Power and Pitfalls of Computational Methods to Identify New Genes Responsible for Acute Liver Failure of Indeterminate Etiology in Adults

Abdelkader Heddar, MD¹,² and Micheline Misrahi, MD, PhD¹,²

Clinical and Translational Gastroenterology 2020;11: e00180. https://doi.org/10.14309/ctg.0000000000000180

We have read with great interest the recent article by Rakela et al. (1) describing genetic alterations detected by whole exome sequencing in a cohort of 26 adult patients with acute liver failure of indeterminate etiology (ALF-IE). The authors conclude that 12 variants of 11 genes were significantly associated with (ALF-IE). The authors probably used the old version of gnomAD, in which all the variants and polymorphisms were not aligned with the new version of the genome and not the updated version of gnomAD. This update issue is also true for the second missense variant of CYP2D6, which has neither a high frequency nor an overall frequency for all ethnic subgroups in the updated gnomAD database (Tables 1 and 2).

Furthermore, SERPINB11 is a unitary pseudogene in human genome (4). The pseudogene refers to a gene that has lost its functional counterpart during evolution because of a mutagenic event resulting in a premature stop codon. These mutagenic effects are often located in the initial portion of the coding gene, as for SERPINB11. Interestingly, SERPINB11 has the particularity of being polymorphic in humans (4). Hence, the rs4940595 (G/T) variant distinguishes between individuals bearing a functional copy of SERPINB11 with a wild-type c.268G allele and those bearing a nonfunctional copy, with the mutated c.268G>T allele, introducing a stop codon. The extremely high frequency of the mutated allele of SERPINB11 in the general population in different ethnic groups makes the link between this variant and ALF-IE in adults highly unlikely.

Finally, 5 of the 9 other variants located in MUC6, OR6f1, AADACL3, CYP2D7, and KIAA1161 also occur frequently in the gnomAD genome, with an AF up to 40%. For the remaining 4 variants, the results can be considered inconclusive because they correspond to multiallelic polymorphisms which must be referenced precisely, theirs ID being insufficient (Table 1). For instance, rs200886831 of RRP36 detected in 9 patients, refers to 4 different alleles and corresponds to in-frame deletion or insertion of either 6 or 12 bp. One of the 4 alleles has an AF of 45% in gnomAD, whereas the other 3 are extremely rare.

Thus, the conclusion raised by Rakela et al. on the significant association of the 12 variants with ALF-IE in adults is incompatible with careful examination of recent databases. A careful processing of exome and genome data is necessary to find out genes causing ALF-IE in adults.

CONFLICTS OF INTEREST

Guarantor of the article: Micheline Misrahi, MD, PhD.

Specific author contribution: A.H. performed reanalysis and interpretation of published computational data and wrote the letter. M.M. supervised the study, wrote, and revised the manuscript.

Financial support: None to report.

Potential competing interests: None to report.

REFERENCES

1. Rakela J, Rule J, Ganger D, et al. Whole exome sequencing among 26 patients with indeterminate acute liver failure: A pilot study. Clin Transl Gastroenterol 2019;10: e00867.
2. Kopanos C, Tsiodras V, Kouris A, et al. VarSome: The human genomic variant search engine. Bioinformatics 2019;35: 1978–80.
3. Karczewski KJ, Francioli LC, Tao G, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv 2019: 531210.
4. Zhang ZD, Frankish A, Hunt T, et al. Identification and analysis of unitary pseudogenes: Historic and contemporary gene losses in humans and other primates. Genome Biol 2010;11: R26.

¹Université Paris Saclay, UMR-S 1193, Faculté de Médecine, Le Kremlin-Bicêtre, France; ²Unité de Génétique Moléculaire des Maladies Métaboliques et de la Reproduction, AP-HP Hôpitaux Universitaires Paris-Saclay, Hôpital Bicêtre, Le Kremlin-Bicêtre, France.

Correspondence: Micheline Misrahi, MD, PhD. E-mail: micheline.misrahi@ap-hp.fr.

© 2020 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of The American College of Gastroenterology
| Gene     | rs                  | N   | chr | GRCh38 | REF | ALT | Canonical RNA transcript | DNA         | Protein          | Variant type | gnomAD E | gnomAD G |
|----------|---------------------|-----|-----|--------|-----|-----|--------------------------|-------------|------------------|--------------|----------|----------|
| SERPINB11 | rs4940595           | 23  | 18  | 63712604 | G   | T   | NM_080475.4               | c.268G>T     | p.Glu90Ter      | Non-sense    | 0.63     |          |
|          | rs4940595           | 0   | 18  | 63712604 | G   | A   | NM_080475.4               | c.268G>A     | p.Glu90Lys      | Missense     | 0.0000279 |          |
|          | rs4940595           | 0   | 18  | 63712604 | G   | C   | NM_080475.4               | c.268G>C     | p.Glu90Gln      | Missense     | 0.00000699 |          |
| CYP2D6   | rs1135840           | 16  | 22  | 42126611 | C   | G   | NM_000106.6               | c.1457G>C    | p.Ser486Thr    | Missense     | 0.577     |          |
| ANTXRL   | rs7091749           | 8   | 10  | 46330066 | C   | A   | NM_001278688.2            | c.1878C>A    | p.Pro626=      | Synonymous    | 0.000007  | 0.194    |
|          | rs7091749           | 10  | 46330066 | G   | C   | NM_001278688.2            | c.1878C>G    | p.Pro626=      | Synonymous    | 0.0103     |          |
| MUC6     | rs776572312         | 9   | 11  | 1016887 | G   | A   | NM_005961.3               | c.5914C>T    | p.Pro1972Ser   | Missense     | 0.00002  | 0.0118   |
| OR6J1    | rs1753430           | 14  | 14  | 22634064 | A   | G   | NM_001348233.1            | c.748T>C     | p.Ser250Pro    | Missense     | 0.41      |          |
| GNAL     | rs201898548         | 3   | 18  | 11689671 | G   | C   | NM_00000334049.11         | c.113_118delTGCCGC c.113_118delTGCCGC | p.Leu38_Ala39del | Insertion | 0.0842   | 0.141    |
|          | rs201898548         | 3   | 18  | 11689671 | T   | G   | NM_00000334049.11         | c.113_118dupTGCCGC | p.Leu38_Ala39dup | Inframe deletion | 0.00014  | 0.0168   |
| AADACL3  | rs3010877           | 7   | 1   | 12719616 | C   | T   | NM_001103170.3            | c.310C>T     | p.Pro104Ser    | Missense     | 0.151     |          |
| MCL1     | rs11580946          | 3   | 1   | 150578851| G   | A   | NM_0021960.5              | c.680C>T     | p.Ala227Val    | Missense     | 0.00842   | 0.0086   |
| CYP2D7   | rs2000754           | 15  | 22  | 42141587 | G   | A   | NM_001348386.3            | c.932C>T     | p.Ser311Leu    | Missense     | 0.848     |          |
|          | rs2000754           | 22  | 42141587 | T   | G   | NM_001348386.3            | c.932C>A     | p.Ser311Ter    | Non-sense     |          |          |
| RRP36    | rs200886831         | 9   | 6   | 43021676 | G   | T   | NM_000003244496.6         | c.43_48delGGGGGCGCC | p.Gly15_Ala16dup | Insertion | 0.000075 | 0.00104  |
|          | rs200886831         | 6   | 43021676 | G   | T   | NM_000003244496.6         | c.43_48delGGGGGCGCC | p.Gly15_Ala16del | Inframe deletion | 0.0419   | 0.285    |
|          | rs200886831         | 6   | 43021676 | G   | T   | NM_000003244496.6         | c.37_48delGGGGGCGGGGGGGCGCC | p.Gly15_Ala16dup | Insertion | 0.000014  |          |
|          | rs200886831         | 6   | 43021676 | G   | T   | NM_000003244496.6         | c.37_48delGGGGGCGGGGGGGCGCC | p.Gly15_Ala16del | Inframe deletion | 0.000187   | 0.000961 |
| KIAA1161 | rs4879782           | 6   | 9   | 34372875 | G   | C   | NM_020702.5               | c.69C>G      | p.Tyr23Ter     | Non-sense     | 0.251     |          |

ALT, alternative allele; Chr, chromosome; gnomAD, Allele frequency according to gnomAD exome database; gnomAD G, Allele frequency according to gnomAD genome database; GRCh38, position of the variant according to the last annotation of the human genome; N, number of patients with the variants in the cohort described by Rakela et al; REF, reference allele; rs: the identifier (ID) of the variant according to the last version of the human database of single nucleotide polymorphism (dbSNP 151)
Table 2. AF in GnomAD genome (V3) database of the 2 variants of SERPINB11 and CYP2D6 reported by Rakela et al. in a cohort of adult patients with acute liver failure of indeterminate etiology

| Population            | SERPINB11: c.268G>T, p.Glu90Ter (rs4940595) | CYP2D6: c.1457G>C, p.Ser486Thr (rs1135840) |
|-----------------------|---------------------------------------------|---------------------------------------------|
|                       | Allele count | Allele number | Homozygotes | AF  | Allele count | Allele number | Homozygotes | AF  |
| African               | 21,033       | 41,932        | 5,329       | **0.502** | 26,148       | 40,822        | 8,881       | **0.641** |
| Amish                 | 604          | 900           | 199         | **0.671** | 466          | 884           | 128         | **0.527** |
| Ashkenazi Jewish      | 2,171        | 3,322         | 723         | **0.654** | 2,138        | 3,312         | 714         | **0.646** |
| East Asian            | 1,206        | 3,124         | 232         | **0.386** | 2,162        | 3,060         | 782         | **0.707** |
| European (Finnish)    | 6,987        | 10,438        | 2,314       | **0.669** | 5,217        | 10,388        | 1,330       | **0.502** |
| European (Non-Finnish)| 46,721       | 64,532        | 16,858      | **0.724** | 25,748       | 64,030        | 10,432      | **0.558** |
| Latino                | 8,124        | 13,646        | 2,430       | **0.595** | 6,615        | 13,548        | 1,734       | **0.488** |
| South Asian           | 1,904        | 3,028         | 597         | **0.629** | 1,663        | 2,972         | 489         | **0.559** |
| Other                 | 1,362        | 2,150         | 438         | **0.633** | 1,223        | 2,126         | 369         | **0.575** |
| Total                 | 90,112       | 143,072       | 29,120      | **0.629** | 81,380       | 141,142       | 24,859      | **0.577** |

AF, allele frequency.

Open Access This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.