

1 *Original article*

2

3 **Epigenetic Evolution of ACE2 and IL-6 Genes as Non-Canonical**
4 **Interferon-Stimulated Genes Correlate to COVID-19 Susceptibility**
5 **in Vertebrates**

6 **Eric R. Sang^{a1}, Yun Tian^{a1}, Laura C. Miller^b, Yongming Sang^{a*}**

7

8 ^aDepartment of Agricultural and Environmental Sciences, College of Agriculture,
9 Tennessee State University, 3500 John A. Merritt Boulevard, Nashville, TN, USA;

10 ^bVirus and Prion Diseases of Livestock Research Unit, National Animal Disease Center,
11 USDA-ARS, Ames, IA, USA

12

13 ¹ These authors contribute equivalently

14 * Correspondence: ysang@tnstate.edu; Tel.: 615-963-5183

15

16

17

18

19 Received: date; Accepted: date; Published: date

20

21

22 **Abstract:** Current novel coronavirus disease (COVID-19) has spread globally within a
23 matter of months. The virus establishes a success in balancing its deadliness and
24 contagiousness, and causes substantial differences in susceptibility and disease
25 progression in people of different ages, genders and pre-existing comorbidities. Since these
26 host factors are subjected to epigenetic regulation, relevant analyses on some key genes
27 underlying COVID-19 pathogenesis were performed to longitudinally decipher their
28 epigenetic correlation to COVID-19 susceptibility. The genes of host angiotensin-
29 converting enzyme 2 (ACE2, as the major virus receptor) and interleukin (IL)-6 (a key
30 immune-pathological factor triggering cytokine storm) were shown to evince active
31 epigenetic evolution via histone modification and cis/trans-factors interaction across
32 different vertebrate species. Extensive analyses revealed that ACE2 and IL-6 genes are
33 among a subset of non-canonical interferon-stimulated genes (non-ISGs), which have been
34 designated recently for their unconventional responses to interferons (IFNs) and
35 inflammatory stimuli through an epigenetic cascade. Furthermore, significantly higher
36 positive histone modification markers and position weight matrix (PWM) scores of key
37 *cis*-elements corresponding to inflammatory and IFN signaling, were discovered in both
38 ACE2 and IL6 gene promoters across representative COVID-19-susceptible species
39 compared to unsusceptible ones. Findings characterize ACE2 and IL-6 genes as non-
40 ISGs that respond differently to inflammatory and IFN signaling from the canonical ISGs
41 and their epigenetic properties may serve as biomarkers to longitudinally predict
42 COVID-19 susceptibility in vertebrates and partially explain COVID-19 inequality in
43 people of different subgroups.

44

45 **Keywords:** COVID-19, Angiotensin Converting Enzyme 2, Interferons, IL-6, Epigenetic
46 regulation

47 1. Introduction

48 First identified in Wuhan, China, last December, the novel coronavirus disease 2019
49 (COVID-19) has spread worldwide and caused over 0.68/17 million confirmed deaths and
50 infected cases across 200 countries by the end of July 2020 [1,2]. COVID-19 stands out as a
51 new zoonotic disease caused by Severe Acute Respiratory Syndrome coronavirus 2 (SARS-
52 CoV2) [3], which, in the view of a virus, obtains an effective balance between its deadliness
53 and contagiousness in humans [4,5]. In line with that, patients with the ages over 45,
54 especially 75 years old had a worse prognosis and 5-10 fold higher mortal rate than
55 younger ones at 0-17 years old, who mostly showed a mild disease or even asymptomatic
56 [6-15]. Similarly, higher mortality rates were observed in males than females, and
57 particularly in the patients who have pre-existing medical conditions (comorbidities)
58 regardless of gender or age [6-15]. These underlying comorbidities include diabetes,
59 cancer, immunodeficiency, hypertension and cardiovascular disease, asthma and lung
60 disease, kidney disease, as well as chronic GI/liver disorders. In addition to predictable
61 symptoms of cough, fever and headache from the lung infection, the virus can spread to
62 almost every organ including the brain, heart, gut, kidneys and skin to cause organ-
63 specific problems [6-15]. Therefore, from the view of the host, SARS-CoV2 susceptibility
64 and disease progression of COVID-19 is a phenomenon of epigenetic regulation, which
65 underlies the diversity of the disease progression throughout the body system and across
66 different patients that share a near identical genetic background [16-19].

67 Zoonosis and reverse zoonosis infer a dynamic exchange of pathogens between humans
68 and animals, particularly domestic and wild vertebrates. This constitutes a major challenge
69 for both public health and animal health, and unites them into ONE ecological health.
70 Therefore, the potential infection of SARS-CoV2 in both wild and domestic animals raises
71 a big public health concern after the COVID-19 prevalence in human society [20,21]. This
72 concern emphasizes: (1) the identification of reservoir animal species that originally
73 passing SARS-CoV2 to humans; and (2) potential risks of infected people passing the virus
74 to animals, particularly domestic species, to form an amplifying zoonotic cycle and
75 exacerbate SARS-CoV2 evolution and cross-species transmission [20,21]. Recent studies
76 provided evidence that domestic cats and dogs could be virally or serologically positive
77 for SARS-CoV2 [22-28], as were several Bronx zoo tigers [29]. Several studies, using
78 experimental inoculations of human SARS-CoV2 isolates, demonstrated that ferrets,
79 hamsters, domestic cats and some non-human primate species were susceptible to human
80 SARS-CoV2 strains; however, pigs, alpacas, and (putatively) cattle are not [22-29].
81 Previously, we and several others have proposed structural simulation models of ACE2
82 and the viral S-Receptor binding domain (S-RBD) to predict SARS-CoV2 susceptibility
83 across representative vertebrates, especially major domestic and wild mammalian species
84 [30-33]. The structural affinity between ACE2 and S-RBD plays a primary role in the viral
85 attachment and accessibility in cells, and the specific early cellular responses that regulate
86 ACE2 expression and signal early immune responses determine the host susceptibility to
87 the virus [34-40]. We propose an integrative model, which incorporates both ACE2-RBD
88 structural affinity (primarily determined by cross-species genetic difference) and
89 epigenetic regulation of key genes during the early phase of the virus-host interaction, to
90 predict host COVID-19 susceptibility and disease progression [30-33].

91 Among the core host factors that determine COVID-19 susceptibility and early disease
92 progression, angiotensin-converting enzyme 2 (ACE2) and interleukin (IL)-6 were focused

93 upon because of their critical roles directly involved in viral infection and host
94 immunopathies [41-45]. In SARS-CoV2 pathogenesis, ACE2 serves as primary receptors
95 for cell attachment and entry [42,43]. Several groups have reported that SARS-CoV2 exerts
96 higher receptor affinity to human ACE2 than other coronaviruses, which may contribute
97 to the high-contagiousness and rapid spread of SARS-CoV2 in humans [42,43]. Being a key
98 enzyme in the body's renin-angiotensin-aldosterone system (RAAS), ACE2 catalyzes
99 angiotensinogen (AGT) to produce the active forms of hormonal angiotensin (Ang) 1-9,
100 which directly regulate the blood volume/pressure, body fluid balance, sodium and water
101 retention, as well as co-opt multiple effects on inflammation, apoptosis, and generation of
102 reactive oxygen species (ROS) [43-45]. In this regard, not only do the virus direct binding
103 and functional impairment of ACE2 enzymatic function serve as a physio-pathological
104 mechanism underlying COVID-19 disease complex, but also epigenetic regulation of
105 ACE2 expression in various tissues/conditions, especially of that related to blood clotting,
106 aneurism and chilblains in infant patients [43-46].

107 SARS-CoV2 seizes ACE2 for cell entry, which is followed by a cytokine-related syndrome,
108 namely acute respiratory distress syndrome (ARDS). Plausibly, the occupancy of the ACE2
109 catalytic domain by the viral Spike protein (S) blocks AGT activation into Ang1-9 and
110 leads to the accumulation of Ang2 in the serum [43-46]. Circulatory increase of Ang2
111 induces inflammatory cytokines, including TNF- α , IL-6, and soluble IL-6 receptor α (sIL-
112 6R α) in pneumocytes and macrophages, through binding Ang1-receptor (AT1R) and
113 activating disintegrin- and metalloprotease 17 (ADAM17)-mediated cascade [41-46]. This
114 process is followed by activation of the IL-6 amplifier (IL-6-AMP), which co-activates NF-
115 κ B and transcription factor STAT3 to enhance inflammatory response and leads to ARDS
116 underlying COVID-19. Ang2-AT1R activation also induces pyroptosis, a highly
117 inflammatory form of programmed cell death accompanying cytotoxicity caused by viral
118 infections [41,45,46]. Aggregately, SARS-CoV2 itself also activates NF- κ B via various
119 pattern recognition receptors (PPRs) [33-40]. Therefore, IL-6 and IL-6 AMP are biomarkers
120 of hyperactivation of inflammatory machinery exacerbated by ACE2 blocking and viral
121 infection, which represent key cytokines in deciphering cytokine-related syndrome and
122 disease progression of COVID-19 [41,45,46].

123 The expression of ACE2 is inter-regulated by multiple physio-pathological factors,
124 including intracellular pathogenic infection, pre-existing inflammatory condition from
125 comorbidities, and inflammatory cytokines including TNF and IFNs [41-46]. Several recent
126 studies demonstrated that human ACE2 gene behaved like an interferon-stimulated gene
127 (ISG) and was stimulated by viral infection and IFN treatment; however, mouse *Ace2* gene
128 was not [47-49]. Canonical ISGs describe over a thousand cellular genes that are induced
129 by IFN simulation via the IFN-JAK-STAT signaling axis [50]. These canonical ISGs are
130 mainly induced by type I and type III IFNs but overlap with those upregulated by type II
131 IFN (i.e. IFN- γ) [47-50]. These ISGs comprise a frontline of antiviral immunity to restrict
132 virus spreading from the initial infection sites [50]. However, based on gene evolution and
133 epigenetic analyses, ACE2 may not be a member of these classical antiviral ISGs, and more
134 likely belong to the non-canonical ISGs (non-ISGs) like IL-6 (a.k.a. IFN- β 2 in humans) [47-
135 51]. These non-ISGs are primed under a pre-inflammatory condition and stimulated by
136 IFN or IFN plus TNF through an epigenetic cascade involving positive histone
137 modification (mainly H3K4me3 and H3K27ac) to increase chromatin accessibility for
138 binding by transcription factors including PU.1, IRFs, and NF- κ B and culminating in non-

139 ISGs expression (Figure 1) [51-54]. To confirm that, we conducted cross-species
140 comparative analysis between IL-6 and ACE2 genes. First, annotation of ENCODE
141 epigenetic datasets discovered similarity of H3K4me3 and H3K27ac markers between IL-6
142 and ACE2 gene promoters in both humans and mice; however, significantly higher Z-
143 scores and enrichment of H3K4me3 and H3K27ac in human IL-6 and ACE2 genes were
144 detected than in their mouse orthologs, respectively [55]. Second, detection of cis-
145 regulatory elements (CREs) that bind core transcription factors of non-ISGs, including
146 PU.1, IRFs, and NF- κ B, in ACE2 and IL-6 gene proximal promoter regions across 25
147 representative animal species [56]. Third, we found that the evolutionary increase of
148 ACE2, and especially the IL-6 genes response to inflammatory and IFN signaling may
149 serve as epigenetic marker for COVID-19 susceptibility in some animal species including
150 humans. Finally, using our non-biased RNA-Seq data, we further categorize some more
151 non-ISGs that resemble the expression pattern of either IL-6 or ACE2 [57]. Notably, we
152 detected two ACE2 isoforms, which differ in both proximal promoters and coding regions,
153 in some livestock species including pigs, dogs and cattle [30]. In pigs, the ACE2 short
154 isoform (ACE2S) has an expression pattern similar to IL-6 than the long isoform (ACE2L).
155 Collectively, our findings characterize ACE2 and IL-6 genes as non-ISGs responding
156 differently to inflammatory and IFN signaling, and their epigenetic properties may serve
157 as biomarkers to predict COVID-19 susceptibility in vertebrates longitudinally and
158 partially explain COVID-19 inequality in people of different subgroups [20,30-33].

159 **2. Results and Discussion**

160 *2.1. Epigenetic processes in induction of non-canonical IFN-stimulated genes (non-ISGs):*

161 Studied mostly in humans and mice, the hundreds of classical ISGs, such as ISG15 and
162 IRF1, contain the main IFN-responsive CREs, including IFN-stimulated regulatory
163 element (ISRE) and γ -activated sequence (GAS), in their promoter regions [47,50]. The
164 tripartite IFN-stimulated gene factor 3 (ISGF3), which is composed of three transcription
165 factors including STAT1, STAT2 and IRF9, is activated downstream of the IFN-JAK-STAT
166 signaling axis to bind ISREs and stimulate canonical ISG expression [47,50]. In addition to
167 this classical axis to induce ISGs, IFNs also co-opt multiple non-canonical signaling
168 pathways to activate these ISGs or other corresponding genes together through various
169 alternative mechanisms [51-54]. These non-canonical IFN signaling pathways involve
170 extensive crosstalk between the signaling pathways mediated by various cellular pathogen
171 pattern-recognition receptors (PRRs) and inflammatory cytokines, notably IL-1, IL-6 and
172 TNF [51-54]. The non-canonical signaling pathways not only diversify mechanisms for
173 inducing ISGs, but also extend the spectrum of IFN-responsive genes, indicating a
174 multifunctional property of IFNs in antiviral and immuno-physiological regulation [50-
175 54]. Recent studies showed that human IL-6 and ACE2 are two candidates for these non-
176 ISGs [47-51]. Figure 1 shows current understanding of the gene activation cascade of
177 human IL-6 (and plausibly ACE2) genes as an example of non-ISGs, whose IFN-inductive
178 property and systemic role recently recognized as underlying multiple inflammatory
179 comorbidities [51,54]. In brief, stimulation of epithelial cells and tissue macrophages by
180 early pro-inflammatory signaling of TNF induces transient expression of TNF-target genes
181 encoding inflammatory mediators, such as IL6 and TNF. This is followed by a transient
182 state that is insensitive to further inflammatory signaling from TLR activation, and thus
183 relevant chromatin containing non-ISGs is not activated (depicted by a grey shade in Fig.

184 1). This transient suppression state, however, can be activated by a co-stimulation with
185 TNF plus IFN- α resulting in increase of positive histone markers (H3K4me3 and H3K27ac)
186 and chromatin accessibility of the gene promoter regions, which sequentially recruit the
187 binding of corresponding transcription factors including IRFs and NF- κ B to activate non-
188 ISG expression [51,54]. Besides IL-6, many tunable ISGs including human ACE2 as
189 demonstrated in recent studies show sustainable response to IFN and pathogenic
190 inflammatory signaling, and share expression patterns involving epigenetic sensation and
191 synergistic IFN-induction as depicted for non-ISGs (Figure 1) [47-54]. However, the cross-
192 species evolutionary characterization of non-ISGs has not been studied. Using IL-6 and
193 ACE2 as examples, extensive epigenetic and expression analyses were performed in this
194 study to determine their epigenetic evolution and potential role as biomarkers to predict
195 the susceptibility and disease progression of COVID-19.

196 2.2. Determine species-specific positive histone markers in human and mouse IL-6 and ACE2 197 gene promoters

198 Epigenetic positive histone modification in a certain chromatin region, mainly including
199 histone H3 with tri-methylation at the 4th lysine residue (H3K4me3) or with the acetylation
200 at the 27th lysine residue (H3K27ac) here, is associated with a higher activation status of
201 adjacent gene transcription, thus defined as active enhancer markers in epigenetic
202 analyses. The enrichment of H3K4me3 and H3K27ac defines one epigenetic feature of non-
203 ISGs post activation [51-55]. Through annotation of Chip-Seq and ATAC-Seq datasets from
204 839 and 157 cell/tissue types of humans and mice through ENCODE
205 (<https://www.encodeproject.org/>) [55], we detected significant and comparative existence
206 of H3K4me3 and H3K27ac markers between IL-6 and ACE2 gene promoters in various
207 humans and mouse samples (Figure 2). However, higher Z-scores and enrichment of
208 H3K4me3 and H3K27ac were found in human IL-6 and ACE2 genes (Figure 2A and 2B)
209 than their mouse orthologs (Figure 2C and 2D). In both distal and proximal regions of the
210 ACE2 gene promoters, the human gene (Figure 2A) was marked by 2-3 fold more of these
211 positive histone medication than the mouse ortholog, indicating higher activation and
212 transcription activity of human IL-6 and ACE2 genes under similar conditions. Because
213 these findings are extracted from the extensive datasets representing systemic sample
214 types, it is convincing that typical epigenetic positive histone modification, H3K4me3 and
215 H3K27ac, is significantly associated with the promoter regions of ACE2 as with IL-6 genes.
216 Specifically, IL-6 genes were shown to have more histone modifications around their
217 proximal promoter regions than ACE2 genes, which had more in distal regions. There
218 were higher Z-score and enrichment of these positive histone markers in the human genes
219 than their mouse orthologs, indicating evolutionary and probably species-specific manner
220 of epigenetic regulation of these non-ISGs [51,54]. This epigenetic difference of key non-
221 ISGs might contribute to disease susceptibility and progression when animals of different
222 species are exposed to same pathogenic pressure.

223 2.3. Cross-species comparison of key cis-regulatory elements (CREs) that mark non-ISG 224 regulation in IL-6 and ACE2 genes

225 After determination of positive histone markers along the IL-6 and ACE2 gene bodies, we
226 examined the existence of cis-regulatory elements (CREs) that interact with typical non-
227 ISGs transcription factors including PU.1 (a.k.a. SPI1), IRFs, and NF- κ B1/2 in the promoter

228 regions of IL-6 and ACE2 gene orthologs [50-56]. We extracted the primal promoter
229 sequences from IL-6 and ACE2 genes from 25 representative vertebrate species, which
230 contain ten previously validated SARS-CoV2-susceptible species and other naturally
231 unsusceptible species based on collected evidence [26-30]. As shown in Figure 3, all three
232 types of CREs (i.e. PU.1, IRFs, and NF- κ B) that mark non-ISG expression were mapped for
233 cross-species existence in the promoter regions of both IL-6 and ACE2 genes. Significant
234 PWM scores ($p < 0.0001$) were determined for their CREs when each was compared with
235 the corresponding human CRE matrix (Figure 3A-3C) [56]. ACE2 genes had a generally
236 lower PWM scores for these CREs than those for IL-6 genes, in particular the PWM scores
237 for NF- κ B2 CRE in ACE2 genes were at 2-8 Log₂units lower (Figure 3D). This indicates
238 that ACE2 genes were less responsive to non-canonical NF- κ B signaling mediated by NF-
239 κ B2 [58,59]. Because dysregulation of non-canonical NF- κ B signaling has been well known
240 for contribution to various autoimmune and inflammatory diseases, the differential role of
241 ACE2 and IL-6 in inflammatory immunopathies are worth further investigation [58,59].
242 Notably, only CRE matrices to IRF1 were shown in Figure 3C, both ACE2 and IL-6 gene
243 promoters actually contain CREs binding IRF2-8 with high PWM scores, except of CREs
244 interacting with IRF5 and IRF9 had low PWM scores in most tested species (Figure 4 and
245 Figure 5). Because IRF9 is a key component of ISGF3 and binding to ISREs to activate
246 canonical ISG expression, this discovery evidently differs ACE2 and IL-6 genes from the
247 classical ISGs such as ISG15 and IRF1 (Figure 4) [50,51]. However, IL-6 genes of eight
248 species maintain their IRF9 binding CREs as for examples in Zebrafish and frogs, only rat
249 ACE2 gene showed a high PWM score for containing an IRF9 binding CRE (Figure 4). This
250 further postulates a species-dependent trend of non-ISG evolution, and warrants further
251 investigation in contributing to host-pathogen interaction.

252 Figure 5 gathers cross-species analyses of mean PWM scores of the CREs, which bind
253 STAT1/2, PU.1 (a.k.a. SPI1), NF- κ B1, NF- κ B2, and multiple IRFs (including, IRF1-4, IRF7,
254 and IRF8) that show significant PWM scores with $p < 0.0001$ under the algorithm's default)
255 in the proximal promoter regions of IL-6 and ACE2 gene orthologs from the 25
256 representative vertebrate species. As shown, these bookmarking CREs for non-ISGs had
257 comparable Log₂(mPWM) scores between ACE2 and IL-6 genes across different species
258 and also showed species-specific variation to some extent. IL-6 genes generally had a
259 higher mPWM scores for more of the tested animal species with CREs that bind STAT1/2
260 and IRFs downstream of IFN signaling (Figure 5A and 5E) [50, 52]. Of significant
261 difference between ACE2 and IL-6 genes was their CREs' PWM scores pertinent to NF- κ B1
262 and NF- κ B2 (Figure 5C and 5D). Whereas ACE2 genes evolved to be slightly more
263 responsive to the canonical NF- κ B1 signaling in most mammalian species (Figure 5C), IL-6
264 genes obtained much higher responsiveness to non-canonical NF- κ B2 signaling (Figure
265 5D). Recent studies showed that defects in non-canonical NF- κ B2 signaling are associated
266 with severe immune deficiencies, and dysregulation of this pathway contributes to the
267 pathogenesis of various autoimmune and inflammatory diseases [58,59]. The epigenetic
268 difference of IL-6 and ACE2 genes downstream of canonical NF- κ B1 and non-canonical
269 NF- κ B2 signaling thus may serve as differential gene markers for inflammatory-related
270 syndromes [58,59].

271 *2.4. Epigenetic evolution of higher PWM scores of non-ISG's core CREs in ACE2 and*
272 *especially IL-6 gene promoters in COVID-19 susceptible species*

273 As previously described, in addition to its core role in physiological regulation of blood
274 volume/pressure and body fluid balance, the RAAS also critically affects inflammation,
275 apoptosis, and other immune reactions. For instance, suppression of ACE2 increases Ang2
276 production to signal pro-inflammatory and apoptotic responses in affected tissues [44-46].
277 When exacerbated by infection of an intracellular pathogen, such as SARS-CoV2 in
278 COVID-19 cases, a high inflammatory form of programmed cell death, known as pyroptosis,
279 is induced accompanying massive production of pro-inflammatory cytokines including IL-
280 1, IL-6, TNF and CXCL10 [41,45,46]. Because the potential clinical relevance to these CREs
281 in COVID-19, we performed a comparative study to determine if the COVID-19
282 susceptible animal species obtain some epigenetic features in these core CREs in regulation
283 of IL-6 and ACE2 expression. Figure 6 compares the mPWM scores of these core non-ISG
284 CREs between two groups: known SARS-CoV2 susceptible species [CoV2(+)] and
285 unsusceptible species [CoV2(-)]. Figure 6 shows that ACE2 and IL-6 genes from CoV2(+)
286 species contain CREs that have significantly higher mPWM scores. This indicates that in
287 some vertebrate species, non-ISGs like ACE2 and especially IL-6 genes evolve to obtain
288 high inductive propensity by inflammatory and IFN signaling [47-54]. Therefore, in
289 addition to the ACE2 structure and affinity to S-RBD, the epigenetic evolution for IL-6 and
290 ACE2 stimulation (reflected by higher mPWM scores), may serve as epigenetic biomarkers
291 (or triggers) for susceptibility prediction of COVID19 and other ARDS longitudinally
292 across vertebrates and horizontally in the subgroups of humans [30,47-54].

293 2.5. Overall comparison of phylogenetic topologies between IL-6 and ACE2 gene promoter 294 sequences

295 In addition to focusing on epigenetic analysis of these non-ISG CREs, we also conducted
296 cross-species comparison of phylogenetic topology between the full proximal promoter
297 sequences of IL-6 and ACE2 genes. Overall, the topology of the phylogenies of IL-6 and
298 ACE2 gene promoters are similar with a comparative topological score of 86.5 % (Figure 7).
299 Sharing a root of low vertebrates (*D. rerio* and/or *X. tropicalis*), the CoV2(+) species were
300 distributed within the clades containing primates, carnivores and glires. In contrast, all the
301 ruminant promoters were clustered into a most phylogenically distant clade and associated
302 with no CoV2(+) species (Figure 7). Comparison of the two phylogenies in detail showed
303 that the major difference came from the location of the chicken, rabbit, guinea pig, and pig.
304 In the IL-6 promoter phylogeny (Figure 7, left panel), chicken IL-6 promoter seems to derive
305 rodent IL-6 gene promoters after evolution from the fish and frog; in the ACE2 promoter
306 phylogeny (Figure 7, right panel), however, the chicken ACE2 promoter serves as a root leaf
307 with the zebrafish. The largest difference is between phylogenetic positions of IL-6 and ACE2
308 gene promoters for pigs and guinea pigs. Whereas in the IL-6 promoter phylogeny, the pig
309 sisters to those of the alpaca and horse, within the Carnivore clade that contains most of the
310 validated CoV2(+) species in addition to the primate clade, porcine ACE2 gene promoter
311 was next to the ruminant clade that has no CoV2(+) species identified so far [20-29]. Guinea
312 pig as a rodent species has its IL-6 promoter surprisingly within the primate clade, but its
313 ACE2 promoter appears more primitive and shares the clade with the frog. Given the
314 primate and carnivore clades contain most identified CoV2(+) species, if pig and guinea pig
315 are proved to be CoV2(+) species, the IL-6 promoter phylogeny may better correlate to
316 CoV2(+) prediction; otherwise, the ACE2 promoter phylogeny correlates better. The rabbit

317 and otter, which occupy similar positions in both IL-6 and ACE2 promoter phylogenies,
318 may have a high potential to be CoV2(+) and COVID-19 susceptible based on this and
319 previous studies, which used epigenetic and structural models, respectively [30-33]. In this
320 regard, pigs and guinea pigs may serve as symbol species to estimate the epigenetic role of
321 non-ISGs in CoV2(+) prediction. No study has tested CoV2/COVID-19 susceptibility in
322 guinea pigs, but studies in pigs concluded the species was unsusceptible [20]. This may
323 indicate that the overall epigenetic feature of ACE2 genes better relates to CoV2(+) status in
324 some mammalian species. However, the study of key CRE scores of non-ISGs in Figure 6
325 indicates that IL-6 gene CRE scores have a higher correlation when compared between the
326 CoV2(+) and CoV2() species. This may reflect an etiological fact that CoV2(+) is necessary
327 but not sufficient for COVID-19 progression; and the latter is indeed dependent on the host
328 immune reaction, particularly the early ISGs and non-ISG responses studied here [51,54]. In
329 that regard, epigenetic evolution/regulation of ACE2 and IL-6 genes may signify two layers
330 of COVID-19 progression, i.e. ACE2 is better for CoV2(+) and IL-6 is better for downstream
331 COVID-19 symptoms [51,54,58,59].

332

333 *2.6. Non-bias transcriptome-based categorization of non-ISGs that resemble to the inductive* 334 *pattern to IL-6 or ACE2 genes*

335 Compared with canonical ISGs, studies of epigenetic regulation and expression of non-
336 ISGs have just started accompanying our understanding of their role in some autoimmune
337 and inflammatory diseases in recent years [50-54]. Although some non-canonical signaling
338 pathways, that are independent of the canonical IFN-JAK-STAT3 axis, play a role in ISG
339 induction, the classification criteria of non-ISGs is not established [50-54]. Using IL-6 and
340 ACE2 genes as examples of non-ISGs, the disparity of their cross-response to
341 inflammatory and IFN signaling could be one way to classify them as IL-6-like or ACE2-
342 like groups. We therefore analyzed a non-biased transcriptome (RNA-Seq) dataset from
343 porcine alveolar macrophages treated with different stimuli and infected with a porcine
344 arterivirus, a respiratory virus belonging to Nidovirales with coronaviruses [57]. We chose
345 to use porcine transcriptome data because of the species-focus of our projects and the
346 anatomy and physiological resemblance between pigs and humans [57]. Figure 8 presents
347 the IL-6-like and ACE2-like groups, which were categorized based on their responsive
348 patterns to LPS and two types of IFNs (i.e. IFN- α or type I and type II IFN- γ) at the early
349 phase of 5 h post the treatment/infection [57]. These clustered IFN responsive genes were
350 mainly from the RAAS, TNF, IL-6, chemokine superfamilies. For IL-6 non-ISG group, all of
351 these genes showed robust stimulation by LPS as well as a weaker response to both IFNs
352 (Figure 8A). In contrast, the ACE2-group genes were insensitive to LPS, but were
353 upregulated significantly by both types of IFNs (Figure 8B). Compared with the canonical
354 group of ISGs (Figure 8C), which shows the highest response to the type I IFN- α , IL-6
355 group had a least increase upon IFN- α and a similar stimulation by IFN- γ as for ISGs; and
356 ACE2 group showed a mid-response to IFN- α but highest to IFN- γ (Figure 8A-8D). Figure
357 8D statistically demonstrates the stimulatory difference among three groups of IFN-
358 responsive genes: (1) for ISGs: IFN- α > IFN- γ > LPS with a higher background expression in
359 PBS, IL-4 and IL-10 treatments; (2) for IL-6-like non-ISGs: LPS > IFN- γ > IFN- α with the
360 lowest background expression; and (3) for ACE2-like non-ISGs: IFN- γ > IFN- α > LPS with a
361 mid-background expression. Therefore, our classification of ISGs and non-ISGs represents

362 a complete scenario of gene response levels (i.e. at low, mid and high levels of responses to
363 LPS and two types of IFNs) to complement each other per their responsive propensity to
364 LPS, IFN- γ , and IFN- α . As previously described, most ISGs especially non-ISGs are inter-
365 regulated through multiple canonical and non-canonical signaling pathways. The cross-
366 talking of signaling pathways mediated by different types of IFNs and inflammatory
367 cytokines is dynamic to form into an intricate regulatory network underlying animal
368 immunity to determine disease pathogenesis in various situations [50-54]. So with the
369 functional extension of physiological genes, such as AGT and ACE2, the new discovery of
370 species-dependent response to viral infections and IFN stimulation, posits them as
371 immunogenetic factors critical to determining COVID-19 disease progression in addition
372 to its role as a major virus receptor [44-49]. Notably, several ACE2 isoforms have been
373 identified in humans and several major livestock species [30]. Our transcriptome analysis
374 also picked up one short porcine ACE2 isoform (ACE2S), its expression pattern actually is
375 more like IL-6 non-ISGs than the consensus ACE2 longer isoform (ACE2L) [30]. In
376 addition to ACE2, the AGT gene of the RAAS also showed a non-ISG property similar to
377 ACE2 (Figure 8A and 8B). Collectively, transcriptomic annotation afforded us to cluster
378 tentative non-ISGs that share expression patterns similar to IL-6 or ACE2 genes.
379 Interestingly, most of them belong to IL-6, TNF and chemokine superfamilies, whose roles
380 in regulation of autoimmune and inflammatory diseases, as well as in COVID-19
381 progression warrant further investigation.

382 3. Conclusions

383 Figure 9 depicts the working summary of this study for epigenetic evolution and
384 regulation of IL-6 and ACE2 as non-ISGs, indicating their potentials as biomarkers for
385 inflammatory syndrome underlying pathogenic viral infection such as of COVID-19. Non-
386 ISGs such as those categorized by resemblance to IL-6 and ACE2 genes were sequentially
387 regulated by TNF, IFN and TLR signaling, which modify chromatin accessibility through
388 activating histone modification and recruitment of transcription factors including PU.1,
389 IRF and NF- κ B binding on the promoter regions of these non-ISGs. In turn, it will amplify
390 the inflammatory loop through IL-6-mediated response and inducing more ACE2
391 expression, which collectively contributes to the occurrence of respiratory and
392 inflammatory syndromes as in COVID-19. Therefore, high expression of non-ISGs such as
393 IL-6 and ACE2 could be biomarkers for the exacerbation of inflammation underlying some
394 viral infections especially those like SARS-CoV2, which dysregulates the physiological
395 function of ACE2 in the RAAS-centric body systems. In addition, the cross-species
396 epigenetic evolution of these key physio-pathological genes may provide a key to decipher
397 molecular mechanisms underlying species-specific susceptibility to COVID-19 from the
398 host side.

399 4. Materials and Methods

400 *Annotation of ENCODE epigenetic datasets:* The profile of epigenetic markers relevant to
401 histone positive modification, mainly H3K4me3 and H3K27ac, were searched using the
402 gene symbols through the ENCODE public domain at <https://www.encodeproject.org/>
403 under the default condition [55]. The ENCODE datasets for generating the epigenetic
404 results include those mainly based on Chip-Seq and ATAC-Seq from 839 and 157
405 cell/tissue types of humans and mice, respectively. The Max Z-Scores and locations of the

406 histone markers on the gene promoter regions were then curated under a permission for
407 academic users, and manually diagrammed.

408 *Promoter sequence extraction and alignment:* The DNA sequences of the proximal promoters
409 of analyzed genes were extracted from NCBI Gene and relevant databases
410 (<https://www.ncbi.nlm.nih.gov/gene>). Both IL-6 and ACE2 genes and corresponding
411 transcripts have been well annotated in most representative vertebrate species. In most
412 cases, the annotations were double verified through the same Gene entries at Ensembl
413 (<https://www.ensembl.org>). The protein and DNA sequences were collected from all non-
414 redundant transcript variants and further verified for expression using relevant RNA-Seq
415 data (NCBI GEO profiles). The proximal promoter region spans ~2.5 kb before the
416 predicted transcription (or translation) start site (TSS). The protein and DNA sequences
417 were aligned using the multiple sequence alignment tools of ClustalW or Muscle through
418 an EMBL-EBI port (<https://www.ebi.ac.uk/>). Other sequence management was conducted
419 using programs at the Sequence Manipulation Suite (<http://www.bioinformatics.org>).
420 Sequence alignments were visualized using Jalview (<http://www.jalview.org>) and MEGAx
421 (<https://www.megasoftware.net>). Sequence similarity calculation and plotting were done
422 using SDT1.2 (<http://web.cbio.uct.ac.za/~brejnev>). Other than indicated, all programs were
423 run with default parameters [30].

424 *Examining transcription factor binding sites in the gene promoters and PWM scoring:* We use
425 two programs/databases to confirm each other for the major CRE predictions. The
426 regulatory elements (and corresponding binding factors) in the ~2.5 kb proximal promoter
427 regions were examined against both human/animal TFD Database using a program Nsite
428 (Version 5.2013, at <http://www.softberry.com>). The mean position weight matrix (PWM) of
429 key cis-elements in the proximal promoters were calculated using PWM tools through
430 <https://ccg.epfl.ch/cgi-bin/pwmtools>, and the binding motif matrices of examined TFs were
431 extracted from MEME-derived HOCOMOCOv11 TF collection affiliated with the PWM
432 tools [56]. The species-specific CRE sequences were then extracted from each promoter
433 sequence for alignments in Fig. 3.

434 *Phylogenetic analysis and topological comparison:* Evolutionary analyses were conducted in
435 MEGA X as described. The evolutionary history was inferred by using the Maximum
436 Likelihood method and Tamura-Nei model. Initial tree(s) for the heuristic search were
437 obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of
438 pairwise distances estimated using the Tamura-Nei model, and then selecting the topology
439 with superior log likelihood value. For topological comparison between phylogenetic trees
440 generated using IL-6 and ACE2 gene proximal promoters, the phylogenies of Newick
441 strings were generated using the MEGA program, and topological comparison between
442 the Newick trees was performed with Compare2Trees at
443 (<http://www.mas.ncl.ac.uk/~ntmwn/compare2trees>) to obtain the overall topological
444 scores. Other than indicated, all programs were run with default parameters as the
445 programs suggested.

446 *RNA-Seq and data analysis:* For expression confirmation, several sets of RNA-Seq data from
447 NCBI Gene databases, and one of ours generated from porcine alveolar macrophages
448 (BioProject with an accession number of SRP033717), were analyzed for categorizing ISGs
449 and non-ISGs accordingly to the expression patterns of IL-6 and ACE2 genes. Significantly

450 and differentially expressed genes (DEGs) between two treatments were called using an
451 edgeR package and visualized using bar charts (RPKM) or heatmaps (Log2 fold ratio) as
452 previously described [57].

453

454 **Author Contributions:** E.R.S. and Y.T. contributed to idea conceptualization, data
455 computation and draft preparation. L.C.M. helped in conception and data discussion;
456 L.C.M. helped for funding acquisition. Y.S. supervises overall conceptualization, data
457 collection & process, computation, draft writing, and funding acquisition.

458

459 **Funding:** This work was supported by USDA NIFA Evans-Allen-1013186 and NIFA 2018-
460 67016-28313 to YS, and in part through reagent sharing of NIFA AFRI 2020-67016-31347 and
461 NSF-IOS-1831988 to YS.

462

463 **Conflicts of Interest:** The authors declare no conflict of interest.

464

465

466 References

- 467 1. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at
468 Johns Hopkins University (JHU). [Cited July 31, 2020]. Available from:
469 <https://coronavirus.jhu.edu/map.html>
- 470 2. Chan JF, Yuan S, Kok KH, et al. A familial cluster of pneumonia associated with the
471 2019 novel coronavirus indicating person-to-person transmission: a study of a family
472 cluster. *Lancet*. 2020;395(10223):514-523.
- 473 3. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease
474 in China [published correction appears in *Nature*. 2020 Apr;580(7803):E7]. *Nature*.
475 2020;579(7798):265-269.
- 476 4. Sanche S, Lin YT, Xu C, Romero-Severson E, Hengartner N, Ke R. High Contagiousness
477 and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. *Emerg Infect*
478 *Dis*. 2020;26(7):1470-1477.
- 479 5. Courtney EP, Goldenberg JL, Boyd P. The contagion of mortality: A terror management
480 health model for pandemics. *Br J Soc Psychol*. 2020;59(3):607-617.
- 481 6. Age, Sex, Existing Conditions of COVID-19 Cases and Deaths. [Cited July 31, 2020].
482 Available from: [https://www.worldometers.info/coronavirus/coronavirus-age-sex-](https://www.worldometers.info/coronavirus/coronavirus-age-sex-demographics/)
483 [demographics/](https://www.worldometers.info/coronavirus/coronavirus-age-sex-demographics/)
- 484 7. Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for
485 Disease Control and Prevention. The Epidemiological Characteristics of an Outbreak of
486 2019 Novel Coronavirus Diseases (COVID-19). *Zhonghua Liu Xing Bing Xue Za Zhi*.
487 2020;41(2):145-151.
- 488 8. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19) [Pdf]
489 - World Health Organization, [Cited July 30, 2020]. Available from:
490 [https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-](https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf)
491 [covid-19-final-report.pdf](https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf)
- 492 9. Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. Considering how biological
493 sex impacts immune responses and COVID-19 outcomes. *Nat Rev Immunol*.
494 2020;20(7):442-447. doi:10.1038/s41577-020-0348-8
- 495 10. Gebhard C, Regitz-Zagrosek V, Neuhauser HK, Morgan R, Klein SL. Impact of sex and
496 gender on COVID-19 outcomes in Europe. *Biol Sex Differ*. 2020;11(1):29.
- 497 11. Jutzeler CR, Bourguignon L, Weis CV, et al. Comorbidities, clinical signs and
498 symptoms, laboratory findings, imaging features, treatment strategies, and outcomes
499 in adult and pediatric patients with COVID-19: A systematic review and meta-analysis
500 [published online ahead of print, 2020 Aug 4]. *Travel Med Infect Dis*. 2020;101825.
- 501 12. Hendren NS, Drazner MH, Bozkurt B, Cooper LT Jr. Description and Proposed
502 Management of the Acute COVID-19 Cardiovascular Syndrome. *Circulation*.
503 2020;141(23):1903-1914.
- 504 13. Sironi M, Hasnain SE, Rosenthal B, et al. SARS-CoV-2 and COVID-19: A genetic,
505 epidemiological, and evolutionary perspective [published online ahead of print, 2020
506 May 29]. *Infect Genet Evol*. 2020;84:104384.
- 507 14. Carter-Timofte ME, Jørgensen SE, Freytag MR, et al. Deciphering the Role of Host
508 Genetics in Susceptibility to Severe COVID-19. *Front Immunol*. 2020;11:1606.
- 509 15. Stawicki SP, Jeanmonod R, Miller AC, et al. The 2019-2020 Novel Coronavirus (Severe
510 Acute Respiratory Syndrome Coronavirus 2) Pandemic: A Joint American College of
511 Academic International Medicine-World Academic Council of Emergency Medicine

- 512 Multidisciplinary COVID-19 Working Group Consensus Paper. *J Glob Infect Dis.*
513 2020;12(2):47-93.
- 514 16. van Dam PA, Huizing M, Mestach G, et al. SARS-CoV-2 and cancer: Are they really
515 partners in crime? [published online ahead of print, 2020 Jul 11]. *Cancer Treat Rev.*
516 2020;89:102068.
- 517 17. El Baba R, Herbein G. Management of epigenomic networks entailed in coronavirus
518 infections and COVID-19. *Clin Epigenetics.* 2020;12(1):118.
- 519 18. Rodríguez Y, Novelli L, Rojas M, et al. Autoinflammatory and autoimmune conditions
520 at the crossroad of COVID-19 [published online ahead of print, 2020 Jun 16]. *J*
521 *Autoimmun.* 2020;102506.
- 522 19. Cicco S, Cicco G, Racanelli V, Vacca A. Neutrophil Extracellular Traps (NETs) and
523 Damage-Associated Molecular Patterns (DAMPs): Two Potential Targets for COVID-19
524 Treatment. *Mediators Inflamm.* 2020;2020:7527953.
- 525 20. Shi J, Wen Z, Zhong G, et al. Susceptibility of ferrets, cats, dogs, and other domesticated
526 animals to SARS-coronavirus 2. *Science.* 2020;368(6494):1016-1020.
- 527 21. Tiwari R, Dhama K, Sharun K, et al. COVID-19: animals, veterinary and zoonotic links.
528 *Vet Q.* 2020;40(1):169-182.
- 529 22. Halfmann PJ, Hatta M, Chiba S, et al. Transmission of SARS-CoV-2 in Domestic Cats
530 [published online ahead of print, 2020 May 13]. *N Engl J Med.*
531 2020;10.1056/NEJMc2013400.
- 532 23. Sia SF, Yan LM, Chin AWH, et al. Pathogenesis and transmission of SARS-CoV-2 in
533 golden hamsters [published online ahead of print, 2020 May 14]. *Nature.*
534 2020;10.1038/s41586-020-2342-5.
- 535 24. Deng J, Jin Y, Liu Y, et al. Serological survey of SARS-CoV-2 for experimental, domestic,
536 companion and wild animals excludes intermediate hosts of 35 different species of
537 animals [published online ahead of print, 2020 Apr 17]. *Transbound Emerg Dis.*
538 2020;10.1111/tbed.13577.
- 539 25. Cleary SJ, Pitchford SC, Amison RT, et al. Animal models of mechanisms of SARS-CoV-
540 2 infection and COVID-19 pathology [published online ahead of print, 2020 May 27]. *Br*
541 *J Pharmacol.* 2020;10.1111/bph.15143.
- 542 26. Sailleau C, Dumarest M, Vanhomwegen J, et al. First detection and genome sequencing
543 of SARS-CoV-2 in an infected cat in France [published online ahead of print, 2020 Jun
544 5]. *Transbound Emerg Dis.* 2020;10.1111/tbed.13659.
- 545 27. Chan JF, Zhang AJ, Yuan S, et al. Simulation of the clinical and pathological
546 manifestations of Coronavirus Disease 2019 (COVID-19) in golden Syrian hamster
547 model: implications for disease pathogenesis and transmissibility [published online
548 ahead of print, 2020 Mar 26]. *Clin Infect Dis.* 2020;ciaa325.
- 549 28. Kim YI, Kim SG, Kim SM, et al. Infection and Rapid Transmission of SARS-CoV-2 in
550 Ferrets. *Cell Host Microbe.* 2020;27(5):704-709.e2.
- 551 29. Wang L, Mitchell PK, Calle PP, et al. Complete Genome Sequence of SARS-CoV-2 in a
552 Tiger from a U.S. Zoological Collection. *Microbiol Resour Announc.* 2020;9(22):e00468-
553 20.
- 554 30. Sang ER, Tian Y, Gong Y, Miller LC, Sang Y. Integrate Structural Analysis, Isoform
555 Diversity, and Interferon-Inductive Propensity of ACE2 to Refine SARS-CoV2
556 Susceptibility Prediction in Vertebrates. *bioRxiv* 2020.06.27.174961.
- 557 31. Luan J, Lu Y, Jin X, Zhang L. Spike protein recognition of mammalian ACE2 predicts
558 the host range and an optimized ACE2 for SARS-CoV-2 infection. *Biochem Biophys Res*
559 *Commun.* 2020;526(1):165-169.

- 560 32. Qiu Y, Zhao YB, Wang Q, et al. Predicting the angiotensin converting enzyme 2 (ACE2)
561 utilizing capability as the receptor of SARS-CoV-2 [published online ahead of print,
562 2020 Mar 19]. *Microbes Infect.* 2020;S1286-4579(20)30049-6.
- 563 33. Sun J, He WT, Wang L, et al. COVID-19: Epidemiology, Evolution, and Cross-
564 Disciplinary Perspectives. *Trends Mol Med.* 2020;26(5):483-495.
- 565 34. Blanco-Melo D, Nilsson-Payant BE, Liu WC, et al. Imbalanced Host Response to SARS-
566 CoV-2 Drives Development of COVID-19. *Cell.* 2020;181(5):1036-1045.e9
- 567 35. O'Brien TR, Thomas DL, Jackson SS, Prokunina-Olsson L, Donnelly RP, Hartmann R.
568 Weak Induction of Interferon Expression by SARS-CoV-2 Supports Clinical Trials of
569 Interferon Lambda to Treat Early COVID-19 [published online ahead of print, 2020 Apr
570 17]. *Clin Infect Dis.* 2020;ciaa453
- 571 36. Hadjadj J, et al. Impaired type I interferon activity and exacerbated inflammatory
572 responses in severe Covid-19 patients. Preprint at medRxiv. 2020 doi:
573 10.1101/2020.04.19.20068015.
- 574 37. Acharya D, Liu G, Gack MU. Dysregulation of type I interferon responses in COVID-19
575 [published online ahead of print, 2020 May 26]. *Nat Rev Immunol.* 2020;1-2.
- 576 38. Zhou Z, Ren L, Zhang L, et al. Heightened Innate Immune Responses in the Respiratory
577 Tract of COVID-19 Patients [published online ahead of print, 2020 May 4]. *Cell Host*
578 *Microbe.* 2020;S1931-3128(20)30244-4.
- 579 39. Deng X, van Geelen A, Buckley AC, et al. Coronavirus Endoribonuclease Activity in
580 Porcine Epidemic Diarrhea Virus Suppresses Type I and Type III Interferon Responses.
581 *J Virol.* 2019;93(8):e02000-18.
- 582 40. Andreaskos E, Tsiodras S. COVID-19: lambda interferon against viral load and
583 hyperinflammation. *EMBO Mol Med.* 2020;12(6):e12465.
584 doi:10.15252/emmm.202012465
- 585 41. Hirano T, Murakami M. COVID-19: A New Virus, but a Familiar Receptor and Cytokine
586 Release Syndrome. *Immunity.* 2020;52(5):731-733.
- 587 42. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 Cell Entry Depends on
588 ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell.*
589 2020;181(2):271-280.e8.
- 590 43. Yuan M, Wu NC, Zhu X, et al. A highly conserved cryptic epitope in the receptor
591 binding domains of SARS-CoV-2 and SARS-CoV. *Science.* 2020;368(6491):630-633.
- 592 44. Verdecchia P, Cavallini C, Spanevello A, Angeli F. The pivotal link between ACE2
593 deficiency and SARS-CoV-2 infection. *Eur J Intern Med.* 2020;76:14-20.
- 594 45. Li Y, Zhou W, Yang L, You R. Physiological and pathological regulation of ACE2, the
595 SARS-CoV-2 receptor. *Pharmacol Res.* 2020;157:104833.
- 596 46. Alifano M, Alifano P, Forgez P, Iannelli A. Renin-angiotensin system at the heart of
597 COVID-19 pandemic. *Biochimie.* 2020;174:30-33.
- 598 47. Ziegler CGK, Allon SJ, Nyquist SK, et al. SARS-CoV-2 Receptor ACE2 Is an Interferon-
599 Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell
600 Subsets across Tissues. *Cell.* 2020;181(5):1016-1035.e19.
- 601 48. Zhuang MW, Cheng Y, Zhang J, et al. Increasing Host Cellular Receptor-Angiotensin-
602 Converting Enzyme 2 (ACE2) Expression by Coronavirus may Facilitate 2019-nCoV (or
603 SARS-CoV-2) Infection [published online ahead of print, 2020 Jun 4]. *J Med Virol.*
604 2020;10.1002/jmv.26139.
- 605 49. Sungnak W, Huang N, Bécavin C, et al. SARS-CoV-2 entry factors are highly expressed
606 in nasal epithelial cells together with innate immune genes. *Nat Med.* 2020;26(5):681-
607 687

- 608 50. Schoggins JW. Interferon-Stimulated Genes: What Do They All Do?. *Annu Rev Virol.*
609 2019;6(1):567-584.
- 610 51. Barrat FJ, Crow MK, Ivashkiv LB. Interferon target-gene expression and epigenomic
611 signatures in health and disease. *Nat Immunol.* 2019;20(12):1574-1583.
- 612 52. Wang W, Xu L, Su J, Peppelenbosch MP, Pan Q. Transcriptional Regulation of Antiviral
613 Interferon-Stimulated Genes. *Trends Microbiol.* 2017;25(7):573-584.
- 614 53. Wang W, Xu L, Brandsma JH, et al. Convergent Transcription of Interferon-stimulated
615 Genes by TNF- α and IFN- α Augments Antiviral Activity against HCV and HEV. *Sci*
616 *Rep.* 2016;6:25482.
- 617 54. Chen K, Liu J, Cao X. Regulation of type I interferon signaling in immunity and
618 inflammation: A comprehensive review. *J Autoimmun.* 2017;83:1-11.
- 619 55. Jou J, Gabdank I, Luo Y, et al. The ENCODE Portal as an Epigenomics Resource. *Curr*
620 *Protoc Bioinformatics.* 2019;68(1):e89.
- 621 56. Ambrosini G, Groux R, Bucher P. PWMScan: a fast tool for scanning entire genomes
622 with a position-specific weight matrix. *Bioinformatics.* 2018;34(14):2483-2484.
- 623 57. Sang Y, Rowland RR, Blecha F. Antiviral regulation in porcine monocytic cells at
624 different activation states. *J Virol.* 2014;88(19):11395-11410.
- 625 58. Cildir G, Low KC, Tergaonkar V. Noncanonical NF- κ B Signaling in Health and Disease.
626 *Trends Mol Med.* 2016;22(5):414-429.
- 627 59. Sun SC. The non-canonical NF- κ B pathway in immunity and inflammation. *Nat Rev*
628 *Immunol.* 2017;17(9):545-558.
- 629
630
631
632
633
634

635 **Figure legends**

636

637 **Figure 1:** Schematic of epigenetic regulation and interferon (IFN) signaling to coordinate induction
638 of non-canonical IFN-stimulated genes (non-ISGs). Stimulation of lung macrophages and epithelial
639 cells with tumor necrosis factor (TNF) induces transient expression of TNF-target genes encoding
640 inflammatory mediators, such as IL6 and TNF, followed by an insensitive state in which signaling
641 responses to TLR ligands are strongly suppressed, and chromatin is not activated (depicted by a grey
642 shade). This transient suppression state can be activated by a co-stimulation with TNF plus IFN- α
643 and results in increase of positive histone markers (mostly H3K4me3 and H3K27ac) and chromatin
644 accessibility, which further coordinate binding of IRFs and NF- κ B transcription factors and lead to
645 non-ISG marker gene (such as IL-6) expression. Many inflammatory genes including angiotensin
646 converting enzyme 2 (ACE2) as demonstrated in recent studies can be among these genes, which are
647 bookmarked with primed chromatin and subsequently exhibit a robust transcriptional response even
648 to very weak proximal TLR-induced signals, which may comprise a critical factor in exacerbation of
649 pulmonary inflammatory and COVID-19 syndrome. Adapted and redrawn from Barrat *et al.* (2019)
650 [51]. Abbreviations: ac, acetyl; me, methylation; Pol, polymerase; PU.1, transcription factor binding
651 to the PU-box, a.k.a SPI1; Non-ISG, non-canonical interferon stimulated genes; GTF, sTF, or TF,
652 general (G), tissue-specific (s) transcription factor (TF); TLR, toll-like receptor; TSS, transcription start
653 site.

654

655 **Figure 2:** Profiling of positive histone markers (H3K4me3 and H3K27ac) indicating chromatin
656 accessibility of RNA polymerase II adjacent to human and mouse ACE2 and IL-6 gene bodies,
657 respectively. Annotation of ENCODE epigenetic datasets (ChIP-Seq and ATAC-Seq from 839 and 157
658 cell/tissue types of humans and mice, respectively from <https://www.encodeproject.org/>).
659 Comparative existence of H3K4me3 and H3K27ac markers was detected between IL-6 and ACE2
660 gene promoters in either humans (A & B) and mice (C & D); however, higher Z-scores and enrichment
661 of H3K4me3 and H3K27ac were found in human IL-6 and ACE2 genes (A & B) than their orthologs
662 in mice (C & D). Distal, >2000 bp before the transcription start sites (TSS), and proximal promoter
663 is within 2000 bp before the TSS. Datasets with Z-score higher than the overall average are shaded
664 with oval shapes.

665

666

667 **Figure 3:** Existence of *cis*-regulatory elements (CREs) that bind typical non-ISGs transcription factors
668 of (A) PU.1 (a.k.a. SPI1), (B) IRF1, and (C & D) NF- κ B1/2 in the promoter regions of IL-6 and ACE2
669 gene orthologs from the representative two SARS-CoV2-unsusceptible species (pigs and mice) and
670 seven susceptible species. All three types of CREs have comparable Log₂(mPWM) scores between
671 ACE2 and IL-6 genes, except NF- κ B2 that mediates non-canonical NF- κ B response (D) has a
672 significant lower mPWN score (2-6 Log₂ units), indicating ACE2 genes are among different non-ISGs
673 group other than IL-6. P/D, proximal or distal regions of promoters; +/- sense or antisense strands.
674 mPWM scores are calculated using tools at <https://ccg.epfl.ch/pwmtools/pwmtools.php> with CRE
675 Matrices are from MEME-derived HOCOMOCOv11 TF collection affiliated with the PWM tools.
676 PWM, position weight matrix.

677

678 **Figure 4:** Lack of ISRE/IRF9 binding site that responds to IFN signaling for ISG expression in analyzed
679 IL-6 and ACE2 genes. Cross-species analysis of mean PWM (mPWM) scores of *cis*-regulatory
680 elements (CREs) that bind ISRE/IRF9 in the proximal promoter regions of IL-6 and ACE2 gene
681 orthologs from the 25 representative vertebrate species. mPWM score is presented in a Log₂(mPWM)
682 scale. It further indicates IL-6 and especially ACE2 genes in most species are non-ISGs. Canonical
683 ISGs of human ISG15 and IRF1 are used as references. mPWM scores are calculated using tools at
684 <https://ccg.epfl.ch/pwmtools/pwmscore.php> with CRE Matrices from MEME-derived
685 HOCOMOCOv11 TF collection affiliated with the PWM tools. PWM, position weight matrix.
686 Abbreviations: D-rerio, *Danio rerio* (Zebrafish); X_trapicalis, *Xenopus trapicalis*; G_monkey, African
687 Green Monkey; h-: human.

688
689

690 **Figure 5:** Cross-species analysis of mean PWM scores of *cis*-regulatory elements (CREs) that bind (A)
691 STAT1/2, (B) PU.1 (a.k.a. SPI1), (C) NF- κ B1, (D) NF- κ B2, and (E) IRFs (including IRF1-9, which show
692 significant PWM scores with $p < 0.0001$) in the proximal promoter regions of IL-6 and ACE2 gene
693 orthologs from the 25 representative vertebrate species. All types of CREs have comparable
694 Log₂(mPWM) scores between ACE2 and IL-6 genes, except NF- κ B2 that mediates non-canonical NF-
695 κ B response (D) has a significant lower mPWN score (2-6 Log₂ units), indicating ACE2 genes are
696 among different non-ISGs group other than IL-6. Canonical ISGs of human ISG15 and IRF1 are used
697 as references. mPWM scores are calculated using tools at
698 <https://ccg.epfl.ch/pwmtools/pwmscore.php> with CRE Matrices from MEME-derived
699 HOCOMOCOv11 TF collection affiliated with the PWM tools. PWM, position weight matrix. Other
700 abbreviations are as in Figure 4.

701
702

703 **Figure 6:** Cross-species correlation of epigenetically regulatory CREs, which associate with
704 inflammatory and IFN signaling, in IL-6 and ACE2 gene promoters as biomarkers for COVID-19
705 susceptibility. Mean PWM (mPWM) scores were generated as described in previous figures, and
706 compared between two groups of known COVID-19 susceptible species [CoV2(+)] and unsusceptible
707 species [CoV2(-)]. This shows that ACE2 and especially IL-6 genes from CoV2(+) species contain the
708 CREs have significantly higher mPWM scores, indicating that in some vertebrate species, non-ISGs
709 like ACE2 and especially IL-6 genes evolved to obtain high inductive propensity by inflammatory
710 and IFN signaling, and may serve as epigenetic biomarkers (or triggers) for susceptibility prediction
711 of COVID19 and other ARD syndrome. Abbreviation: H_Bat, Great horseshoe bat, and other
712 abbreviations are as in Figure 4.

713

714 **Figure 7:** Cross-species phylogenic and topological comparison of IL-6 and ACE2 gene
715 promoters. Evolutionary analyses were conducted in MEGA X. The evolutionary history was
716 inferred by using the Maximum Likelihood method and Tamura-Nei model. The tree with the highest
717 log likelihood (-52755.39) is shown. The percentage of trees in which the associated taxa clustered
718 together is shown next to the branches. Initial tree(s) for the heuristic search were obtained
719 automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances
720 estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood

721 value. For topological comparison between phylogenetic trees generated using IL-6 and ACE2 gene
722 proximal promoters, the phylogenies of Newick strings were generated using the MEGA, and
723 topological comparison between the Newick trees was performed with Compare2Trees at
724 (<http://www.mas.ncl.ac.uk/~ntmwn/compare2trees>) to obtain the overall topological scores. Orange
725 circle: marking COVID-19 susceptible species. Arrows: other tentative marker species to determine
726 which group (IL-6 or ACE2) of non-ISGs are more determined for COVID-19 susceptibility.
727 Abbreviations are as in Figure 4.

728

729 **Figure 8:** Genome-wide categorizing non-ISGs based on the similarity of inductive pattern to IL-6
730 and ACE2 genes. The non-biased genome-wide transcriptomic data was generated using a RNA-Seq
731 procedure in porcine lung macrophages stimulated with each of activation stimulator of IL-4, IL-10,
732 LPS, IFN- α or IFN- γ at 20 ng/ml and infected by porcine arterivirus virus for 5 h, using an Illumina
733 procedure as previously described [57]. Significantly differentially expressed genes (DEGs) in renin-
734 angiotensin system (RAS), interleukin (IL)-6, TNF and chemokine super-families were annotated and
735 grouped using heatmaps according to their inductive expression patterns similar to: (A) IL-6, (B)
736 ACE2; (C) Examples of canonical ISGs as reference; (D) Averaged transcriptomic expression levels
737 (normalized at Reads Per Kilobase of transcript per Million mapped reads, RPKM) of the grouped
738 ISGs or non-ISGs above. Indicated by arrows, pigs have two ACE2 isoforms, namely ACE2L and
739 ACE2S, which have different expression patterns, ACE2S similar to IL-6 was showing less responsive
740 to IFN- α but highly responsive to LPS and IFN- γ . In contrast, ACE2L and another key gene, AGT, in
741 RAS were categorized together with other non-ISGs (B), which is more like the expression pattern of
742 canonical ISGs (C) than the IL-6 group (A).

743

744 **Figure 9:** Working summary for IL-6 and ACE2 as non-ISGs biomarkers and contribution to COVID-
745 19 susceptibility. Epigenetic regulation of non-ISGs such as IL-6 and ACE2 was sequentially regulated
746 by such as TNF, IFN and TLR signaling, which modify chromatin accessibility through activating
747 histone modification and recruitment of transcription factors including PU.1, IRF and NF- κ B binding
748 on promoter regions of IL-6 and ACE2 genes. In turn, it will amplify inflammatory loop through IL-
749 6-mediated response and inducing more ACE2 expression, which collectively contribute to the
750 occurrence of respiratory distress syndrome as in COVID-19. Therefore, high expression of non-ISGs
751 such as IL-6 and ACE2 could be biomarkers to determine COVID-19 susceptibility and disease
752 development in different animal species. Abbreviations: non-ISG, non-canonical interferon
753 stimulated genes; GTF, sTF, or TF, general (G), tissue-specific (s) transcription factor (TF); TLR, toll-
754 like receptor; TSS, transcription start site.

755

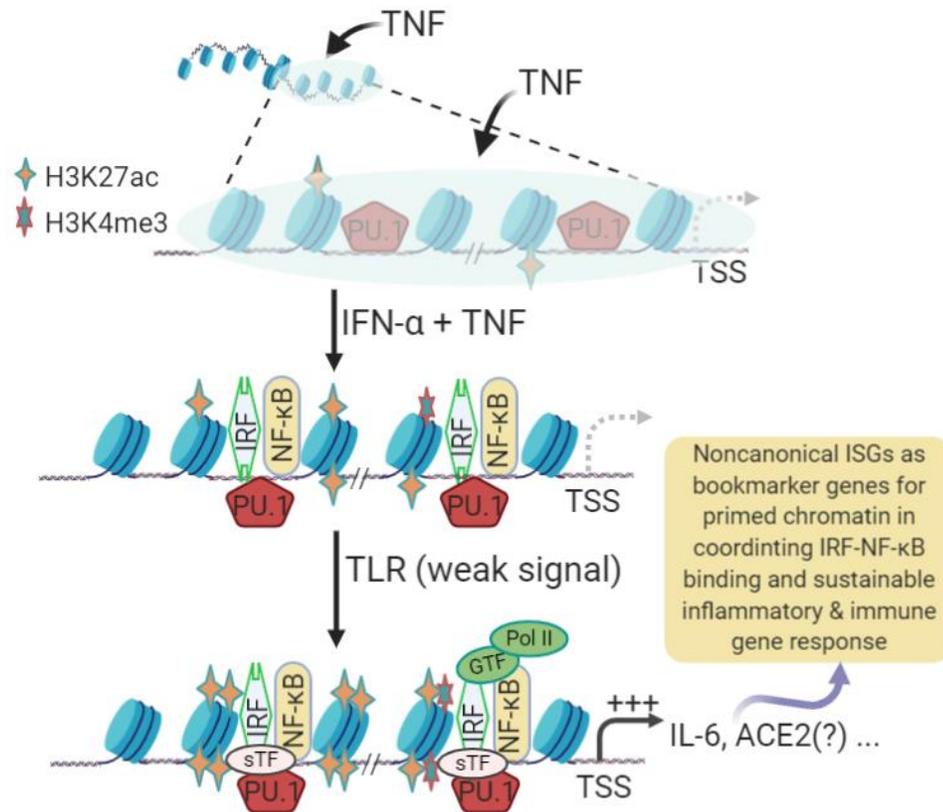


Figure 1: Schematic of epigenetic regulation and interferon (IFN) signaling to coordinate induction of non-canonical IFN-stimulated genes (non-ISGs). Stimulation of lung macrophages and epithelial cells with tumor necrosis factor (TNF) induces transient expression of TNF-target genes encoding inflammatory mediators, such as IL6 and TNF, followed by an insensitive state in which signaling responses to TLR ligands are strongly suppressed, and chromatin is not activated (depicted by grey shading). This transient suppression state can be activated by a co-stimulation with TNF plus IFN- α and results in an increase of positive histone markers (mostly H3K4me3 and H3K27ac) and chromatin accessibility, which further coordinate binding of IRFs and NF- κ B transcription factors and lead to non-ISG marker gene (such as IL-6) expression. Many inflammatory genes, including angiotensin converting enzyme 2 (ACE2) as demonstrated in recent studies, can be among these genes bookmarked with primed chromatin and subsequently exhibit a robust transcriptional response even to very weak proximal TLR-induced signals, which may comprise a critical factor in exacerbation of pulmonary inflammatory and COVID-19 syndrome. Adapted and redrawn from Barrat *et al.* (2019) [51]. Abbreviations: ac, acetyl; me, methylation; Pol, polymerase; PU.1, transcription factor binding to the PU-box, a.k.a SPI1; Non-ISG, non-canonical interferon stimulated genes; GTF, sTF, or TF, general (G), tissue-specific (s) transcription factor (TF); TLR, toll-like receptor; TSS, transcription start site.

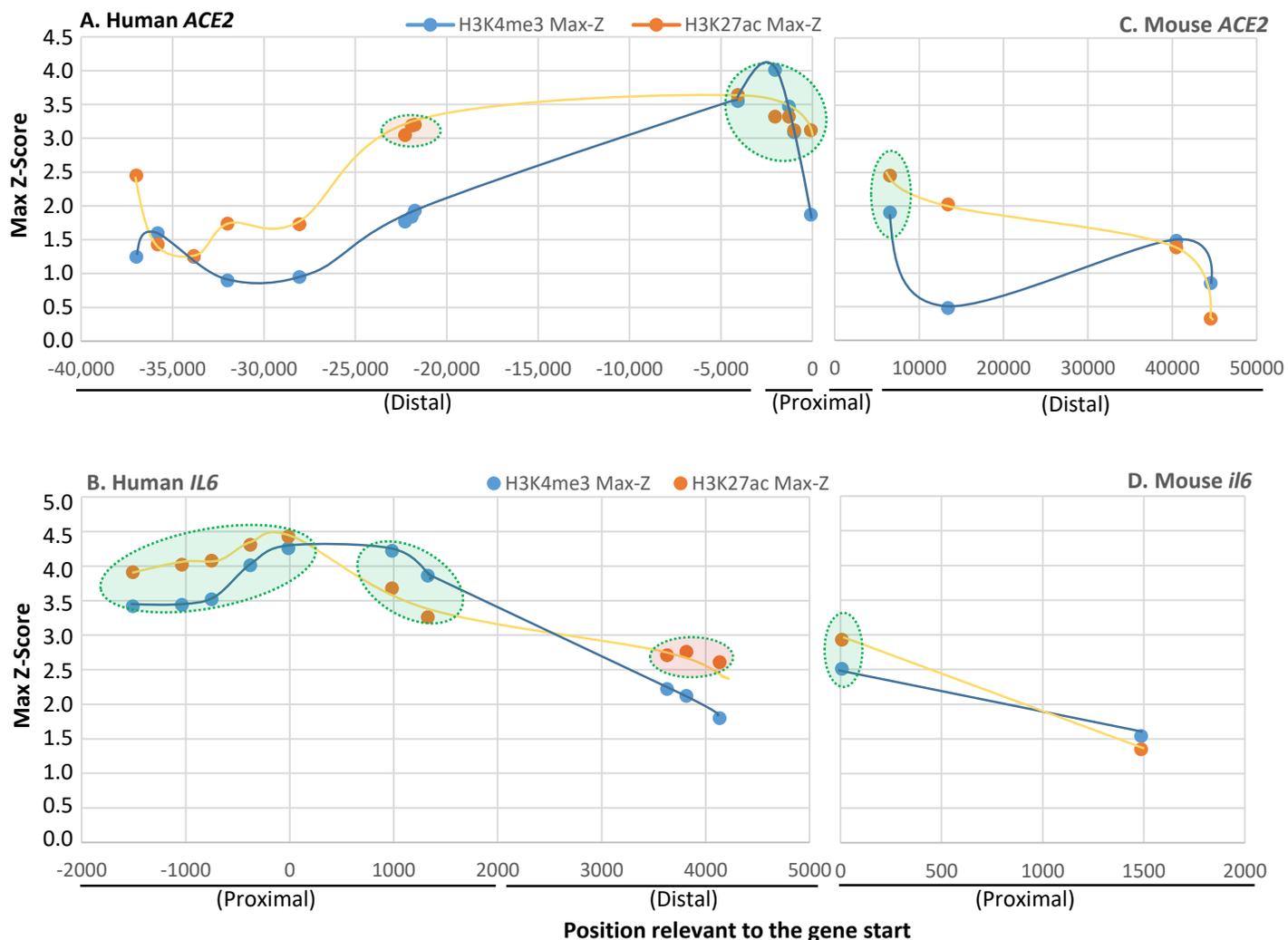


Figure 2: Profiling of positive histone markers (H3K4me3 and H3K27ac) indicating chromatin accessibility of RNA polymerase II adjacent to human and mouse *ACE2* and *IL-6* gene bodies, respectively. Annotation of ENCODE epigenetic datasets (Chip-Seq and ATAC-Seq from 839 and 157 cell/tissue types of humans and mice, respectively, from <https://www.encodeproject.org/>). Comparative existence of H3K4me3 and H3K27ac markers was detected between *IL-6* and *ACE2* gene promoters in either humans (A & B) and mice (C & D): higher Z-scores and enrichment of H3K4me3 and H3K27ac were found in human *IL-6* and *ACE2* genes (A & B) than their orthologs in mice (C & D). Distal, >2000 bp before the transcription start sites (TSS), and proximal promoter is within 2000 bp before the TSS. Datasets with Z-score higher than the overall average are shaded with oval shapes.

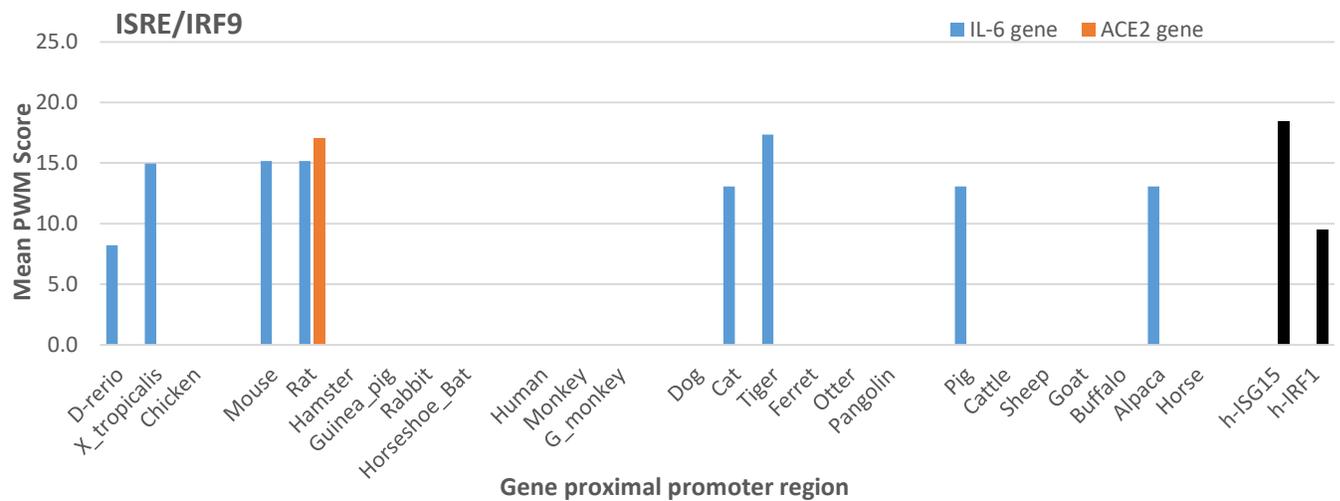


Figure 4: Lack of the ISRE/IRF9 binding site that responds to IFN signaling for ISG expression in analyzed IL-6 and ACE2 genes. Cross-species analysis of mean PWM (mPWM) scores of *cis*-regulatory elements (CREs) that bind ISRE/IRF9 in the proximal promoter regions of IL-6 and ACE2 gene orthologs from the 25 representative vertebrate species. mPWM score is presented in a $\text{Log}_2(\text{mPWM})$ scale. It further indicates IL-6 and especially ACE2 genes in most species are non-ISGs. Canonical ISGs of human ISG15 and IRF1 are used as references. mPWM scores are calculated using tools at <https://ccg.epfl.ch/pwmttools/pwmscore.php> with CRE Matrices from MEME-derived HOCOMOCOv11 TF collection affiliated with the PWM tools. PWM, position weight matrix. Abbreviations: D-rerio, Danio rerio (Zebrafish); X_tropicalis, Xenopus tropicalis; G_monkey, African Green Monkey; h-: human.

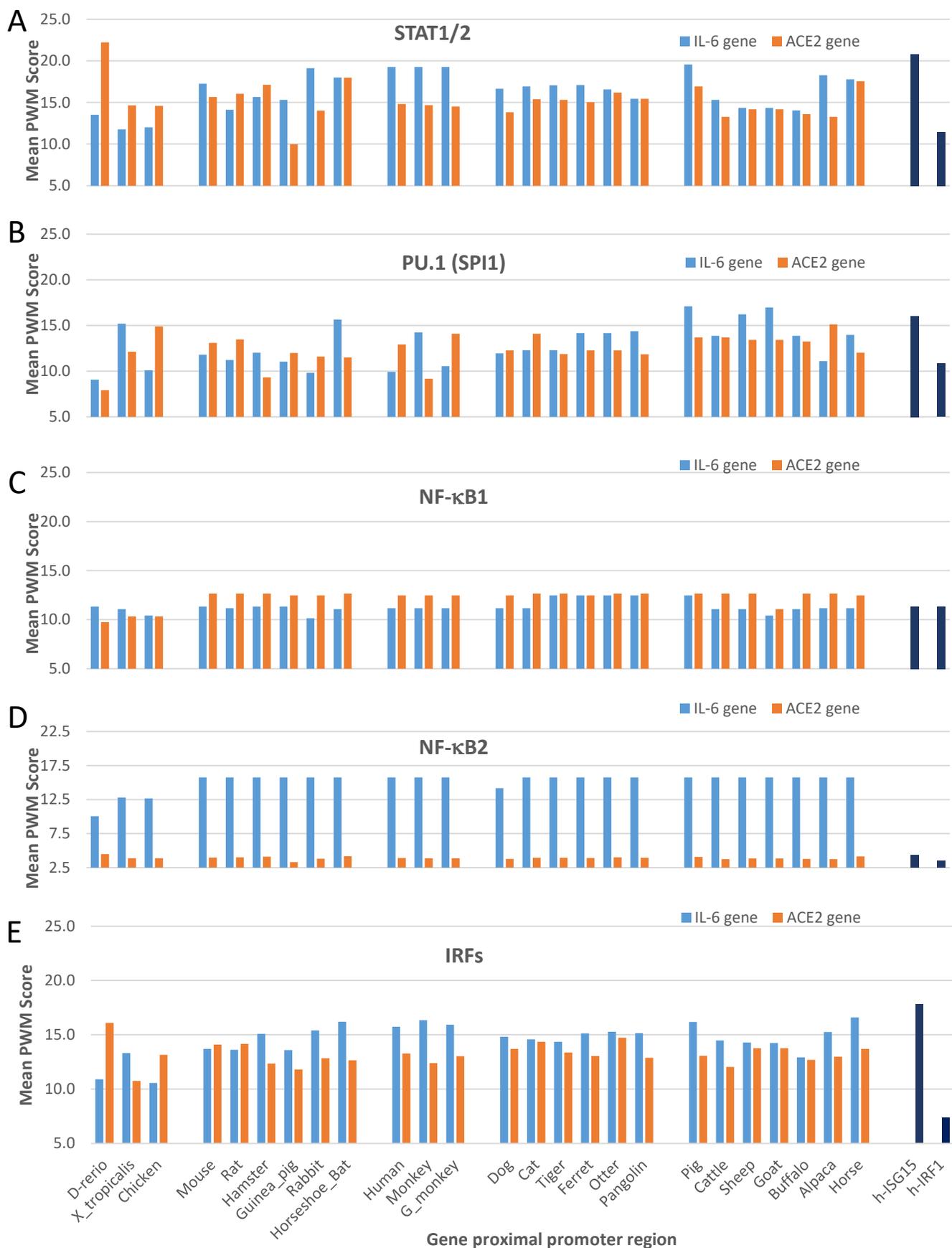
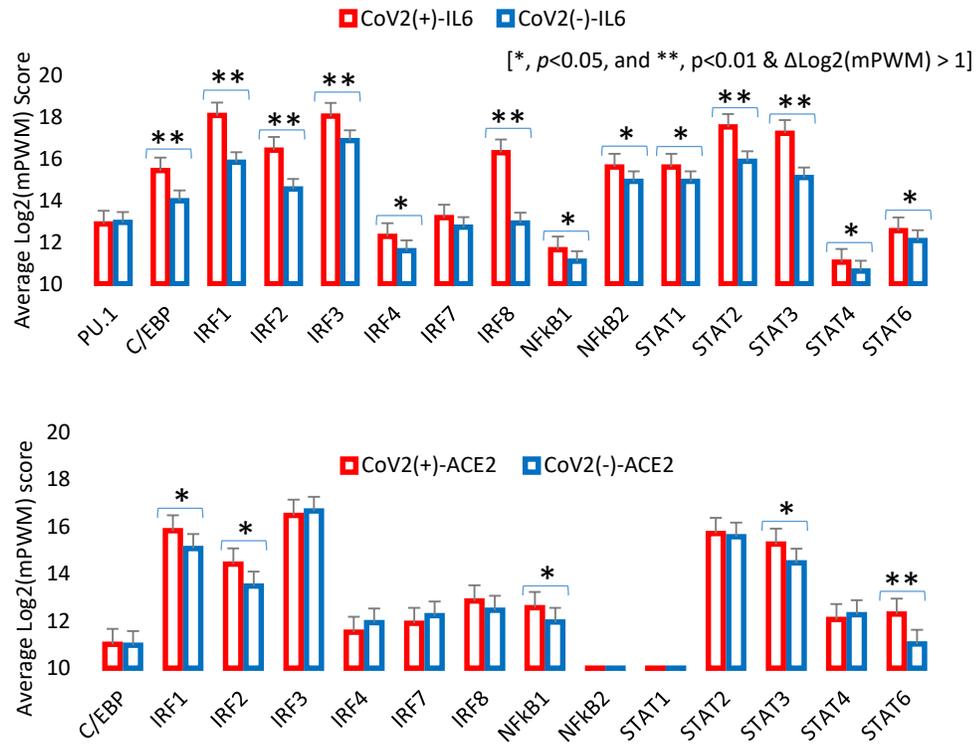


Figure 5: Cross-species analysis of mean PWM scores of *cis*-regulatory elements (CREs) that bind (A) STAT1/2, (B) PU.1 (a.k.a. SPI1), (C) NF-κB1, (D) NF-κB2, and (E) IRFs (including IRF1-9, which show significant PWM scores with $p < 0.0001$) in the proximal promoter regions of IL-6 and ACE2 gene orthologs from the 25 representative vertebrate species. All types of CREs have comparable $\text{Log}_2(\text{mPWM})$ scores between ACE2 and IL-6 genes, except NF-κB2 that mediates non-canonical NF-κB response (D) has a significant lower mPWM score (2-6 Log_2 units), indicating ACE2 genes are among different non-ISGs group other than IL-6. Canonical ISGs of human ISG15 and IRF1 are used as references. mPWM scores are calculated using tools at <https://ccg.epfl.ch/pwmtools/pwmtools.php> with CRE Matrices from MEME-derived HOCOMOCOv11 TF collection affiliated with the PWM tools. PWM, position weight matrix. Other abbreviations are as in Figure 4.



CoV2 (+)	Human	Monkey	G_monkey	Hamster	Dog	Cat	Tiger	Ferret	H_Bat	Pangolin	
CoV2 (-)	Mouse	Rat	Guinea_pig	Rabbit	Pig	Cattle	Sheep	Goat	Buffalo	Alpaca	Horse

Figure 6: Cross-species correlation of epigenetically regulatory CREs, which associate with inflammatory and IFN signaling, in IL-6 and ACE2 gene promoters as biomarkers for COVID-19 susceptibility. Mean PWM (mPWM) scores were generated as described in previous figures, and compared between two groups of known COVID-19 susceptible species [CoV2(+)] and unsusceptible species [CoV2(-)]. This shows that ACE2 and especially IL-6 genes from CoV2(+) species contain CREs with significantly higher mPWM scores, indicating that in some vertebrate species, non-ISGs like ACE2 and especially IL-6 genes evolved to obtain high inductive propensity by inflammatory and IFN signaling, and may serve as epigenetic biomarkers (or triggers) for susceptibility prediction for COVID-19 and other ARD syndromes. Abbreviation: H_Bat, Great horseshoe bat, and other abbreviations are as in Figure 4.

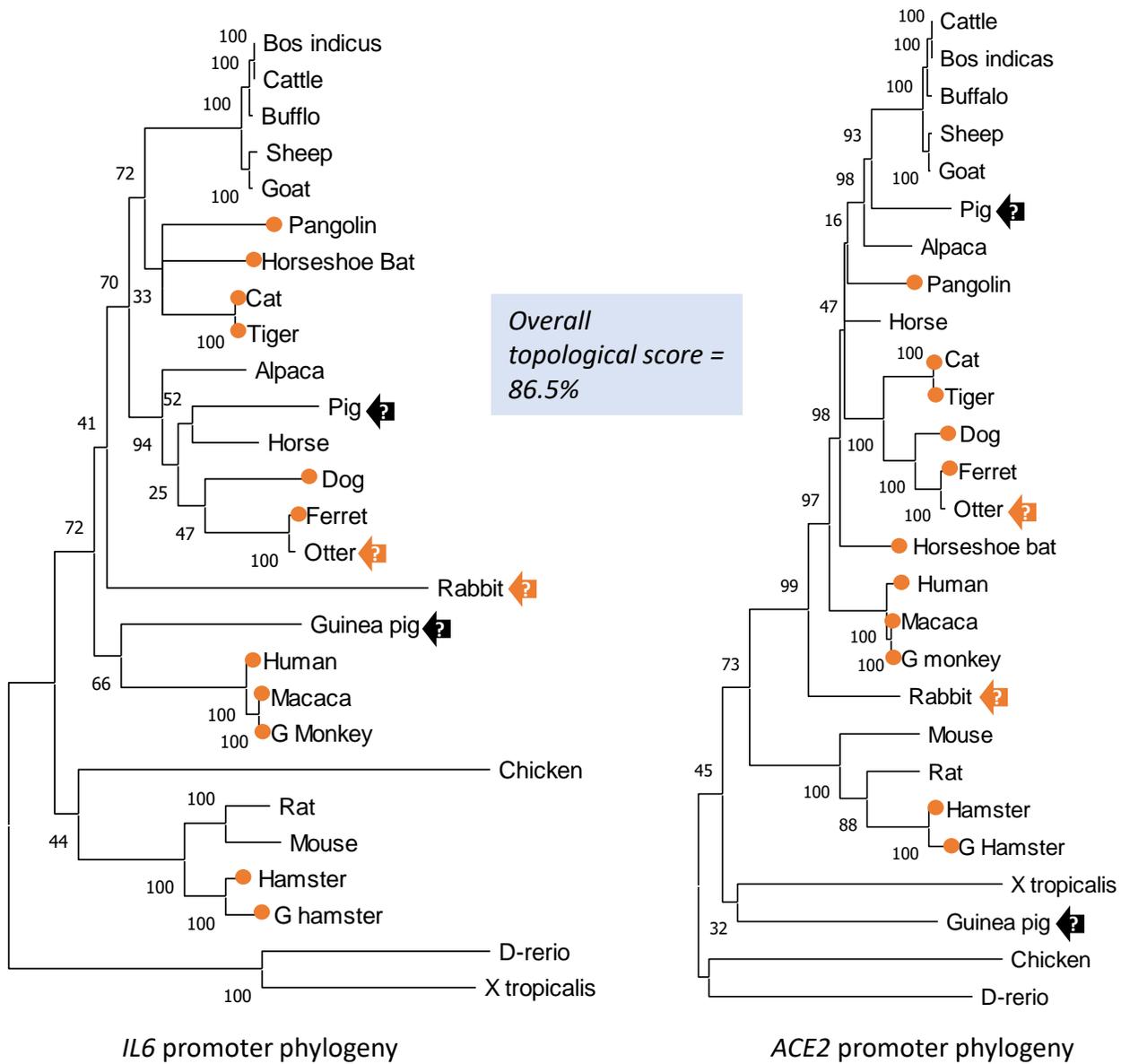


Figure 7: Cross-species phylogenetic and topological comparison of IL-6 and ACE2 gene promoters. Evolutionary analyses were conducted in MEGA X. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model. The tree with the highest log likelihood (-52755.39) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood value. For topological comparison between phylogenetic trees generated using IL-6 and ACE2 gene proximal promoters, the phylogenies of Newick strings were generated using the MEGA, and topological comparison between the Newick trees was performed with Compare2Trees at (<http://www.mas.ncl.ac.uk/~ntmwn/compare2trees>) to obtain the overall topological scores. Orange circle: COVID-19 susceptible species. Arrows: other tentative marker species to determine which group (IL-6 or ACE2) of non-ISGs are more determined for COVID-19 susceptibility. Abbreviations are as in Figure 4.

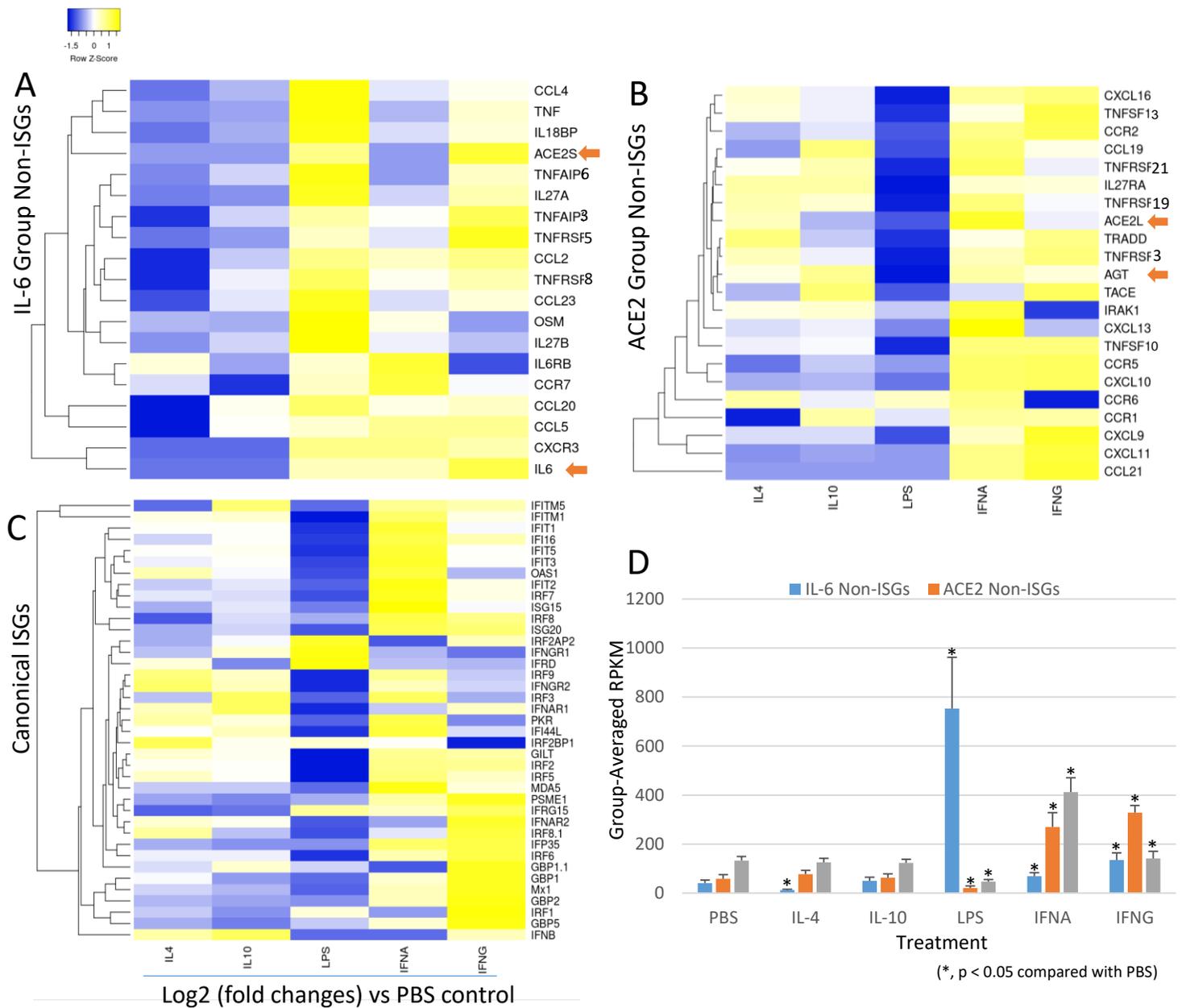


Figure 8: Genome-wide categorizing of non-ISGs based on the similarity of inductive pattern to IL-6 and ACE2 genes. The non-biased genome-wide transcriptomic data was generated using RNA-Seq of porcine lung macrophages activated with stimuli of IL-4, IL-10, LPS, IFN- α or IFN- γ at 20 ng/ml and infected by porcine arterivirus virus for 5 h, using an Illumina procedure as previously described [57]. Significantly differentially expressed genes (DEGs) in renin-angiotensin system (RAS), interleukin (IL)-6, TNF and chemokine super-families were annotated and grouped using heatmaps according to their inductive expression patterns similar to: (A) IL-6, (B) ACE2; (C) Examples of canonical ISGs as reference; (D) Averaged transcriptomic expression levels (normalized at Reads Per Kilobase of transcript per Million mapped reads, RPKM) of the grouped ISGs or Non-ISGs above. Indicated by arrows, pigs have two ACE2 isoforms, namely ACE2L and ACE2S, which have different expression patterns, ACE2S similar to IL-6 was shown to be less responsive to IFN- α but highly responsive to LPS and IFN- γ . In contrast, ACE2L and another key gene, AGT, in RAS were categorized together with other non-ISGs (B), which is more like the expression pattern of canonical ISGs (C) than the IL-6 group (A).

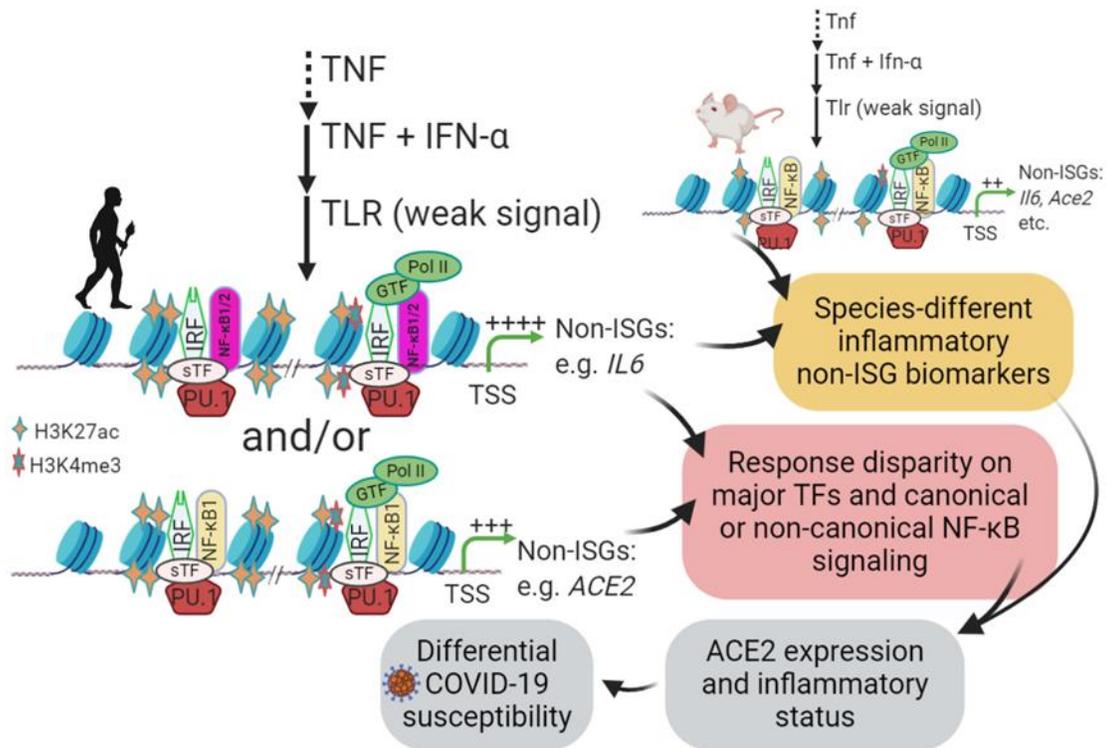


Figure 9: Working summary for IL-6 and ACE2 as non-ISGs biomarkers and contribution to COVID-19 susceptibility. Epigenetic regulation of non-ISGs such as IL-6 and ACE2 was sequentially regulated by such as TNF, IFN and TLR signaling, which modify chromatin accessibility through activating histone modification and recruitment of transcription factors including PU.1, IRF and NF-κB binding on promoter regions of IL-6 and ACE2 genes. In turn, it will amplify the inflammatory loop through the IL-6-mediated response and induce greater ACE2 expression, which collectively contributes to the occurrence of respiratory distress syndrome as in COVID-19. Therefore, high expression of non-ISGs such as IL-6 and ACE2 could be biomarkers to determine COVID-19 susceptibility and disease development in different animal species. Abbreviations: non-ISG, non-canonical interferon stimulated genes; GTF, sTF, or TF, general (G), tissue-specific (s) transcription factor (TF); TLR, toll-like receptor; TSS, transcription start site.