. Biophotonics*. 7, 37–48. doi: 10.1002/jbio.201200108
- Akai, J., Kimura, A., and Hata, R. I. (1999). Transcriptional regulation of the human type I collagen  $\alpha 2$  (*COL1A2*) gene by the combination of two dinucleotide repeats. *Gene* 239, 65–73. doi: 10.1016/S0378-1119(99)00380-7
- Audebert, A. (1990). Endométriase externe: histogénèse, étiologie et évolution naturelle. *Rev. Praticien*. 40, 1077–1081.
- Baird, D. D., Dunson, D. B., Hill, M. C., Cousins, D., and Schectman, J. M. (2003). High cumulative incidence of uterine leiomyoma in black and white women: ultrasound evidence. *Am. J. Obstet. Gynecol.* 188, 100–107. doi: 10.1067/mob.2003.99
- Barrett, E. S., Tran, V., Thurston, S. W., Frydenberg, H., Lipson, S. F., Thune, I., et al. (2015). Women who are married or living as married have higher salivary estradiol and progesterone than unmarried women. *Am. J. Hum. Biol.* 27 (4), 501–507. doi: 10.1002/ajhb.22676
- Berisavac, M., Sparic, R., and Argirovic, R. (2009). Contraception: modern trends and controversies. *Srp. Arh. Celok. Lek.* 137 (5–6), 310–319. doi: 10.2298/SARH0906310B
- Chiapparino, F. (1999). Diet and uterine myomas. *Obstet. Gynecol.* 94 (3), 395–398. doi: 10.1016/S0029-7844(99)00305-1
- Cho, A., Howell, V. M., and Colvin, E. K. (2015). The extracellular matrix in epithelial ovarian cancer—a piece of a puzzle. *Front. Oncol.* 5, 1–16. doi: 10.3389/fonc.2015.00245
- Commandeur, A. E., Styer, A. K., and Teixeira, J. M. (2015). Epidemiological and genetic clues for molecular mechanisms involved in uterine leiomyoma development and growth. *Hum. Reprod. Update.* 21 (5), 593–615. doi: 10.1093/humupd/dmv030
- Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resources*. 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x
- Fields, K. R., and Neinstein, L. S. (1996). Uterine myomas in adolescents: case reports and a review of the literature. *J. Pediatr. Adolesc. Gynecol.* 9, 195–198. doi: 10.1016/S1083-3188(96)70030-X
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 95–98.
- Hauptman, N., Boštjancic, E., Zljajpah, M., Rankovic, B., and Zidar, N. (2018). Bioinformatics analysis reveals most prominent gene candidates to distinguish colorectal adenoma from adenocarcinoma. *BioMed. Res. Intern.* 2018, article 9416515, 10. doi: 10.1155/2018/9416515
- Jayes, F. L., Ma, X., Flannery, E. M., Moutos, F. T., Guilak, F., and Leppert P. C. (2013). Biomechanical evaluation of human uterine fibroids after expo-sure to purified clostridial collagenase. In: *Supplement to Biology of Reproduction for the 46th Annual Meeting of the Society for the Study of Reproduction*; 2013 July 22–26; Montreal, Canada.
- Kirkland, S. C. (2009). Type I collagen inhibits differentiation and promotes a stem cell-like phenotype in human colorectal carcinoma cells. *BJC.* 101 (2), 320–326. doi: 10.1038/sj.bjc.6605143
- Krasny, L., Shimony, N., and Tzukert, K. (2010). An in-vitro tumour microenvironment model using adhesion to type I collagen reveals Akt-dependent radiation resistance in renal cancer cells. *Neph. Dial. Transpl.* 25 (2), 373–380. doi: 10.1093/ndt/gfp525
- Kumar, S., Stecher, G., and Tamura, K. (2016). Molecular evolution genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874.
- Lei, S. F., Denga, F. Y., Xiaoa, S. M., Chena, X. D., and Deng, H. W. (2005). Association and haplotype analyses of the *COL1A2* and *ER- $\alpha$*  gene polymorphisms with bone size and height in Chinese. *Bone* 36, 533–541. doi: 10.1016/j.bone.2004.11.002
- Leppert, P. C., Baginski, T., Prupas, C., Catherino, W. H., Pletcher, S., and Segars, J. H. (2004). Comparative ultrastructure of collagen fibrils in uterine leiomyomas and normal myometrium. *Fertil. Steril.* 82 (3), 1182–1187. doi: 10.1016/j.fertnstert.2004.04.030
- Librado, P., and Rozas, J. (2009). Dna SP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452.
- Parazzini, F., Negri, E., La Vecchia, C., Fedele, L., Rabaiotti, M., and Luchini, L. (1992). Oral contraceptive use and risk of uterine fibroids. *Obstet. Gynecol.* 79, 430–433. doi: 10.1097/00006250-199203000-00021
- Rogers, R., Norian, J., Malik, M., Abu-Asab, M., Christman, G., Malik, M., et al., et al. (2008). Mechanical homeostasis is altered in uterine leiomyoma. *Am. J. Obstet. Gynecol.* 198 (4), 1–474. doi: 10.1016/j.ajog.2007.11.057
- Ross, R. K., Pike, M. C., Vessey, M. P., Bull, D., Yeates, D., and Casagrande, J. T. (1986). Risk factors for uterine fibroids: reduced risk associated with oral contraceptives. *Br. Med. J.* 293, 359–362. doi: 10.1136/bmj.293.6543.359
- Rossert, J., Terraz, C., and Dupont, S. (2000). Regulation of type I collagen genes expression. *Nephrol. Dial. Transplant.* 15, 66–68. doi: 10.1093/ndt/15.suppl\_6.66
- Sandberg, A. A. (2005). Updates on the cytogenetics and molecular genetics of bone and soft tissue tumors: leiomyoma. *Cancer Genet. Cytogenet.* 158, 1–26. doi: 10.1016/j.cancergencyto.2004.08.025
- Stewart, E. A., Friedman, A. J., Peck, K., and Nowak, R. A. (1994). Relative overexpression of collagen type I and collagen type III messenger ribonucleic acids by uterine leiomyomas during the proliferative phase of the menstrual cycle. *J. Clin. Endocrin. Metabol.* 79 (3), 900–906. doi: 10.1210/jcem.79.3.8077380
- Tenni, R., Cetta, G., Dyne, K., Rossi, A., Quacci, D., Lenzi, L., et al. (1988). Type I procollagen in the severe non-lethal form of osteogenesis imperfecta. Defective pro- $\alpha 1(I)$  chains in a patient with abnormal proteoglycan metabolism and mineral deposits in the dermis. *Hum. Genet.* 79, 245–250. doi: 10.1007/BF00366245
- Thiaw, M. (2018). Diversité génétique de la population Sénégalaise: comparaison entre groupes ethniques. Mémoire de Diplôme de Master en Biologie Animale (mémoire non publié). Dakar, Sénégal: Faculté des Sciences et Techniques, Université Cheikh Anta Diop.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22 (22), 4673–4680. doi: 10.1093/nar/22.22.4673
- Trojanowska, M. (2002). Molecular aspects of scleroderma. *Front. Biosci.* 7, 608–618. doi: 10.2741/A798
- Yasul, W., Oue, N., and Ito, R. (2004). Search for new biomarkers of gastric cancer through serial analysis of gene expression and its clinical implications. *Cancer Sci.* 95, 385–392. doi: 10.1111/j.1349-7006.2004.tb03220.x
- Vevelib, P., Wingo, P. A., Xia, Z., Wilcox, L. S., and Peterson, H. B. (1995). Rate of hospitalization for gynecologic disorders among reproductive-age women in the United States. *Obstet. Gynecol.* 86, 764–769. doi: 10.1016/0029-7844(95)00252-M
- Wise, L. A., and Laughlin-Tommaso, S. K. (2016). Epidemiology of uterine fibroids from menarche to menopause. *Clin. Obstet. Gynecol.* 59 (1), 2–24. doi: 10.1097/GRF.0000000000000164
- Wise, L. A., Ruiz-Narvaez, E. A., Palmer, J. R., Cozier, Y. C., Tandon, A., Patterson, N., et al. (2012). African ancestry and genetic risk for uterine leiomyomata. *Am. J. Epidemiol.* 176, 1159–1168. doi: 10.1093/aje/kws276

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Kénémé and Sembène. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.