

1 **Title: C allele of rs479200 of the host EGLN1 gene - a risk factor for severe COVID-19**
2 **(pilot study)**

3 Authors: Renuka Harit¹, Sajal De^{#2}, Piyooash Kumar Singh^{#1}, Deepika Kashyap¹, Manish
4 Kumar¹, Dibakar Sahu², Chander Prakash Yadav¹, Mradul Mohan¹, Vineeta Singh¹,
5 Ram Singh Tomar¹, Kailash C Pandey^{1,3*}, Kapil Vashisht^{1*}

6 ¹ ICMR-National Institute of Malaria Research, New Delhi

7 ² All India Institute of Medical Sciences, Raipur, Chhattisgarh

8 ³ Academy of Scientific and Innovative Research (AcSIR), Ghaziabad

9

10 #Authors contributed equally

11 Correspondence: kapil@mrcindia.org, kailash.pandey@icmr.gov.in

12

13 **Abstract**

14 **Background**

15 Coronavirus disease-2019 (COVID-19) symptoms can range from asymptomatic, moderate to
16 severe manifestations that result in an overall global case fatality rate of 2-7 %. While each
17 variant has had its challenges, and some variants are more severe than others, risk factors of
18 severe COVID-19 are still under investigation. In this context, the host genetic predisposition
19 is also a crucial factor to investigate. In the present study, we investigated host genotypes of
20 the SNP rs479200 of the host EGLN1 gene, previously implicated in high altitude pulmonary
21 edema (HAPE), some of whose symptoms such as hypoxia profoundly overlap with severe
22 COVID-19.

23 **Methods**

24 After informed consent, 158 RT-PCR confirmed COVID-19 patients (March 2020 to June
25 2021) were enrolled in the study. Based on their clinical manifestations, disease severity was
26 categorized by the clinical team. Blood samples were drawn and DNA was extracted from the
27 clot to infer different genotypes of the SNP rs479200 of the host EGLN1 gene. PCR-RFLP
28 analysis of the SNP rs479200 (C > T) was performed with an amplicon size of 367 bp. Various
29 genotypes (TT, TC and CC) were assigned based on the presence/absence of a restriction site
30 (T/GTACA) for restriction enzyme *BsrGI*. Allele frequencies, Hardy-Weinberg Equilibrium
31 (HWE) and multinomial logistic regression were performed using statistical tool SPSS version
32 23 (IBM).

33

34

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

1 **Findings**

2 We observed that the severe COVID-19 category was composed of comparatively younger
3 patients with mean age (34.9 ± 15.6), compared to asymptomatic and moderate categories whose
4 mean age was 49.7 ± 17.9 and 54.3 ± 15.7 , respectively. Preponderance of males and high
5 heterozygosity (TC) was observed across the clinical categories. Notably, the frequency of C
6 allele (0.664) was 2-fold higher than the T allele (0.336) in severe COVID-19 patients, whereas
7 the allele frequencies were similar in asymptomatic and moderate category of COVID-19
8 patients. Multinomial logistic regression showed an association of genotypes with increasing
9 clinical severity; odds ratio (adjusted OR- 11.414 (2.564-50.812)) and (unadjusted OR- 6.214
10 (1.84-20.99)) for the genotype CC in severe category of COVID-19. Interestingly, the TC
11 genotype was also found to be positively associated with severe outcome (unadjusted OR-5.816
12 (1.489-22.709)), indicating association of C allele in imparting the risk of severe outcome.

13 **Interpretation**

14 The study provides strong evidence that the presence of C allele of SNP (rs479200) of the
15 EGLN1 gene associates with severity in COVID-19 patients. Thus, the presence of C allele
16 may be a risk factor for COVID-19 severity. This study opens new avenues towards risk
17 assessment that include EGLN1 (rs479200) genotype testing and identifying patients with C
18 allele who might be prioritized for critical care.

19 **Keywords**

20 EGLN1; COVID-19; PCR-RFLP; Genotype; SNPs, rs479200.

21

1 Introduction

2 Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is the causal agent of
3 coronavirus disease-2019 (COVID-19). The continuous evolution of the virus has generated
4 different variants (Alpha, Beta, Gamma, Delta & Omicron) with their characteristic
5 transmission and virulence profile. The clinical disease symptoms vary drastically, majority of
6 the infections remain asymptomatic or moderate, often progressing to severe complications or
7 mortality ranging from 2-7 % ¹. Symptoms of severe COVID-19 pneumonia may include acute
8 respiratory distress symptoms- hypoxia, dyspnea, hypocapnia or hypercapnia ². However, the
9 precise determinants of asymptomatic infections turning into severe disease are elusive.
10 Interestingly, the COVID-19 hypoxia partially overlaps with high altitude pulmonary edema
11 (HAPE) with some atypical features such as preserved lung compliance, intrapulmonary shunt
12 and mild dyspnea. These features indicate towards loss of the homeostatic oxygen-sensing
13 system which regulates oxygen uptake and systemic delivery. At molecular level, the
14 mitochondrial proteins had been altered during SARS-CoV-2 infection leading to plausible
15 disruptions in oxygen-sensing mechanisms ³. Several risk factors are being explored- both from
16 the virus and host origin; these broadly include age, gender, BMI and pre-existing
17 comorbidities ⁴. Thorough investigations of the host genetic predisposition is crucial as several
18 genes have been reported to be associated with inflammation and immune responses and are
19 considered as risk factors ⁴. Moreover, inter-individual variations in the effector or regulator
20 genes might play role in progression to severe disease.

21 Focusing on the narrative of COVID-19 hypoxia, EGLN1 is one of the gene that has been
22 comprehensively studied in context of HAPE hypoxia ⁵. The EGLN1 gene product- (2-
23 oxoglutarate-dependent dioxygenase) regulates the hypoxia-inducible factor (HIF- α) *via* prolyl
24 hydroxylation. Whereas in hypoxic conditions inactivation of EGLN1 leads to increased HIF-
25 α and subsequent dimerization with HIF- β activate the hypoxia response elements (HRE) ^{6,7}.
26 The HRE response may include, but is not limited to alterations in metabolic, cellular and
27 systemic responses inside the cells. Numerous studies in the literature have correlated the inter-
28 individual variations of the EGLN1 gene (SNPs) in HAPE susceptibility ^{5,8,9}, aerobic capacity
29 ¹⁰, predisposition to preeclampsia ¹¹ in different ethnicities. The TT genotype of the rs479200
30 SNP of the EGLN1 gene has been shown to associate with HAPE susceptibility ⁵. The EGLN1
31 gene on chromosome 1 translates into a 426 amino acid protein ¹². Among the several
32 genotypes of the EGLN1 gene, six of them (rs2472261, rs480902, rs479200, rs2808611,
33 rs2153364, rs2491405) are present in the various ethnic populations of India ¹³. Specifically,

1 rs479200 and rs480902 SNPs have been repeatedly implicated for significant associations with
2 HAPE patients^{5,14}. Mishra et al.¹⁵ have observed a multi-fold increase in EGLN1 expression in
3 HAPE patients and significant association with the risk genotypes, their haplotypes and
4 interacting genotypes. The arterial oxygen saturation (SaO₂) levels were also found to be
5 inversely correlated with pulmonary artery systolic pressure (PASP) and EGLN1 expression.
6 SNPs from candidate genes have revealed significant differences of gene expression in HAPE
7 patients and healthy controls- ACE (rs4309), HSP70 (rs1008438), SP-A2 (rs1965708), PAI-1
8 (rs1799889), NOS (rs199983) and EGLN1 (rs480902) etc.¹⁶. The seven polymorphisms of
9 EGLN1 gene- rs1538664, rs479200, rs2486729, rs2790879, rs480902, rs2486736 and
10 rs973252; were found to be significantly different in HAPE patients and HAPE controls¹⁴. A
11 significant number of patients in severe COVID-19 mimic HAPE hypoxia¹⁷.

12 The lung transcriptome of COVID-19 patient implicated interaction of viral proteins ORF8 and
13 OS9 with EGLN1 and EGLN2 genes with potential role in disrupting the EGLN1-HIF oxygen
14 sensor mechanism^{18,19}. It is worth noting that EGLN1 gene acts as a natural regulator of the
15 HIF- α protein and hence corroborates the hypothesis that inter-individual variations in the
16 EGLN1 gene might play role in the disease severity and subsequently risk assessment.
17 Interestingly, Roxadustat- HIF inhibitor shown to reduce ACE-2 expression, thereby inhibiting
18 entry and replication of SARS-CoV-2 inside lung epithelial cells, thus argued as potential
19 therapeutic agents for COVID-19²⁰. In the present study, categorized COVID-19 patients were
20 analyzed for the SNP rs479200 of the EGLN1 gene and potential associations of the genotype
21 with disease severity were investigated.

22

23 **Material and methods**

24 **Sample collection and DNA isolation**

25 Adult patients (age \geq 18 years) with RT-PCR or Rapid Antigen Test positive for COVID-19
26 were included in the study. The patients were classified and defined by the clinical team as
27 asymptomatic COVID-19 cases (no hospitalization); moderate COVID-19 infection (requiring
28 oxygen support during hospitalization) and severe COVID-19 infection (required
29 invasive/non-invasive ventilator support for $>$ 24hrs) . Clotted blood samples of the COVID-
30 19 patients were procured from the biorepository at ICMR-NIMR, Delhi and AIIMS, Raipur.
31 The study was approved by the Institutional Ethics Committees (IECs), AIIMS, Raipur
32 [1379/IEC-AIIMSRPR/2020] & ICMR-NIMR, Delhi [PHB/NIMR/EC/2020/145]. Nucleic

1 acid extractions were performed at the COVID-19 testing facility at ICMR-NIMR, Delhi, using
2 appropriate precautionary measures and personal protective equipment (PPE). DNA isolation
3 from small fractions of the clotted blood was performed using DNeasy Blood and Tissue kit,
4 Qiagen, Germany. Isolated DNA samples were quantified and purity was assessed using
5 Quawell UV-Vis Spectrophotometer Q5000.

6 **Amplification for rs479200 and RFLP analysis**

7 The EGLN1 gene sequence was retrieved from NCBI using Gene ID: 54583 for the egl-9
8 family hypoxia-inducible factor 1 [Homo sapiens (human)]. The reference SNP (rs479200)
9 was identified from dbSNP Reference (rs) SNP report at the chromosomal position
10 chr1:231408034 (GRCh38.p13) with alleles G>A / G>C. A 367 bp segment of the EGLN1
11 gene covering the SNP rs479200 was amplified using primers- FP-
12 5`CTCCCAGCACATCTGTGAAT3` and RP- 5`TCGGATGGAAAGGTGGTAAAG3`. PCR
13 amplification of the SNP region rs479200 with the above primers was performed using
14 DreamTaq PCR Mastermix (2x) (Thermo Scientific, USA) in following steps- 1) initial
15 denaturation [95 °C / 3 minutes, 2) denaturation 95 °C / 30 seconds, 3) annealing 55 °C / 30
16 seconds, 4) extension 72 °C / 30 seconds, 5) steps 2-4 for 34 cycles and 5) final extension 72
17 °C / 5 minutes. The amplified products were visualized on 2% agarose gel and stored at 4 °C
18 until further use.

19 For RFLP analysis, the 367 bp amplicon containing the rs479200 was cut with restriction
20 enzyme *BsrGI* (T/GTACA)¹¹. The reaction comprised- 10x Cut Smart Buffer, 5 U *BsrGI*-HF
21 (New England Biolabs, USA), 100 ng of PCR amplicon containing rs479200 as template and
22 incubated at 37 °C for 30 minutes. The digested products were visualized on 2% agarose gel to
23 decipher the various genotypes of SNP rs479200- homozygous (TT & CC), heterozygous
24 (TC).

25 **Statistical analysis**

26 The study data was fed in Microsoft Excel 365 and cross checked by two independent
27 researchers. The statistical and association analysis was performed using the SPSS version 23
28 (IBM). The statistical analysis included the chi-square, genotype and allele frequency, Hardy-
29 Weinberg analysis. Further the multinomial regression was employed to test the association of
30 demographic and genetic factors.

31

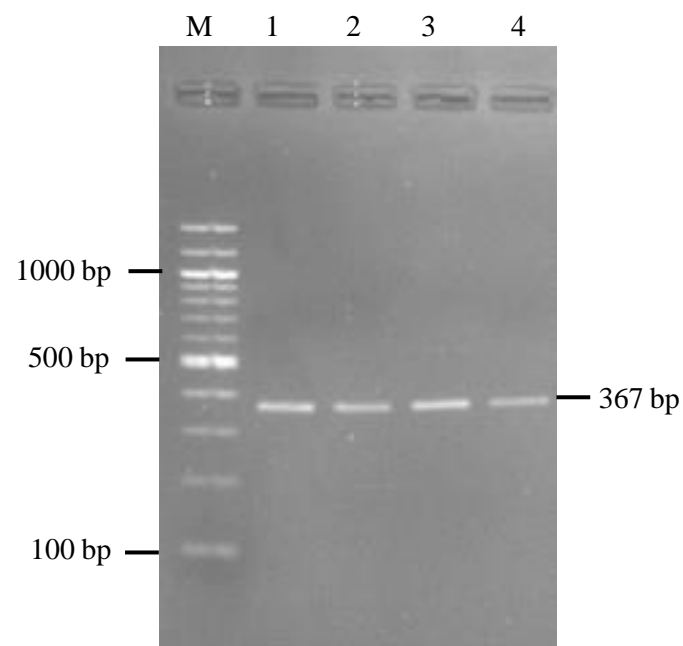
1 Results

2 Blood samples from 158 COVID-19 patients were tested by PCR-RFLP analysis to assess the
 3 genotypes of SNP rs479200 of the EGLN1 gene, using restriction enzyme *BsrGI* (T/GTACA).
 4 **Table 1** shows the demographic profiles of the patients included in the study, classified
 5 according to the disease severity. In our study, the mean age of the total number of participants
 6 was 45.9 ± 18.3 , however the mean age of severe COVID-19 patients was 34.9 ± 15.6 , contrary
 7 to the notion of prevalent severe complications in older patients. The mean age of COVID-19
 8 patients in the asymptomatic and moderate disease was 49.7 ± 17.9 and 54.3 ± 15.7 , respectively.
 9 In addition, preponderance of male COVID-19 patients was observed compared to females,
 10 across the clinical categories and also in a total number of participants, but it is unclear if this
 11 distribution is an attribute to social or biological causes.

	Asymptomatic		Moderate		Severe		Total		
	N	%	N	%	N	%	N	%	
Age Mean \pmSD	49.7 \pm 17.9		54.3 \pm 15.7		34.9 \pm 15.6		45.9 \pm 18.3		
Gender	Female	9	23.70%	11	17.70%	18	31.00%	38	22.75
	Male	29	76.30%	51	82.30%	40	69.00%	120	71.86

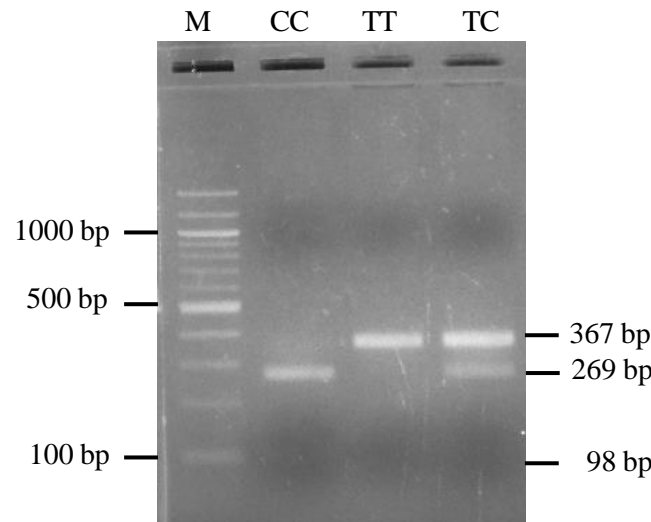
12 **Table 1:** Demographic profiles of the COVID-19 patients classified according to disease
 13 severity.

14 We then successfully amplified the EGLN1 gene segment (367 bp), flanking the SNP
 15 rs479200 (367 bp) from the DNA isolated from clotted blood of the COVID-19 patients (**Fig.**
 16 **1**).



17 **Figure 1:** SNP rs479200 of the EGLN1 gene amplified from DNA isolated from clotted blood samples.
 18

1 Further, PCR-RFLP analysis revealed the genotypes of the SNP rs479200 of the EGLN1 gene
2 (**Fig. 2**)- *TT* genotype due to absence of ‘C’ at restriction site and hence no digestion (367 bp)
3 by the restriction enzyme; *CC* genotype due to the presence of ‘C’ at restriction site leading to
4 complete digestion (269 + 98 bp); *TC* genotype due to the presence of ‘C’ at restriction site,
5 only in one allele and incomplete digestion conferring heterozygous allele (367 + 269 + 98 bp).



6
7 **Figure 2:** PCR-RFLP analysis of SNP rs479200 of EGLN1 gene with enzyme *BsrGI* (T/GTACA).
8 The genotypes, allele frequencies of the EGLN1 gene rs479200 (T/C) and Hardy-Weinberg
9 Equilibrium (HWE) of the study population are summarized in **Table 2**. The data showed high
10 heterozygosity across the clinical categories and in the total number of study patients. It was
11 observed that both the alleles (T & C) were uniformly present in the total study patients with a
12 high proportion of ancestral T allele, except for the severe category patients. The frequency of
13 ‘C’ allele (0.664) was observed to be 2-fold higher in proportion than the ‘T’ allele (0.336).
14 The allele frequencies were in HWE status across the disease categories.

		Gender		Classification				
		Female	Male	Asymptomatic	Moderate	Severe	Total	
Genotypes	TT	N	11	21	12	16	7	35
		%	28.20%	17.10%	31.60%	25.80%	10.40%	20.96%
	TC	N	20	61	18	33	31	82
		%	51.30%	49.60%	47.40%	53.20%	46.30%	49.10%
	CC	N	8	41	8	13	29	50
		%	20.50%	33.30%	21.10%	21.00%	43.30%	29.94%
Chi square (p-value)		3.484 (.175)		13.274 (.010*)				
Alleles	T	0.538	0.419	0.553	0.524	0.336	0.455	
	C	0.462	0.581	0.447	0.476	0.664	0.545	
HW equilibrium		0.039	0.044	0.067	0.278	0.093	0.01662	
p-value		0.843	0.835	0.796	0.598	0.761	0.89741	

Table 2: Genotypes, allele frequencies and Hardy–Weinberg Equilibrium (HWE) of the COVID-19 study patients.

Multinomial logistic regression was performed in an adjusted and unadjusted model to test the association of genotype variables with clinical severity. Moderate and severe COVID-19 patients were compared with the asymptomatic patients, separately. In the unadjusted model the effects of genotype and disease severity were inferred, while the adjusted model, the genotypes were tested in isolation for their association with moderate and severe clinical outcomes. **Table 3** depicts the association analysis data, highlighting the association of genotype (CC) and genotype (TC) with severe outcome in COVID-19. The highest odds ratio (11.414 (2.564-50.812)) was observed with genotype (CC) in the adjusted model in asymptomatic vs severe COVID-19 subject's comparison; however, the unadjusted model genotype (CC) shows the odds ratio of 6.214 (1.84-20.99). The heterozygous genotype (TC) was also found to be positively associated with disease severity (unadjusted OR-5.816 (1.489-22.709)). The genotypes did not demonstrate any association in the comparison scenario of moderate vs asymptomatic patients.

Variable (genotypes)		B- coefficient	Wald (Sig.)	Odds Ratio (CI)
Asymptomatic Vs Moderate				
Genotype (CC)	Adjusted	0.077	0.016 (0.9)	1.08 (0.326-3.576)
	Unadjusted	0.198	0.113 (0.737)	1.219 (0.384-3.871)
Genotype (TC)	Adjusted	0.199	0.162 (0.688)	1.22 (0.463-3.215)
	Unadjusted	0.318	0.438 (0.508)	1.375 (0.535-3.532)
Asymptomatic Vs Severe				
Genotype (CC)	Adjusted	2.435	10.213 (0.001)	11.414 (2.564-50.812)
	Unadjusted	1.827	8.653 (0.003)	6.214 (1.84-20.99)
Genotype (TC)	Adjusted	1.761	6.417 (0.011)	5.816 (1.489-22.709)
	Unadjusted	1.083	3.733 (0.053)	2.952 (0.984-8.854)

1 **Reference genotype (TT)**

2 **Table 3:** Multinomial logistic regression to infer the association of genotypes with disease
3 severity.

4 **Discussion**

5 Whilst there is a consensus that severe COVID-19 patients present acute respiratory distress,
6 mimicking attributes of high altitude pulmonary edema (HAPE) ². However, the similarities in
7 the pathophysiology of severe COVID-19 and HAPE are debatable ²¹. Failure of respiratory
8 homeostasis in acute respiratory distress syndrome (ARDS), severe COVID-19 and HAPE
9 include variable hypoxia, dyspnoea, hypoxic pulmonary vasoconstriction (HPV),
10 Ventilation/perfusion (V/Q) mismatch, pulmonary hypertension, pulmonary capillary pressure
11 and reduced lung compliance ¹. HIF- α is shown to be central to the oxygen-sensing mechanism
12 recruited by cells in response to hypoxia, whose expression is regulated by 2-oxoglutarate-
13 dependent dioxygenase (2ODD- a product of EGLN1 gene); acting as a natural inhibitor of
14 HIF- α . In normoxic conditions, a balanced proteasomal degradation of constitutively expressed
15 HIF-1 α induced by 2ODD abrogates the translocation of HIF-1 α and HIF- β dimer to activate
16 HRE transcription. However, in a hypoxic scenario the balance of 2ODD and HIF-1 α skews
17 towards overproduction of HIF-1 α , leading to dimerization with HIF- β and transcriptional
18 activation of HRE²². Higher levels of HIF- α have also been shown to induce higher expression
19 of angiotensin-converting enzyme-2 (ACE-2) receptor on cell surface facilitating the viral entry
20 in the case of SARS-CoV-2. Further, complicating the situation are the interindividual
21 variations of the EGLN1 gene, that have been observed to be responsible for varying
22 capabilities in coping with the anti-hypoxic responses.

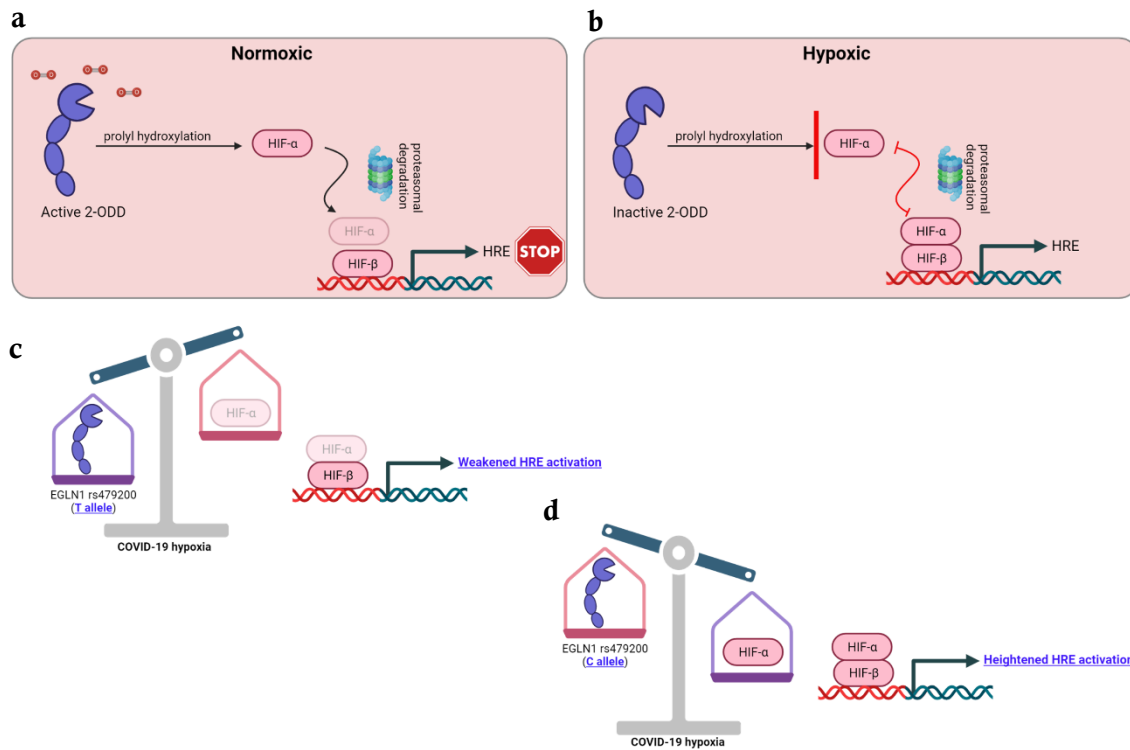
23 The present study investigated the EGLN1 genotypes in asymptomatic, moderate and severe
24 COVID-19 patients at the time of diagnosis. Particularly, SNP rs479200 of the EGLN1 gene
25 that was previously found to be associated with HAPE susceptibility, aerobic capacity and

1 predisposition to pre-eclampsia in multiple studies was explored further; where in SNP
2 rs479200 (C>T) results in three distinct genotypes- homozygous (CC & TT) and heterozygous
3 (TC), with C & T as major and minor alleles, respectively. Statistical analysis SNP rs479200
4 revealed that both the alleles are present in equal proportion in asymptomatic and moderate
5 categories of COVID-19 patients, but overrepresentation of C allele (2-fold higher than T
6 allele) in severe COVID-19 was peculiar. The CC genotype was significantly overrepresented
7 in severe COVID-19 patients (p-value = 0.047). Multinomial logistic regression was performed
8 to evaluate the association of different genotypes with disease severity, which estimated (OR-
9 11.414 (2.564-50.812)- adjusted) and (OR- 6.214 (1.84-20.99)- unadjusted) for genotype CC,
10 in asymptomatic vs severe categories of COVID-19. Moreover, the TC genotype was also
11 found to be strongly associated with severe COVID-19, indicating the role of the C allele as a
12 risk factor.

13 Transcription of the hypoxia response elements (HRE) in normoxic conditions is regulated and
14 mediated by 2ODD- a gene product of EGLN1 gene, however in hypoxic conditions, inactive
15 2ODD cannot inhibit HIF- α and HIF- β dimerization leading to activation of HRE (**Fig. 3a, b**)
16 Considering the genotype data of the EGLN1 gene, TT genotype is correlated with higher
17 expression of EGLN1 gene resulting in overproduction of 2ODD ⁵ and conversely, the CC
18 genotype results in lower expression of EGLN1 gene producing less 2ODD. Accordingly,
19 Aggarwal et al. ⁵ previously showed that individuals residing at low altitudes with TT genotype
20 do not respond well in hypoxic conditions succumbing to HAPE at high altitudes.
21 Overproduction of 2ODD in the presence of T allele (**Fig. 3c**) results in minimal HIF- α levels,
22 rendering the individuals hypo-responsive towards hypoxia. In contrast, the ancestral CC
23 genotype in highlanders imparts a physiological adaptation to respond well in hypoxic
24 conditions. Notably, C allele- the identified risk factor in severe COVID-19, producing inactive
25 2ODD in hypoxic conditions, accumulates high levels of HIF- α making them hyper-responsive
26 towards hypoxia with activation of HRE (**Fig. 3d**). The hyper-responsiveness towards hypoxia
27 might explain the severe COVID-19 manifestations- pulmonary vascular alterations,
28 overproduction of inflammatory cytokines, fluid accumulation in alveoli and the cytokine
29 storm.

30 Thus, the hypothesis above postulates that individuals with severe COVID-19 symptoms and
31 the C allele (risk factor), make them hyper-responsive in hypoxic conditions and progress
32 towards worsening the respiratory hemodynamics and inflammatory pathways. To strengthen
33 the hypothesis of C allele as a risk factor a larger cohort of categorized COVID-19 patients

1 must be investigated in different settings. It is worth noting that deducing the genotype of SNP
 2 rs479200 can be easily accomplished in 4-6 hours' time, equivalent to a COVID-19 RT-PCR
 3 test. Therefore, the presence of C allele as a risk factor in severe COVID-19 might serve as a
 4 feasible risk assessment prognostic marker and prioritization of limited medical infrastructure
 5 in view of COVID-19 waves, as0020witnessed globally002E



6 **Figure 3:** Schematic representation of interplay between HIF- α and 2ODD (product of EGLN1
 7 gene), inhibiting activation of HRE in normoxic (a) and activation of HRE via
 8 dimerization with HIF- β in hypoxic (b) conditions. The postulated effect of T-allele
 9 (c) in COVID-19 hypoxia weakening the HRE activation and effect of C-allele (d)
 10 in the heightened HRE activation. (Created by BioRender trial version)

11 (2ODD- 2-oxoglutarate-dependent dioxygenase; HIF- α - Hypoxia inducible factor- α ; HIF- β -
 12 Hypoxia inducible factor- β ; HRE- Hypoxic response elements.)

13

1 **Acknowledgement**

2 This work was funded by intramural funding support from ICMR-NIMR, Delhi. We thank Drs.
3 Amit Sharma and O. P. Singh for their critical feedback in the study. We thank the NIMR
4 biorepository team members for procuring the COVID-19 infected patient samples during the
5 first wave in India. The publication bears the NIMR publication committee no. RIC-11/2022.
6 We thank Mr. Puneet Saini and Ramakant Pal from ICMR-NIMR, stores for timely
7 procurement of reagents in unprecedented COVID-19 lockdown situation.

8 **Conflict of Interest**

9 Authors declare no conflict of interest.

10 **Author contributions**

11 Conceived and designed the experiments: KV, KCP, SD

12 Data generation and analysis: RH, PKS, DK, MK, DS, CPY, MM, VS, RST, KV, KCP

13 Data interpretation, manuscript writing and review: KV, KCP, PKS, CPY, VS, MM, SD

1 Bibliography

- 2 1. Archer, S. L., Sharp, W. W. & Weir, E. K. Differentiating COVID-19 Pneumonia from
3 Acute Respiratory Distress Syndrome and High Altitude Pulmonary Edema:
4 Therapeutic Implications. *Circulation* **142**, 101–104 (2020).
- 5 2. Gattinoni, L. *et al.* COVID-19 pneumonia: different respiratory treatments for different
6 phenotypes? *Intensive Care Med.* **46**, 1099–1102 (2020).
- 7 3. Lai, C. C. *et al.* Proteomic analysis of up-regulated proteins in human promonocyte
8 cells expressing severe acute respiratory syndrome coronavirus 3C-like protease.
9 *Proteomics* **7**, 1446–1460 (2007).
- 10 4. Velavan, T. P. *et al.* Host genetic factors determining COVID-19 susceptibility and
11 severity. *EBioMedicine* **72**, (2021).
- 12 5. Aggarwal, S. *et al.* EGLN1 involvement in high-altitude adaptation revealed through
13 genetic analysis of extreme constitution types defined in Ayurveda.
14 doi:10.1073/pnas.1006108107.
- 15 6. Semenza, G. L. Regulation of oxygen homeostasis by hypoxia-Inducible factor 1.
16 *Physiology* vol. 24 97–106 (2009).
- 17 7. Jahani, M., Dokaneheifard, S. & Mansouri, K. Hypoxia: A key feature of COVID-19
18 launching activation of HIF-1 and cytokine storm. *J. Inflamm. (United Kingdom)* **17**,
19 1–10 (2020).
- 20 8. Lü, P., Fan, J. & Xi, H. J. Association of polymorphisms of rs479200 and rs480902 in
21 EGLN1 gene with hypoxia adaptation in high altitude in Tibetans. *Acta Anat. Sin.* **44**,
22 419–422 (2013).
- 23 9. Yasukochi, Y., Nishimura, T., Motoi, M. & Watanuki, S. Association of EGLN1
24 genetic polymorphisms with SpO₂ responses to acute hypobaric hypoxia in a Japanese
25 cohort. *J. Physiol. Anthropol.* **37**, 1–12 (2018).
- 26 10. Brutsaert, T. D. *et al.* Association of EGLN1 gene with high aerobic capacity of
27 Peruvian Quechua at high altitude. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 24006–24011
28 (2019).
- 29 11. Kumar K. S., P., Arcot, M., Munisamaiah, M. & Balakrishna, S. Novel association of
30 SNP rs479200 in EGLN1 gene with predisposition to preeclampsia. *Gene* **705**, 1–4
31 (2019).
- 32 12. Tweedie, S. *et al.* Genenames.org: The HGNC and VGNC resources in 2021. *Nucleic
33 Acids Res.* **49**, D939–D946 (2021).
- 34 13. IGVDB.
35 http://www.igvdb.res.in/searchcard.php?page_no=6&choice=200&sort=GENESYMBOL
36 OL.
- 37 14. Horiuchi, T. *et al.* Gene-Gene Interaction between EPAS1 and EGLN1 in Patients with
38 High-Altitude Pulmonary Edema. in *European Respiratory Journal* vol. 52 PA3060
39 (European Respiratory Society (ERS), 2018).
- 40 15. Mishra, A., Mohammad, G., Thinlas, T. & Pasha, M. A. Q. EGLN1 variants influence
41 expression and Sao2 levels to associate with high-altitude pulmonary oedema and

- 1 adaptation. *Clin. Sci.* **124**, 479–489 (2013).
- 2 16. Wu, A. *et al.* Correlation between single nucleotide polymorphisms in hypoxia-related
3 genes and susceptibility to acute high-altitude pulmonary edema. *Genet. Mol. Res.* **14**,
4 11562–11572 (2015).
- 5 17. Gattinoni, L. *et al.* Intensive Care Medicine COVID-19 pneumonia: different
6 respiratory treatment for different phenotypes? doi:10.1007/s00134-020-06033-2.
- 7 18. Bashar, A., Khademul Islam, M. M. & Khan, A.-A.-K. Lung transcriptome of a
8 COVID-19 patient and systems biology predictions suggest impaired surfactant
9 production which may be druggable by surfactant therapy. (2020) doi:10.1038/s41598-
10 020-76404-8.
- 11 19. Ivan, M. & Kaelin, W. G. The EGLN-HIF O₂-Sensing System: Multiple Inputs and
12 Feedbacks. *Mol. Cell* **66**, 772–779 (2017).
- 13 20. Wing, P. A. C. *et al.* Hypoxic and pharmacological activation of HIF inhibits SARS-
14 CoV-2 infection of lung epithelial cells. *Cell Rep.* **35**, (2021).
- 15 21. Luks, A. M. & Swenson, E. R. COVID-19 lung injury and high altitude pulmonary
16 edema: A false equation with dangerous implications. *Ann. Am. Thorac. Soc.* **17**, 918–
17 921 (2020).
- 18 22. Poloznikov, A. A. *et al.* HIF Prolyl Hydroxylase Inhibitors for COVID-19 Treatment:
19 Pros and Cons. *Front. Pharmacol.* **11**, 2522 (2021).
- 20