# Analysis of *ACE2* Genetic Variability Among Populations Highlights A Possible Link With COVID19-Related Neurological Complications

**Claudia Strafella** ( claudia.strafella@gmail.com )
Medical Genetics Laboratory, Tor Vergata University, Rome; Genomic Medicine Laboratory UILDM, IRCCS Santa Lucia Foundation, Rome https://orcid.org/0000-0003-1334-0920

**Valerio Caputo**
Medical Genetics Laboratory, Tor Vergata University, Rome; Genomic Medicine Laboratory UILDM, IRCCS Santa Lucia Foundation, Rome https://orcid.org/0000-0002-3503-3318

**Andrea Termine**
Genomic Medicine Laboratory UILDM, IRCCS Santa Lucia Foundation, Rome https://orcid.org/0000-0003-4374-7430

**Shila Barati**
Genomic Medicine Laboratory UILDM, IRCCS Santa Lucia Foundation, Rome

**Stefano Gambardella**
Neuromed Institute IRCCS, Pozzilli; Department of Biomolecular Sciences, University of Urbino "Carlo Bo", Urbino https://orcid.org/0000-0002-3727-4502

**Paola Borgiani**
Medical Genetics Laboratory, Tor Vergata University, Rome https://orcid.org/0000-0003-0859-4328

**Carlo Caltagirone**
Department of Clinical and Behavioral Neurology, IRCCS Fondazione Santa Lucia, Rome https://orcid.org/0000-0002-0189-4515

**Giuseppe Novelli**
Medical Genetics Laboratory, Tor Vergata University, Rome; Neuromed Institute IRCCS, Pozzilli https://orcid.org/0000-0002-7781-602X

**Emiliano Giardina**
Medical Genetics Laboratory, Tor Vergata University, Rome; Genomic Medicine Laboratory UILDM, IRCCS Santa Lucia Foundation, Rome https://orcid.org/0000-0002-2741-5009

**Raffaella Cascella**
Medical Genetics Laboratory, Tor Vergata University, Rome; Department of Biomedical Sciences, Catholic University Our Lady of Good Counsel , Tirana https://orcid.org/0000-0002-2148-0758

---

**Research Article**

**Keywords:** ACE2, SARS-COV-2, COVID19, eQTLs, bioinformatic analysis, neurological symptoms

**Posted Date:** May 15th, 2020

**DOI:** https://doi.org/10.21203/rs.3.rs-28871/v1



---

**Version of Record:** A version of this preprint was published at Genes on May 15th, 2020. See the published version at https://doi.org/10.3390/genes11070741.

# Abstract

The Angiotensin-converting enzyme 2 (ACE2) has been recently recognized as the entry receptor of the novel pathogenic Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-Cov-2). The presence of structural and sequence variants in *ACE2* gene may affect its expression in different tissues and determine a differential response to SARS-Cov-2 infection and COVID19-related phenotype. The present study investigated the genetic variability of *ACE2* in terms of Single Nucleotide Variants (SNVs), Copy Number Variations (CNVs) and expression Quantitative Loci (eQTLs) in a cohort of 268 individuals representative of the Italian general population. The analysis identified 5 SNVs (rs35803318, rs41303171, rs774469453, rs773676270, rs2285666) which displayed a significantly different frequency distribution in the Italian cohort compared to the worldwide populations. The analysis of eQTLs located in and targeting *ACE2,* revealed a high distribution of eQTL variants in different brain tissues, suggesting a possible link between the genetic variability of *ACE2* and the neurological complications in patients with COVID19. Further research is needed to clarify the possible relationship between *ACE2* expression and the susceptibility to neurological complications in patients with COVID19. In fact, patients at higher risk of neurological involvement may need different monitoring and treatment strategies in order to prevent severe, permanent brain injury.


# Introduction

The Angiotensin-converting enzyme 2 (ACE2) has recently caught the attention of the scientific community, since it has been recognized as the entry receptor of the novel pathogenic Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-Cov–2) [1]. ACE2 is a protein encoded by its homologous gene (*ACE2),,* which maps on chromosome X (Xp2.22) and consists of 18 exons. ACE2 is classified as an ectoenzyme with an extracellular catalytic domain, which is able to cleave circulating peptides [2]. ACE2 is expressed in several tissues, including airway epithelia, small intestine, heart, kidney, lung, testis and brain [2,3]. Concerning the function of ACE2, it takes part in the Renin Angiotensin System (RAS), by converting the Ang I to Ang-(1–9) and Ang II to Ang-(1–7) peptides, respectively. The function of Ang-(1–9) peptide is still unknown, whereas Ang-(1–7) peptide act as a vasodilator. Moreover, ACE2 is able to cleave other peptides, including apelin, kinins and morphins [2]. By cleavage and subsequent inactivation of Angiotensin II, ACE2 has been recognized as an essential determinant in the regulation of blood pressure, vascular function, metabolism and in the protection against atherosclerosis, heart, kidney, brain and lung injuries [2,3]. However, a pathophysiological function has also be assigned to ACE2 that is the entry receptor for SARS-Cov–1 and the novel SARS-Cov–2 viruses, which have been responsible of the Sever Acute Respiratory Syndrome (SARS) epidemic in 2003 and the current Coronavirus Infectious Disease (COVID19) pandemic, respectively [1]. Thanks to high binding affinity between the viral Spike (S) protein and ACE2 receptor, SARS-Cov–2 is able entering into the human cells and hijacks the cellular mechanisms for replicating itself and invading other cells [4]. Although ACE2 is utilized by SARS-Cov–2 as entry receptor, an accessory protein is necessary to prime the S protein, which is the Transmembrane Serine Protease 2 (TMPRSS2), which is highly expressed in tissues of the respiratory tract, including

bronchus, pharyngeal mucosa and lung and it is known to be critical for viral spread and pathogenesis in the infected host [4–6]. The lung is the most affected organ by SARS-Cov–2 infection, which has been shown to cause severe cytopathic effects into lung alveolar infected cells and induce an acute immuno-inflammatory response that, ultimately, result in lung injury [7]. It is important to note that ACE2 expression normally protects from lung injury and it has been found downregulated upon SARS-Cov–1 infection [4]. ACE2 downregulation has been regarded as an innate immune defence mechanism to impair the viral entry into the host cells [8]. On the other hand, the downregulation of ACE2 has also been shown to be a mechanism induced by the virus to fasten its cell-to-cell spreading by triggering the injury of the host lung tissue [8]. On this subject, whether SARS-Cov–2 interferes with ACE2 expression has to determined, although it is an interesting matter of investigation. The SARS-Cov–2 infection results in COVID–19 symptomatology that is characterized by a range of symptoms including dry cough, fever, myalgia or fatigue, dyspnoea and pneumonia [1,9]. In most severe cases, several clinical complications (Acute Respiratory Distress Syndrome, Sepsis, acute cardiac injury, acute kidney injury and secondary infections) have been described and associated with a more systemic response to the infection, which results to be fatal [9,10]. Over systemic and respiratory symptoms, 14–36% of severe patients with COVID19 show neurological signs including dizziness, headache, taste and smell impairment, impaired consciousness, encephalitis, seizures, ataxia, and stroke [11,12]. This finding, together with the evidence of SARS-Cov–2 in the cerebrospinal fluid of patients with COVID19, suggested that the neurotropic effects of SARS-Cov–2 may contribute to the morbidity and mortality caused by COVID19 [3]. Given the multiple roles of ACE2 and its expression pattern in different tissues, the alteration of its expression levels may have an impact on the susceptibility, symptomatology and outcome of COVID19. In particular, the presence of structural and sequence variants in *ACE2* gene may affect its expression in different tissues and determine a differential response to SARS-Cov–2 infection. In this context, the present study aimed to investigate the genetic variability of *ACE2* in terms of Single Nucleotide Variants (SNVs), Copy Number Variations (CNVs) and expression Quantitative Loci (eQTLs) in a cohort of 268 Italian individuals.

# Results And Discussion

The final goal of the study has been the research of variants potentially affecting ACE2 expression and function, which may contribute to SARS-Cov–2 spreading among worldwide populations and/or may have a clinical significance regarding the clinical variability and outcome displayed by patients with COVID19.

The analysis of CNVs in *ACE2* did not report any significant variation in our study cohort ruling out that frequent copy number variations could potentially impact *ACE2* expression. Concerning SNVs instead, the screening of the168 Italian samples for the 34 SNVs of interest revealed the presence of 5 variants: rs35803318 (C/T), rs41303171 (T/C), rs774469453 (A/-), rs773676270 (T/C), rs2285666 (C/T). Interestingly, these variants presented a significant differential distribution in Italian cohort with respect to worldwide populations (Table 1).

The rs35803318 (C/T) is a synonymous variant, whose allelic frequency resulted to be significantly different from the frequencies observed in African and Asian populations, whereas it overlapped with frequencies recorded in the European and American groups (Table 1 and Table 2).

In fact, the rs35803318 appeared to be more frequent in Italian, European and American populations compared to the very low frequency observed in the African and Asian cohorts (Table 1). Concerning rs41303171, the Italian population showed overlapping frequencies with American, European and Asian populations but not with the African cohort, in which the variant allele was almost absent (Table 1 and Table 2). The frequency of rs774469453 in the Italian population overlaps with all investigated populations except for the American group, which reported a slightly higher frequency (Table 1 and Table 2). The rs773676270 did not display significant differences in terms of frequency distribution of wild-type and variant alleles between Italian and the other worldwide populations (Table 2). The variant allele was extremely rare among all populations, with a slight increase in the frequency observed in the Italian cohort (Table 1).

Concerning the rs2285666, this is the only variant showing a higher frequency with respect to other variants previously discussed that appeared to be very rare among the investigated populations (Table 1). Concerning the frequency observed in the Italian population, it was found to be significantly different compared to African, Americans, Asian and even European population (Table 2). In particular, the variant allele of rs2285666 showed the lowest frequency in the Italian cohort in comparison to the other populations (Table 1). The frequency of the 5 SNVs were similarly distributed among male and female patients either in each population as well as comparing different populations, indicating that there is not gender effects underlying the frequency distribution of *ACE2* variants. Altogether, the variable frequencies of the SNVs analysed in the Italian population in comparison to the worldwide population frequencies suggested possible population-specific effects which may contribute to the variable susceptibility to Sars-Cov–2 infection. Supporting this thesis, a similar study on Chinese population investigated the genetic variability of *ACE2* in their population, finding a different frequency distribution of ACE2 variants with respect to the other populations [13]. Moreover, they found a higher allelic frequency of eQTLs variants associated with higher ACE2 expression in tissues, suggesting a different susceptibility or response to Sars-Cov–2 infection with respect to other populations under similar conditions [13]. However, these data are not sufficient to demonstrate a direct evidence of relationship between *ACE2* genetic variants and a differential susceptibility to Sars-Cov–2 infection among populations. In fact, data obtained on Italian and Chinese populations should be replicated in larger cohorts and in case/control studies in order to assess their potential association with the susceptibility to Sars-Cov–2 infection. Moreover, further functional studies should be carried out to demonstrate and explain these eventual associations.

Concerning the functional analysis of the 5 SNVs variants identified in the Italian cohort, most results were inconclusive or not significant to predict the functional impact of such variants on the resulting proteins. The rs35803318 is a synonymous variant, so that it does not result in an aminoacid change and thus it is unlikely to affect the protein function. Moreover, it is located in a region coding for the transmembrane portion of the protein, which normally does not interact with the Sars-Cov–2 S protein.

The rs41303171 is a missense variant resulting in the aminoacid change from Asparagine (Asn) with a neutral side chain to Aspartate carrying a negative charged side chain, which is therefore more hydrophilic. However, interrogation of Varsite showed that the change from Asn to Asp side change is not a large one, indicating that it may or may not result in a change to protein's function. The prediction analysis of the impact of rs41303171 on protein function was therefore inconclusive so that it is not possible to predict the functional impact of this variant at the moment. However, it is interesting to note that the rs41303171 is localized in the region coding for the extracellular portion of ACE2, which normally interact with the Sars-Cov–2 S protein. It will be therefore interesting to evaluate the impact of this variant by functional experiments in future. The rs774469453 variant is a single nucleotide deletion and it is located in a splicing intronic region. Therefore, it has been subjected to HSF analysis in order to test the variant for a potential alteration of splicing.

The HSF interrogation showed that the variant allele of rs774469453 may create an intronic Exonic Splicing Silencer (ESS) site but it is not significant and therefore it probably does not affect splicing. The rs773676270 is a synonymous variant localized in a region encoding the extracellular portion of the protein, which interact with the S protein of Sars-Cov–2. Interestingly, interrogation of HSF reported that this variant may affect the splicing by activating an exonic cryptic acceptor site or altering an Exonic Splicing Enhancer (ESE) site. These findings suggested that rs773676270 should be further investigated together with other variants with a larger effect on ACE2 function.

The rs2285666 variant was a variant located in the splice site region of *ACE2*. However, the prediction analysis by HSF did not reveal significant splicing alterations. Successively, all previously discussed variants were evaluated as potential eQTLs on GTEx database and only rs2285666 was classified as significant eQTL in several brain tissues, namely amygdala, anterior cingulate cortex, basal ganglia, cortex, cerebellum, hippocampus and hypothalamus (Fig.1).

As shown in Figure 1, the homozygous genotype for the variant allele may increase the expression of *ACE2* in multiple brain tissues and, consequently, may affect ACE2 functions in the brain. This finding suggested that the genetic variability of *ACE2* may have a greater impact on COVID19-related symptoms and Sars-Cov–2 tissue tropism rather than on the susceptibility to Sars-Cov–2 infection. On this subject, evidence of genetic variants in *TMPRSS* and high presence of eQTLs in the lung may suggest that the genetic variability of TMPRSS might have a role in determining the different susceptibility to Sars-Cov–2 infection among populations. However, these are preliminary observations, which have to be confirmed by further investigations.

Considering the broader effect of rs2285666 in multiple brain tissues, we decided to look at the distribution of the eQTL variants located in and targeting *ACE2* in the different brain tissues, which may thereby affect ACE2 expression and, in turn, contribute to the neurological symptoms and complications observed in patients with COVID19. Interestingly, literature studies highlighted the crucial role of ACE2 in brain physiology and pathophysiology, including marked regulatory effects on blood pressure, cardiac hypertrophy, stress response, anxiety, cognition and brain injury [2,14,15]. The eQTL analysis on GTex

portal and Biomart allowed identifying 29 significant eQTLs, which have been predicted to affect *ACE2* expression in brain at different levels (Supplementary Table S1). Most of them (23 eQTLs) have a significant effect on multiple brain tissues, suggesting that they may affect ACE2-related brain functions as a whole. Six eQTLs, instead, showed a more tissue-specific effect, indicating that they may be involved in the alteration of brain functions regulated by restricted areas of the brain. Interestingly, the mostly enriched tissues with significant ACE2-associated eQTLs were the basal ganglia, cortex, hypothalamus and substantia nigra, whereas amygdala and cerebellum appeared to be less affected. These findings suggested that the alteration of *ACE2* expression may be involved in different neurological symptoms (seizures, stroke, encephalitis, dizziness, headache, confusion, alteration of body temperature, anosmia, ataxia) observed in COVID19 patients in relation to the brain-affected area. These findings raised the need of further investigation on the role of *ACE2* genetic variability in the susceptibility and clinical outcome of patients with COVID19, especially concerning neurological symptoms. Indeed, these studies will be useful to identify patients at higher risk of neurological complications, which may need different monitoring and treatment strategies in order to prevent fatal outcomes or severe, permanent brain injury.

# Methods

The study was performed by utilizing 268 DNA samples representative of the Italian general population, which were partially available at the Genomic Medicine Laboratory of Santa Lucia Foundation Hospital and partially derived from international databases. The Italian cohort of samples was composed of 100 samples analysed by array Comparative Genomic Hybridization (aCGH) for assessing the presence of structural genomic variations and 168 samples utilized for identifying common and rare variants located in the coding or splice site regions of the genome. The genetic data referred to these samples were partially derived by Whole Exome Sequencing (WES) available at the Genomic Medicine Laboratory of Santa Lucia Foundation Hospital and partially obtained by Ensembl database [16–18]. Italian patients had an average age of 46±15 and a female:male ratio of 45:55. The research was approved by the Ethics Committee of Santa Lucia Foundation of Rome (CE/PROG.650 approved on 01/03/2018) and was performed according to the Declaration of Helsinki. The participants provided signed informed consent.

The CNV analysis was performed by Chromosome Analysis Suite (ChAS) 3.1 (Affymetrix, Santa Clara, CA, USA) using the Cytoscan750k_Array Single Sample analysis: NA33_hg19 as reference file. Concerning SNVs, we decided to analyzed SNVs (namely SNPs and indels) located within the coding and splice site DNA regions, because they are most likely to affect protein function. We therefore selected the SNV of interest by extracting the variants localized in the exonic and splice site regions of *ACE2,* whose frequency data were reported on 1000 Genomes database [19] that is accessible from Ensembl database [15]. This approach allowed selecting 34 putative variants located within *ACE2* sequence (Supplementary Table S2). The Ensembl database was also utilized to extract the frequency data of the 34 SNV of interest in the worldwide populations, in order to compare the frequencies between the Italian cohort and the frequencies observed in African, American, Asian and European populations. The presence of the 34 SNVs in the Italian samples was evaluated by analyzing the output file derived by WES and Ensembl database. The variant caller files (vcf) obtained by WES analysis were firstly scanned with vcfR [20] and

then subjected to GARFIELD-NGS analysis [21] in order to confirm the real existence of the detected variants. The allelic frequency distribution of the detected SNVs and the existence of significant differences among Italian and worldwide populations were calculated by statistical tools. All statistical analyses were performed in R environment [22]. To this purpose, two sided Fisher's Exact Test and a *p-value* (*p)* were calculated in order to assess the different allelic distributions of detected SNVs in the Italian cohorts with respect to the other populations. The significance threshold was set at $p < 0.05$. Considering that *ACE2* maps to the X chromosome, statistical analysis was also performed by stratifying the cohorts according to the gender to evaluate sex-related effects. Moreover, the SNVs detected in the Italian population were subjected to bioinformatic predictive analysis to assess their potential impact on ACE2 protein function. To this purpose, Mutation Taster [23], Human Splicing Finder [24], VarSite [25] and Uniprot database [26] were interrogated. Concerning eQTL analysis, GTex database [27] was utilized to retrieve the eQTLs variants with a significant effect on *ACE2* expression in different tissues and Biomart tool [28] to extract the significant eQTLs distributed on the basis of the affected tissue. The significance threshold for eQTLs analysis was set at $p < 0.05$.

# Declarations

Competing interests: The authors declare no competing interests

# References


1. Lan, J. et al. Structure of the SARS-CoV–2 spike receptor-binding domain bound to the ACE2 receptor. Nature. 10.1038/s41586–020–2180–5 (2020).
2. Alenina, N. & Bader, M. ACE2 in Brain Physiology and Pathophysiology: Evidence from Transgenic Animal Models. Neurochem. Res. 44, 1323–1329 (2019).
3. Baig, A.M., Khaleeq, A., Ali, U. & Syeda, H. Evidence of the COVID–19 Virus Targeting the CNS: Tissue Distribution, Host-Virus Interaction, and Proposed Neurotropic Mechanisms. ACS Chem. Neurosci. 11, 995–998 (2020).
4. Hoffmann, M. et al. SARS-CoV–2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell. 10.1016/j.cell.2020.02.052 (2020).
5. Iwata-Yoshikawa, N. et al. TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. J. Virol. 93, e01815-e01818 (2019).
6. Bertram, S. et al. TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium. J. Virol. 87, 6150–6160 (2013).
7. Zhang, H., Penninger, J. M., Li, Y., Zhong, N. & Slutsky A. S. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV–2 receptor: molecular mechanisms and potential therapeutic target. Intensive Care Med. 46, 586–590 (2020).

8. Guzzi, P. H., Mercatelli, D., Ceraolo, C. & Giorgi, F. M. Master Regulator Analysis of the SARS-CoV–2/Human Interactome. J. Clin. Med. 9, 982 (2020).
9. Huang, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 395, 497–506 (2020).
10. Mehta, P. et al. HLH Across Speciality Collaboration, UK. COVID–19: consider cytokine storm syndromes and immunosuppression. Lancet. 395, 1033–1034 (2020).
11. Li, Y. C., Bai, W. Z. & Hashikawa, T. The neuroinvasive potential of SARS-CoV2 may play a role in the respiratory failure of COVID–19 patients. J. Med. Virol. 10.1002/jmv.25728 (2020).
12. Helms, J. et al. Neurologic Features in Severe SARS-CoV–2 Infection. N. Engl. J. Med. 10.1056/NEJMc2008597 (2020).
13. Cao, Y. et al. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV–2) receptor ACE2 in different populations. Cell Discov. 10.1038/s41421–020–0147–1 (2020).
14. Wu, Y. et al. Nervous system involvement after infection with COVID–19 and other coronaviruses. Brain Behav. Immun. S0889–1591, 30357–3 (2020).
15. Natoli, S., Oliveira, V., Calabresi, P., Maia, L. F. & Pisani, A. Does SARS-Cov–2 invade the brain? Translational lessons from animal models. Eur. J. Neurol. 10.1111/ene.14277 (2020).
16. Strafella, C. et al. Limb-Girdle Muscular Dystrophies (LGMDs): The Clinical Application of NGS Analysis, a Family Case Report. Front. Neurol. 10, 619 (2019).
17. Cascella, R. et al. Assessing individual risk for AMD with genetic counseling, family history, and genetic testing. Eye (Lond.). 32, 446–450 (2018).
18. Cunningham, F. et al. Ensembl 2019. Nucleic Acids Res. 47(D1), D745-D751 (2019).
19. 1000 Genomes Project Consortium, et al. A global reference for human genetic variation. Nature. 526, 68–74 (2015).
20. Knaus, B. J. & Grunwald, N. J. VCFR: a package to manipulate and visualize variant call format data in R. Mol. Ecol. Resour. 17, 44–53 (2017).
21. Ravasio, V., Ritelli, M., Legati, A. & Giacopuzzi, E. Garfield-ngs: Genomic variants filtering by deep learning models in ngs. Bioinformatics. 34, 3038–3040 (2018).
22. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (2019).
23. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep- sequencing age. Nat. Methods. 11, 361–362 (2014).
24. Desmet, F. O. et al. Human splicing finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 37, e67 (2009).
25. Laskowski, R. A., Stephenson, J. D., Sillitoe, I., Orengo, C. A. & Thornton, J. M. VarSite: Disease variants and protein structure. Protein Sci. 29,111–119 (2020).
26. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 8, D506-D515 (2019).

27. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 348, 648–660 (2015).
28. Smedley D. et al. The BioMart community portal: an innovative alternative to large, centralized data repositories. Nucleic Acids Res. 43, W589–598 (2015).

# Tables

**Table 1.** Comparison of *ACE2* allelic variants between the Italian and worldwide populations. Allele counts and frequencies of the variants of interest for each population are shown. Italians (168 subjects); AFR: 661 African 1000 Genomes subjects; Americans: 446 subjects composed by the 99 CEU (Utah residents with Northern and Western European ancestry) subjects and the 347 AMR (Latino Americans) subjects from 1000 Genomes; Europeans: 297 subjects composed by the 107 IBS (Iberian population in Spain) subjects, the 91 GBR (British in England and Scotland) and the 99 FIN (Finnish in Finland) subjects from 1000Genomes; EAS: 504 East-Asian 1000Genomes subjects; SAS: 489 South-Asian 1000Genomes subjects. ACs: Allele Counts; AFs: Allele Frequencies.

| SNP | Italians | | AFR | | Americans | | Europeans | | EAS | | SAS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ACs | AFs | ACs | AFs | ACs | AFs | ACs | AFs | ACs | AFs | ACs | AFs |
| rs35803318 C/T | C: 231 T: 13 | C: 0.947 T: 0.053 | C: 1002 T: 1 | C: 0.999 T: 0.001 | C: 629 T: 44 | C: 0.935 T: 0.065 | C: 430 T: 26 | C: 0.943 T: 0.057 | C: 764 T: 0 | C: 1.000 T: 0.000 | C: 718 T: 0 | C: 1.000 T: 0.000 |
| rs2285666 C/T | C: 209 T: 35 | C: 0.857 T: 0.143 | C: 791 T: 212 | C: 0.789 T: 0.211 | C: 458 T: 215 | C: 0.680 T: 0.32 | C: 345 T: 111 | C: 0.757 T: 0.243 | C: 354 T: 410 | C: 0.463 T: 0.537 | C: 374 T: 344 | C: 0.521 T: 0.479 |
| rs41303171 T/C | T: 242 C: 2 | T: 0.992 C: 0.008 | T: 1003 C: 0 | T: 1.000 C: 0.000 | T: 668 C: 5 | T: 0.993 C: 0.007 | T: 447 C: 9 | T: 0.980 C: 0.020 | T: 764 C: 0 | T: 1.000 C: 0.000 | T: 717 C: 1 | T: 0.999 C: 0.001 |
| rs774469453 A/- | A: 242 -: 2 | A: 0.992 -: 0.008 | A:1002 -: 1 | A: 0.999 -: 0.001 | A: 650 -: 23 | A: 0.966 -: 0.034 | A: 456 -: 0 | A: 1.000 -: 0.000 | A: 757 -: 7 | A: 0.999 -: 0.001 | A: 716 -: 2 | A: 0.997 -:0.003 |
| rs773676270 T/C | T: 243 C: 1 | T: 0.996 C: 0.004 | T: 1003 C: 0 | T: 1.000 C: 0.000 | T: 673 C: 0 | T: 1.000 C: 0.000 | T: 456 C: 0 | T: 1.000 C: 0.000 | T: 764 C: 0 | T: 1.000 C: 0.000 | T: 717 C: 1 | T: 0.999 C: 0.001 |

**Table 2.** Statistical results of the 5 SNVs in the Italian population versus (vs) worldwide populations. Two sided Fisher Exact Test's *p-value* (*p*) showing the statistical significance of differences of the frequency distributions of the 5 variants detected in the Italian cohort with respect to the frequencies observed in the other populations. ns: not significant.

| SNP | *p* Italians vs AFR | *p* Italians vs Americans | *p* Italians vs Europeans | *p* Italians vs EAS | *p* Italians vs SAS |
|---|---|---|---|---|---|
| rs35803318 C/T | $5.48*10^{-9}$ | ns | ns | $7.65*10^{-9}$ | $1.41*10^{-8}$ |
| rs2285666 C/T | 0.016 | $4.1*10^{-8}$ | 0.002 | $2.11*10^{-29}$ | $3.55*10^{-22}$ |
| rs41303171 T/C | 0.038 | ns | ns | ns | ns |
| rs774469453 A/- | ns | 0.036 | ns | ns | ns |
| rs773676270 T/C | ns | ns | ns | ns | ns |

# Figures

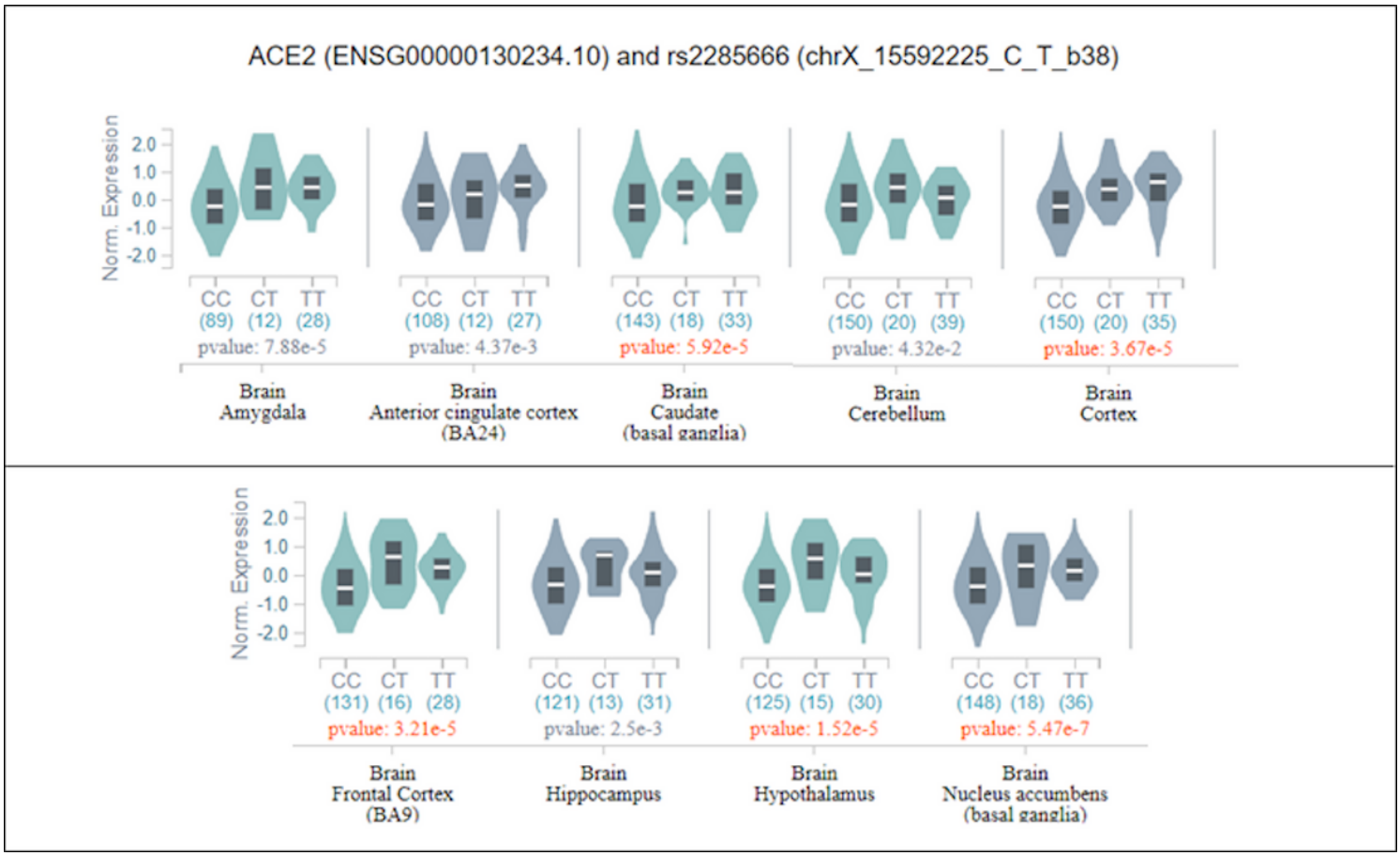


**Figure 1**

Violin plots represent the correlation between rs2285666 genotypes and ACE2 mRNA expression in different brain tissues. The reports on the x-axis the genotypes and the correspondent counts in brackets. On the y-axis, the normalized expression of ACE2 is reported. Moreover, a p-value of for the variant significance is reported. The figure has been obtained by GTex [27].

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1.xlsx
- SupplementaryTableS2.docx