CD4 T-cell transcriptome expression reversal of the lncRNA-mRNA co-expression network in elite controller *vs*. normal-process HIV patients

Short title: LncRNA-mRNA co-expression network in patients with HIV

Chaoyu Chen[1], Xiangyun Lu[1], Nanping Wu[1*]

[1]State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China

[*]Corresponding Author

Chaoyu Chen ccy112359@163.com

Xiangyun Lu present81@163.com

Nanping Wu[*] flwnp@zju.edu.cn

Keywords: Transcriptome, CD4 T-cell, network, elite controller, HIV

Abstract

Elite controller refers to a patient with human imunodeficienvcy virus infection with an undetected viral load without anti-viral treatment. Studies on gene expression and regulation in these individuals are limited but significant. We enrolled 196 patients and collected CD4 T-cell samples from two elite controllers, two normal-process infected patients, and two healthy controls to perform second-generation transcriptome sequencing. Using the Cuffdiff model, we identified differentially expressed mRNAs and long non-coding RNAs with corrected P value < 0.05, and constructed a protein-protein interaction network as well a long non-coding RNA-mRNA co-expression network based on the Pearson correlation coefficient. Interestingly, some interactions within the networks were identified as associated with viral infections and immune responses. This was the first study to examine gene transcription in elite controllers and to study their functional relationships. Our results provide a reference for subsequent functional verification at the molecular or cellular level.

Author Summary

Some individuals can spontaneously inhibit HIV replication after infection with HIV, and thus lack any symptoms. Studies on these patients, termed elite controllers (ECs) will help researchers and clinicians to understand the interrelationship between HIV and the host. In the present study, we focused on the interactions and functional relationships between significantly differentially expressed long non-coding RNAs

(lncRNAs) and mRNAs in ECs *vs*. normal-process patients (NPs). RNA-sequencing was performed for six representative samples of CD4 T-cells. Using the Pearson correlation test, an lncRNA-mRNA co-expression network was constructed. Several new regulatory relationships between transcripts were revealed that might be closely related to the ability of ECs to maintain a low viral load for long periods without anti-viral treatment. For example, lncRNA *C3orf35* was upregulated in ECs *vs*. NPs and was positively related to downregulation of *GNG2* mRNA (encoding G protein subunit gamma 2), which functions in chemokine signaling pathways and HIV-1 infection. Overall, we identified certain interesting genetic interactions that will provide information about the mechanism of host suppression of viral replication.

Introduction

Human imunodeficienvcy virus (HIV) has been studied for more than 30 years, and research has found that HIV invades host CD4+ cells, integrates its own DNA into the host genome, and establishes a reservoir in the early stage, followed by massive replication, which destroys the normal immune system function of the host, and triggers various concurrent symptoms [1-2]. Ultimately, HIV infection can cause the host to die. Although anti retroviral therapy (ART) treatment can partially reconstitute the immune function of the host, it cannot eradicate the latent HIV pool in the host, which always maintains a low level of virus replication [3]. However, among HIV-infected individuals, a small number of patients are found to be inherently resistant to viral replication, and spontaneously induce suppression of the latent pools

without any anti-viral therapy, resulting in an undetectable viral load in plasma for a long period [4-5]. These patients, termed elite controllers (ECs), have attracted significant research interest. Studying the characteristics of the autoimmune factors of ECs is anticipated to identify important factors that control virus replication. Such valuable information could lead to novel methods to treat and alleviate HIV infection.

Long-noncoding RNAs (lncRNAs) are transcribed RNA molecules greater than 200 nt in length that regulate gene expression by diverse, but as yet not completely understood, mechanisms [6]. Although the function of most lncRNAs is unknown, several have been shown to regulate gene expression at multiple levels from DNA to phenotype [7]. Through a variety of means, including *cis* (near the site of lncRNA production) or *trans* (co-expressed with their target gene) mechanisms, lncRNAs play a vital role in many biological processes [8]. Studies on lncRNAs have become a hotspot in current non-coding RNA research.

Recent studies focused on the role of lncRNAs in HIV pathogenesis, especially the relationship between lncRNA regulation of gene expression and viral infection, replication, and latency [9-12]. A few lncRNAs have been characterized and proven to be closely associated HIV, for example, the knockdown of host lncRNA *NEAT1* enhanced virus production by increasing nucleus-to-cytoplasm export of Rev-dependent instability element (INS)-containing HIV-1 mRNAs [13]. The knockdown of lncRNA *NRON* enhanced HIV-1 replication through increased activity of nuclear factor of activated T-cells (NFAT) and the viral long terminal repeat (LTR) [14]. Similarily, lncRNA *MALAT1* releases epigenetic silencing of HIV-1 replication

by displacing the polycomb repressive complex 2 from binding to the LTR promoter [15]. Furthermore, an HIV-encoded antisense lncRNA, *ASP-L*, was proven to promote latency HIV [16-17]. Although, the mechanism of lncRNA function is sometimes elusory and unpredicted, many current studies, which are limited to the screening and functional prediction stage based on chip or sequencing results, also provide us with valuable information and a basis for mechanistic research.

In the present study, we conducted a transcriptomics investigation for HIV elite controllers, to identify differences in the transcriptional expression profiles between elite controllers and normal-process HIV patients as compared with healthy controls. The obtained lncRNA-mRNA co-expression network revealed the possible role of the functional relationships between lncRNAs and mRNAs in the ability of elite controllers to inhibit viral replication.

Results

Subjects

A total of 196 individuals infected with HIV were enrolled in our cohort, including patients under treatment (179, 91.3%), untreated normal-process patients (15,7.7%) and elite controllers (2, 1.0%). In addition, we recruited two healthy individuals from the out-patient department of the hospital. Most patients were infected with the virus through heterosexual contact, followed by intravenous infection, and homosexual transmission. Based on the basic principle of intra-group identity, we determined three men and three women who were assigned to each of the three study groups as study

111 subjects. Their ages ranged from 39 to 54 and were free of other diseases, such as
112 tuberculosis, diabetes, or hepatitis. Specific details of the clinical characteristics of the
113 participants are shown in S1 Table. Descriptive statistics are reported as counts
114 (percentage) for dichotomous and categorical variables, and the median (the 25th
115 percentile and 75th percentile) for continuous variables. The experimental design and
116 analysis process of this research are shown in a flowchart (Fig 1).

117 **Identification of differentially expressed (DE) transcripts between health**
118 **controls (HCs) and normal-process infected patients (NPs)**

119 First, we scanned the expression profiles of all transcripts in CD4 T-cells from two
120 HCs, two ECs and two NPs. In all, more than 82000 mRNA transcripts and more than
121 20000 corresponding genes were identified. Likewise nearly 15000 lncRNAs and
122 1000 corresponding gene IDs were identified. Regular grouping between HCs and
123 normal-process patients was first considered. Altoghter, 3602 mRNA transcripts and
124 383 lncRNA transcripts were differently expressed according to the threshold of fold
125 change > 2 and p value < 0.05, among which 2313 mRNAs and 207 lncRNAs were
126 upregulated and 1289 mRNAs and 176 lncRNAs were downregulated. We similarly
127 compared gene expression differences between HCs and ECs though more attention
128 was paid to ECs *vs.* NPs. The comparation with HCs helped to provide a basic
129 reference background. The lists of fragments per kilobase of transcript per million
130 mapped reads (FPKM) values and regulatory situation of all DE mRNAs and
131 lncRNAs in HCs *vs.* NPs and HCs *vs.* ECs are shown in S2 Table and S3 Table
132 respectively.

133 **Identification of DE mRNAs between ECs and NPs**

134 Among all transcripts, 2936 transcripts were differentially expressed according to the
135 threshold of fold change > 2 and p value < 0.05. There were 1327 mRNAs that were
136 upregulated and 1609 that downregulated in ECs *vs*. NPs. Gene ontology (GO) and
137 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analysis was carried
138 out to annotate the function of these DE mRNAs. Certain terms and pathways that
139 might be associated with infection and immunity were significantly enriched (S1 Fig).
140 These terms and pathways included negative regulation of interferon-beta production
141 (GO:0032688), cellular component disassembly involved in execution phase of
142 apoptosis (GO:0006921), regulation of defense response to virus by virus
143 (GO:0050690), negative regulation of type I interferon production (GO:0032480),
144 RIG-I-like receptor signaling pathway (hsa04622), and antigen processing and
145 presentation (hsa04612).

146 **Identification of DE lncRNAs between ECs and NPs**

147 As same, among all transcripts, 3543 official gene symbols that matched in the
148 NONCODE database had corresponding lncRNA transcripts. Three main categories
149 of lncRNAs accounted for the majority, including antisense lncRNAs (2537/14853,
150 17.1%), intronic lncRNAs (8152/14853, 54.9%), and long intergenic noncoding
151 RNAs (lincRNAs; 3344/14853, 22.5%). There were 151 upregulated lncRNAs and
152 160 downregulated lncRNAs in ECs *vs*. NPs according to the threshold of fold change
153 > 2 and p < 0.05. In both the mRNA and lcnRNA clusters (Fig 2), the samples showed
154 marked intra-group correlations and inter-group differences, which further confirmed

155 the specificity of the samples. In terms of protein-coding mRNAs encoding proteins
156 with specific biological functions, the HC and EC groups showed closer clustering
157 and were relatively distant from the NP group. The list of fragments per kilobase of
158 transcript per million mapped reads (FPKM) values and regulatory situation of all DE
159 mRNAs and lncRNAs are shown in S4 Table.

160 **PPI network of DE mRNAs**

161 Fig 3 shows the top 50 hub genes playing an important role in interaction of DE
162 mRNAs, in which many genes functioned centrally and could be frequently regulated
163 by other transcripts. Considering the differences in the outputs of different algorithms,
164 we identified those genes that were commonly detected by the algorithms and focused
165 on those genes that were regulated by lncRNAs and their functional enrichment
166 analysis. S5 Table shows the intersection list of the top 50 hub-mRNAs calculated
167 using 12 algorithms.

168 **LncRNA-mRNA co-expression network analysis**

169 Almost 0.9 million co-expression relationships were discovered, and the exact form
170 was expressed as a linear orientation whereby one lncRNA targets one mRNA. In all
171 orientations, 1109 DE lncRNAs targeting DE mRNAs were identified. Fig 4a shows
172 the network constructed using these relationships, in which 142 DE lncRNAs were
173 co-expressed with 127 DE mRNAs. Moreover, we found 16 of 24 intersecting hub
174 mRNAs in the diagram, which helped us to quickly locate the target protein coding
175 genes and easily access the enriched functions of these genes. Next, we selected 15
176 annotated lncRNAs and their associated mRNAs from Fig 4a and constructed Fig 4b.

We then identified the top 10 GO terms and pathways most closely related to infection and immunity by combining the function of the hub genes. Many significant functions are co-enriched, such as immune response-regulating signaling pathway (GO:0002764) and apoptosis (hsa04215) (Fig 5).

Immune regulation and infection associated lncRNA-mRNA networks

We further investigated the specific functions of the co-expressed mRNAs shown in Fig 4b and found several interesting regulation relationships. Certain genes were enriched to GO terms and pathways that are closely related to immune regulation and infection. These genes and corresponding pathways included *GNG2* (Human cytomegalovirus infection, hsa05163; Chemokine signaling pathway, hsa04062; Kaposi sarcoma-associated herpesvirus infection, hsa05167; Human immunodeficiency virus 1 infection, hsa05170), *RFX5* (Primary immunodeficiency, hsa05340; Tuberculosis, hsa05152; Antigen processing and presentation, hsa04612), *HDAC6* (Viral carcinogenesis, hsa05203), *MAPK8* (Apoptosis - multiple species, hsa04215; NOD-like receptor signaling pathway, hsa04621; Fc epsilon RI signaling pathway, hsa04664; RIG-I-like receptor signaling pathway, hsa04622; Inflammatory mediator regulation of TRP channels, hsa04750; Toll-like receptor signaling pathway, hsa04620), *MAPK4* (IL-17 signaling pathway, hsa04657), and *RARA* (Th17 cell differentiation, hsa04659). Corresponding lncRNAs including *C3orf35*, *TMEM254-AS1*, *ARRDC3-AS1*, *LINC00202-1*, *TRAF3IP2-AS1*, *PSMD5-AS1*, *AC093375.1*, and *SVIL-AS1* were co-expressed with these mRNAs, which suggested several new regulatory relationships that might play an important role in controlling

199 virus replication.

200

201 ## Dicussion

202 In the present study, we systematically analyzed the expression profiles of lncRNAs
203 and mRNAs from CD4+ T cells from ECs, NPs, and HCs, and constructed a
204 co-expression network based on the relationships among DE genes and database
205 annotations. The functions of most lncRNAs are not well characterized; therefore, the
206 gene functions in this research mainly referred to those of mRNAs and their
207 expression products. The potential functions of candidate lncRNAs and the
208 lncRNAs-associated biological processes in ECs were predicted using the
209 lncRNA-mRNA network and functional enrichment analysis. We identified certain
210 genes that play a significant role in the ability of ECs to inhibit the replication of
211 latent virus. Most of these genes were downregulated in the HCs and ECs and
212 upregulated in the NPs, which was consistent with their clinical characteristics, in that
213 ECs maintain a similar gene expression pattern to the HCs, which distinguishes them
214 from the NPs. Among these rigorously selected genes, some hub genes were closely
215 related to microbial infections and host immune processes, which are reported to
216 greatly affect the progress of HIV infection [18-19].
217 To the best of our knowledge, the present study was the first to predict
218 lncRNA-mRNA interactions for HIV ECs, in which many of the reported genes had
219 not been previously reported to function in HIV-associated biological process. The
220 *MAPK* family has been associated with the reactivation and replication of the virus in

many studies [20-21], and the HIV-1 virus inhibition activity of MAPK pathway inhibitors was the result of the negative regulation of HIV-1 LTR promoter activity [22]. In addition, HDAC6 gene expression product was reported to mediate HIV-1 tat-induced proinflammatory responses by regulating the MAPK-NF-kappaB/AP-1 pathways in astrocytes [23]. In the present study, we also found that HDAC6 interacts with MAPK8, MAPK4, and MAPK9 (Fig 3). This suggested their related lncRNAs including *C3orf35*, *TMEM254-AS1*, *ARRDC3-AS1*, *LINC00202-1*, *TRAF3IP2-AS1*, *PSMD5-AS1*, and *AC093375.1* may control HIV replication by regulating the expression of these genes. In addition, we identified genes that had not been studied before, whose functions were enriched into the following processes: Virus infection, immunodeficiency, apoptosis, and cytokines. Previous studies had shown that these functions are closely related to viral replication [24-25]. For instance, *GNG2* (encoding G protein subunit gamma 2) showed a strong correlation with HIV in the KEGG pathway analysis. In Kaposi sarcoma-associated herpesvirus infection, *GNG2* activates Ras pathways and PI3K-Akt pathways to affect the production of cellular cytokines such as IL-6, IL-8, VEGF, and COX2, and other inflammatory and cellular life cycle-related biological processes. HIV infection and its concurrent Kaposi sarcoma are also closely related to the occurrence of other cancers [26-27].

ART treatment cannot eradicate the latent virus in the host; therefore, the genes identified in the present study may play a non-negligible regulatory role in the relationship between the latent virus and the host. Latent virus is the biggest obstacle to the elimination of the disease; thus, understanding the processes that contribute to

its persistence, such as inflammation and immune activation, are crucial for the remission and cure of HIV [28-29].

There are also some limitations to this study. First, because of the limited amount of RNA in the samples, the sample size of this study was relatively small. Although we selected eight genes for verification using qRT-PCR in 12 samples and found that the expression trends of four of them were consistent with the RNA sequencing results, these results are not statistically significant because of the small sample size. Second, we only assessed gene expression of CD4+ T cells, which, although they are the main targets of HIV, are only one of the many cell types that are infected by HIV. Further research should be performed to validate the function and mechanism of the genes in a larger sample and in more cell types.

In summary, this study identified several important differentially expressed genes associated with the EC phenomenon, which could form the basis for subsequent cellular and molecular studies, and could provide new targets for gene-targeted therapy in the future.

Methods

Ethics statement

All participating medical institutions provided local institutional review board approval, and the participants provided informed consent for this study.

Cohort and patients

In total, 196 patients were recruited into our cohort, from different provinces of

China, who were diagnosed in Qingchun Hospital of Zhejiang Province from January 2000 to January 2018, and received or voluntarily rejected highly active antiretroviral therapy (HAART) treatment. The viral loads of the patients were detected using the Cobas Taqman system (Roche). The HCs were randomly selected from hospital clinics and were not suffering from any diseases. We chose two ECs, two NPs, and two HCs of similar ages, gender, and traditional risk factors to perform second generation transcriptome sequencing. ECs were defined as HIV-infected patients with and undetectable viral load without any anti-viral treatment for nearly 10 years.

Sample collection and preparation

Peripheral blood mononuclear cells (PBMCs) were isolated from 3–5 mL EDTA-K2+ anticoagulant venous blood using density gradient centrifugation. Primary CD4+ T cells were purified through negative selection using a CD4+ T cell isolation Kit (Miltenyi). According to flow cytometry analysis, the purity of the separated CD4+ T cells was more than 90%. Total RNA was extracted using the TRIzol reagent (Life Technologies). RNA purity was checked using a NanoPhotometer® spectrophotometer (IMPLEN). The RNA concentration was measured using a Qubit® RNA Assay Kit in Qubit® 2.0 Fluorometer (Life Technologies) and its integrity was assessed using an RNA Nano 6000 Assay Kit for the Bioanalyzer 2100 system (Agilent Technologies).

RNA library construction and RNA sequencing

Ribosomal RNA was removed using a Epicentre Ribo-zero™ rRNA Removal Kit (Epicentre), and the rRNA free residue was cleaned up using ethanol precipitation.

Subsequently, sequencing libraries were generated using the rRNA depleted RNA by NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB) following manufacturer's recommendations. After cDNA synthesis and purification, clustering of the index-coded samples was performed on a cBot Cluster Generation System using a TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. The libraries were then sequenced on an Illumina Hiseq 4000 platform and raw data were generated.

RNA sequencing analysis

Clean data were obtained by removing reads containing adapters, reads only containing poly-N, and low quality reads from the raw data. After quality control, paired-end clean reads were aligned to the reference genome downloaded from website using HISAT2 (v2.0.4). Cuffdiff (v2.1.1) was used to calculate the FPKM values of both lncRNAs and mRNAs in each sample [30]. Gene FPKM values were computed by summing the FPKM values of transcripts in each gene group. The Cuffdiff model provides statistical routines to determine differential expression in the digital transcript or gene expression data using a model based on the negative binomial distribution. Transcripts with a corrected P value < 0.05 were assigned as differentially expressed. We plotted a heatmap to observe the clustering between the samples and the genes using Multi Experiment Viewer (v4.9.0).

LncRNA-mRNA co-expression network analysis

Trans role is lncRNA binding to target DNA as a RNA:DNA heteroduplex, as RNA:DNA:DNA triplex, or RNA recognition of specific chromatin-like complex

surfaces[31]. We calculated the expression correlation between lncRNAs and mRNAs using the Pearson correlation test for target gene prediction and the results were expressed as Pearson correlation coefficients. To narrow the scope of investigation and the number of genes, we first assessed the interaction among all DE mRNAs using the STRING database (v11.0) and constructed a PPI network using the Hubba package in Cytoscape (v3.7.1). After considering 12 synthetic algorithms, the intersection of the top 50 genes yielded 24 hub genes, which occupy a central position in the network of all DE mRNAs. After identifying and retaining the DE mRNAs and lncRNAs in all lncRNA-mRNA relations, we screened 1109 pairs of exact relationships between the two groups and constructed a co-expression network based on the Pearson relation coefficient (all > 0.95). Then, we identified the annotated lncRNAs and associated mRNAs for further functional enrichment analysis. PPI and lncRNA-mRNA network diagrams were also drawn using the Cytoscape software.

Gene ontology and KEGG pathways analysis

By jointly using the David database (v6.8) and KOBAS database (v3.0), gene GO enrichment and KEGG pathways analysis of all DE mRNAs and DE lncRNA co-expressed DE mRNAs were implemented to interpret the biological meaning of the transcripts. GO terms and KEGG pathways with corrected P values less than 0.05 were considered significantly enriched by the differentially expressed RNAs.

Acknowledgements

We would like to thank all of the participants in the cohort. In addition, we would like to thank the native English speaking scientists of Elixigen Company (Huntington Beach, California) for editing our manuscript.

Author Contributions

Chaoyu Chen and Nanping Wu were involved in the conception and design of the study. Chaoyu Chen and Xiangyun Lu were involved in sample preparation, data acquisition, and organization. Chaoyu Chen conducted the data analysis, and all authors were responsible for data interpretation. Chaoyu Chen was the major contributor to the writing of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported in part by grants from the Mega-Project for National Science and Technology Development under the "Study on comprehensive prevention and treatment of AIDS in demonstration areas, 13th, Five-Year Plan of China" (NO.2017ZX10105001-005) and "Research and application of appropriate treatment and prevention strategies for children with AIDS, 13th, Five-Year Plan of China" (NO.2018ZX10302-102). The funders had no role in study design, data collection and

351 analysis, decision to publish, or preparation of the manuscript.

352

353 ## References

354 1. Lee C A. The natural history of HIV infection. Aids, 2013, 1(2):177-179.

355 2. Eisele E, Siliciano RF. Redefining the viral reservoirs that prevent HIV-1

356 eradication. Immunity 37(3), 377 – 388 (2012).

357 3. Murray A J, Kwon K J, Farber D L, et al. The Latent Reservoir for HIV-1: How

358 Immunologic Memory and Clonal Expansion Contribute to HIV-1 Persistence. The

359 Journal of Immunology, 2016, 197(2):407-417.

360 4. Long-term nonprogressive disease among untreated HIV-infected individuals:

361 clinical implications of understanding immune control of HIV. JAMA 304(2),

362 194 – 201 (2010).

363 5. Gonzalogil E, Ikediobi U, Sutton R E. Focus: Infectious Diseases: Mechanisms of

364 Virologic Control and Clinical Characteristics of HIV+ Elite/Viremic Controllers.

365 Yale J Biol Med, 2017, 90(2):245-259.

366 6. Wilusz, J.E., Sunwoo, H., and Spector, D.L. (2009). Long noncoding RNAs:

367 functional surprises from the RNA world. Genes Dev. 23, 1494 – 1504.

368 7. Wapinski, O., and Chang, H.Y. (2011). Long noncoding RNAs and human disease.

369 Trends Cell Biol. 21, 354 – 361.

370 8. Wang, K. C. & Chang, H. Y. Molecular Mechanisms of Long Noncoding RNAs.

371 Mol. Cell 43, 904 – 914 (2011).

372 9. Trypsteen W, Mohammadi P, Van Hecke C, et al. Differential expression of

lncRNAs during the HIV replication cycle: an underestimated layer in the HIV-host interplay. Scientific Reports, 2016, 6:36111.

10. Lazar D C, Morris K V, Saayman S. The emerging role of long non-coding RNAs in HIV infection. Virus Research, 2015:S0168170215300277.

11. Saayman S, Ackley A, Turner A M W, et al. An HIV-Encoded Antisense Long Noncoding RNA Epigenetically Regulates Viral Transcription. Molecular Therapy, 2014, 22(6):1164-1175.

12. Nair M, Sagar V, Pilakka-Kanthikeel S. Gene-expression reversal of lncRNAs and associated mRNAs expression in active vs latent HIV infection. Scientific Reports, 2016, 6:34862.

13. Zhang Q H, Chen C Y, Yedavalli V S R K, et al. NEAT1 Long Noncoding RNA and Paraspeckle Bodies Modulate HIV-1 Posttranscriptional Expression. Mbio, 2013, 4(1):00596-12.

14. Imam, H., Shahr Bano, A., Patel, P., Holla, P. & Jameel, S. The lncRNA NRON modulates HIV-1 replication in a NFAT-dependent manner and is differentially regulated by early and late viral proteins. Sci. Rep. 5, 8639 (2015).

15. Qu Di, Sun Wei-Wei, Li Li et al. Long noncoding RNA MALAT1 releases epigenetic silencing of HIV-1 replication by displacing the polycomb repressive complex 2 from binding to the LTR promoter. Nucleic Acids Res., 2019.

16. Kobayashi-Ishihara, M. et al. HIV-1-encoded antisense RNA suppresses viral replication for a prolonged period. Retrovirology 9, 38 (2012)

17. Saayman, S. et al. An HIV-Encoded Antisense Long Noncoding RNA

395 Epigenetically Regulates Viral Transcription. Mol. Ter. 22, 1164–1175 (2014).

396 18. Bell L C K, Noursadeghi M. Pathogenesis of HIV-1 and Mycobacterium
397 tuberculosis co-infection. Nature Reviews Microbiology, 2017.

398 19. Sumner R P, Thorne L G, Fink D L, et al. Are Evolution and the Intracellular
399 Innate Immune System Key Determinants in HIV Transmission?. Frontiers in
400 Immunology, 2017, 8:1246-.

401 20. Wang X, Sun B, Mbondji C, et al. Differences in Activation of HIV‐1
402 Replication by Superinfection With HIV‐1 and HIV‐2 in U1 Cells. Journal of
403 Cellular Physiology, 2016, 232(7).

404 21. Prasad S, Tyagi A K. Curcumin and its analogues: A potential natural compound
405 against HIV infection and AIDS. Food Funct. 2015, 6(11):10.1039.C5FO00485C.

406 22. Gong J, Shen X H, Chen C, et al. Down-regulation of HIV-1 infection by
407 inhibition of the MAPK signaling pathway. Virologica Sinica, 2011, 26(2):114-122.

408 23. Youn G S, Ju S M, Choi S Y, et al. HDAC6 mediates HIV-1 tat-induced
409 proinflammatory responses by regulating MAPK-NF-kappaB/AP-1 pathways in
410 astrocytes. Glia, 2015, 63(11):1953-1965.

411 24. Khan S Z, Hand N, Zeichner S L. Apoptosis-induced activation of HIV-1 in
412 latently infected cell lines. Retrovirology,12,1(2015-05-16), 2015, 12(1):42.

413 25. Jacobs E S, Keating S M, Abdel-Mohsen M, et al. Cytokines Elevated
414 in HIV Elite Controllers Reduce HIV Replication In Vitro and Modulate HIV
415 Restriction Factor Expression. J Virol, 2017 Feb 28;91(6).

416 26. Gonçalves PH, Uldrick T S, Yarchoan R. HIV-associated Kaposi sarcoma and

417 related diseases. AIDS, 2017:1.

418 27. Gonçalves P H, Ziegelbauer J, Uldrick T S, et al. Kaposi sarcoma
419 herpesvirus-associated cancers and related diseases. Current Opinion in Hiv & Aids,
420 2016, 12(1):47.

421 28. Dahabieh M S, Battivelli E, Verdin E. Understanding HIV Latency: The Road to
422 an HIV Cure. Annual Review of Medicine, 2015, 66(1):407-421.

423 29. Massanella M, Fromentin R, Chomont N. Residual inflammation and viral
424 reservoirs: alliance against an HIV cure. Curr Opin HIV AIDS 2016;11:234-241.

425 30. Trapnell C, Williams B A, Pertea G, et al. Transcript assembly and quantification
426 by RNA-Seq reveals unannotated transcripts and isoform switching during cell
427 differentiation. Nature Biotechnology, 2010, 28(5):511-515.

428 31. Hung T , Chang H Y . Long noncoding RNA in genome regulation. RNA
429 Biology, 2010, 7(5):582-585.

430 Supporting information

431 **S1 Figure. Gene ontology (GO) terms and Kyoto Encyclopedia of Genes and**
432 **Genomes (KEGG) pathways enrichment of all differentially expressed (DE)**
433 **mRNAs.**

434 **S1 Table. Clinical characteristics of the participants.**

435 **S2 Table. Lists of differentially expressed (DE) mRNAs and long non-coding**
436 **RNAs (lncRNAs) in health controls (HCs) *vs*. normal-process infected patients**
437 **(NPs).**

438 **S3 Table. Lists of DE mRNAs and lncRNAs in HCs *vs*. elite controllers (ECs).**

S4 Table. Lists of DE mRNAs and lncRNAs in ECs *vs*. NPs.

S5 Table. The intersection list of the top 50 hub-mRNAs calculated using 12 algorithms.

Fig 1. Flowchart of the analysis procedure. lncRNA, long non-coding RNA; EC, elite controller; NP, normal-process infected patient; DE, differentially expressed; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Fig 2. Heatmaps of differentially expressed (DE) mRNAs and long non-coding RNAs (lncRNAs).

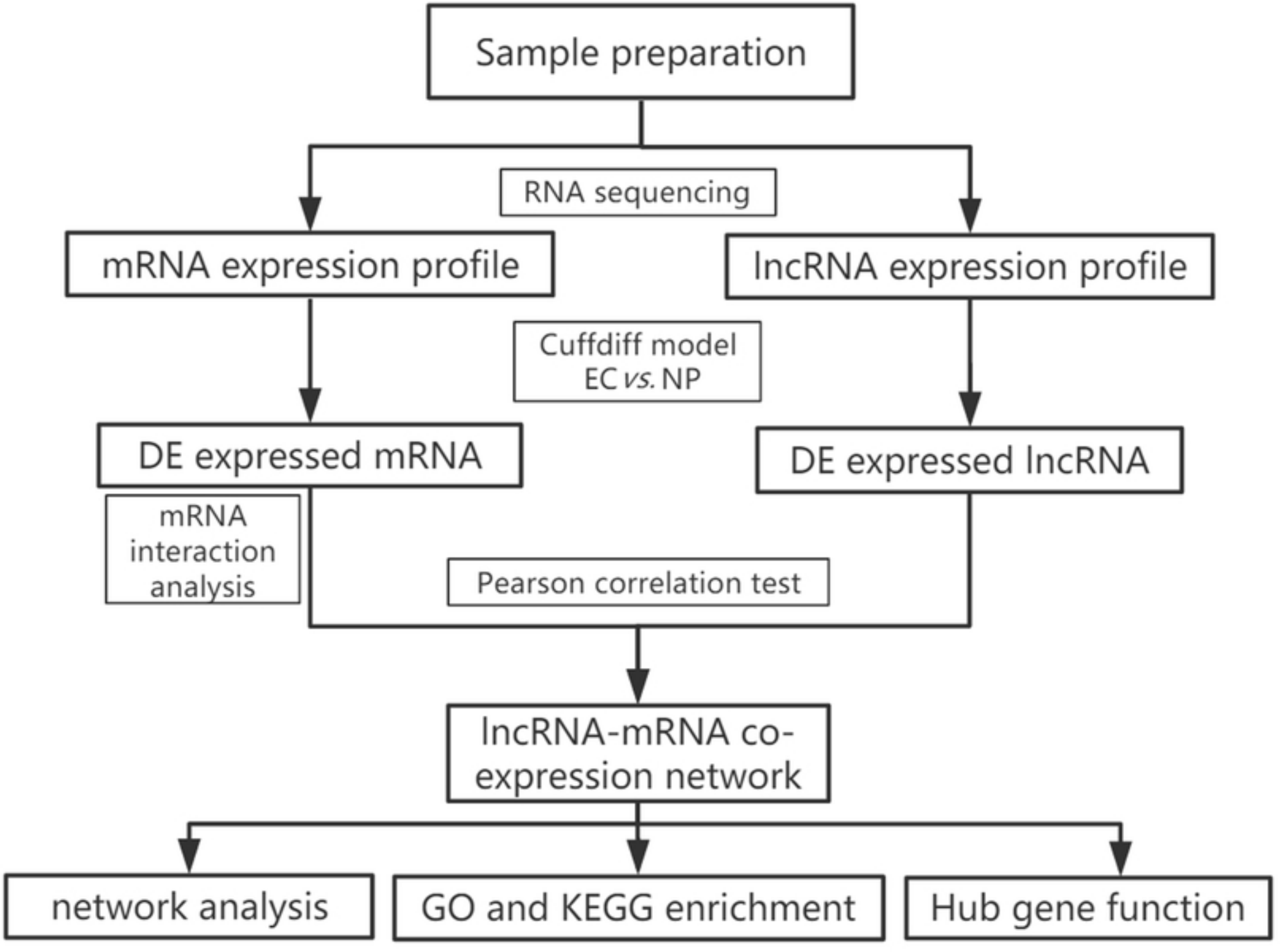
Hierarchical clustering analysis of sequencing detected mRNAs and lncRNAs with their expression abundance. Patient groups are on the horizontal axis, and mRNA and lncRNA genes are grouped along the vertical axis. Pinkish-purple bars indicate increased abundance of the corresponding genes, and blue indicates decreased abundance. White bars indicate that the corresponding genes were not detected.

Fig 3. Protein-protein interaction (PPI) network of differentially expressed (DE) mRNAs.

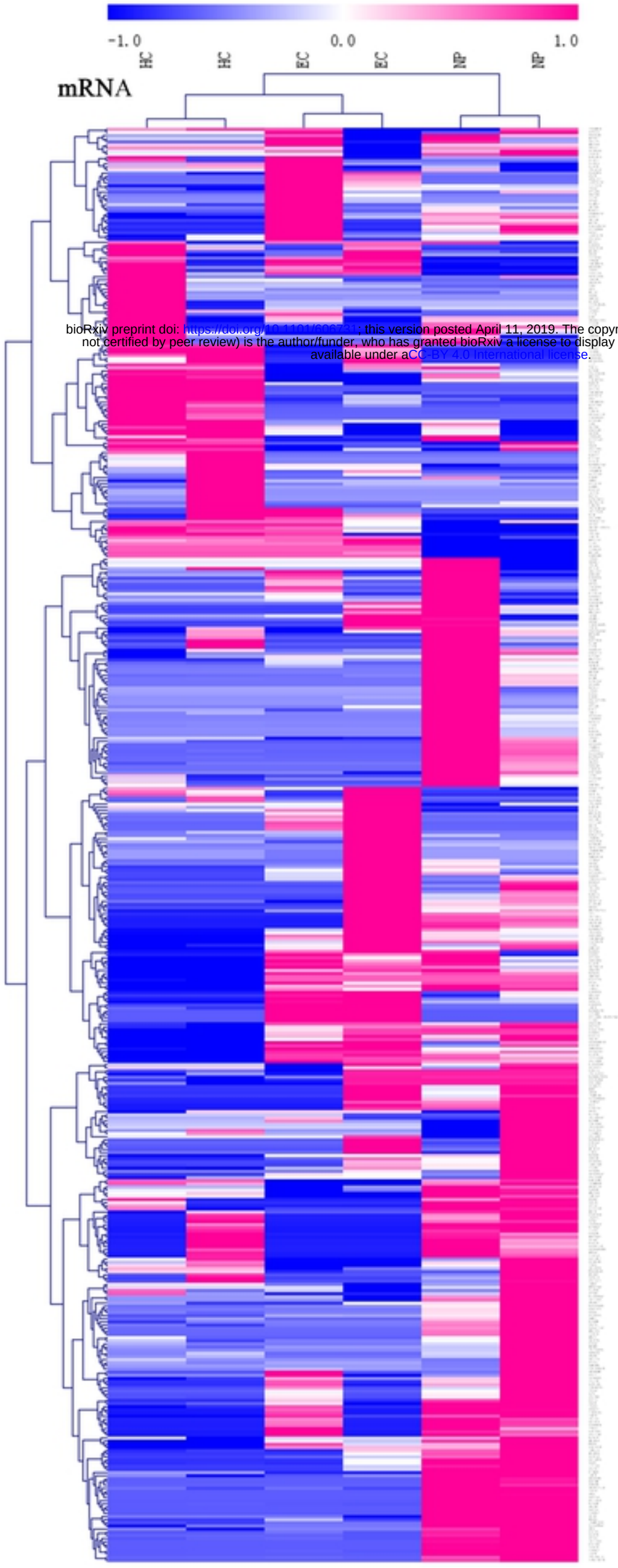
Top 50 hub mRNAs were included in using EcCentricity algorithm. The intensity of the color indicates the density of the correlation. The closer to red, the more central the gene is. The size of the circle represents the significance of a gene.

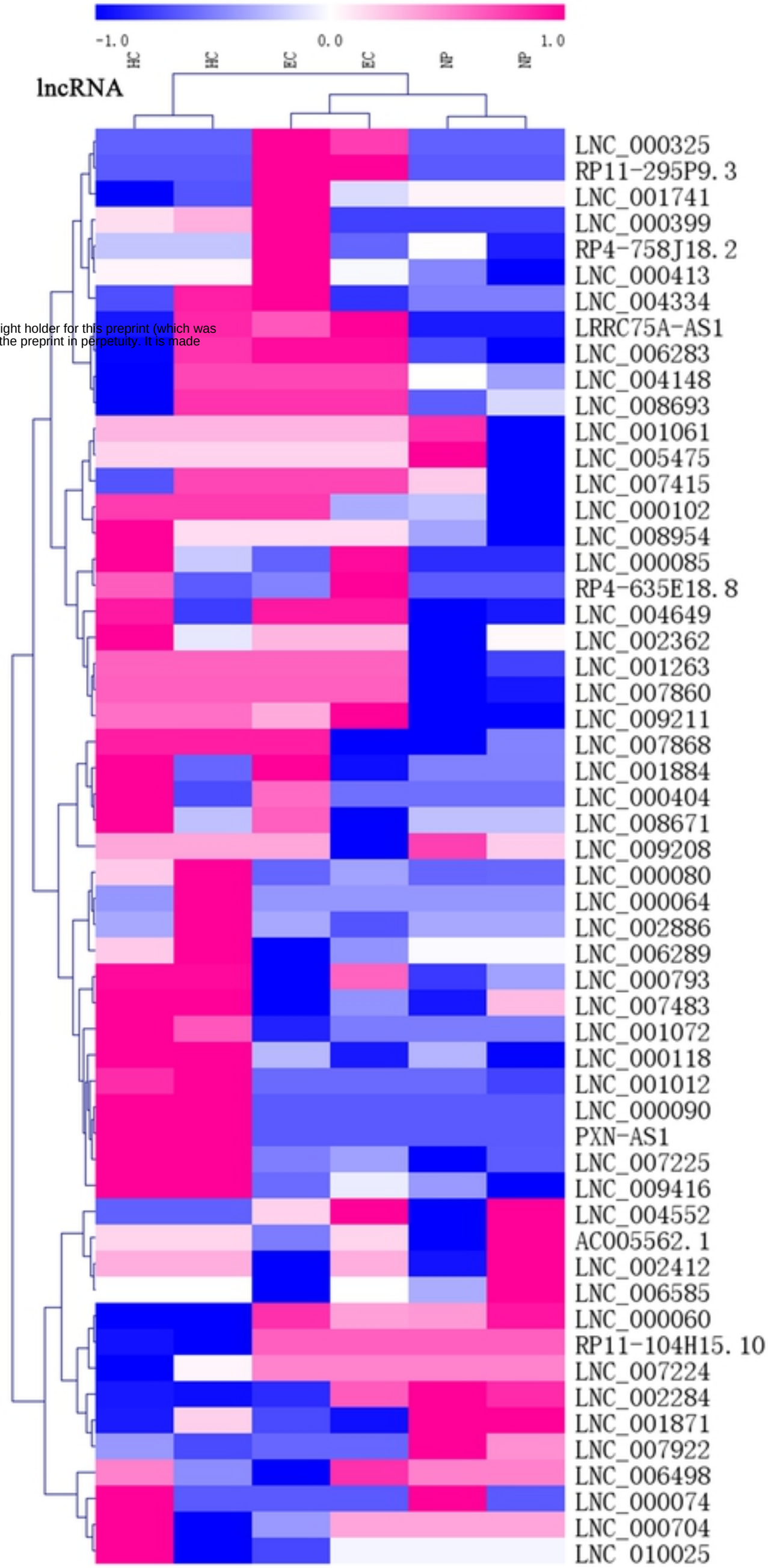
Fig 4. Long non-coding RNA (lncRNA)-mRNA co-expression network in elite controllers (ECs) and normal-process infected patients (NPs). The network was based on the Pearson relation coefficient of an lncRNA targeting an mRNA. Round shapes correspond to mRNAs, and squares to lncRNAs. The size of the shape is positively related to the P value. Red means upregulation and green means downregulation. A solid line indicates a positive correlation and a dotted line indicates a negative correlation. All hub-mRNAs and annotated lncRNAs are tagged with large labels. (a) All lncRNAs and co-expressed mRNAs. (b) Annotated lncRNAs and co-expressed mRNAs.

Fig 5. Gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment classification of the predicted targeting mRNAs of the long non-coding RNA (lncRNA)-mRNA network. The 10 most significantly enriched KEGG pathways and the 10 most significantly enriched GO terms in the biological process, cellular components, and molecular function categories.

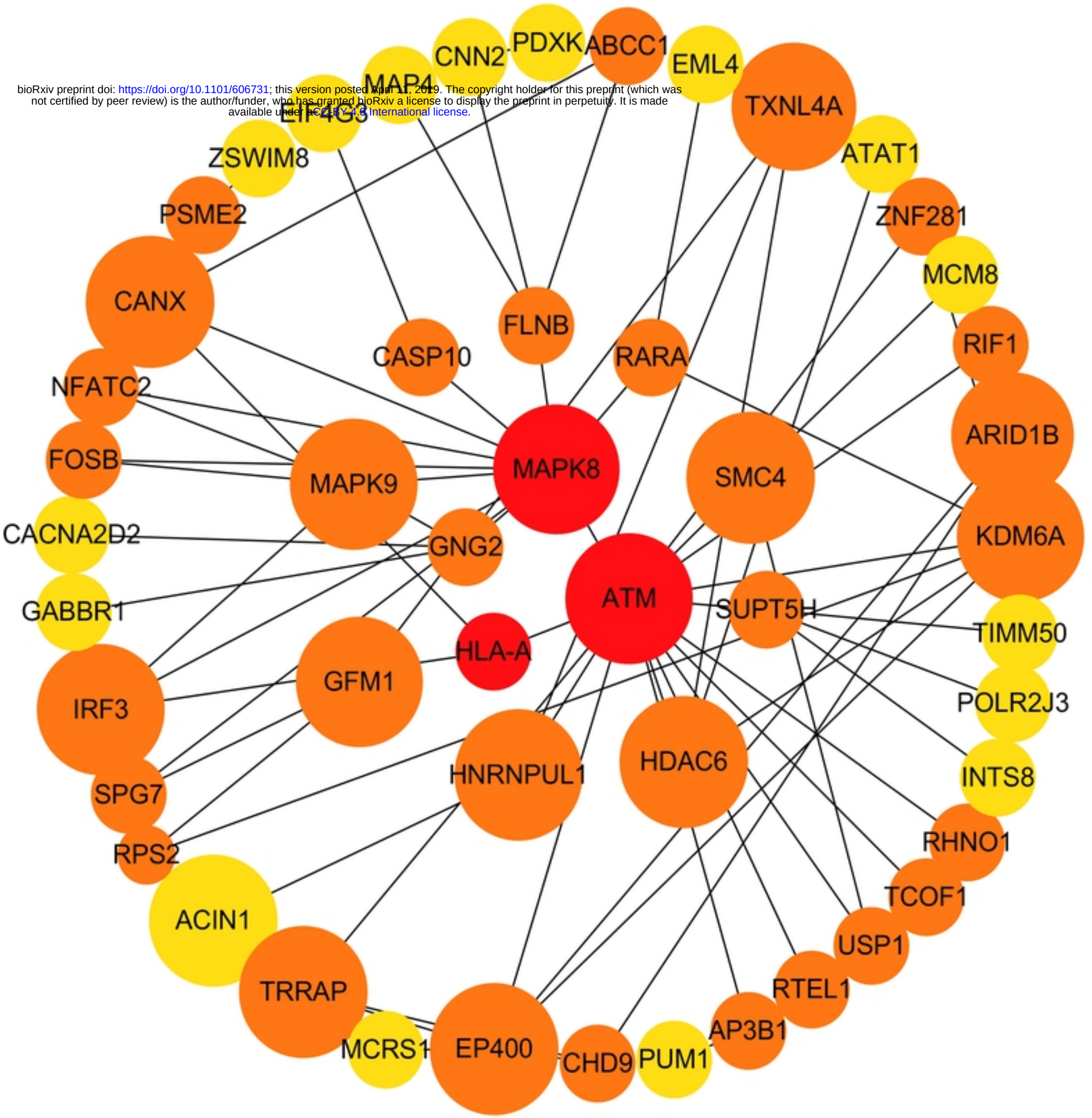

Sample preparation
RNA sequencing
mRNA expression profile
lncRNA expression profile
Cuffdiff model
EC vs. NP
DE expressed mRNA
DE expressed lncRNA
mRNA interaction analysis
Pearson correlation test
lncRNA-mRNA co-expression network
network analysis
GO and KEGG enrichment
Hub gene function

Figure



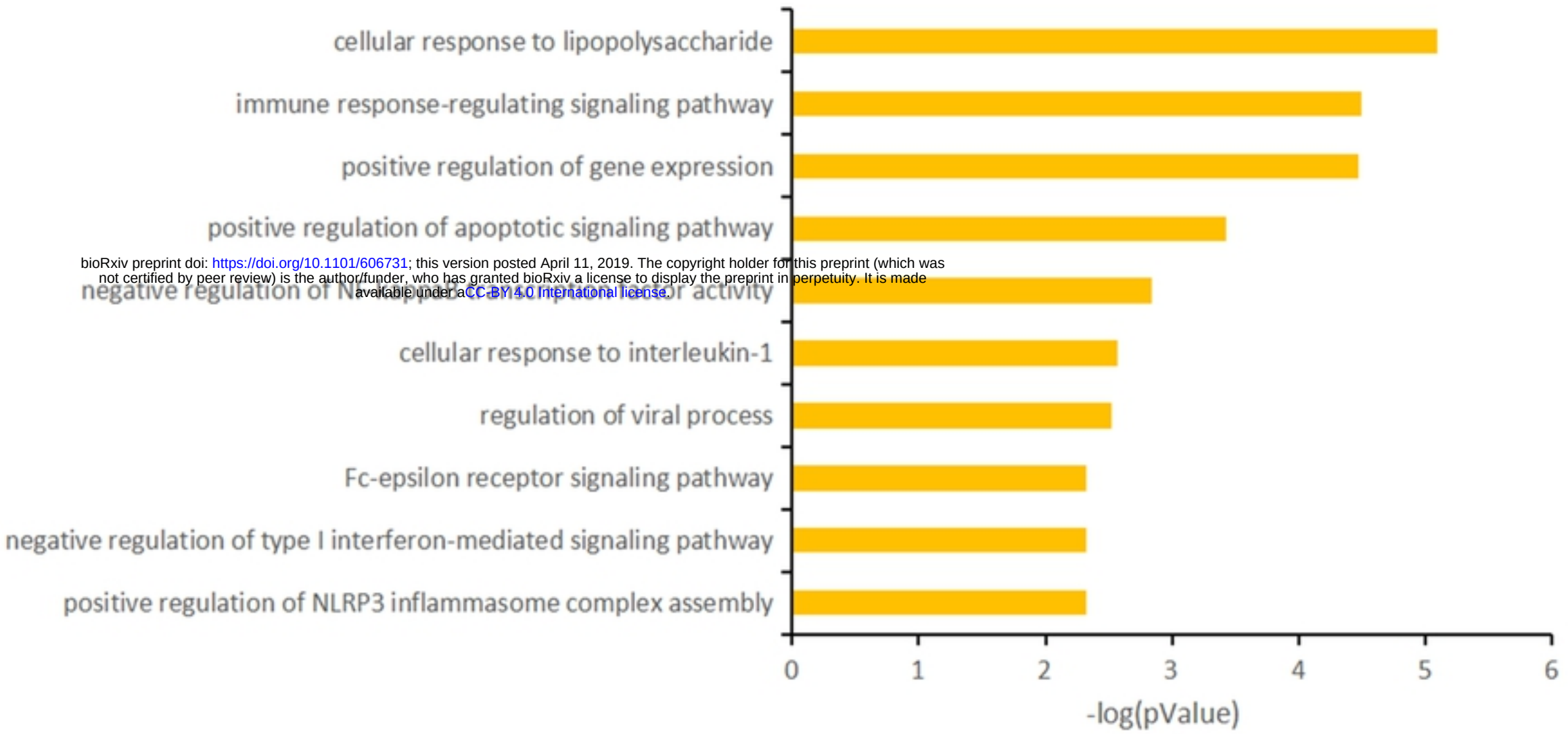


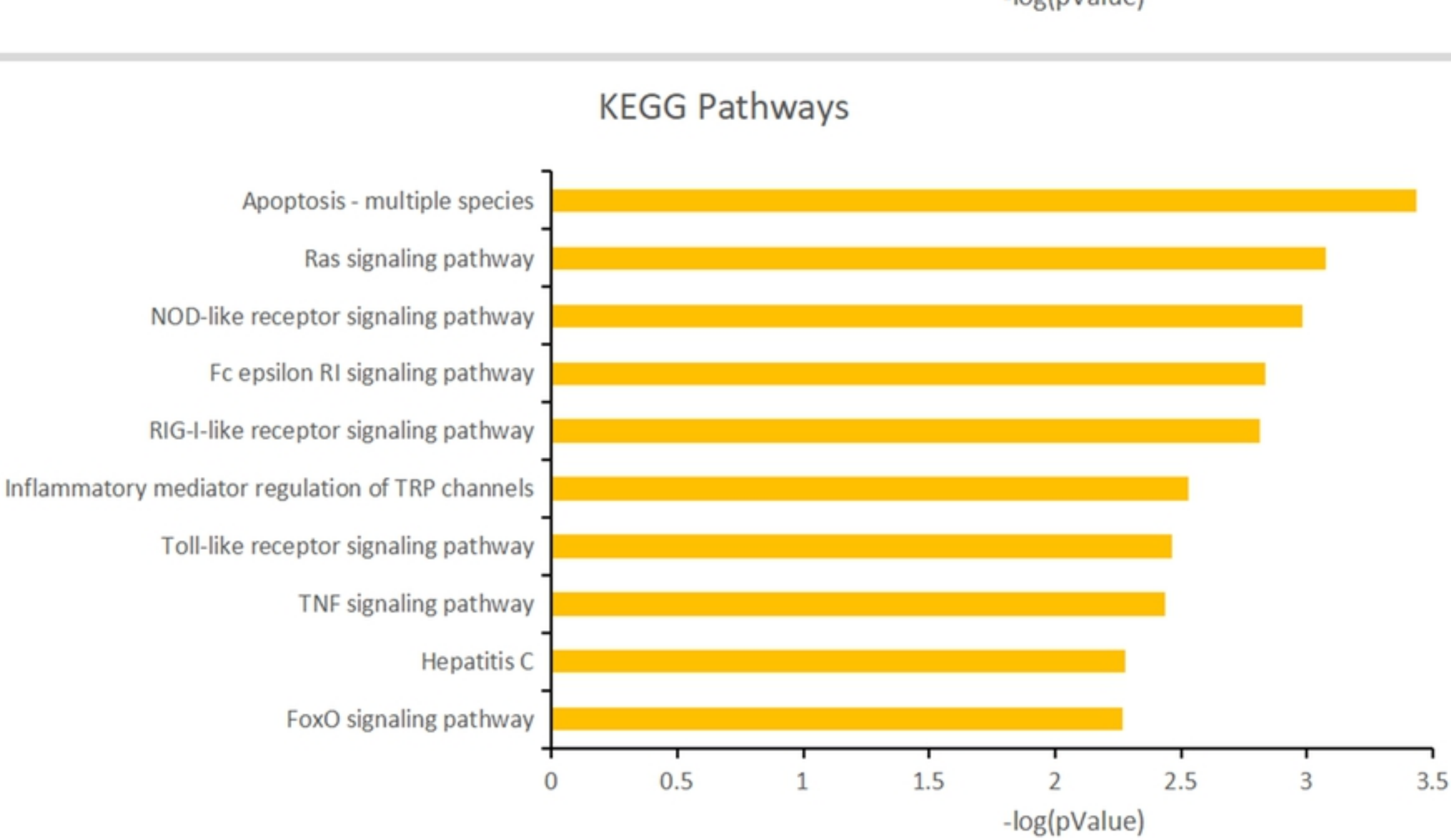
Figure



Figure

Figure





Figure