Prof Jing Yu **19th Sep 2016**
Science Editor, World Journal of Gastroenterology

Attn: Submission of the revised manuscript / Reference number: 27954

Dear Professor Yu,

Hereby we would like to send you the revised version of our manuscript entitled *'Aging connected methylation influences the gene expression of colorectal cancer and adenoma linked key control genes.'* We would like to thank you for the aware and extensive review. We agree with the comments of the Reviewers and the manuscript has been modified following the suggestions:

1. The detailed explanation of healthy normal samples has been attached to Materials and methods section.
2. The English version of both IRB and informed consent was attached.
3. The summarized demographic and clinical data of samples involved in *in silico* gene expression analysis was attached as Supplementary Table 3.
4. Material and Methods section has been completed with brief description of method for identification the differentially methylated CpG sites or gene promoters.
5. Description and results of further SFRP1 immunohistochemical analysis have been inserted in the revised version of the manuscript (Figure 4).
6. The Results section of the manuscript has been completed with the more detailed findings about the DNA methylation alterations of age-related CpG site in CRC.
7. The language revision of the manuscript was performed by a native English speaker scientist.
8. The description of pre-tests and the explanation of applied statistical tests and corrections have been inserted in the Material and methods section.

All changes in the revised version of the manuscript are marked with blue.

Looking forward to your kind reply about the hopefully acceptance of our response to the Reviewers and hope that the above modifications, enhancements make this manuscript acceptable.

Sincerely yours,

Orsolya Galamb Ph. D.
Semmelweis University
2nd Department of Medicine
Szentkiralyi str 46
1088 Budapest
Hungary
Tel: + 36 1 266 0926
Fax: + 36 1 266 0816

Dear Reviewer 1,

We would like to thank you for the aware and extensive review, and for drawing our attention to important topics which can improve our manuscript.

Hereunder, we would like to reply your specific comments and to give details about the modifications in our manuscript according to your useful suggestions:

(i) Altogether 27 tissue samples (19 from children and 8 from adults) with normal histology were involved in SFRP1 MS-HRM study. These children and adults were referred to the outpatient clinic with rectal bleeding, constipation or chronic abdominal pain. Ileocolonoscopy was part of their diagnostic work-up to exclude organic disease and the biopsy specimens showed normal macroscopic appearance and histology[Ref 28]. This was the reason why all these samples (19 from children and 8 from adults) were called as healthy normal colonic tissue samples. In concert of your well-founded critical comment, in order to clarify the mentioned confusion, this explanation was inserted in the revised version of the manuscript.

(ii) In case of children, the informed consent has been written by parents or tutelaries.
The English version of both IRB and informed consent was attached.

All changes in the revised version of the manuscript are marked with blue.

We thank for your extensive and positive evaluation and hope that the above changes, modifications, enhancements make this manuscript acceptable.

Sincerely yours,

Orsolya Galamb PhD.
Semmelweis University
2nd Department of Medicine
Szentkiralyi str 46
1088 Budapest
Hungary
Tel: + 36 1 266 0926
Fax: + 36 1 266 0816

Dear Reviewer 2,

We would like to thank you for the aware and extensive review, and for drawing our attention to important topics which can improve our manuscript.

Hereunder, we would like to reply your specific comments and to give details about the modifications in our manuscript according to your useful suggestions:

1. The Table 1. contains the demographic data (age, gender) of tissue samples involved in SFRP1 MS-HRM study. The demographic data of samples applied in *in silico* gene expression analysis are available in the text or in the supplementary information of the following research articles:

- in Table S1 of Galamb O et al. PLoS One 2012; 7: e48547 (Ref 25)
- in Supplementary Table S1 of Galamb O et al. Cancer Epidemiol Biomarkers Prev 2008; 17: 2835-2845 (Ref 26)
- in Supplemental Table 1 of Galamb O et al. Dis Markers 2008; 25: 1-16 (Ref 27)
- in Table 1 of Leiszter K et al. PlosOne 2013; PLoS One 2013; 8: e74140 (Ref 28)

In case of colonic tissue samples involved in *in silico* DNA methylation analysis, only the integrated demographic data are obtainable /in Table 1 of Luo Y et al. Gastroenterology 2014; 147: 418-429.e8 (Ref 24)/

The summarized demographic and clinical data of samples involved in *in silico* gene expression analysis was attached as Supplementary Table 3.

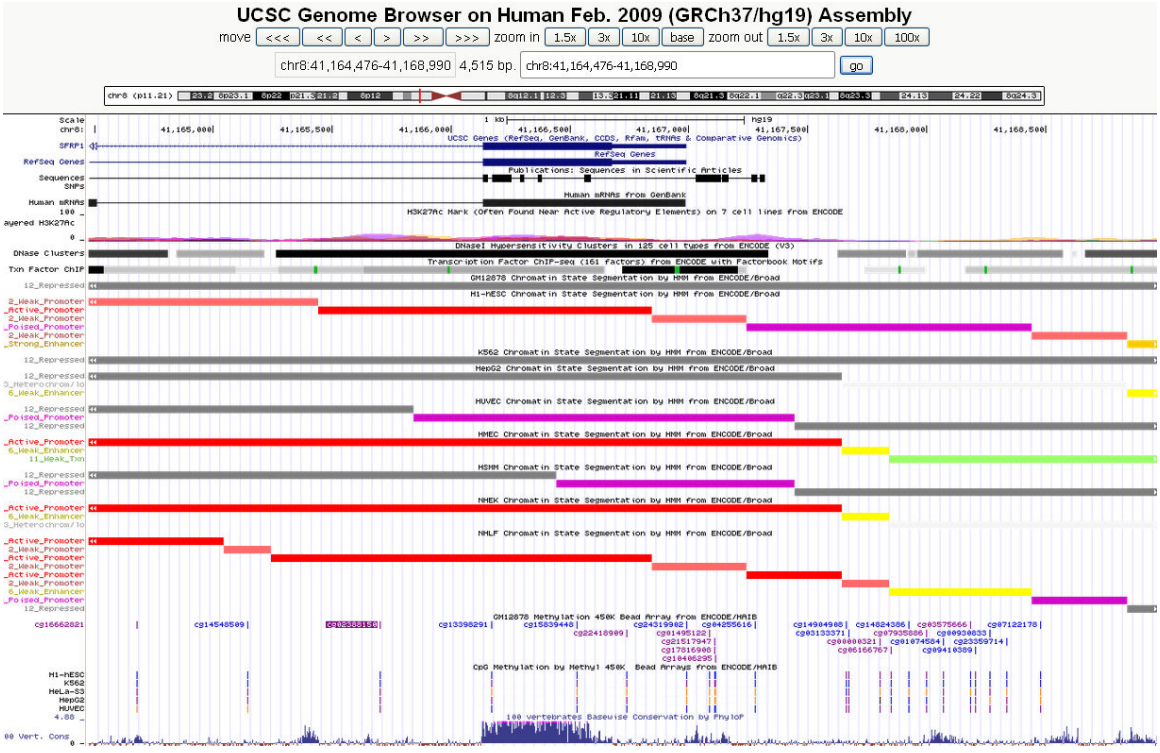
2. In UCSC Genome Browser, CpG sites (marked with cg IDs) of Illumina BeadChip 450K array system can be searched and localized using GRCh37/hg19 genomic version (https://genome-euro.ucsc.edu/cgi-bin/hgGateway?redirect=manual&source=genome.ucsc.edu). Also in hg19 genomic version of UCSC, special tracks such as Encode ChromHMM can be installed. This track displays a chromatin state segmentation for each of nine human cell types. ChromHMM is a common set of states across the cell types were learned by computationally integrating ChIP-seq data for nine factors plus input using a Hidden Markov Model (HMM). In total, fifteen states were used to segment the genome, and these states were then grouped and colored to highlight predicted functional elements such as active promoter, weak promoter, insulator etc.

The genomic position was considered as active promoter, if it was marked with active promoter in at least 1 from the 9 analyzed cell lines. The genomic position was considered as weak promoter, if it was marked with weak promoter in at least 1 from the 9 analyzed cell lines.

ChromHMM categories:

- State 1 - Bright Red - Active Promoter
- State 2 - Light Red -Weak Promoter
- State 3 - Purple - Inactive/poised Promoter
- State 4 - Orange - Strong enhancer
- State 5 - Orange - Strong enhancer
- State 6 - Yellow - Weak/poised enhancer
- State 7 - Yellow - Weak/poised enhancer
- State 8 - Blue - Insulator
- State 9 - Dark Green - Transcriptional transition
- State 10 - Dark Green - Transcriptional elongation
- State 11 - Light Green - Weak transcribed
- State 12 - Gray - Polycomb-repressed
- State 13 - Light Gray - Heterochromatin; low signal
- State 14 - Light Gray - Repetitive/Copy Number Variation
- State 15 - Light Gray - Repetitive/Copy Number Variation

SFPR1 promoter region with ChromHMM categories and Illumina BeadChip 450 K cg IDs (UCSC Genome Browser, hg19):



3. Differentially methylated genes were identified as described earlier (Galamb et al, Epigenetics 2016; 11(8): 588-602. Ref12)

In case of Illumina BeadChip data analysis, differences between average methylation values of the compared diagnostic groups (Δβ-values) and P values were determined for each CpG sites (cg IDs).

'In case of methyl capture sequencing data analysis, Bowtie2[62] with default settings was used to map the 100 bp paired and 50 bp unpaired reads to the hg19 human genome reference assembly.[63] The generated bam files were sorted and indexed by samtools.[64] Data were processed by the MEDIPS[65] bioconductor R package. After quality control, unpaired reads were extended to the length of the average fragment length of the paired samples (250 bp). PCR duplicates were removed, then the coverage data was binned with 100 bp window size. Methylation probabilities (β-values hereafter) were calculated with respect to genome wide CpG density dependent Poisson distributions.' (Galamb O et al, Epigenetics, 2016;11(8):588-602. Ref 12).

Along with your suggestions, Material and Methods section of the revised version of the manuscript has been completed with this reference and brief description of method for identification the differentially methylated CpG sites (in case of Illumina BeadChip 450K) or gene promoters (in case of methyl capture sequencing).

4. Description and results of further SFRP1 immunohistochemical analysis have been inserted in the revised version of the manuscript.

5. As cancer is one of the main age-related diseases, age-related and cancer-related epigenetic alterations and mechanisms are not completely divided (Finkel T et al. Nature 2007; 448: 767-774 Ref 3), especially in sporadic cancers. In this manuscript we mentioned the age-related epigenetic changes published by Steve Horvath (epigenetic clock including 353 CpG sites (DNA methylation

levels of 193 positively and of 160 negatively correlating with age). In the fifth column (medianByCpGOld-medianByCpGYoung) of the Supplementary Table 1, the age-related (found by Steve Horvath) methylation differences of the given CpG sites are represented, while methylation in CRC vs. N samples can be seen in the sixth column (Methylation in CRC) of the same table. According to these data, there are CpG sites with similar methylation pattern in CRC as in aging (eg. SFRP1, SYNE1, DKK3, AKT3, ADHFE1) or opposite methylation pattern in CRC as in aging (eg. CEMIP, RPL31, FXN) and with no methylation changes in CRC vs N (eg. BIK, FZD9, NHLRC1). In summary, 137 (38.8%) from the 353 CpG sites considered as age-related by Horvath were found to be significantly differentially methylated in CRC tissue samples compared to normals. Approximately two third of these CpG sites had similar methylation changes in CRC samples as during aging, while one third of these CpG sites showed opposite alterations in CRC tissue as during aging.
The Results section of the manuscript has been completed with the above findings.

6. In conjunction with your well-founded critical comment**,** SFRP1 expression was further studied at protein level by immunohistochemistry. As the reduced SFRP1 protein expression in CRC compared to normal colonic tissue was previously described (Valcz G, etal. PLoS One 2014; 9: e106143), we focused on SFRP1 expression analysis during aging.

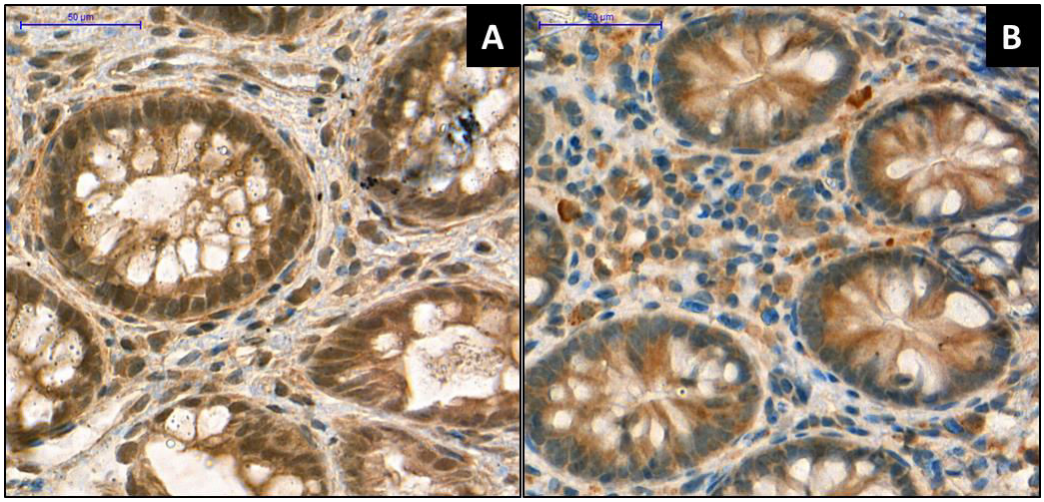


Figure 4. Strong/moderate cytoplasmic and nuclear SFRP1 expression both in epithelial and stromal compartments of healthy children (**A**) and healthy adult (**B**) samples. Digital microscopic images; 40x magnification; scale: 50 µm.

Not significantly, but remarkably lower SFRP1 protein expression were detectable both in epithelial and stromal component of healthy normal adults compared to healthy normal children samples.

The description and results of this analysis have been inserted in the revised version of the manuscript (Figure 4.).

7. The language revision of the manuscript was performed by a native English speaker scientist (Theo deVos PhD, Epigenomics AG, USA, Seattle, Washington) and the grammatical and stylistic errors have been corrected.

All changes in the revised version of the manuscript are marked with blue.

We thank for your extensive and positive evaluation and hope that the above changes, modifications, enhancements make this manuscript acceptable.

Sincerely yours,

Orsolya Galamb PhD.
Semmelweis University
2nd Department of Medicine
Szentkiralyi str 46
1088 Budapest
Hungary
Tel: + 36 1 266 0926
Fax: + 36 1 266 0816

Dear Reviewer 3,

We would like to thank you for the aware and extensive review, and for drawing our attention to important topics which can improve our manuscript.

Hereunder, we would like to reply your specific comments and to give details about the modifications in our manuscript according to your useful suggestions:

a. For statistical evaluation, normal distribution was checked using Kolmogorov-Smirnov test. Hence normal distribution was observed in any cases, Student's t-test with Benjamini and Hochberg correction was applied for paired group comparisons. For gene expression analysis, normal distribution was found using Kolmogorov-Smirnov test, therefore Student's t-test (in case of differentiation of two groups with equal variances) or Welch-test (in case of differentiation of two groups with unequal variances) and ANOVA (when more than two groups were compared) were applied. For paired comparisons Benjamini and Hochberg correction was applied. In case of ANOVA, Tukey HSD post-test was used in order to find out which group refers to the differentiation if any. Significance criteries were $P<0.05$ in any cases.
In conjunction with your well-founded critical comment, the description of pre-tests and the explanation of applied statistical tests and corrections have been inserted in the Material and methods section.

b. Along with your suggestions, on page 9 the sentence has been completed with description of other genomic regions as following: 'The 'epigenetic clock' of Horvath contains 353 CpG sites (DNA methylation levels of 193 positively and of 160 negatively correlating with age)[2] belonging to different genes, gene promoters and other genomic regions such as enhancers, insulators, Polycomb-repressed regions.'

c. The above-mentioned mistakes and clerical errors have been rectified in the revised version of the manuscript and the language revision of the whole manuscript was performed by a native English speaker scientist (Theo deVos PhD, Epigenomics AG, USA, Seattle, Washington), the grammatical and stylistic errors have been corrected.

All changes in the revised version of the manuscript are marked with blue.

We thank for your extensive and positive evaluation and hope that the above changes, modifications, enhancements make this manuscript acceptable.

Sincerely yours,

Orsolya Galamb PhD.
Semmelweis University
2nd Department of Medicine
Szentkiralyi str 46
1088 Budapest
Hungary
Tel: + 36 1 266 0926