Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Epigenetic/Genetic variations in CG-rich elements of immune-related genes contribute to food allergy development during childhood

Maria Kostara[1], Vasiliki Chondrou[2], Vassilis Fotopoulos[2], Argyro Sgourou[2], and Sophia Tsabouri[3]

[1]University General Hospital of Ioannina
[2]Hellenic Open University
[3]University of Ioannina

February 9, 2022

Abstract

Background: Genetic areas of FOXP3 TSDR, HLA-G upstream of CpG island 96, CpG41 and CpG73 islands of the HLA-DRB1 and HLA-DQB1 genes respectively, previously documented to display immune modulatory properties, were subjected to epigenetic/genetic analysis to assess their influence in IgE-mediated food allergy (FA) development in children. Methods: 64 orally challenged and IgE- tested food allergic subjects together with 44 controls were recruited. Targeted pyrosequencing analysis, to detect DNA methylation status and genetic variations was utilized and experimental results obtained were analysed by statistical software platform and correlated to clinical data. Also, transcription factor (TF) binding sites at study areas were unmasked by the JASPAR prediction database. Results: Parents' smoking was significantly correlated with aberrant methylation patterns, regardless food allergic or control status. HLA-G promoter region showed a trend for hypomethylation in food allergic subjects, with one of the CG sites displaying significantly decreased methylation values. Rs1233333, residing within HLA-G promoter region preserved a protective role towards DNA methylation. Variable methylation patterns were recorded for CpG41 of the HLA-DRB1 gene and hypermethylation of the region was significantly correlated with the presence of (Single Nucleotide Polymorphisms) SNPs. TFs' recognition sites, located at studied genetic areas and exerting pivotal regulatory biological roles, are potentially affected from divergent DNA methylation status. Conclusions: We propose that HLA-G expression is triggered by food derived allergens, providing a TregFoxP3-/HLA-G+ subpopulation generation and direct immune-tolerance. Furthermore, clear evidence is provided for the underlying co-operation of genetic polymorphisms with epigenetic events, mainly at CpG41 island of HLA-DRB1 gene, which need an extended investigation and elucidation.

Title: " Epigenetic/Genetic variations in CG-rich elements of immune-related genes contribute to food allergy development during childhood"

Running title: "Epigenetic/Genetic variations' contribution to food allergy"

Authors

Maria Kostara[1++], Vasiliki Chondrou[2++], Vassilis Fotopoulos[3], Argyro Sgourou[2*], Sophia Tsabouri[4*]

1: Department of Paediatrics, Ioannina University Hospital, Ioannina, Greece

2: Laboratory of Biology, School of Science and Technology, Hellenic Open University,Patras, Greece

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

3: Digital Systems Laboratory, School of Science and Technology, Hellenic Open University, Patras, Greece

4: Department of Paediatrics, Faculty of Medicine, School of Health Sciences University of Ioannina, Ioannina, Greece

++: equal contribution

***: Corresponding authors: sgourou@eap.gr, tsabouri@uoi.gr**

Abstract

Background: Genetic areas of *FOXP3* TSDR, *HLA-G*upstream of CpG island 96, CpG41 and CpG73 islands of the*HLA-DRB1* and *HLA-DQB1* genes respectively, previously documented to display immune modulatory properties, were subjected to epigenetic/genetic analysis to assess their influence in IgE-mediated food allergy (FA) development in children.

Methods: 64 orally challenged and IgE- tested food allergic subjects together with 44 controls were recruited. Targeted pyrosequencing analysis, to detect DNA methylation status and genetic variations was utilized and experimental results obtained were analysed by statistical software platform and correlated to clinical data. Also, transcription factor (TF) binding sites at study areas were unmasked by the JASPAR prediction database.

Results: Parents' smoking was significantly correlated with aberrant methylation patterns, regardless food allergic or control status. *HLA-G* promoter region showed a trend for hypomethylation in food allergic subjects, with one of the CG sites displaying significantly decreased methylation values. Rs1233333, residing within *HLA-G* promoter region preserved a protective role towards DNA methylation. Variable methylation patterns were recorded for CpG41 of the *HLA-DRB1* gene and hypermethylation of the region was significantly correlated with the presence of (Single Nucleotide Polymorphisms) SNPs. TFs' recognition sites, located at studied genetic areas and exerting pivotal regulatory biological roles, are potentially affected from divergent DNA methylation status.

Conclusions: We propose that HLA-G expression is triggered by food derived allergens, providing a $Treg^{FoxP3-/HLA-G+}$subpopulation generation and direct immune-tolerance. Furthermore, clear evidence is provided for the underlying co-operation of genetic polymorphisms with epigenetic events, mainly at CpG41 island of*HLA-DRB1* gene, which need an extended investigation and elucidation.

Keywords: *FOXP3* TSDR methylation, *HLA-G* promoter region, *HLADQB1/HLADR* B1 CpG islands, SNP detection, epigenetic background, genetic/epigenetic interplay

Key message

Epigenetic combined with genetic background of immune-related genetic loci, mainly *HLA-G* promoter region and CpG41 island of*HLA-DRB1* highlighted their potential to FA appearance.*HLA-G* hypomethylation and consequent generation of HLA-G positive/FOXP3 negative Tregs emerge as an immediate modulator of the immune cells' phenotype urging an early desensitization process. Also, rs1233333 residing within*HLA-G* gene's promoter region, shows significant results in safeguarding the area from DNA methylation spreading. Methylation levels of *HLA-DRB1* CpG island 41 and co-occurrence of SNPs pronounce their cooperation and their modulatory role towards FA development. Finally, aberrant background methylation patterns at *HLA-G* gene's promoter region and*FOXP3* TSDR are attributed to parents' smoking habit.

Introduction

IgE-mediated food allergy (FA) development during childhood appears as a complex disorder, implicating immune dysfunction and influenced by genetic variants, epigenetic modifications, parental habits and environmental contributors or combinations thereof[1-5]. Among factors mediating gene-environment interactions is the epigenetic remodelling of genome, a critical mechanism for immune system adaptation to environmental exposures[6].

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

FOXP3-expressing Tregs hold a leading role in the immune system homeostasis and effectively protect against allergies[7]. High levels of*FOXP3* methylation inhibiting gene's expression, have been attributed to extensive exposure to air pollutants, an effect pronounced in asthmatic compared to non-asthmatic children[8].

Other immune related genes, encompassed in the major histocompatibility complex (MHC) system, modulate adaptive immune responses. The human leukocyte antigen-G (*HLA-G*) is a MHC class I gene, epigenetically regulated by methylation[9] and has been recognized as an important mediator of immune tolerance at the fetal–maternal interface of human placenta[10,11]. In healthy adults, thymus generates a proportion of 1-3% HLA-G-positive/FOXP3-negative $CD4^+$ and $CD8^+$ T-cells,[12] with regulatory properties (Tregs) [1314] and increased frequency at sites of inflammation.

MHC class II cluster includes *HLA-DQ* and *HLA -DR*genes, whose expression is limited to the professional antigen-presenting cells. Cell-processed antigens' presentation by MHC class II to $CD4^+$ T lymphocytes, stimulates the antigen-specific hypersensitivity type I allergic reaction. GWAS studies and review analysis have provided convincing evidence that *HLA-DR*and *HLA-DQ* genes' polymorphic regions harbor significant genetic and epigenetic risk for FA[15,16].

The present study aimed to accomplish deeper insights for the genetic/epigenetic and environmental associations influencing the predisposition to FA development.

Methods

2.1 Research strategy. CpGs and SNPs selection and localization

CpGs of the study were selected by setting specific criteria such as their involvement in immune-related responses in FA in a genetic or epigenetic context. Several studies suggest that DNA methylation profile of the *FOXP3* , mainly at Treg-specific demethylated region (TSDR) and HLA locus could mediate genetic susceptibility to immune-mediated disorders such as FA. We thus, evaluated and compared the DNA methylation patterns of immune-related genetic areas among children exhibiting IgE-mediated food allergy versus healthy subjects. We further correlated methylation data with genetic variants residing in the same CpG regions and parental habits.

CpG islands and regulatory CpGs residing as discrete sites were retrieved by the UCSC genome browser (https://genome.ucsc.edu/). Single nucleotide polymorphisms (SNPs) across studied genetic areas were also identified. Characteristic MHCI and MHCII genetic loci with CpGs, including predicted transcription factor binding (TF) motifs (JASPAR CORE 2020), are illustrated in Figure 1. Transcription factors' recognition sites were detected, aiming to observe sequence alterations ascribed to SNPs or to CG biochemical alterations due to methylation.

Methylation profile at Treg-specific demethylated region (TSDR), located on the 2nd conserved non-coding sequence of *FOXP3* gene, was assessed in both DNA strands. PCR and sequencing primers designed (Table 2), encompassed well characterized CpGs of TSDR in either direction. The immediate upstream region of *HLA-G* CpG96 island was subjected to methylation analysis to uncover its potential role in allergy via demethylation. Genotypic analyses for rs1233333 (G/A), located within studied region was also conducted to reveal its effect in DNA methylation spreading.

Finally, CpG41 and CpG73 within exon 2 of the *HLA-DRB1* and*HLA-DQB1* genes respectively, were studied. Also, sequencing of a 70bp core sequence revealed a number of SNPs residing within CpG41 (Figure 1), in a variety of combinations. Methylation status of CpG41 co-occurring with unique SNPs in the extended gene locus region have been involved in multiple sclerosis disease development[17,18].

2.2. Sample collection. Clinical criteria for food allergy definition and diagnostic tests

Patients with FA were recruited at Paediatric Allergy Department of the University Hospital of Ioannina, from July 2019 to May 2021 and blood samples were collected following the World Medical Association Declaration of Helsinki. Informed consent was given from parents and the conducted study has been approved by the Ethical and Scientific Committee (protocol No1290), of University Hospital, Ioannina.

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Inclusion criteria were the presence of definite IgE-mediated FA, according to the criteria set by the European Academy of Allergy and Clinical Immunology (EAACI)[19]. Samples from children with chronic diseases and those on long-term treatment with systemic steroids were excluded from the study. Control group consisted of blood samples from healthy children who visited the outpatient Paediatric department for routine check- up. FA cases and controls were matched for gender and age. During study period, 64 subjects with IgE-mediated FA were enrolled and 44 clinically healthy children, as controls.

Based on guidelines provided by EAACI[19] IgE-mediated FA was diagnosed when all clinical criteria were fulfilled. A diagnostic open oral food challenge (OFC)[20] and SPTs were also performed; subjects under study were not administrated any antihistamines for a period of ten days before participating in OFC.

Specific IgEs to food allergens, including milk, egg, nuts, cereal, and fish were measured with ImmunoCAP (Phadia AB, Uppsala, Sweden) and 0.35 kU/L value was set as the minimum limit. Sensitization was classified as negative, mono-sensitization or poly-sensitization, based on the number of sensitized food allergens.

Demographic and clinical characteristics of all subjects were recorded. A structured questionnaire regarding breastfeeding, parental smoking, family history of atopy (including the presence of allergic rhinitis, asthma, atopic dermatitis, and FA was completed by the parents and data are presented at Table 1.

TABLE 1: *Demographic and clinical data of the studied population.*

Demographic data	**Cases (n=64)**	**Controls (n=44)**	**Controls (n=44)**	**Total (n=108)**
Age range	6 months-15 years	2 months-15 years	2 months-15 years	
Gender (M/F)	39/25	22/22	**61/47**	**61/47**
History	**Cases (n=64)**	**Control (n=44)**	**Total (n=108)**	**Total (n=108)**
Breast feeding (Y/N)	51/13	29/15	**80/28**	**80/28**
Parental Smoking (Y/N)	29/35	18/26	**47/61**	**47/61**
History of allergy (Y/N)	32/32	2/42	**34/74**	**34/74**
Allergens	**Cases (n=64)**			
Milk	15/64			
Egg	15/64			
Nuts	7/64			
Fish	2/64			
Fruit	1/64			
Polysensitized	24/64			

Y: Yes, N:No

DNA extraction, SNP detection by pyrosequencing

Genomic DNA was extracted from whole blood samples using the PureLink Genomic DNA kit (Invitrogen) according to manufacturer's protocol.

Genotyping of SNPs at promoter region of *HLA-G* and CpG41 of*HLA-DRB1* was performed with Pyrosequencing methodology (AQ and SQA assays), according to manufacturer's protocols. PCR and sequencing primers used are presented in Table 2.

2.4 Bisulfite conversion and CpG pyrosequencing analysis

The methylation status of single CpG sites along the studied sequence was analysed by the Pyrosequencing CpG assay methodology. Bisulfite conversion reactions for each DNA sample conducted prior to Pyrosequencing for the complete conversion of the unmethylated Cytosines (existing as CpG dinucleotides) to Uraciles, with Epitect Bisulfite Kit (Qiagen).

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Sequencing and PCR primers were designed with the PyroMark Assay Design Software version 2.0 (Qiagen) (Table 2).

TABLE 2: *Pyrosequencing primer sets*

Gene	Region	Assay	Sequence 5'-3'	Prime
FoxP3 reverse strand			TGGGTTAAGTTTGTTGTAGGATAGG	Forwa
	TSDR	CpG	CCCAAACACATATAAAATAACCTAACTCAA	Revers
			AGGATAGGGTAGTTAGTT	Forwa
FoxP3 forward strand			TGTGGGATGGTTTGATTTAGTAAAGTAT	Forwa
	TSDR	CpG	CACCCATATCACCCCACCT	Revers
			GGGTTTTGTTGTTATAGT	Forwa
HLA-DRB1			AGGATGAGTTTTTGGGTTGTAG	Forwa
	CpG41	CpG	CCTACTCCAAAATATCCTTCTAACTATTC	Revers
			GGAGTGTTATTTTTTTAATGGG	Forwa
HLA-DQB1			TTTTTGGATGTAGGAAGGTAGAT	Forwa
	CpG73	CpG	CCTCCAAAACTTCCTTCTAACTATT	Revers
			TTAAGGGTATGTGTTATTTTATTAA	Forwa
HLA-G			GGGAGGTAGGGAGTTTAGTTTA	Forwa
	Immediately upstream CpG96	CpG	CCATAACCACCATCCTTAAC	Revers
			GGTTAGGGAGAAGTTTTA	Forwa
HLA-DRB1			TTCGCCTCAGGAAGACGG	Forwa
	CpG41	SQA	CCTTCTGGCTGTTCCAGTACTCA	Revers
			CGGAGCGGGTGCGGT	Forwa
HLA-G			CAGAACGCTTGGCACAAGAGTAG	Forwa
	Immediately upstream CpG96	AQ	TGTCACCTAATGGGAGTGAGAACT	Revers
			CACAGGCGGTGTATG	Forwa

2.5. Statistical analyses

All analyses were performed with SPSS software version 20.0[21]. Statistical significance was defined as a p-value <0.05. Shapiro-Wilk test was used to determine whether variables were normally distributed. To assess statistical differences between cases and controls independent t-test or Mann-Whitney test, for non-normal distributed data, was used. T-test or Mann-Whitney was also used to identify the differences among different genotypes. Two-way ANOVA test with Bonferroni correction was conducted to explore the interaction effect between parents' smoking and FA on methylation levels.

Results

3.1. Hypermethylation of *FOXP3* and distinctive outspread of *HLA-G* methylation among FA samples vs controls suggests a potential tolerance acquisition promoted by Treg$^{HLA\text{-}G+/FOXP3-}$

3.1.1. *FOXP3* TSDR methylation profile

CpG sites within TSDR of *FOXP3* gene were analysed at both DNA directions (forward and reverse DNA strands) and revealed a clear deviation of methylation pattern between them. Among FA and control samples, *FOXP3* TSDR CpGs showed almost 100% methylation across forward strand and a variable demethylation rate in the reverse strand (Figure 2A), indicating that only reverse DNA strand methylation pattern accounts for the gene's transcriptional regulation.

FOXP3 gene is located on human chromosome X (Xp11.23) and consequently, epigenetic discrepancies between different sexes may arise from the inactivation process of X chromosome in females. To avoid artifacts during methylation data analysis, male and female donors (both controls and food allergic children)

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

were analysed independently, however insignificant gender differences were observed in methylation profiles (data not shown).

Subsequent analysis revealed marginal variations (Figure 2B) in food allergic children compared to controls. Also, the average methylation level of all TSDR CpG sites tested was not statistically significant between food allergic children and control groups, providing poor evidence for the TSDR demethylation.

3.1.2. *HLA-G* promoter methylation profile

Methylation analysis of CpG sites located in *HLA-G* promoter region, immediately upstream from CpG96 island (Figure 1) revealed that CpG site 2 was significantly demethylated (p=0.004) and also, all CpGs in the area presented an obvious trend (T) for hypomethylation in the food allergic group, however, not reaching statistical significance (p=0.05), compared to control group (Figure 2C).

Conclusively, insignificant differences in methylation status of*FOXP3* TSDR between food allergic children and controls, suggests that Tregs, positive to FOXP3 expression, are distributed equivalently between the studied groups. Food allergic subjects participating in the study were untreated, though they had already abstained from food allergens causing hypersensitivity type I reaction and therefore, sample collection was performed under an evolving desensitization process. The experienced allergic reaction herein, reinforces the $\text{Treg}^{HLA\text{-}G+/FOXP3-}$ subgroup generation as an immediate immune response to partially restore immune tolerance to the allergen and/or stimulating the next levels of desensitization.

3.2 Heterogeneity in methylation status along genetic loci reflects inherent genetic/epigenetic background differences

3.2.1. Epigenetic variations influenced by genetic polymorphisms and parents' smoking habit

We explored the potential moderating role of rs1233333, located within*HLA-G* promoter region (Figure 1), on DNA methylation. Genotyping food allergic and control samples indicated only AA and GG homozygotes and statistical analyses performed uncovered a significant association between rs1233333 (A/A) genotype and the lower levels of methylation at CpGs covering the *HLA-G* promoter region. More specifically, CpG sites 3, 4, 5 and 6, as well as average methylation of the total analysed region were found to differ with statistical significance (p=0.014, p=0.002, p=0.010, p=0.004 and p=0.010 respectively) between homozygous AA and homozygous GG individuals regardless of food allergic or control status (Figure 2D). Furthermore, food allergic children exhibited 2,5 fold elevated frequency in AA genotype compared to control group, implying a contribution of AA genotype both in DNA methylation status and FA development.

Differences in *FOXP3* TSDR and *HLA-G* methylation profiles based on parents' smoking habit among FA and control groups were also evaluated. A two-way ANOVA analysis was conducted to examine the co-effect of FA and smoking on methylation levels of *FOXP3* TSDR. Smoking phenotype was significantly associated with higher methylation at CpG sites 1, 5, 6 and 9, both in control and food allergic groups (p=0.004, p=0.004, p=0.011 and p=0.023 respectively) (Figure 3A), reflecting previously documented conclusions that an exposure-response epigenetic effect appears in the offspring for many smoking-associated DMPs (Differentially Methylated Positions), which increase the risk for allergic reactions development[22].

Additionally, a two-way ANOVA analysis was conducted for the co-effect of FA and smoking on methylation profile of *HLA-G* promoter. Analysis indicated that food allergic children with smoking parents had significantly lower methylation levels than controls at CpG site 3 (F=6.28 df=1, 33 p<0.001), CpG site 4 (F=5.268, df=1, 33 p=0.028), CpG site 5 (F=7.752, df=1, 33 p=0.009), CpG site 6 (F=11.888 df=1, 33 p=0002), CpG site 7 (F=7.773, df=1, 33 p=0.036) and throughout the whole CpG region tested (F=9.597, df=1, 33 p=0.004) (Figure 3B).

Association of breastfeeding with methylation was not possible due to the small sample size of children that were not breastfed.

3.2.2. Methylation status of*HLA-DRB1* CpG41 and co-existence of genetic variants emerge as risk factors for FA development

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

HLA-DRB1 and *HLA-DQB1* (encoding for the β-chains) belong to MHC class II, are closely linked within the genetic loci, share common regulatory features in their promoter regions[23] and common CpG islands within exon 2[18]. Also, genetic variants within *HLA-DQB1* and *HLA-DRB1* sequences have strongly been associated with predisposition to FA[15,16], but the linkage between DNA methylation and single-nucleotide polymorphisms (SNPs) across these loci has not yet been elucidated. CpG73 and CpG41 also include DNAse hypersensitive sites, indicating essential trans-acting factor occupancy.

Methylation analyses of CpG73 within exon 2 of the *HLA-DQB1* gene clearly showed hypomethylation of the region in both study groups (data not shown), implying a nonculpable epigenetic frame during FA development. In contrast, methylation profile of *HLA-DRB1* CpG41 showed highly variable patterns and a strong correlation of hypermethylation and coexistence of sequence polymorphisms. Sequencing of a 70bp hot polymorphic region within CpG41 was preceded to methylation analysis. 25.6 % of controls and 16.9% of food allergic samples were not polymorphic. Polymorphic samples exhibited the following SNPs: rs1059572 (G/A/C/T), rs1059575 (G/C/T), rs3175105 (T/A/C), rs17882300 (A/C/T), rs17882603 (C/A/T), rs11554462 (A/C/G), rs1064664 (A/C/G), rs707957 (G/A/C/T), rs17878951 (A/G), and rs150747106 (C/G/T) in irregular combinations (Figure 1). Absence of SNPs was accompanied with almost complete demethylation of CpG41 (Figure 4).

CpG41 of *HLA-DRB1* gene resides within sequence of the antigen-binding groove, explaining the high rate of polymorphisms obtained along this region and its association with autoimmune disorders, such as multiple sclerosis and rheumatoid arthritis[17,18,24,25]. Accordingly, during FA development certain MHC alleles are possibly more efficient at presenting important food allergic components, leading to increased immune responses and fast progression of allergy symptoms. Failure to correlate any of the genetic variants detected among FA and control samples with CpG41 methylation levels, indicates that other variants located within the extended genetic loci potentially acquire guiding roles for DNA the methylation spreading throughout this genetic area.

3.3. Alterations in CpGs methylation pattern potentially disrupt binding affinity of transcription factors

Lengthwise sequences including CpGs at *HLA-G* promoter and CpG41 of *HLA-DRB1* gene, transcription factor (TF) binding sites were appointed to specified positions utilizing the JASPAR CORE open-source database 2020 [26], a collection of all most recently defined transcription factor binding sites for eukaryotes. Figure 1 illustrates curated TFs occupying either forward (+) or reverse (-) DNA strand, limited only to those comprising CpG sites in their consensus motif. Methylation status at specific Cytosines could potentially affect their recognition capacity and/or binding affinity and to uncover such disruptions we have emphasized only on the CpG areas of genes that manifested statistically significant results in the present study.

Analysis revealed different sets of TFs, mainly composed of factors with zinc finger motifs (*ZIC1, ZIC4, ZIC5, ZBTB7A, ZNF281*), subunits of NFKB factor (*REL, RELA* , p100 and p105 subunits) and TFs from the Helix-Loop-Helix family (*NHLH2, TCF3*), known to influence essential biological functions.

Related epigenetic research has shown that *HLA-G* is demethylated and expressed in various types of tumors and further correlated with immune tolerance to malignancy, tumor escape and poor prognosis[27]. Predicted TF binding sites within the *HLA-G* promoter include zinc finger proteins ZIC1, ZIC4 and ZIC5 and NF-kappa-B p100 and p105 subunits, among which ZIC5, a known transcriptional repressor recognizes the CpG site 2 exhibiting the highest demethylating value with statistical significance among the FA group (Figure 1).

ZBTB7A and ZNF281 acquiring zinc finger motifs, REL and RELA subunits of NFKB factor and NHLH2 (Helix-Loop-Helix) are among TFs with recognition binding sites residing across CpG41 of *HLA-DRB1* gene (Figure 1). Also, co-existence of genetic variants clearly showed an impact towards DNA methylation spread (Figure 4). These results reflect a regulatory role for this gene's area, located within *HLA-DRB1* gene body. Potential roles for the TFs binding are downstream enhancer and/or insulator and/or nucleosome modulatory properties.

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Discussion

DNA methylation signatures across specific genetic-loci emerge as essential diagnostic and disorder-monitoring tools for FA. Furthermore, DNA methylation could distinguish allergy-risk epigenetic profiles, predict responses to therapy or inception of the natural desensitization process.

Previous studies have validated DNA methylation across the top-ten gene loci and showed altered hyper- or hypo-methylation profiles among children suffering from cow's milk allergy (CMA) (IgE and non IgE-mediated), relative to controls. Methylation disruptions were sex-dependent and disappeared upon developing tolerance to CMA[28], confirming the plasticity of epigenetic events. Results were not replicated by other epigenetic studies in CMA[29,30], herein differences can be attributed to cell source subjected to analyses and the sort of FA exhibited, either IgE- or non IgE-mediated CMA. Also, specific DNA methylation signatures have been underscored in twelve gene loci between peanut allergic versus nonallergic participants. DNA methylation status of *BDNF* and *CXCL12*genes exhibited the most superior diagnostic potential compared to serum peanut–specific IgE detection for discriminating allergy[31].

Our study revealed similar methylation profiles of *FOXP3* TSDR and CpG73 of *HLA-DQB1* in both IgE-mediated food allergic samples and controls; *FOXP3* TSDR exhibited a hypermethylated profile and*HLA-DQB1* an unmethylated (Figures 2A, 2B). In this context, results from *HLA-DQB1* gene are not further discussed. Contrariwise, results from CpGs at *HLA-G* promoter region and CpG island 41 of *HLA-DRB1* gene in food allergic samples were deviated from controls (Figures 2C, 2D, 4) and high methylation status was significantly assisted by the presence of SNPs located at CpG island 41. On the contrary, rs1233333, located within *HLA-G* promoter region had an apparent effect on DNA methylation protection.

Tolerance acquisition in children with IgE-mediated CMA has been shown to involve epigenetic regulation of *FOXP3* gene[28,32], which has been also suggested as a biomarker of oral tolerance[33]. Data derived from our study showed insignificant methylation level differences in*FOXP3* TSDR between the 64 subjects with IgE-mediated FA versus the 44 clinically healthy children (in both male and female donors), underlying the importance and the great bias potential of epigenetic, tissue and sex related heterogeneity, which should be accounted for the contradictory results obtained between parallel epigenetic studies. However, higher methylation at *FOXP3* TSDR CpG sites 1, 5, 6 and 9, both in control and food allergic groups (p=0.004, p=0.004, p=0.011 and p=0.023 respectively) was attributed to parents' smoking habit (Figure 3A), confirming results from other pediatric disorders deriving from parental burdened epigenetic background[22]. Analogous results were displayed from methylation analyses at the *HLA-G* promoter region (Figure 3B).

Methylation status of CpG41, located at exon 2 of *HLA-DRB1* and CpGs lying at promoter region of *HLA-G* manifested their contingent nature to FA predisposition. Genetic and epigenetic contribution of *HLA-DRB1* has been established in multiple sclerosis, and rheumatoid arthritis[2425], diseases with evident autoimmune features. Recent findings from GWAS in an Australian population trying to decode the genetic background association between IgE-mediated peanut allergy and *HLA-DRB1,* have established an amino acid variant located at position 71 within the peptide-binding groove [34]. In the present study, methylation status and SNPs residing within GpC island 41 highlighted their capacity to influence IgE-mediated FA development, however SNPs detected (Figure 1) failed to be significantly corelated with DNA methylation status along the same genetic area. Nevertheless, it was significantly demonstrated that high DNA cytosine methylation is accompanied with genetic SNPs (Figure 4). Our results suggest that critical SNPs that would differentiate the challenge-proven food allergic samples from controls, potentially reside within the genetic loci surroundings and a broader genetic area needs to be investigated to endorse more accurate results.

The attempt to investigate possible disruptions caused in transcription factor (TF) binding affinity by the differential methylation status of CpG 41, revealed a variety of critical factors predicted to have recognition sites (Figure 1). These TFs potentially co-operate with factors acting as downstream enhancers (probably the NFKB subunits) and/or may accomplish insulator power.

To the best of our knowledge results from this study are the first which show the potential involvement of CpG rich areas of *HLA-G*promoter and *HLA-DRB1* , with simultaneous coexistence and cooperation of

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

genetic variants, in FA. Demethylation of the *HLA-G* promoter region among FA samples potentially induces differentiation of a subset of $CD4^+$T cells towards $Treg^{HLA\text{-}G+/FOXP3-}$ with potent immunosuppression functions[14], similar to the natural, thymus-derived Tregs, as a premature immune response towards the desensitization process. Another important result from the present study was the significant correlation of rs1233333, residing in*HLA-G* promoter region, with DNA methylation levels, exerting its modulatory effect in gene's expression by preventing DNA methylation spreading (Figure 2).

Allergic reactions should be considered in terms of complex pathways' cooperation anticipating to decode synergistic effects towards FA onset and thus, an augmented analysis should be conducted. Obstacles to overcome are the insufficient quantities of pediatric samples, which pose limitations for extended analysis requirements. Other limitations are the inconsistent time points of specimen collection, the wide range of FA symptoms' manifestation or the drug regimens utilized in clinical research studies, underlying discrepancies in results among different research conducted. Nevertheless, targeted research holds great promise for decoding the FA complex genetic/epigenetic background and offer new insights on molecular events contributing to FA development.

Acknowledgments: Authors wish to thank all patients and donors for providing samples for the completion of this research.

Funding

This research was financially supported by PENER_EAP grant of the Hellenic Open University. VC was also supported by a two year-fellowship by the State Scholarships Foundation of Greece.

Conflicts of interest

Authors have nothing to declare.

Author contribution

Maria Kostara and **Sophia Tsabouri** recruited food allergic patients and healthy controls.**Vasiliki Chondrou** carried out the experiments and data collection. **Vassilis Fotopoulos**with support from **Vasiliki Chondrou** , performed the statistical analysis. **Maria Kostara** and **Vasiliki Chondrou** wrote the first draft of the manuscript. **Argyro Sgourou** and**Sophia Tsabouri** conceived the study, assisted in data interpretation and finalized the manuscript. All authors reviewed, edited and approved the final manuscript.

References

1. Holt PG, Sly PD. Viral infections and atopy in asthma pathogenesis: new rationales for asthma prevention and treatment. *Nature medicine.* 2012;18(5):726-735.

2. Hong X, Wang X. Early life precursors, epigenetics, and the development of food allergy.*Seminars in immunopathology.* 2012;34(5):655-669.

3. Anacker C, O'Donnell KJ, Meaney MJ. Early life adversity and the epigenetic programming of hypothalamic-pituitary-adrenal function. *Dialogues in clinical neuroscience.* 2014;16(3):321-333.

4. Andersson NW, Hansen MV, Larsen AD, Hougaard KS, Kolstad HA, Schlunssen V. Prenatal maternal stress and atopic diseases in the child: a systematic review of observational human studies. *Allergy.* 2016;71(1):15-26.

5. Ong PY, Leung DY. Bacterial and Viral Infections in Atopic Dermatitis: a Comprehensive Review.*Clinical reviews in allergy & immunology.* 2016;51(3):329-337.

6. Hong X, Wang X. Epigenetics and development of food allergy (FA) in early childhood. *Current allergy and asthma reports.* 2014;14(9):460.

7. Schaub B, Liu J, Hoppler S, et al. Maternal farm exposure modulates neonatal immune mechanisms through regulatory T cells. *The Journal of allergy and clinical immunology.* 2009;123(4):774-782 e775.

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

8. Hew KM, Walker AI, Kohli A, et al. Childhood exposure to ambient polycyclic aromatic hydrocarbons is linked to epigenetic modifications and impaired systemic immunity in T cells. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology.* 2015;45(1):238-248.

9. Carosella ED, HoWangYin KY, Favier B, LeMaoult J. HLA-G-dependent suppressor cells: Diverse by nature, function, and significance. *Human immunology.*2008;69(11):700-707.

10. Kovats S, Main EK, Librach C, Stubblebine M, Fisher SJ, DeMars R. A class I antigen, HLA-G, expressed in human trophoblasts. *Science.* 1990;248(4952):220-223.

11. Castelli EC, Veiga-Castelli LC, Yaghi L, Moreau P, Donadi EA. Transcriptional and posttranscriptional regulations of the HLA-G gene. *Journal of immunology research.*2014;2014:734068.

12. Pankratz S, Bittner S, Herrmann AM, et al. Human CD4+ HLA-G+ regulatory T cells are potent suppressors of graft-versus-host disease in vivo. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology.* 2014;28(8):3435-3445.

13. Lazana I, Zoudiari A, Kokkinou D, et al. Identification of a novel HLA-G+ regulatory population in blood: expansion after allogeneic transplantation and de novo HLA-G expression at graft-versus-host disease sites. *Haematologica.*2012;97(9):1338-1347.

14. Stamou P, Marioli D, Patmanidi AL, et al. Simple in vitro generation of human leukocyte antigen-G-expressing T-regulatory cells through pharmacological hypomethylation for adoptive cellular immunotherapy against graft-versus-host disease. *Cytotherapy.* 2017;19(4):521-530.

15. Hong X, Hao K, Ladd-Acosta C, et al. Genome-wide association study identifies peanut allergy-specific loci and evidence of epigenetic mediation in US children. *Nature communications.* 2015;6:6304.

16. Kostara M, Chondrou V, Sgourou A, Douros K, Tsabouri S. HLA Polymorphisms and Food Allergy Predisposition.*Journal of pediatric genetics.* 2020;9(2):77-86.

17. Kular L, Jagodic M. Epigenetic insights into multiple sclerosis disease progression. *Journal of internal medicine.* 2020;288(1):82-102.

18. Kular L, Liu Y, Ruhrmann S, et al. DNA methylation as a mediator of HLA-DRB1*15:01 and a protective variant in multiple sclerosis. *Nature communications.*2018;9(1):2397.

19. Muraro A, Werfel T, Hoffmann-Sommergruber K, et al. EAACI food allergy and anaphylaxis guidelines: diagnosis and management of food allergy. *Allergy.*2014;69(8):1008-1025.

20. Calvani M, Bianchi A, Reginelli C, Peresso M, Testa A. Oral Food Challenge. *Medicina.*2019;55(10).

21. IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.

22. Keil T, Lau S, Roll S, et al. Maternal smoking increases risk of allergic sensitization and wheezing only in children with allergic predisposition: longitudinal analysis from birth to 10 years. *Allergy.* 2009;64(3):445-451.

23. Ting JP, Trowsdale J. Genetic control of MHC class II expression. *Cell.* 2002;109 Suppl:S21-33.

24. Sawcer S, Hellenthal G, Pirinen M, et al. International Multiple Sclerosis Genetics, Consortium.Wellcome Trust Case Control, ConsortiumGenetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature.*2011;476(7359):214-219.

25. Raychaudhuri S, Sandor C, Stahl EA, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis.*Nature genetics.* 2012;44(3):291-296.

26. Fornes O, Castro-Mondragon JA, Khan A, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic acids research.*2020;48(D1):D87-D92.

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

27. Liu L, Wang L, Zhao L, He C, Wang G. The Role of HLA-G in Tumor Escape: Manipulating the Phenotype and Function of Immune Cells. *Frontiers in oncology.* 2020;10:597468.

28. Petrus NCM, Henneman P, Venema A, et al. Cow's milk allergy in Dutch children: an epigenetic pilot survey.*Clinical and translational allergy.* 2016;6:16.

29. Berni Canani R PL, Nocerino R, Cosenza L, Pezzella V, Di Costanzo M, Capasso M, Del Monaco V, D'Argenio V, Greco L, Salvatore F. . Differences in DNA methylation profile of Th1 and Th2 cytokine genes are associated with tolerance acquisition in children with IgE-mediated cow's milk allergy. *Clinical epigenetics.* 2015;7(1).

30. Bonnelykke K, Sparks R, Waage J, Milner JD. Genetics of allergy and allergic sensitization: common variants, rare mutations. *Current opinion in immunology.*2015;36:115-126.

31. Zhou X, Han X, Lyu SC, et al. Targeted DNA methylation profiling reveals epigenetic signatures in peanut allergy. *JCI insight.* 2021;6(6).

32. Lluis A, Depner M, Gaugler B, et al. Increased regulatory T-cell numbers are associated with farm milk exposure and lower atopic sensitization and asthma in childhood. *The Journal of allergy and clinical immunology.*2014;133(2):551-559.

33. Paparo L, Nocerino R, Cosenza L, et al. Epigenetic features of FoxP3 in children with cow's milk allergy.*Clinical epigenetics.* 2016;8:86.

34. Martino DJ, Ashley S, Koplin J, et al. Genomewide association study of peanut allergy reproduces association with amino acid polymorphisms in HLA-DRB1. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology.* 2017;47(2):217-223.

Figure legends

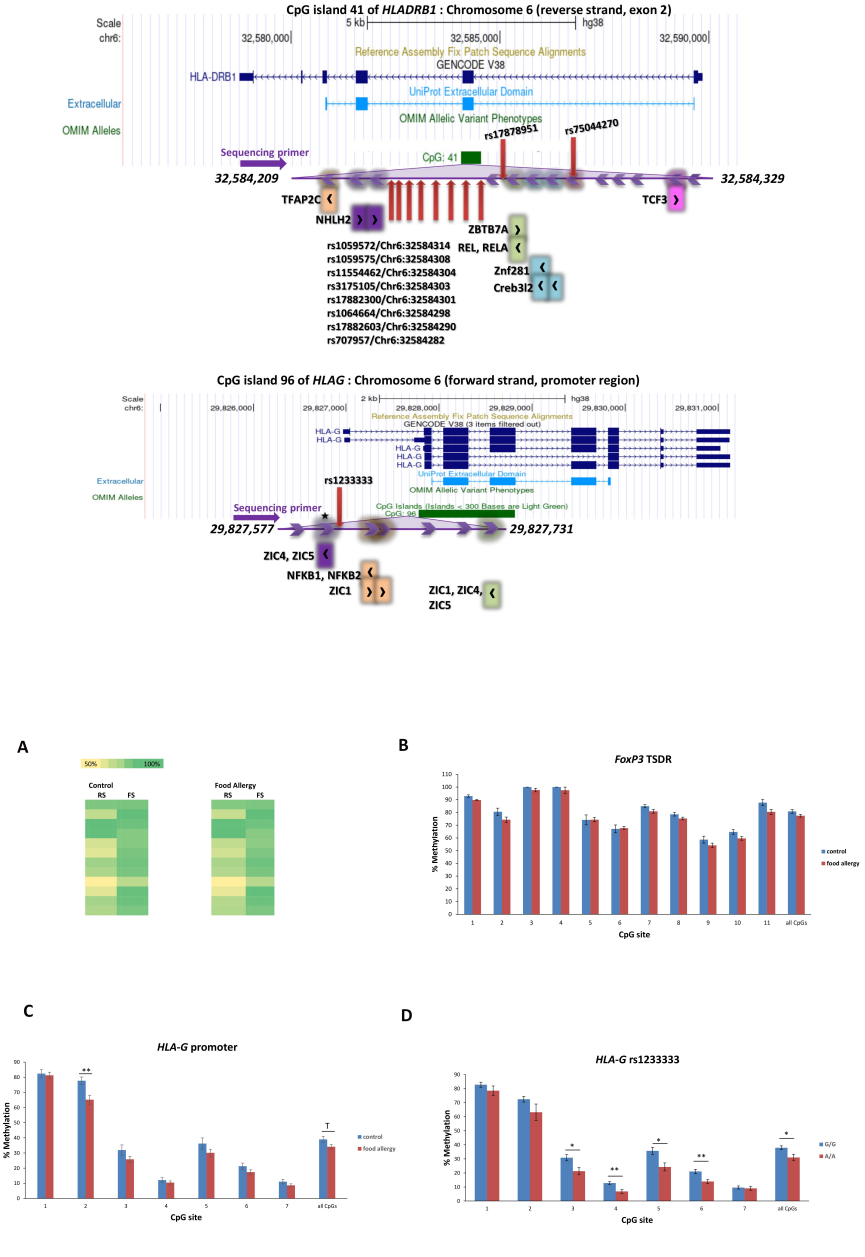
FIGURE 1: *HLA-DRB1* and *HLA-G* genes' chromosomal location, transcripts (dark blue), protein domains (light blue) and CpG islands (green boxes) from GENCODE v.38 (exons are presented as boxes and introns as lines, arrows indicate the 5'-3' direction). Analyzed genomic areas with CpGs are displayed as purple arrows. Different colors correlate with transcription factors' binding sites, potentially affected from methylation status of studied CpGs. SNPs and their location across the analyzed regions are displayed as red arrows. rs17878951 is involved in the creation of a new CpG site, while rs75044270 is involved in the deletion of a CpG site within*HLA-DRB1* region. Sequencing primers of each genetic region are also presented.

FIGURE 2: Methylation levels of *FOXP3* TSDR and HLA-G promoter. (A) Methylation pattern of reverse (R5) and forward (F5) DNA strand of *FOXP3* TSDR among control and FA sample group. Differential methylation profiles were observed between the two DNA strands. (B, C) Methylation status of specific CpG sites lying within*FOXP3* TSDR and *HLA-G* promoter and total values from all CpGs studied across each region. *HLA-G* promoter's CpG site 2 showed statistically significant differences between control and food allergy group (**p=0.004). (D) Effect of rs1233333 on DNA methylation levels of *HLA-G* promoter region. Homozygous genotype AA showed decreased methylation compared to homozygous GG. Data are presented as mean ± SE. CpG sites across studied region are numbered relative to 5'-3' direction.

* Indicates p value <0.05, **indicates p value < 0.01

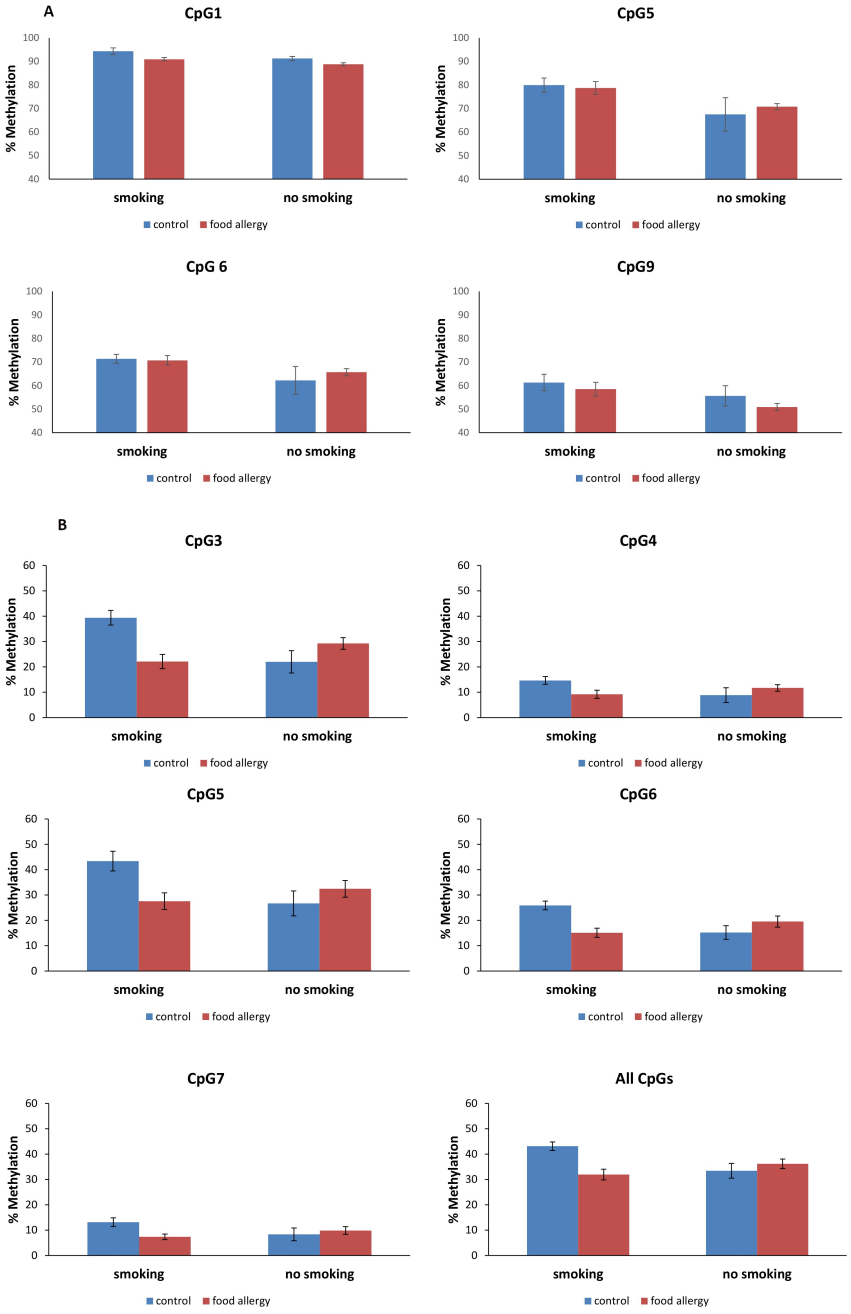
FIGURE 3 : DNA methylation levels of food allergic children and non-allergic controls with and without parental smoking history. (A) Parental smoking habit significantly influenced methylation levels of CpG site 1 (p=0.004), CpG site 5 (p=0.004), CpG site 6 (p=0.011) και CpG site 9 (p=0.023) of *FOXP3* TSDR regardless of FA or healthy status. (B) Parental smoking habit in combination with FA significantly influenced methylation levels of specific CpG sites in *HLA-G*promoter: CpG site 3 (p<0.001), CpG site 4 (p=0.028), CpG site 5 (p=0.009), CpG site 6 (p=0.002), CpG site 7 (p=0.036) and whole tested genetic region (p=0.004).

FIGURE 4 : Methylation levels of a 70bp highly polymorphic region within CpG41 of *HLA-DRB1* gene. Absence of SNPs is accompanied with complete demethylation (<5%), while the presence of SNPs causes simultaneous high methylation levels (>70%). Data are presented as mean ± SE.



Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.



Posted on Authorea 9 Feb 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164444724.45371782/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.