Clínica e Investigación en Arteriosclerosis 34 (2022) 33–35

Sociedad Española de Arteriosclerosis

CLÍNICA E INVESTIGACIÓN EN ARTERIOSCLEROSIS

www.elsevier.es/arterio

EDITORIAL

Why is it important to know DNA methylation patterns in people with hypertriglyceridaemia?*

¿Por qué es importante conocer los patrones de metilación del ADN en personas con hipertrigliceridemia?

Dolores Corella[a,b,*]

[a] *Department of Preventive Medicine and Public Health, School of Medicine, University of Valencia, Valencia, Spain*
[b] *CIBER Fisiopatología de la Obesidad y Nutrición, Instituto de Salud Carlos III, Madrid, Spain*

We are currently entering what is known as the era of personalized or precision medicine.[1] Although this concept is still emerging, the realization that ''one size does not fit all'' in the prevention and treatment of disease has led to the study of new biomarkers. These markers can be used to learn more about the molecular bases of the different phenotypes analysed, and they also enable better prediction of patient response to different preventive and/or therapeutic interventions within the context of precision or personalized medicine.[2] It is therefore crucial to search for these new biomarkers for each phenotype of interest (lipid concentrations in plasma, arterial blood pressure, fasting glycaemia, diabetes and cardiovascular diseases, etc.). Different types of biomarkers may be used. Interest centred at first on genomic markers based on fundamental changes in the DNA sequence.[3] These changes in the sequence run from polymorphisms of a single nucleotide, better known by their acronym SNPs (single nucleotide polymorphisms) to changes in a larger number of bases in the DNA sequence, such as short insertions/deletions or polymorphisms of longer fragments such as those known as CNV (copy number variations). In the early years of this century, after the termination of the Human Genome Project, DNA genetic variation analysis technology was still very expensive and slow.[4] Thus when it was used to study genetic susceptibility to different diseases in general, and hypertriglyceridaemia in particular, a few polymorphisms in candidate genes were studied. These works included the first studies by Hegele et al.,[5] which identified common and rare variants in the LPL and APOC3 genes. Technological progress then made it possible to genotype many thousands of polymorphisms along the entire genome, using massive genotyping chips. The number of genetic variants studied rose, and Genome-wide association studies (GWAS) were performed.[4] These GWAS made it possible to discover new genetic variations associated with hypertriglyceridaemia. GWAS and GWAS meta-analyses have grown exponentially in number, including hundreds of thousands of participants and producing new results every day.[6–11] The GWAS undertaken in different populations have most often included the study of the SNPs of the following genes, among others: MLXIPL ANGPTL3, GCKR, RPL26P19-HAVCR1, LPL, XKR6-AMAC1L2, FADS1-FADS2-FADS3, TRIB1 and APOA5-APOA4-APOC3-APOA1. Nevertheless, the study of genetic susceptibility to hypertriglyceridaemia based on analysis of the main genetic polymorphisms identified in GWAS has not proven very informative. Complementary alternatives have been suggested, such as analysis of genetic risk scores), while also considering the cumulative contribution of several SNPs; direct sequencing; or even studying genetic-environmental interactions.[12] In spite of this, these additional strategies have also been found to be incomplete and insufficient for use in the newly emerging precision medicine, even when ethnic differences in different populations are taken into account.[13]

We are currently entering what is known as the era of personalized or precision medicine.1Although this concept is still emerging, the realization that ''one size does not fit all'' in the prevention and treatment of disease has led to the study of new biomarkers. These markers can be used to learn more about the molecular bases of the different phenotypes analysed, and they also enable better prediction of patient response to different preventive and/or therapeutic interventions within the context of precision or personalized medicine.[2] It is therefore crucial to search for these new biomarkers for each phenotype of interest (lipid concentrations in plasma, arterial blood pressure, fasting glycaemia, diabetes and cardiovascular diseases, etc.). Different types of biomarkers may be used. Interest centred at first on genomic markers based on fundamental changes in the DNA sequence.[3] These changes in the sequence run from polymorphisms of a single nucleotide, better known by their acronym SNPs (single nucleotide polymorphisms) to changes in a larger number of bases in the DNA sequence, such as short insertions/deletions or polymorphisms of longer fragments such as those known as CNV (copy number variations). In the early years of this century, after the termination of the Human Genome Project, DNA genetic variation analysis technology was still very expensive and slow.[4] Thus when it was used to study genetic susceptibility to different diseases in general, and hypertriglyceridaemia in particular, a few polymorphisms in candidate genes were studied. These works included the first stud-

* Please cite this article as: Corella D. ¿Por qué es importante conocer los patrones de metilación del ADN en personas con hipertrigliceridemia? Clin Investig Arterioscler. 2022;34:33–35.

* Corresponding author:
E-mail address: dolores.corella@uv.es

ies by Hegele et al.,[5] which identified common and rare variants in the LPL and APOC3 genes. Technological progress then made it possible to genotype many thousands of polymorphisms along the entire genome, using massive genotyping chips. The number of genetic variants studied rose, and Genome-wide association studies (GWAS) were performed.[4] These GWAS made it possible to discover new genetic variations associated with hypertriglyceridaemia. GWAS and GWAS meta-analyses have grown exponentially in number, including hundreds of thousands of participants and producing new results every day.[6–11] The GWAS undertaken in different populations have most often included the study of the SNPs of the following genes, among others: MLXIPL ANGPTL3, GCKR, RPL26P19-HAVCR1,LPL, XKR6-AMAC1L2, FADS1-FADS2-FADS3,TRIB1 and APOA5-APOA4-APOC3-APOA1. Nevertheless, the study of genetic susceptibility to hypertriglyceridaemia based on analysis of the main genetic polymorphisms identified in GWAS has not proven very informative. Complementary alternatives have been suggested, such as analysis of genetic risk scores), while also considering the cumulative contribution of several SNPs, direct sequencing, or even studying genetic-environmental interactions.[12] In spite of this, these additional strategies have also been found to be incomplete and insufficient for use in the newly emerging precision medicine, even when ethnic differences in different populations are taken into account.[13]

Parallel to genome analysis based on DNA sequence changes, another new approach is epigenomic study. The epigenome is composed of functional elements that do not involve any change in the DNA sequence, although they are able to regulation gene expression and thereby influence different disease phenotypes.[4] Epigenetic modifications take different forms, and of these DNA methylation is among those that have been studied the most. This is caused by the enzymatic addition of a methyl group to the carbon 5 of the cytosine by methyltransferase action. The majority of 5-methylcytosines (5mC) are present in the dinucleotides -CpG-. Other demethylases would carry out the reverse process of eliminating methyl groups. In general the hypothesis is that there is an inverse correlation between DNA methylation levels and gene expression. However, there are exceptions to this, and it also depends on cell type and the stage of development. Further research is therefore required into the relationship between hypermethylation or hypomethylation and levels of the expression and functionality of each gene.[14] It is therefore necessary to increase research into new hypertriglyceridaemia biomarkers based on the study of the methylation of different genes in functionally relevant CpG locations. However, it is harder to study epigenetic methylation biomarkers than it is to analyse SNPs in the genome. The first difficulty consisted of the high cost of the technology and its under-developed nature in comparison with genotyping chips.[4] Moreover, unlike the genome the epigenome is different for each cell type, and the conclusions of studies in terms of which genes are differently methylated may vary widely. This depends on whether the biological sample analysed consist of peripheral blood, adipose tissue, skin, saliva, liver or other tissues.[4] Even if blood is used, the results may vary depending on its composition in the different types of leucocytes. In methylation studies of blood it is therefore necessary to take the precaution of measuring the different types of leucocyte in the fresh sample for the analysis of results, or to use computational algorithms which permit the indirect derivation of the proportions of leucocytes in the blood of each participant in the study, using the methylation data of different epigenome locations. Of these algorithms, the one proposed by Houseman et al. is the most widely used.[15] Although this adjustment based on different cell types is widely used, it has to be applied with caution because of the multiple collinearity which has been described. This means that the adjustment for these highly correlated variables within the statistical models may distort associations and give rise to false positives.[16] As is the case in SNP analysis, massive genotyping chips are used to analyse GWAS, and methylation chips are used for EWAS (epigenome-wide methylation study). These chips have more technical problems than the ones used in genotyping, and they require a process of delicate adjustment due to batch effects, which may also lead to false positives or negatives. Methylation may vary from 0% to 100%. Additionally, the different versions of these chips are not completely comparable, so that differences may arise between studies that use the earliest versions (27 K, or 450 K) and the most modern and complete version (EPIC chips that analyse 850,000 methylation locations).[17]

In spite of all these problems, in this edition of the journal *Clínica e Investigación en Ateriosclerosis*, we emphasise the publication of an EWAS that was undertaken by Guardiola et al.[18] This analysed the differentially methylated CpG locations at complete epigenome level in patients with severe hypertriglyceridaemia (n = 16, with an average of total triglycerides of 1687 mg/d) in comparison with 16 controls (average triglycerides 106 mg/dL), using the 850K chip. After carefully performing the quality controls necessary for leucocyte-based chip processing, the Houseman et al.[15] correction and those for multiple comparisons, together with the relevant statistical analysis, the authors identified 31 differentially methylated cytosines between the cases and controls. Among the most relevant locations they underline cg03636183 in the F2RL3 gene, which was found to be hypomethylated in individuals with hypertriglyceridaemia. This gene is very well-known for its association with ageing and cardiovascular risk. The Cg13824500 in the gene in VTI1A also stood out as functionally relevant. This is involved in the transit of quilomicrons within the enterocyte, and it is hypomethylated in individuals with hypertriglyceridaemia. Other locations with a relevant differential methylation in cases of hypertriglyceridaemia identified in this work were Cg26468118-RAB20 (hypomethylated) and cg21560722 - SBF2 (33% hypermethylated). Although the authors also identify other locations, they have not found any previous reference to support these findings. Although one limitation of this work is its small sample size, the advantage for use in our context is that the study took place in the Spanish population. We now know that geographical (and not only ethnic) differences in studied populations are highly relevant in genetics as well as in epigenetics, and that may not be possible to extrapolate the results obtained in a population for specific SNPs or CpG to another different population.[13,14] Thus although the EWAS and triglyceride studies published in other populations[19,20] are important and have to be taken into account as starting points, the clinical usefulness of biomarkers in precision medicine will depend on their validation as biomarkers in each specific population. Within this context it is indispensable to conduct more research into the population in our own environment, broadening and validating the results in the same way as the pioneering work presented here by Guardiola et al.[18]

Parallel to genome analysis based on DNA sequence changes, another new approach is epigenomic study. The epigenome is composed of functional elements that do not involve any change in the DNA sequence, although they are able to regulation gene expression and thereby influence different disease phenotypes.[4] Epigenetic modifications take different forms, and of these DNA methylation is among those that have been studied the most. This is caused by the enzymatic addition of a methyl group to the carbon 5 of the cytosine by methyltransferase action. The majority of 5-methylcytosines (5mC) are present in the dinucleotides –CpG–. Other demethylases would carry out the reverse process of eliminating methyl groups. In general, the hypothesis is that there is an inverse correlation between DNA methylation levels and gene expression. However, there are exceptions to this, and it also depends on cell type and the stage of development. Further research is therefore required into the relationship between hypermethylation or hypomethylation and levels of the expression and functionality of each gene.[14] It is therefore necessary to increase research into new hypertriglyceridaemia biomarkers based on the study of the methylation of different genes in functionally relevant CpG locations. However, it is harder to study epigenetic methylation biomarkers than it is to analyse SNPs in the genome. The first difficulty consisted of the high cost of the technology and its under-developed nature in comparison with genotyping chips.[4] Moreover, unlike the genome the epigenome is different for each cell type, and the conclusions of studies in terms of which genes are differently methylated may vary widely. This depends on whether the biological sample analysed consist of peripheral blood, adipose tissue, skin, saliva, liver or other tissues.[4] Even if blood is used, the results may vary depending on its composition in the different types of leucocytes. In methylation studies of blood it is therefore necessary to take the precaution of measuring the different types of leucocyte in the fresh sample for the analysis of results, or to use computational algorithms which permit the indirect derivation of the proportions of leucocytes in the blood of each participant in the study, using the methylation data of different epigenome locations. Of these algorithms, the one proposed by Houseman et al. is the most widely used.[15] Although

this adjustment based on different cell types is widely used, it has to be applied with caution because of the multiple collinearity which has been described. This means that the adjustment for these highly correlated variables within the statistical models may distort associations and give rise to false positives.[16] As is the case in SNP analysis, massive genotyping chips are used to analyse GWAS, and methylation chips are used for EWAS (epigenome-wide methylation study). These chips have more technical problems than the ones used in genotyping, and they require a process of delicate adjustment due to batch effects, which may also lead to false positives or negatives. Methylation may vary from 0% to 100%. Additionally, the different versions of these chips are not completely comparable, so that differences may arise between studies that use the earliest versions (27 K, or 450 K) and the most modern and complete version (EPIC chips that analyse 850,000 methylation locations).[17] In spite of all these problems, in this edition of the journal Clínica e Investigación en Aterioscleorsis, we emphasise the publication of an EWAS that was undertaken by Guardiola et al.[18] This analysed the differentially methylated CpG locations at complete epigenome level in patients with severe hypertriglyceridaemia (n = 16, with an average of total triglycerides of 1687 mg/d) in comparison with 16 controls (average triglycerides 106 mg/dL), using the 850 K chip. After carefully performing the quality controls necessary for leucocyte-based chip processing, the Houseman et al.[15] correction and those for multiple comparisons, together with the relevant statistical analysis, the authors identified 31 differentially methylated cytosines between the cases and controls. Among the most relevant locations they underline cg03636183 in the F2RL3 gene, which was found to be hypomethylated in individuals with hypertriglyceridaemia. This gene is very well-known for its association with ageing and cardiovascular risk. The Cg13824500 in the gene in VTI1A also stood out as functionally relevant. This is involved in the transit of quilomicrons within the enterocyte, and it is hypomethylated in individuals with hypertriglyceridaemia. Other locations with a relevant differential methylation in cases of hypertriglyceridaemia identified in this work were Cg26468118-RAB20 (hypomethylated) and cg21560722 - SBF2 (33% hypermethylated). Although the authors also identify other locations, they have not found any previous reference to support these findings. Although one limitation of this work is its small sample size, the advantage for use in our context is that the study took place in the Spanish population. We now know that geographical (and not only ethnic) differences in studied populations are highly relevant in genetics as well as in epigenetics, and that may not be possible to extrapolate the results obtained in a population for specific SNPs or CpG to another different population.[13,14] Thus although the EWAS and triglyceride studies published in other populations[19,20] are important and have to be taken into account as starting points, the clinical usefulness of biomarkers in precision medicine will depend on their validation as biomarkers in each specific population. Within this context it is indispensable to conduct more research into the population in our own environment, broadening and validating the results in the same way as the pioneering work presented here by Guardiola et al.[18]

References

1. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015;372(9):793–5.
2. Olivier M, Asmis R, Hawkins GA, Howard TD, Cox LA. The need for multi-omics biomarker signatures in precision medicine. Int J Mol Sci. 2019;20(19):4781.
3. Kamatani Y, Nakamura Y. Genetic variations in medical research in the past, at present and in the future. Proc Jpn Acad Ser B Phys Biol Sci. 2021;97(6):324–35.
4. Corella D, Ordovas JM. Basic concepts in molecular biology related to genetics and epigenetics. Rev Esp Cardiol (Engl Ed). 2017;70(9):744–53.
5. Hegele RA, Dron JS. 2019 George Lyman Duff Memorial Lecture: three decades of examining DNA in patients with dyslipidemia. Arterioscler Thromb Vasc Biol. 2020;40(9):1970–81.
6. Kathiresan S, Manning AK, Demissie S, D'Agostino RB, Surti A, Guiducci C, et al. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. BMC Med Genet. 2007;8 Suppl 1:S17.
7. Chasman DI, Paré G, Zee RY, Parker AN, Cook NR, Buring JE, et al. Genetic loci associated with plasma concentration of low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, apolipoprotein A1, and Apolipoprotein B among 6382 white women in genome-wide analysis with replication. Circ Cardiovasc Genet. 2008;1(1):21–30.
8. Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet. 2009;41(1):56–65.
9. Kurano M, Tsukamoto K, Kamitsuji S, Kamatani N, Hara M, Ishikawa T, et al. Genome-wide association study of serum lipids confirms previously reported associations as well as new associations of common SNPs within PCSK7 gene with triglyceride. J Hum Genet. 2016;61(5):427–33.
10. Siewert KM, Voight BF. Bivariate genome-wide association scan identifies 6 novel loci associated with lipid levels and coronary artery disease. Circ Genom Precis Med. 2018;11(12):e002239.
11. Graham SE, Clarke SL, Wu KH, Kanoni S, Zajac GJM, Ramdas S, et al. The power of genetic diversity in genome-wide association studies of lipids. Nature. 2021;600(7890):675–9, http://dx.doi.org/10.1038/s41586-021-04064-3. Epub 2021 Dec 9. PMID: 34887591; PMCID: PMC8730582.
12. Dron JS, Wang J, McIntyre AD, Cao H, Hegele RA. The polygenic nature of mild-to-moderate hypertriglyceridemia. J Clin Lipidol. 2020;14(1):28–34.e2, http://dx.doi.org/10.1016/j.jacl.2020.01.003. Epub 2020 Jan 14. PMID: 32033914.
13. Mostafavi H, Harpak A, Agarwal I, Conley D, Pritchard JK, Przeworski M. Variable prediction accuracy of polygenic scores within an ancestry group. Elife. 2020 Jan;30(9):e48376, http://dx.doi.org/10.7554/eLife.48376. PMID: 31999256; PMCID: PMC7067566.
14. Tang H, Zeng Z, Shang C, Li Q, Liu J. Epigenetic regulation in pathology of atherosclerosis: a novel perspective. Front Genet. 2021;12:810689, http://dx.doi.org/10.3389/fgene.2021.810689. PMID: 34976029; PMCID: PMC8714670.
15. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86, http://dx.doi.org/10.1186/1471-2105-13-86. PMID: 22568884; PMCID: PMC3532182.
16. Barton SJ, Melton PE, Titcombe P, Murray R, Rauschert S, Lillycrop KA, et al. In Epigenomic studies, including cell-type adjustments in regression models can introduce multicollinearity, resulting in apparent reversal of direction of association. Front Genet. 2019;10:816.
17. Vanderlinden LA, Johnson RK, Carry PM, Dong F, DeMeo DL, Yang IV, et al. An effective processing pipeline for harmonizing DNA methylation data from Illumina's 450K and EPIC platforms for epidemiological studies. BMC Res Notes. 2021;14(1):352.
18. Guardiola M, Ibarretxe D, Plana N, Masana L, Ribalta J. DNA methylation pattern of hypertriglyceridemic subjects. Clin Investig Arterioscler. 2021;S0214-9168(21), 00135-2. English, Spanish.
19. Pfeiffer L, Wahl S, Pilling LC, Reischl E, Sandling JK, Kunze S, et al. DNA methylation of lipid-related genes affects blood lipid levels. Circ Cardiovasc Genet. 2015;8(2):334–42.
20. Gomez-Alonso MDC, Kretschmer A, Wilson R, Pfeiffer L, Karhunen V, Seppälä I, et al. DNA methylation and lipid metabolism: an EWAS of 226 metabolic measures. Clin Epigenetics. 2021;13(1):7.