Integrated Multi-omics Analysis Identifies Molecular Signatures of Regular Physical Activity in Human Peripheral Blood

Xuemei Song[1,2]†, Jingzhi Lv[3]†, Siliang Ge[4,5]†, Sihan Xu[6], Yisha Wu[1,7], Yuhui Zheng[1,8], Wenwen Zhou[1,8], Yan Zhang[1], Jiaqiang Zhang[9,10], Pengbin Yin[4,5]*, Jianhua Yin[1,11,12]*, Chuanyu Liu[1,11,12,13]*

[1]State Key Laboratory of Genome and Multi-omics Technologies, BGI Research, Shenzhen 518083, China

[2]College of Life Sciences and Technology, Huazhong Agricultural University, Wuhan 430070, China

[3]State Key Laboratory of Molecular Oncology, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Peking University Cancer Hospital & Institute, Beijing 100021, China.

[4]Department of Orthopedics, Chinese PLA General Hospital, Beijing 100853, China.

[5]National Clinical Research Center for Orthopedics, Sports Medicine & Rehabilitation, Beijing 100853, China.

[6]Henan Institute of Medical and Pharmaceutical Sciences & BGI College, Zhengzhou University, Zhengzhou 450052, China

[7]Department of Otolaryngology, Head and Neck Surgery, Second Hospital, Shanxi Medical University, Taiyuan 030001, China

[8]College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

[9]Department of Anesthesiology and Perioperative Medicine, People's Hospital of Zhengzhou University, Henan Provincial People's Hospital, People's Hospital of Henan Academy of Innovations in Medical Science, Zhengzhou, China

[10]Institute of Electrophysiology, Henan Academy of Innovations in Medical Science, Zhengzhou, China

[11]Shenzhen Proof-of-Concept Center of Digital Cytopathology, BGI Research, Shenzhen, China

[12]Shanxi Medical University-BGI Collaborative Center for Future Medicine, Shanxi Medical University, Taiyuan 030001, China

[13]Zhangzhou Affiliated Hospital of Fujian Medical University, Zhangzhou, China

*Corresponding author. Email: yinpengbin@gmail.com (Y.P.); yinjianhua@genomics.cn (Y.J.); liuchuanyu@genomics.cn (C.L.);

†These authors contributed equally to this work.

Abstract

Regular physical activity is well established to protect against metabolic disorders and bolster immunity; yet, the underlying molecular and cellular mechanisms remain incompletely understood. Here, leveraging the Chinese Immune Multi-omics Atlas (CIMA), we integrated plasma metabolomics and lipidomics with single-cell transcriptomic and chromatin accessibility profiles (scRNA-seq and scATAC-seq) to decode the systemic impact of physical activity on human immunity and metabolism. Our data reveal that regular physical activity is linked to a coordinated metabolic signature marked by enhanced fatty acid oxidation and antioxidant defense. In circulating immune cells, it reprograms myeloid chromatin accessibility, activating antigen-presentation pathways while dampening inflammatory programs. Concurrently, effector functional programs in T cells and natural killer cells show extensive enhancement. Moreover, cellular communication analysis highlights a coordinated reinforcement of antigen presentation and chemotaxis signaling. Together, these findings elucidate molecular mechanisms underlying exercise-induced health benefits and offer a theoretical basis for enhancing public health and preventing chronic diseases.

Introduction

Regular physical activity is a powerful, modifiable driver of health, influencing chronic disease risk, cancer outcomes, infection resistance, aging processes, and lifespan[1–6]. Decades of mechanistic research have shown that exercise exerts its effects by inducing broad alterations in molecular signaling across multiple levels of the body, including specific tissues, the whole organism, and circulating body fluids[7–9]. However, most clinical benefits are observed in people who have been naturally and habitually active over many years, whereas much mechanistic evidence arises from short-term, tightly controlled interventions in selective populations or model organisms[10–13]. This mismatch not only obscures our understanding of the true molecular and cellular mechanisms that underlie the clinical benefits of exercise, but also creates gaps in external validity, and limits our ability to directly translate mechanical findings into realistic, population-level public health strategies. A key unmet need is comprehensive, multi-level comparison of habitually active and sedentary individuals, using deep cellular and molecular profiling to better understand how real-world regular exercise shape biology and health.

Notably, metabolism and immunity are widely regarded as two central pillars through which exercise exerts its health-promoting effects[14,15]. Metabolite and lipid profiling studies have identified candidate "exercise mimetics" or exercise-responsive metabolites, such as betaine[12], 12,13-dihydroxy-9Z-octadecenoic acid[16] (12,13-diHOME), β-hydroxybutyrate[17,18], and others, In parallel, flow cytometry–based and related immunophenotyping approaches have characterized how exercise influences immune cell populations, both in terms of their abundance, activation state, and effector functions[19–22]. Building on these important advances, there is now a growing need to move beyond isolated, bulk-level measurements toward a more integrated, systems-level understanding of exercise-induced remodeling of metabolic and immune networks. In particular, we still lack an integrated view of the heterogeneous responses within and across specific immune cell subsets, their dynamic interactions with one another, and how these processes are orchestrated through epigenetic regulation and metabolic reprogramming in response to exercise. Addressing these questions with coordinated multi-omics

85 approaches will be essential to construct a more holistic atlas of how exercise reshapes
86 immunometabolic circuits and, ultimately, to translate these insights into more precise and
87 scalable health interventions.

88 Here, we conducted a comprehensive multi-omics study leveraging data from the
89 Chinese Immune Multi-omics Cohort[23] (CIMA), a large-scale population-based cohort. By
90 integrating plasma targeted metabolomics and lipidomics with single-cell multi-omics
91 (scRNA-seq and scATAC-seq) of peripheral blood mononuclear cells (PBMCs)(Figure 1A),
92 we systematically mapped the molecular correlates of habitual physical activity in naturally
93 active individuals compared with sedentary controls. Using both modality-specific and
94 integrative analyses, we characterized exercise-associated alterations in circulating
95 metabolites and lipids, as well as immune cell type–resolved transcriptional and
96 epigenetic landscapes. Our data reveal that long-term regular exercise is linked to a
97 coordinated metabolic signature marked by enhanced fatty-acid oxidation and antioxidant
98 defenses, together with widespread remodeling of immune cell states mediated by
99 epigenetic reprogramming. Collectively, these results provide, to our knowledge, the first
100 multi-dimensional atlas of molecular adaptations to long-term spontaneous exercise,
101 spanning metabolism, chromatin regulation, and immunity. Beyond deepening
102 mechanistic insight into how physical activity promotes health, this work highlights
103 candidate biomarkers and regulatory pathways that may inform prevention and
104 management of chronic diseases, underscoring the translational relevance of our findings.
105

106 **Results**

107

108 **Physiological Adaptations Induced by EX**

109 This study is based on a secondary analysis of data from the Chinese Immune
110 Multi-omics Cohort (CIMA). We categorized 428 adults according to the World Health
111 Organization (WHO) physical activity guidelines[24]. The regularly exercising group (EX,
112 n=40) was defined as individuals who engaged in physical activity ≥3 times per week,
113 accumulating ≥150 minutes of moderate-intensity or ≥75 minutes of vigorous-intensity
114 activity. The sedentary control group (SED, n=46) was defined as individuals with no

exercise habits whatsoever whose daily life was predominantly sedentary.

The study population had a mean age of 28.2 years (95% CI: 27.2–29.2) and a mean body mass index of 21.3 kg/m² (SD ± 3.0) (Table S1). Analysis of baseline characteristics showed no significant differences between the two groups in age, diet, smoking, alcohol consumption, recent medication history, or sleep duration. However, the sex distribution differed markedly: SED was heavily skewed toward females, whereas EX had a relatively balanced sex ratio. This disparity led to a significantly higher unadjusted mean BMI in EX compared to SED (Figure S1A). Given that sex and BMI are key covariates for metabolic and immune phenotypes, all subsequent analyses employed multiple linear regression models adjusted for sex and BMI to independently assess the effects of exercise.

To validate the physiological relevance of the exercise-based classification, we first compared routine blood biochemical indices. Of the 19 indices tested, eight differed significantly between EX and SED (Figure 1B). These differences were specifically reflected in glucose and lipid metabolism indicators: compared to SED, EX showed significantly lower fasting blood glucose(GLU) and triglyceride(TG) levels, along with significantly higher high-density lipoprotein cholesterol(HDL-C) levels (Figure 1C). This finding is consistent with a substantial body of prior research concluding that exercise confers clear cardiometabolic benefits[25–27].

Beyond these classic signatures, EX showed significantly elevated levels of creatinine, uric acid, and bilirubin, all within normal clinical ranges, whereas key markers of hepatocellular injury and function did not differ between groups (Figures 1B, S1B, S1C). This profile suggests that the observed elevations reflect exercise-induced physiological adaptations, including increased muscle mass and metabolic turnover[28], as well as enhanced systemic antioxidant defense[29,30], rather than underlying organ pathology.

In summary, classification based on exercise habits effectively captured distinct physiological states. EX demonstrated both a healthier metabolic profile and features indicative of positive physiological adaptation, providing a phenotypic foundation for subsequent molecular investigations.

Multi-omics Profiling of Plasma and Circulating Immune Cells in EX

145 To systematically elucidate the regulatory effects of long-term regular exercise on
146 systemic metabolism and the molecular characteristics of immune cells, we conducted a
147 multi-omics analysis on 86 subjects (40 in EX and 46 in SED). This analysis integrated
148 targeted plasma metabolomics, lipidomics, single-cell RNA sequencing (scRNA-seq), and
149 single-cell assay for transposase-accessible chromatin sequencing (scATAC-seq) of
150 peripheral blood mononuclear cells (PBMCs) (Figure 1A). First, to dissect the impact of
151 exercise on systemic metabolism, we analyzed targeted metabolomics and lipidomics
152 data from all plasma samples. Following standardized pre-processing, we quantified 321
153 metabolites and 718 lipid molecules (complete list in Table S2). We observed extensive
154 differences in metabolite and lipid levels between EX and SED (Figure 1C and Table S2),
155 indicating that regular exercise significantly remodels the human metabolic network.

156 Subsequently, to investigate the impact of exercise on the immune system, we
157 proceeded with single-cell multi-omics analysis. Rigorous quality assessment confirmed
158 high technical reproducibility and the absence of significant batch effects (Figure S2A–C).
159 After quality control, 1,317,665 cells from scRNA-seq and 644,045 cells from scATAC-seq
160 were retained for downstream analysis (Figure 1B, 1C). Based on the scRNA-seq data,
161 we manually annotated cells using key lineage marker genes, identifying six major
162 immune cell populations: myeloid cells, B cells, $CD8^+$ T cells, $CD4^+$ T cells, NK cells,
163 and non-canonical T cells. These were further stratified into 34 distinct cell subsets (Figure
164 1D, 1E, S2E). For the scATAC-seq data, we employed a mapping annotation strategy
165 aligned with the scRNA-seq data, ultimately resolving 6 major classes and 30 cell subsets
166 (Figure 1F). To validate the reliability of these annotations, a marker gene heatmap based
167 on scATAC-seq gene activity scores demonstrated that chromatin accessibility in each
168 subset was highly specific and consistent with the transcriptomic data (Figure S2D).
169 Visualization of chromatin openness at key lineage gene promoter regions confirmed
170 significant accessibility at expected loci: B cells (MS4A1), T cells (IL7R), $CD4^+$ T cells
171 (CD4), $CD8^+$ T cells (CD8B), NK cells (KLRB1), and non-canonical T cells (RORC)
172 (Figure 1G, S2F). These chromatin accessibility patterns in known immune marker gene
173 promoters further verified the accuracy of our annotation. Furthermore, combining our
174 scRNA-seq data with MiloR[31] analysis revealed that the proportion of myeloid cells in EX

175 was significantly higher than in SED (Figure 1H-I), suggesting that exercise habits may
176 influence the cellular composition of the human immune system.
177
178 **Plasma Metabolic Changes in EX**
179 To decipher the impact of regular physical activity on systemic metabolic homeostasis, we
180 performed targeted metabolomics and lipidomics analyses on plasma samples from the
181 EX and SED groups. A total of 62 differentially abundant molecules were identified (Table
182 S2, Figure 2A). Notably, the majority of differential metabolites (86%) were upregulated in
183 EX, whereas the majority of differential lipids (94%) were downregulated, suggesting that
184 habitual exercise induces a global rewiring of the metabolic network.
185 Pathway enrichment analysis revealed that the upregulated metabolites were
186 significantly enriched in core energy metabolism pathways such as the tricarboxylic acid
187 cycle, as well as amino acid conversion pathways including phenylalanine/tyrosine
188 metabolism, indicating more active energy metabolism and amino acid turnover in the
189 exercising population (Figure 2B). Concurrently, pathways related to cofactor metabolism
190 and glutathione metabolism—a key antioxidant defense system—were also significantly
191 enriched, underscoring the role of redox homeostasis in exercise adaptation.
192 At the molecular level, we first focused on fatty acid oxidation (FAO), a central
193 energy-yielding pathway during exercise (Figure 2C). The significant upregulation of
194 O-adipoylcarnitine, a crucial carrier for mitochondrial long-chain fatty acid import, directly
195 reflects an increased flux of fatty acid transport[32]. Concurrently, elevated levels of
196 3-hydroxybutyrate, a hallmark of ketone body metabolism, indicate activation of hepatic
197 fatty acid β-oxidation and ketogenesis[33]. Furthermore, accumulation of acetic acid, the
198 end product of short-chain fatty acid oxidation, further supports a comprehensive
199 mobilization of FAO pathways across varying fatty acid chain lengths[34]. The coordinated
200 elevation of these three core metabolites demonstrates that regular exercise
201 systematically enhances the body's fatty acid oxidative capacity through multiple
202 complementary mechanisms.
203 Given that heightened energy metabolism is often accompanied by increased
204 reactive oxygen species (ROS) production[35], we further examined exercise-induced

205 modulation of the antioxidant defense system (Figure 2D). In concert with elevated energy
206 flux, several metabolites involved in redox homeostasis were synchronously upregulated
207 in EX. Betaine, an important methyl donor, showed significantly higher plasma levels in
208 exercisers, consistent with its recently identified role in mediating exercise-induced
209 anti-inflammatory and metabolic benefits[12]. The tryptophan derivative 3-hydroxyanthranilic
210 acid, which possesses free radical-scavenging activity[36], was also elevated, potentially
211 contributing to exercise-associated antioxidant adaptation. Moreover, increased levels of
212 2-hydroxybutyrate, an indirect marker of glutathione synthesis, suggest that exercise may
213 bolster endogenous antioxidant system function[37]. Notably, other metabolites with
214 immunomodulatory potential followed a similar trend, including xanthurenic acid (another
215 tryptophan-kynurenine pathway product) and the neuro-immunomodulator γ-aminobutyric
216 acid (GABA) (Table S2). These molecules can regulate the function and inflammatory
217 status of immune cells such as T cells and monocytes/macrophages through the aryl
218 hydrocarbon receptor or GABAergic signaling pathways[38,39], implying that exercise may
219 co-ordinate antioxidant and immune homeostasis via metabolic reprogramming.

220 Furthermore, exercise-associated changes extended to the host–microbiota interface
221 (Figure 2E). Levels of microbially derived aromatic amino acid metabolites—phenyllactic
222 acid and indoleacrylic acid—were significantly elevated in EX; the latter has been shown
223 to enhance intestinal barrier function and modulate innate immunity, thereby influencing
224 systemic inflammation[40]. In parallel, levels of two secondary bile acids, glycodeoxycholic
225 acid and glycochenodeoxycholic acid, were reduced. Bile acids are important signaling
226 molecules that regulate intestinal and hepatic immunity[41], and their altered composition
227 may influence local and systemic immune activity.

228 At the lipid level, a broad downregulation was observed in EX (Figure 2F-G).
229 Triacylglycerols, the primary energy-storage form, exhibited widespread reduction,
230 consistent with the well-established effect of exercise-promoted lipolysis and fatty acid
231 utilization[42]. Multiple membrane phospholipid subclasses (e.g., LPE, PC, PG, PI) also
232 showed decreased abundance, reflecting extensive remodeling of the lipid metabolic
233 network. However, finer-grained comparison revealed molecule-specific regulation; for
234 instance, distinct species within the same phosphatidylethanolamine subclass displayed

235 opposite abundance changes (Figure 2H). Overall, these lipid alterations, together with
236 enhanced fatty acid oxidative capacity, constitute a characteristic metabolic adaptation to
237 exercise.
238
239 **Multi-omic Characterization of Myeloid Cells Heterogeneity in EX**
240 Myeloid cells, especially monocytes and macrophages, are recognized as key cell types
241 that detect metabolic signals and regulate systemic inflammation and immune
242 homeostasis[43]. In conditions of obesity and other metabolic disorders, the chronic
243 low-grade inflammation driven by these cells is a key driver of pathological processes,
244 including insulin resistance[44]. Classic studies have demonstrated that acute or short-term
245 exercise can rapidly mobilize neutrophils and monocytes[45]. However, Whether and how
246 regular physical activity remodels the composition, functional states, and epigenetic
247 landscape of myeloid cell subsets, thereby modulating systemic immune homeostasis,
248 remains to be systematically elucidated.We first performed fine-grained clustering of
249 myeloid cells. Based on classical lineage marker genes, cells were manually annotated
250 into five functional subsets: classical monocytes (cMono), intermediate monocytes
251 (intMono), non-classical monocytes (ncMono), classical dendritic cells (cDC), and
252 plasmacytoid dendritic cells (pDC) (Figure 3A, S3A, S3F). Although conventional
253 abundance analysis revealed no significant inter-group differences at the population level
254 (Figure S4A), differential abundance testing using the MiloR algorithm at the local
255 neighborhood level revealed significant increases in intMono, cMono, and ncMono in EX,
256 with median $\log_2$ fold changes of +0.37, +0.31, and +0.25, respectively (Figure 3B). This
257 suggests that long-term regular physical activity specifically remodels peripheral myeloid
258 cells, with a pronounced effect on the composition of monocyte subsets.
259 To uncover the functional reprogramming underlying these compositional changes,
260 we performed differential expression gene (DEG) analysis (Table S3) and pathway
261 enrichment analysis across subsets. Analysis revealed that in cMono, ncMono, cDC, and
262 pDC from the EX, upregulated pathways were consistently enriched in antigen processing
263 and presentation, cell adhesion, and chemokine signaling, whereas downregulated
264 pathways were concentrated in oxidative stress and hypoxia responses (Figure 3C).

265 Pathway scoring analysis further confirmed enhanced antigen presentation-related
266 signals and a general attenuation of stress-related pathways across multiple subsets in
267 EX (Figure S4B, S4C). At the molecular level, the expression of core antigen presentation
268 genes (e.g., MHC class II molecules) was elevated, alongside increased expression of
269 migration-associated receptors such as *CX3CR1* in cDC and *CXCR3* in pDC (Figure 3D),
270 suggesting a concurrent enhancement of myeloid cell immune surveillance and
271 tissue-homing potential.

272 To investigate the epigenetic mechanisms underpinning these functional changes, we
273 performed an in-depth analysis of chromatin accessibility using scATAC-seq. Analysis of
274 differentially accessible regions (DARs) between the two groups (Table S4) revealed that
275 94.3% of DARs across all myeloid subsets were specifically enriched in cMono (Figure
276 S4D); therefore, subsequent investigations focused on this subset. Gene Ontology (GO)
277 enrichment analysis of DARs in cMono showed that regions with increased accessibility
278 were enriched for pathways such as antigen presentation and T cell activation, while
279 regions with decreased accessibility were associated with stress response pathways
280 (Figure 3E). Transcription factor binding motif enrichment analysis identified significant
281 enrichment for motifs of KLF2, KLF4, SP1, SPI1 (PU.1), IRF1, and STAT2 (Figure 3F).
282 Further motif deviation analysis confirmed that the accessibility of binding sites for these
283 transcription factors was significantly upregulated in EX, whereas accessibility for motifs
284 of AP-1 family members (e.g., JUN, FOS) and BACH1 was downregulated (Figure 3G).
285 Notably, IRF1 regulates the expression of CTSS, which participates in MHC class II
286 antigen presentation[46]. SPI1, a master regulator of myeloid differentiation, can synergize
287 with IRF4 to drive CIITA expression, thereby activating the antigen presentation pathway[47].
288 Upregulated STAT2 activity suggests cells may be in an “anti-viral pre-activated” state.
289 KLF2 and KLF4, important regulators of immune homeostasis highly expressed in
290 monocytes and macrophages, can inhibit pro-inflammatory activation and promote cellular
291 residency and migration[48,49]. Conversely, transcription factors with significantly
292 downregulated binding site accessibility, such as AP-1 and BACH1, are typically
293 associated with acute/chronic stress, pro-inflammatory responses, and oxidative stress
294 responses[50,51].

To elucidate how exercise enhances antigen presentation function at the epigenetic level, we focused on genes linked to differentially accessible peaks. The results showed that the promoter regions of multiple key antigen presentation genes (e.g., *HLA-DQA1*, *HLA-DQB1*) exhibited significantly enhanced chromatin accessibility in EX (Figure 3H), consistent with their upregulated mRNA expression (Figure 3D). Concurrently, the chromatin accessibility of the promoter region of the pro-inflammatory gene *CXCL8* was significantly reduced in EX (Figure 3I), with a corresponding downregulation in its mRNA expression (Figure 3J). Furthermore, other inflammation-related genes such as *CCL2* and *CCL7* showed a consistent trend of decreased chromatin accessibility (Figure S4E), further indicating that physical activity may suppress the inflammatory propensity of myeloid cells through epigenetic regulation.To validate these findings at the single-cell transcriptomic level, we assessed transcription factor activity in cMono cells using the SCENIC algorithm. The results demonstrated that the activities of transcription factors such as SPI1, STAT1, IRF1, and KLF4 were significantly higher in EX than in SED (Figure 3J), highly consistent with the scATAC-seq data indicating upregulated accessibility of the corresponding motifs (Figure 3F and 3G). To deeply analyze how these key transcription factors synergistically regulate the functional state of cMono cells, we constructed their core regulatory network (Figure 3K). Network analysis revealed that SPI1, IRF1, KLF4, and STAT1 form a synergistic regulatory module, targeting immune-related genes across multiple functional modules, including antigen processing and presentation (e.g., *HLA-DQA1, HLA-DQB1, CD74*), inflammation and cytokine signaling (e.g., *IL1B, IL10*), and chemotaxis and cell migration (e.g., *CCL20, CCL2, ITGB1*).In summary, regularly physical activity remodels the epigenetic landscape and transcriptional regulatory network of cMono cells, synergistically enhancing antigen presentation function while inhibiting inflammatory responses, thereby shaping a “high-function, low-inflammation” functional phenotype in these cells.

Multi-omic Characterization of T Cell Heterogeneity in EX

Although substantial evidence suggests that regular exercise remodels T cell composition and enhances their function[52], how exercise drives these functional changes via

325 epigenetic reprogramming at the single-cell level remains to be systematically elucidated.
326 Using classical marker genes, we identified 17 T cell subsets, including ten $CD8^+$ T cell
327 subsets and several non-canonical T cell subsets: $CD8^+$ naive T cells (CD8_Tn), central
328 memory T cells (CD8_Tcm), effector memory T cells (CD8_Tem), cytotoxic T cells
329 (CD8_CTL), MAIT cells, γδT1 cells, γδT2 cells, proliferating T cells, and pre-T-like cells
330 (Figure 4A, S3B, and S3G). Additionally, we identified seven $CD4^+$ T cell subsets (Figure
331 S5A, S3C, and S3H). MiloR differential abundance analysis revealed that the proportion
332 of CD4_CTL cells was significantly lower in the exercise group (median log_2FC = -0.706)
333 (Figure S5B), while the proportion of pre-T-like cells was significantly increased (median
334 log_2FC = +0.315) (Figure 4B).

335 To dissect the impact of exercise on the functional state of T cells from an epigenetic
336 perspective, we performed DAR analysis across all subsets (Table S4). Quantitative
337 analysis of upregulated DARs indicated that the most significant changes occurred in
338 CD8_CTL, MAIT, CD8_Tn, NKT, CD8_Tem, and γδT2 cells. Notably, the number of
339 upregulated DARs in CD8_CTLs was significantly higher than in other subsets (Figure
340 4C), with multiple DARs shared between two or more subsets. GO enrichment analysis
341 showed that genes associated with these upregulated DARs were significantly enriched in
342 pathways related to $CD8^+$ T cell effector functions, such as cell killing and defense
343 against infection (Figure 4D). Feature gene set scoring of the transcriptomic data further
344 supported these findings. In the exercise group, scores for gene modules related to
345 "cytotoxicity," "type II interferon response," and "oxidative phosphorylation" were
346 significantly upregulated across multiple T cell subsets (Figure 4E and S5C). Conversely,
347 scores for hypoxia response and inflammatory response gene modules were markedly
348 downregulated (Figure S5D). This suggests that regular exercise may not only enhance
349 the effector functions and metabolic adaptations of $CD8^+$ T cells but also attenuate their
350 inflammatory signaling.

351 Transcription factors (TFs) serve as core hubs regulating immune cell effector
352 functions. To resolve the transcriptional regulatory basis of these functional changes, we
353 performed motif bias analysis based on scATAC-seq data. The results showed that in
354 CD8_CTLs, which possessed the highest number of upregulated DARs, the accessibility

355 of binding motifs for TFs closely related to cytotoxic differentiation and effector
356 function—including EOMES, RUNX3, RUNX1, and TBX21—was significantly upregulated.
357 Among these, TBX21 (T-bet) and EOMES are master regulators driving $CD8^+$ T cell
358 differentiation into effector cells, and their upregulation is directly linked to enhanced
359 cytotoxic function[53,54]. Similarly, in MAIT cells, motif accessibility for key TFs involved in
360 activation, cytokine response, and functional regulation (e.g., RUNX3, RORC, STAT1, and
361 STAT4) was also significantly upregulated (Figure 4F and S5E). STAT1 and STAT4 are
362 core mediators of the interferon-γ signaling pathway[55], their upregulated activity
363 corresponds with the enhanced IFN-γ response pathway described above, providing a
364 transcriptional mechanism for the improved effector function of MAIT cells. Furthermore,
365 SCENIC analysis based on scRNA-seq data demonstrated that the regulatory activities of
366 key TFs such as RUNX3 and TBX21 were significantly elevated in CD8_CTL cells in the
367 exercise group (Figure 4G), consistent with the epigenetic findings. To further unveil the
368 downstream regulatory networks of these key TFs, we constructed a target gene
369 regulatory network for RUNX3 (Figure 4H). The analysis indicated that RUNX3 targets
370 and regulates genes such as GZMB and TBX21, which collectively drive the
371 enhancement of CD8_CTL effector functions. Further analysis revealed that in CD8_CTL
372 cells, the chromatin accessibility of distal regulatory regions for the *TBX21* and *GZMB* loci
373 was significantly enhanced in the exercise group (Figure 4I); in MAIT cells, the promoter
374 region of the core effector gene *PRF1* exhibited a similar trend of increased accessibility
375 (Figure 4J). Collectively, both single-cell transcriptomic and ATAC-seq results suggest that
376 regular exercise enhances T cell effector functions.
377
378 **Multi-omic Characterization of NK and B Cell Heterogeneity in EX**
379 Extensive research demonstrates that regular aerobic exercise effectively bolsters
380 immune function and defense against infection[56]. As core effector cells of the innate
381 immune system, natural killer (NK) cells not only serve as the first line of defense against
382 viral infections and in tumor immunosurveillance but also play a pivotal role in bridging
383 innate and adaptive immune responses[57]. However, the specific molecular mechanisms
384 by which exercise enhances NK cell effector function remain incompletely understood. We

385 first classified NK cells into six subsets based on classical marker genes: Mature NK cells,
386 Transitional NKs, type 2 innate lymphoid cells (ILC2), $CD56^+$ NK cells, Terminal NK cells,
387 and Cycling NK cells (Figure 5A, S3D, and S3I). Notably, the proportion of $CD56^+$ NK
388 cells was significantly reduced in the exercise group (Figure S6A), a finding consistent
389 with MiloR algorithm results (Figure 5B). To evaluate functional changes in NK cells, we
390 identified differentially expressed genes (DEGs) (Table S3) and performed pathway
391 scoring analysis (Figure 5C and Figure S6B). The results showed elevated scores for
392 killing function-related pathways across multiple NK cell subsets in the exercise group,
393 while scores for inflammation, oxidative stress, and apoptosis signaling pathways were
394 generally reduced.Epigenetically, we identified differentially accessible regions (DARs)
395 (Table S4), with Mature NK cells exhibiting the highest number of upregulated DARs
396 (Figure S6C). Gene Ontology (GO) enrichment analysis of these upregulated DARs
397 further revealed that pathways related to "immune defense and killing," "immune
398 regulation," and "immune cell mobilization" were significantly enriched in Mature NK cells,
399 whereas inflammation and stress-related pathways were downregulated (Figure S6D),
400 consistent with the transcriptomic findings. To deeply investigate the transcriptional
401 regulatory basis, we performed motif bias analysis based on scATAC-seq data (Figure
402 S6E and 5E). Results showed that in the exercise group, motif accessibility was
403 significantly enhanced for TFs closely related to NK cell effector function, including key
404 regulators driving cytotoxic differentiation such as RUNX3, EOMES, and TBX21[58] (T-bet),
405 as well as ETS1, which is involved in innate lymphoid cell development and activation[59].
406 Conversely, motif accessibility for inflammation and stress-related TFs (RELA, RELB,
407 NFKB1, JUND, FOS, FOSL2, BACH1)[60] was generally reduced. Footprint analysis
408 confirmed the decreased accessibility of RELA and RELB transcription factor binding sites
409 (Figure 5F).

410 To investigate whether common regulatory mechanisms exist for the cytotoxic
411 functions of adaptive and innate immunity under exercise, we compared transcription
412 factor activities between CD8_CTL (adaptive immunity) and Mature_NK (innate immunity)
413 cells (Figure 5G). Venn diagram analysis revealed that both shared upregulation of a
414 series of core transcription factors, including RUNX3, EOMES, TBX21, ETS2, and

415 FOXO3, with FOXO3 involved in maintaining T cell memory and NK cell homeostasis[61].
416 This provides a unified transcriptional regulatory basis for the synergistic enhancement of
417 cytotoxic functions across different immune branches by exercise. Concurrently, each
418 population displayed distinct activation patterns: CD8_CTLs specifically upregulated
419 RUNX1, STAT6, and BCL11B, which are associated with T cell lineage commitment[62–64];
420 whereas Mature_NK cells uniquely activated factors such as ETS1, IRF1, STAT1, and
421 STAT2, which are closely related to interferon signaling responses and rapid NK effector
422 functions[65–67]. Finally, through track analysis of specific loci, we provided direct evidence
423 for the epigenetic regulation of key genes. In Mature_NK cells, chromatin accessibility
424 was significantly enhanced in the exercise group at the intronic regulatory region of *STAT1*,
425 a core mediator of type I interferon signaling, and the promoter region of *CXCR4*, a known
426 exercise-responsive factor[68] (Figure 5H). Furthermore, accessibility was significantly
427 reduced at the promoter regions of *RELB*, associated with non-canonical NF-κB pathway
428 activation, and *CCL5*, an important inflammatory chemokine[69] (Figure 5J). In summary,
429 natural killer cells in individuals who exercise regularly exhibit higher effector potential and
430 lower inflammatory characteristics.

431 We also investigated changes in B cells in the regular exercise population, given their
432 crucial role in humoral immunity. B cells were annotated into eight subtypes (Figure S7A,
433 S3E, and S3J): Switched memory B cells (Switched_Bm), Naive B cells, pre-Switched
434 memory B cells, Unswitched memory B cells (Unswitched_Bm), Atypical memory B cells
435 (Atypical_Bm), Plasma cells, and Plasmablast cells. No significant differences in cell
436 proportions were observed across all subsets (Figure S7B). We identified DEGs (Table S3)
437 and performed enrichment analysis, which showed upregulation of pathways related to
438 antigen processing and presentation, humoral immune response, and chemotaxis across
439 multiple B cell subsets (Figure S7C). Notably, the expression of key genes encoding MHC
440 Class II molecules was consistently upregulated in multiple B cell subsets, alongside
441 increased expression of inflammation-regulatory genes and certain chemokine genes
442 (Figure S7D). To resolve the epigenetic regulatory basis, we performed differential
443 accessibility region analysis (Table S4). Results showed that changes in chromatin
444 accessibility were highly concentrated in Naive B cells and Switched Bm cells (Figure

445 S7E). GO enrichment analysis of upregulated DARs indicated that Naive B cells
446 upregulated pathways such as T cell costimulation, antigen presentation, and apoptosis
447 inhibition, while Switched_Bm cells upregulated pathways including lymphocyte activation
448 and antigen presentation (Figure S7F).

449 To further determine the molecular mechanisms behind these B cell functional
450 changes, we identified upregulated TFs in these two subsets. Naive B cells showed
451 upregulation of BCL11A, RUNX1, SPI1 (PU.1), and EBF1 (Figure S7G), with SPI1 and
452 EBF1 being master regulators of early B cell development and identity[70,71]. Switched Bm
453 cells showed upregulation of NF-κB family members, IRF4, as well as SPI1 and EBF1
454 (Figure S7H); NF-κB and IRF4 are key drivers of memory B cell formation, class switching,
455 and plasma cell differentiation[72,73]. We focused our analysis on the regulatory regions of
456 genes upregulated in the transcriptome. The results showed that in the exercise group,
457 chromatin accessibility at the promoter region of the key antigen presentation gene
458 *HLA-DQB1* was significantly enhanced in both Naive B cells (Figure S7I) and
459 class-switched memory B cells (Figure S7J). Additionally, in Switched Bm, promoter
460 accessibility of the chemokine gene *CCL5* also showed a synchronous increase (Figure
461 S7K). Overall, this study reveals that while regular exercise does not significantly affect B
462 cell subset proportions, it may enhance their potential for antigen presentation and
463 immune activation through chromatin reprogramming and transcriptional changes in key
464 subsets.

465

466 **Altered Intercellular Communication**

467 Immune cells rely on ligand-receptor interactions for communication to coordinate
468 immune responses[74]. However, the impact of exercise on this interaction network remains
469 unclear. To address this, we employed the CellChat algorithm[75]. The results indicated a
470 significant increase in the number of interactions involving $CD8^+$ T cells in the exercise
471 group, alongside a general strengthening of interaction intensity among B cells, $CD8^+$ T
472 cells, and myeloid cells (Figure 6A–B). Comparison at the pathway level revealed that
473 information flow for inflammation and immune suppression-related pathways (THBS,
474 RESISTIN, IL6) was higher in the sedentary group, whereas signals for chemotaxis

475 migration, antigen presentation, and cell adhesion pathways (CCL, MHC-II, PECAM1)
476 were significantly enhanced in the exercise group (Figure 6C).

477 Human resistin is a key regulator directing inflammatory responses in various
478 immune cells, including macrophages and peripheral blood mononuclear cells[76]. The cell
479 communication network showed that myeloid cells were the core signal source for this
480 pathway, with classical monocytes being the primary senders, and their signal output was
481 lower in the exercise group (Figure 6D and S8B). The interaction intensity of key
482 ligand-receptor axes in this pathway, *RETN*–CAP1 and *RETN*–TLR4, was weakened in
483 the exercise group, with a synchronous downregulation of *RETN* gene expression in
484 myeloid cells (Figure 6E–F). Given that the pro-inflammatory effects of Resistin rely
485 partially on NF-κB pathway activation[77], to explore changes in the downstream
486 inflammatory network, we analyzed pathway activity scores in myeloid cell subsets. The
487 results showed that in classical monocytes, multiple core inflammation-related pathways,
488 including TNFA_SIGNALING_VIA_NFKB, INFLAMMATORY_RESPONSE, and
489 IL6_JAK_STAT3_SIGNALING, were significantly suppressed in the exercise group
490 (Figure S8A).

491 Regarding adaptive immune regulation, antigen presentation-related signals were
492 significantly enhanced in the exercise group. The MHC-II signaling pathway between B
493 cells/myeloid cells and $CD4^+$ T cells was elevated in the exercise group, involving
494 multiple ligand-receptor pairs (Figure 6G–H). Subset analysis revealed that interactions
495 between classical monocytes, naive B cells, and memory B cells with naive and helper
496 $CD4^+$ T cells were particularly enhanced in the exercise group (Figure S8C).
497 Furthermore, MHC-I signaling pathways between multiple cell subsets and $CD8^+$ T cells
498 were also activated in the exercise group, especially the enhancement of the
499 *HLA-G*–CD8A/B axis (Figure 6I), with interactions between classical monocytes and naive
500 $CD8^+$ T cells contributing most significantly (Figure S8D). IFN-γ, a member of the type II
501 interferons, plays a critical role in host defense and immune surveillance . In EX, IFN-γ
502 signal connections significantly strengthened with NKT cells acting as the source and
503 multiple myeloid and B cell subsets as targets (Figure 6J). Concurrently, interactions
504 mediated by the *IFNG*–(IFNGR1+IFNGR2) axis, with cytotoxic T cells ($CD4^+$ CTL and

505 $CD8^+$ CTL) as sources and myeloid and B cell subsets as targets, also exhibited specific
506 enhancement (Figure 6K).
507 Notably, exercise also potentiated signaling networks related to tissue repair and
508 immune microenvironment regulation. On one hand, the COLLAGEN signaling pathway
509 was dominated by non-canonical T cells (Unconv_T cells) in the exercise group, with
510 upregulation of both the interaction of the key ligand-receptor axis *COL6A2*–CD44 and
511 ligand gene expression (Figure S8E–G). Collagen signaling is known to participate in cell
512 adhesion, migration, and tissue homeostasis maintenance[78]. On the other hand, the
513 chemokine CCL signaling pathway was also significantly enhanced in the exercise group,
514 primarily manifested as increased signaling from multiple immune cell subsets to myeloid
515 cells (Figure S8H-I). CCL chemokines are known to play a core role in immune cell
516 recruitment and tissue localization[79]. These results suggest that exercise may enhance
517 adaptive immunity and tissue repair-related signal transmission while specifically inhibiting
518 myeloid cell-mediated inflammatory pathway activity by influencing ligand-receptor
519 communication networks among immune cells.
520
521 **Discussion**
522 Extensive research indicates that exercise exerts its effects by inducing widespread
523 alterations in molecular signaling across multiple levels of the body, including specific
524 tissues, the entire organism, and circulating fluids. Regular physical activity serves as an
525 effective means to improve metabolic health, delay aging, reduce chronic disease risk,
526 and enhance immune function[1–6]. However, most mechanistic studies originate from
527 short-term, rigorous exercise interventions or animal models, potentially differing from the
528 molecular profiles observed in populations that spontaneously maintain long-term
529 exercise habits in real-world settings[10–13]. This discrepancy limits our understanding of the
530 true molecular and cellular mechanisms underlying exercise's clinical benefits and
531 impedes the effective translation of basic research findings into public health strategies.
532 Therefore, conducting comprehensive, multi-level comparisons between individuals who
533 engage in regular physical activity and sedentary populations is crucial for understanding
534 the molecular mechanisms through which exercise reshapes physiology and health in the

real world. Adopting a multi-omics strategy, this study performed a multidimensional analysis of systemic metabolism and single-cell immunity, thereby delineating a molecular panorama of the coordinated adaptations that drive the metabolic and immune systems under long-term spontaneous exercise.

The study first analyzed targeted metabolomics and lipidomics data from all plasma samples, identifying 62 differentially abundant molecules between the two groups, including metabolites (72.6%) and complex lipids (27.4%). This revealed a remodeling of metabolic networks in the plasma of long-term regular exercisers. Notably, the upregulation of fatty acid oxidation (FAO)-related metabolites (e.g., O-hexanoylcarnitine, 3-hydroxybutyrate) corresponded with a widespread decrease in triglycerides (TG). This confirmed that regular exercise systematically enhances the body's FAO capacity by boosting fatty acid transport, ketone body production, and short-chain fatty acid utilization. This finding is highly consistent with previous studies demonstrating exercise-induced lipolysis and energy metabolism shifts[34]. Simultaneously, we observed synchronous upregulation of multiple antioxidant defense-related metabolites, including betaine, 3-hydroxy-o-aminobenzoic acid, and 2-hydroxybutyrate. This suggests that while improving energy metabolism efficiency, exercise may synergistically activate endogenous antioxidant networks to counteract potential oxidative stress, further substantiating the hypothesis that exercise enhances the body's antioxidant capacity[35]. Research over the past decade in the field of immunometabolism has established that circulating metabolites serve as pivotal signaling molecules and epigenetic modulators that critically regulate immune cell function and differentiation[80]**.** This conceptual framework offers a novel perspective for interpreting how long-term exercise-induced systemic metabolic reprogramming may influence immune homeostasis. In this study, we found that several significantly upregulated metabolites after long-term exercise possess well-defined immunomodulatory properties. For instance, β-hydroxybutyrate (BHB) has been demonstrated in seminal work to function as an endogenous histone deacetylase (HDAC) inhibitor, which directly suppresses the assembly and activation of the NLRP3 inflammasome in macrophages, thereby exerting a core anti-inflammatory effect[81]**.** Concurrently, the significantly altered betaine, serving as a crucial methyl donor, can

565 modulate methylation metabolism within immune cells such as T cells, thereby influencing
566 their differentiation and functional responses[82]. Furthermore, γ-aminobutyric acid (GABA)
567 can directly regulate lymphocyte proliferation and inflammatory cytokine secretion through
568 functional receptors expressed on the surface of immune cells[83].These findings
569 collectively suggest that long-term exercise shapes a "systemic metabolic
570 microenvironment" constituted by specific immunomodulatory metabolites, which may
571 continuously deliver pro-homeostatic signals to the immune system. Based on this, we
572 propose a central hypothesis: these characteristic metabolites constitute a critical
573 metabolic interface linking exercise adaptation to immune optimization and are potential
574 messengers driving the observed beneficial phenotypes, such as the epigenetic
575 reprogramming of monocytes. Consequently, the central objective of future research is to
576 establish the causal relationship between these metabolites and specific immune
577 phenotypes through functional experiments, thereby completely delineating the
578 physiological map of the "exercise-metabolism-immunity" axis at the mechanistic level.

579 We then provides deep insights into exercise-induced remodeling of the immune
580 system through single-cell multi-omics analysis, revealing that regular exercise profoundly
581 impacts immune function via epigenetic reprogramming. We found that in classical
582 monocytes (cMono), the exercise group exhibited increased chromatin accessibility in
583 regions associated with antigen presentation and migration functions, along with
584 upregulation of transcription factor motif accessibility. Conversely, motif accessibility for
585 stress and inflammation-related factors such as AP-1 and BACH1 was downregulated.
586 This provides an epigenetic explanation for why exercise-trained monocytes exhibit
587 enhanced antigen-presenting potential and reduced inflammatory propensity. Regarding
588 adaptive immunity, $CD8^+$ cytotoxic T cells (CD8_CTL) and MAIT cells exhibited
589 significant epigenetic activation. Their gene module scores for cell killing and interferon
590 responses were upregulated, and the accessibility of binding motifs for TFs closely
591 associated with cytotoxic differentiation and effector functions—including EOMES,
592 RUNX3, RUNX1, and TBX21—was markedly increased. Similarly, NK cells (particularly
593 mature NK subsets) within the innate immune system exhibited analogous pro-effector
594 epigenetic signatures. These findings suggest that exercise may synergistically enhance

595 the killing and defense capabilities of effector cells across distinct immune branches while
596 suppressing their inflammatory states through a conserved set of epigenetic regulatory
597 mechanisms.

598 Cell communication analysis further substantiated the relationship between
599 metabolic regulation and immune remodeling, establishing a systematic view of exercise
600 adaptation. We found that exercise enhances signaling in immune cell pathways
601 represented by MHC-I/II and chemokines (e.g., CCL), including chemotaxis, antigen
602 presentation, and cell adhesion. Simultaneously, it specifically attenuates the
603 pro-inflammatory network dominated by myeloid cells and centered on retinoic acid-like
604 protein (RETN) signaling. The reduction in *RETN* levels depends on the inhibition of its
605 downstream NF-κB pathway activity. This suggests exercise may influence immune
606 cell-mediated ligand-receptor communication networks, enhancing adaptive immune and
607 tissue repair signaling while specifically suppressing myeloid cell-mediated inflammatory
608 pathway activity. This dynamic metabolic-immune cell interaction network provides crucial
609 clues and foreshadows future investigations into exercise benefit mechanisms.

610 In summary, this study employs multi-omics integration to describe exercise-related
611 changes in circulating metabolites and lipids, alongside cell-type-resolved transcriptional
612 and epigenetic landscapes of immune cells. These are characterized by enhanced lipid
613 oxidation and antioxidant metabolic capacity, alongside functionally optimized immune
614 states guided by epigenetic reprogramming. We provide the first multidimensional map of
615 long-term spontaneous exercise-induced molecular adaptations across metabolic,
616 epigenetic, and immune dimensions, thereby deepening our understanding of
617 exercise-mediated health promotion mechanisms and identifying potential biomarkers and
618 regulatory pathways for chronic disease prevention and management. This study also has
619 limitations: First, physical activity assessment relied on self-report. While this is a common
620 and valid approach in large-scale epidemiological research, it remains susceptible to
621 recall bias and social desirability bias. Second, our cohort consisted of relatively young
622 and healthy individuals. This facilitates the identification of core exercise adaptation
623 signatures in a less confounded context but concurrently raises uncertainty regarding the
624 direct extrapolation of the discovered molecular profiles to older, obese, or clinical

populations. Third, although the sample size was adequate for in-depth single-cell multi-omics profiling, a sex imbalance across groups was present. While statistical adjustments were applied, residual confounding effects may persist. Future studies with expanded cohorts would enhance statistical power for more granular subgroup analyses (e.g., by sex or exercise modality) and improve the balancing of baseline characteristics. Furthermore, the study focused exclusively on peripheral blood, leaving the local immunomodulatory effects of exercise in key tissues such as muscle and adipose tissue unaddressed—contexts where tissue-specific immune microenvironments may obey distinct regulatory principles. Finally, and most critically, the absence of functional experimental validation for predictive networks, such as the "metabolite-driven immune remodeling" hypothesis, means the underlying mechanistic relationships await confirmation in subsequent investigations.

Methods

Data Source

We utilized data from the Chinese Immune Multi-omics Atlas (CIMA) project[23]. This public cohort initially enrolled 428 adults aged 20–77 and systematically collected multi-dimensional information, including questionnaire surveys, clinical biochemical indices, and multi-omics data. The present secondary analysis utilized detailed physical activity questionnaires, blood biochemical markers, targeted lipidomics data, as well as the raw data matrices generated by paired single-cell RNA and ATAC sequencing performed on peripheral blood mononuclear cells (PBMCs).

Assessment of Physical Activity, Grouping, and Final Study Population

Within the CIMA cohort, physical activity was assessed using its standardized questionnairesacross three dimensions: intensity, frequency, and duration. First, exercise intensity was evaluated with the question “What intensity of exercise do you usually engage in?” (options: low intensity [e.g., walking], moderate intensity [e.g., brisk walking, jogging], high intensity [e.g., running, ball sports], or almost none). Second, exercise frequency and single-session duration were obtained from the items “Exercise habit”

(options: ≥5 times/week to almost none) and “What is the average daily exercise duration?” (options: >2 hours to <30 minutes and almost none). To provide a conservative estimate, the lower bounds of the frequency and duration intervals were used for calculation (e.g., “3–4 times/week” was counted as 3, “1–2 hours” as 60 minutes). The total weekly moderate-to-vigorous physical activity (MVPA) duration was then derived for each participant. Participants were grouped strictly according to World Health Organization (WHO) recommendations[24]. The regular exercise group (EX) was defined as meeting both of the following criteria: (1) engaging in moderate- or high-intensity exercise ≥3 times per week, and (2) achieving a total weekly duration of ≥150 minutes of moderate-intensity activity or ≥75 minutes of vigorous-intensity activity. The sedentary control group (SED) was required to satisfy both conditions: (1) not meeting any of the above exercise criteria, and (2) self-reporting a daily living pattern dominated by sedentary behavior.

To reduce potential confounding from age and chronic disease status and to establish a relatively homogeneous analytical sample, the following inclusion and exclusion criteria were applied: Inclusion criteria: age 20–40 years. Exclusion criteria: (1) self-reported major chronic diseases (e.g., cardiovascular and cerebrovascular diseases, diabetes, autoimmune diseases, malignancies) or obesity; (2) missing key physical-activity questionnaire items or omics data. After applying these screening steps, the above grouping criteria were implemented. Finally, 86 participants were clearly classified and included in all subsequent analyses, comprising 40 in EX and 46 in SED.

Differential Analysis of Plasma Metabolites and Lipids

The raw targeted metabolomics and lipidomics data matrices were preprocessed and normalized using the MetaboAnalyst 6.0 online platform[84]. Metabolites or lipid species missing in more than one-third of samples were removed. Remaining missing values were imputed using the minimum value detected for each molecule. The imputed data were then subjected to Auto-Scaling (unit variance scaling), i.e., each molecule was mean-centered and divided by its standard deviation, to eliminate scale differences and give equal analytical weight to all molecules.To identify molecules associated with exercise habits, a combined screening strategy was applied. First, linear models were

685 fitted using the limma package in R, adjusting for sex and BMI, and molecules with $p <$
686 0.05 and $|Log2FC| \geq 0.15$ between exercise groups were selected. In parallel, PLS-DA
687 was performed using the MetaboAnalystR package[85] (version 4.0) to calculate variable
688 importance in projection (VIP) scores. Ultimately, metabolites and lipids that satisfied both
689 criteria—VIP scores ≥ 1 and the linear-model thresholds—were defined as
690 exercise-associated differential molecules (complete list provided in Supplementary Table
691 S2).
692
693 **Pathway Analysis of Plasma Metabolites**
694 To uncover the potential biological processes underlying the differential metabolites,
695 KEGG pathway enrichment analysis was conducted for the exercise-associated
696 differential metabolites using the MetaboAnalyst 6.0 platform[84]. Enrichment significance
697 was evaluated with the hypergeometric test, and p-values were corrected for multiple
698 testing using the false discovery rate (FDR). To identify key metabolic alterations, we
699 focused on pathways that passed a significance threshold of FDR-corrected p-value <
700 0.05 in our enrichment analysis.
701
702 **Single-cell RNA-seq data processing**
703 This study is based on processed single-cell gene expression count matrices, with all data
704 preprocessing and analysis performed using the Scanpy toolkit[86] (version 1.10.3). First,
705 stringent quality control was applied: cells were retained only if they exhibited a total gene
706 count between 500 and 6000, a total UMI count between 1000 and 25000, and a
707 mitochondrial gene expression proportion below 10%(Fig S1A). Subsequently, doublets
708 were identified and removed using the Scrublet algorithm[87]. This method generates
709 simulated doublets as a reference and computes a doublet score for each cell using a
710 k-nearest neighbor classifier. Key parameters were set as follows: expected doublet rate =
711 0.06 (expected_doublet_rate=0.06), minimum counts per gene = 3 (min_counts=3),
712 minimum cells per gene = 3 (min_cells=3), with log-transformation enabled
713 (log_transform=True). The detection procedure focused on the top 85% most variable
714 genes (min_gene_variability_pctl=85) and used the first 30 principal components

715 (n_prin_comps=30) to construct the feature space for discrimination. Cells with a doublet
716 score above the threshold of 0.2 were classified as doublets and excluded from
717 downstream analysis.Following doublet removal, the top 3000 highly variable genes were
718 selected from the quality-controlled data using the sc.pp.highly_variable_genes function.
719 Expression data were then normalized by total count (sc.pp.normalize_total,
720 target_sum=1e4) and log-transformed (sc.pp.log1p). Only highly variable genes were
721 scaled (sc.pp.scale), and linear dimensionality reduction was performed via principal
722 component analysis (sc.tl.pca). Finally, to correct for batch effects across samples, the
723 Harmony algorithm[88] (sc.external.pp.harmony_integrate) was applied to integrate the
724 principal components, yielding a batch-corrected low-dimensional embedding for
725 subsequent clustering and differential expression analysis.
726
727 **Unsupervised clustering and subclustering**
728 Following batch-effect correction, the integrated low-dimensional embedding was used for
729 unsupervised cell clustering. Cell communities were identified using the Louvain
730 algorithm[89] (implemented via sc.tl.louvain in Scanpy) with a resolution parameter set to
731 0.75. The nonlinear dimensionality reduction method UMAP (sc.tl.umap) was employed to
732 visualize the clusters in two-dimensional space. Differential expression analysis within
733 each cluster was performed using the sc.tl.rank_genes_groups function (with default
734 parameters), and the resulting genes were compared with known cell-type-specific
735 markers to assign preliminary cell identities. Based on the expression patterns of
736 canonical marker genes, six major cell populations were initially annotated, including
737 $CD4^+$ T cells, $CD8^+$ T cells, unconventional T cells, B cells, myeloid cells, and NK
738 cells.To further resolve the heterogeneity within each major population, subsets
739 corresponding to each of the above categories (e.g., myeloid, $CD4^+$ T, $CD8^+$ T, NK, and
740 B cells) were extracted as separate AnnData objects. For each subset, we reapplied
741 normalization, highly variable gene selection, PCA-based dimensionality reduction, and
742 sub-clustering using the Louvain algorithm at a resolution tailored to the size and internal
743 complexity of the subset. Cell-state identities of the sub-clusters were again determined
744 by differential expression analysis and comparison with established subtype-specific

745 markers. Ultimately, 34 distinct cell subpopulations were identified across the entire
746 dataset.
747
748 **Differential abundance analysis using scRNA-seq data**
749 To systematically evaluate the influence of exercise habits on immune cell composition,
750 we performed differential abundance analysis using the MiloR package
751 (https://github.com/MarioniLab/MILOR[31]) on single-cell RNA-sequencing data. First, To
752 capture the neighborhood relationships among cells, we constructed a k-nearest neighbor
753 graph (k=30) using the first 50 principal components. Cells were then partitioned into
754 overlapping neighborhoods using the makeNhoods function (prop = 0.2, k = 30, d = 50,
755 refined = TRUE). For cell counts in each neighborhood, We used MiloR to perform
756 differential abundance testing. This method fits a negative-binomial generalized linear
757 model, and we applied the trimmed mean of M-values (TMM) method for cross-sample
758 normalization.To isolate the independent effect of exercise status, the model included the
759 exercise group (EX vs. SED) as the primary variable while adjusting for sex and BMI as
760 covariates. Hypothesis testing was carried out via the testNhoods function. We classified
761 neighborhoods as exhibiting significant differential abundance if their false discovery rate
762 (FDR)-adjusted p-value was less than 0.05. The log2 fold-changes of these significant
763 neighborhoods were then used for subsequent visualization and biological interpretation.
764
765 **Differential Expression Analysis of Cell Subpopulations**
766 For the comparison of gene expression differences between the EX and SED groups
767 across cell subpopulations, we employed the linear model-based limma package[90] for
768 statistical testing. For each subpopulation, we constructed a linear model in which gene
769 expression was the dependent variable, exercise group was the primary predictor, and
770 sex and BMI were included as covariates. From the fitted model, the log2 fold-change
771 (log2FC) and raw p-value were extracted for every gene. To control for false positives, we
772 adopted the Benjamini-Hochberg (BH) procedure for multiple testing correction of the
773 p-values. In our analysis, a gene was considered a significant DEG within a given
774 subpopulation if it showed $|\log2FC| \geq 0.25$ together with an adjusted p-value < 0.05.

775

776 **Module scoring using scRNA-seq data**

777 To evaluate activity differences in specific biological processes across immune cell
778 subpopulations, Scoring cells for specific gene modules was based on the Scanpy
779 sc.tl.score_genes method, which computes the difference between the average
780 expression of target genes and that of a matched random background per cell. For our
781 analysis, we sourced gene sets from the Molecular Signatures Database[91] (MSigDB).To
782 assess statistical differences in module scores between the EX and SED groups, we first
783 computed the mean module score per subpopulation. We then fitted linear models
784 comparing groups, adjusting for sex and BMI.

785

786 **Gene regulatory network analysis based on SCENIC**

787 We employed the pySCENIC toolkit (v0.12.1)[92] to infer potential transcriptional regulatory
788 networks across distinct immune cell subpopulations. Our analysis began by randomly
789 sampling a subset of 10,000 cells to serve as the training set. Using a comprehensive
790 transcription factor list, we then applied the GRNBoost2 algorithm[93] to infer co-expression
791 relationships between each transcription factor and all potential target genes. This step
792 generated the initial gene regulatory network. Subsequently, the preliminary network was
793 pruned using the cisTarget database. Only target genes whose transcription start sites
794 (TSS) fell within ±10 kb of a corresponding transcription factor binding motif were retained,
795 yielding high-confidence regulons. Next, the AUCell module of pySCENIC was used to
796 calculate the activity level (represented as the area under the curve, AUC) of each regulon
797 in all cells, followed by z-score normalization. To identify regulons with differential activity
798 between the exercise (EX) and sedentary (SED) groups, we compared the AUC z-scores
799 of regulons across cell types using the limma package. Regulons meeting the thresholds
800 of adjusted p-value < 0.05 and $|\log2FoldChange| \geq 0.25$ were defined as differentially
801 active.

802

803 **scATAC-seq quality control.**

804 We used ArchR[94] (version 1.0.2) to analyze single-cell ATAC-seq data. First, Arrow files

805 were generated for each sample from the raw sequencing fragments following the default
806 settings. Subsequently, stringent cell-level quality control was applied. Our cell quality
807 control was based on the parameters below: TSS enrichment ≥ 10, number of unique
808 fragments between 1×10^3 and 1×10^5, nucleosome signal < 4, and blacklist ratio < 0.05.
809 Doublets were identified and removed using ArchR's default parameters, resulting in a
810 high-quality chromatin accessibility matrix for downstream analysis.
811
812 **scATAC-seq dimensionality reduction and clustering.**
813 To construct a single-cell chromatin accessibility landscape, we first performed
814 dimensionality reduction on the cell-by-peak matrix using latent semantic indexing (LSI) to
815 capture dominant patterns of chromatin accessibility variation. To mitigate technical
816 variability introduced during sample preparation, we applied the Harmony[88] algorithm to
817 integrate the reduced-dimension embeddings across samples.Using the integrated
818 low-dimensional representation, we partitioned cells into communities via shared nearest
819 neighbor (SNN)-based graph clustering. An initial clustering step was performed at a
820 uniform resolution of 0.8 to delineate major cell categories. To further resolve cellular
821 heterogeneity and enhance the detection of small-scale cell states, we carried out
822 sub-clustering on selected populations derived from the initial clustering, with resolutions
823 adapted between 0.3 and 1.5 according to the intrinsic structure of each cell group. As the
824 final step in our single-cell ATAC-seq analysis, all cells were projected into a
825 two-dimensional space for visualization using uniform manifold approximation and
826 projection (UMAP), providing an intuitive representation of cell-state distributions and
827 community relationships.
828
829 **scATAC–seq gene activity scores.**
830 Gene activity was quantified with ArchR[94] (v1.0.2) , which computes scores by correlating
831 accessibility across gene-linked regions. Following this, the MAGIC algorithm[95] was
832 applied to smooth the scores and mitigate noise inherent in sparse scATAC-seq data.
833
834 **Integrating scATAC-seq and scRNA-seq Data to Infer Cell Identity**

835 To accurately annotate cell types in the scATAC-seq data, we performed an integrative
836 analysis against a matched scRNA-seq reference. This was achieved using the
837 addGeneIntegrationMatrixfunction in ArchR. This process was conducted in two
838 sequential steps: first, an unconstrained integration was carried out to broadly map each
839 scATAC-seq cell cluster to the major cell types defined in the scRNA-seq data.
840 Subsequently, based on this initial mapping, A second, more constrained integration was
841 then carried out by re-executing addGeneIntegrationMatrix with updated broad group
842 labels. This refinement enabled finer multimodal alignment and produced final annotations
843 at high resolution, ensuring consistency with the transcriptomic profiles.To independently
844 assess the reliability of the annotations, we employed the Gene Activity Score (GAS)
845 matrix derived from the scATAC-seq data for validation. By comparing the distribution of
846 annotated cells in the gene-activity space with that of the corresponding scRNA-seq
847 reference cells in gene-expression space, we further confirmed the accuracy of the
848 cross-modal annotation.
849
850 **scATAC–seq pseudobulk replicate generation and peak calling.**
851 To support differential analysis across cell clusters and conditions, we first constructed
852 pseudobulk samples. This involved running the addGroupCoverages function with tailored
853 parameters to create non-overlapping replicates for each cell group. These replicates
854 were then consolidated via addReproduciblePeakSet to generate a consensus peak
855 matrix. For precise peak identification, we used MACS2[96] . The resulting unified peak set
856 provided the foundation for all subsequent differential accessibility testing.
857
858 **scATAC–seq genomic regions annotation.**
859 For the identified chromatin-accessible regions (peaks), We annotated the called peaks
860 by applying the annotatePeak function from the ChIPseeker package [97] package with
861 default parameters, in order to determine the nearest associated genes and genomic
862 features for each peak.
863
864 **scATAC motif enrichment and motif deviation analysis**

We carried out transcription factor motif enrichment followed by chromatin accessibility deviation analysis using the unified pseudobulk peak set. For motif annotation, we utilized the CIS-BP motif database (via ChromVAR[98]), JASPAR2020[99], and HOMER[100]. Additionally, motif accessibility deviation scores were calculated for each cell using the chromVAR algorithm implemented in ArchR, enabling systematic assessment of changes in transcription factor binding activity across different cellular states.

Identification of differential peaks

To detect differentially accessible chromatin regions between groups, we employed two complementary approaches. First, we used the getMarkerFeatures function with parameters useMatrix = "PeakMatrix", bias covariates including "TSSEnrichment" and "log10(nFrags)", and testMethod = "wilcoxon" to identify differential peaks between groups, applying thresholds of FDR ≤ 0.1 and log2FC ≥ 0.25. Second, for differential transcription factor activity analysis, we applied the same function with useMatrix = "MotifMatrix" and thresholds of FDR ≤ 0.01 and MeanDiff ≥ 0.05.

Gene Ontology Enrichment Analysis of differential peaks

To interpret the potential biological functions of differential chromatin accessible regions, we performed Gene Ontology (GO) enrichment analysis using the R package rGREAT[101]. This package associates genomic regions with nearby genes and evaluates their enrichment significance across Biological Process, Molecular Function, and Cellular Component terms using a hypergeometric test. The analysis was conducted with the human genome as background, applying default parameters, We selected terms passing the significance threshold of an FDR-adjusted p-value < 0.05 as significantly enriched.

scATAC–seq TF Foot-print analysis

We performed transcription factor footprinting analysis based on Tn5 transposase cleavage profiles across genomic motif regions. First, the sequence bias of Tn5 was subtracted from the raw cleavage signal to obtain a corrected footprint profile. Motif positions were determined using a combination of the CIS-BP database[102] (from

895 chromVAR human_pwms_v1) and the JASPAR2020 database. Subsequently, footprint
896 signals were normalized using the average insertion density within ±200–250 bp regions
897 centered on each motif. The normalized results were aggregated by pseudo-replicate, and
898 their means and standard deviations were calculated and visualized. Finally, to assess
899 differences in footprint signals between groups, We chose the nonparametric Wilcoxon
900 rank-sum test for comparisons between two groups, setting the significance level at $p <$
901 0.05.
902
903 **scATAC-seq Peak to gene linkage analysis**
904 To identify peak-to-gene links, we performed prediction analysis using the
905 addPeak2GeneLinks function in ArchR. The analysis was based on the batch-corrected
906 low-dimensional embedding (specified via the reducedDims parameter) and computed the
907 correlation between chromatin accessibility and gene expression. A correlation threshold
908 of corCutOff = 0.4 was applied to screen for high-confidence linkages. The resulting
909 significant associations were output as a GRanges object for subsequent visualization
910 and functional interpretation.
911
912 **Cell-cell interaction analysis**
913 We performed a systematic analysis of potential signaling communication among all
914 immune cell subpopulations using the CellChat R package[75] (v.1.5.0). This approach
915 leverages the built-in ligand–receptor pair databases of CellChat to identify and
916 characterize potential signaling interaction networks between different immune cell
917 subtypes.
918
919 **Acknowledgements**
920 We express our sincere appreciation to the China National GeneBank (CNGB).
921
922 **Funding**
923 Funding for this research was provided by the National Key Research and Development
924 Program of China (2025YFC3409300, 2022YFC3400400) and the National Science and

Technology Major Project (2024ZD0530500).

Author contributions

X.S., J.L., S.G., J.Y., P.Y., and C.L. conceived the idea; C.L., P.Y., J.Y., Y.Z., and J.Z. supervised the work; X.S., J.L., S.G., S.X., Y.W., Y.Z., and W.Z. analyzed the data and performed visualization; X.S., J.L., S.G., P.Y., J.Y., and C.L. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Data availability

The single-cell RNA-seq and single-cell ATAC-seq data, plasma metabolomics and lipidomics data, and blood biochemistry data presented in this study were sourced from the CIMA cohort, hosted on the TrueBlood database within the CNGBdb portal (https://db.cngb.org/trueblood/cima/).

Code availability

The code can be obtained from the corresponding author upon reasonable request. Information on publicly available software used is provided in the Methods.

Reference

1. Schmid, D. & Leitzmann, M. F. Association between physical activity and mortality among breast cancer and colorectal cancer survivors: a systematic review and meta-analysis. *Ann Oncol* **25**, 1293–1311 (2014).

2. Moore, S. C. *et al.* Association of Leisure-Time Physical Activity With Risk of 26 Types of Cancer in 1.44 Million Adults. *JAMA Intern Med* **176**, 816–825 (2016).

3. Patel, A. V. *et al.* American College of Sports Medicine Roundtable Report on Physical Activity, Sedentary Behavior, and Cancer Prevention and Control. *Med Sci Sports Exerc* **51**, 2391–2402 (2019).

4. Im, L. *et al.* Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet (London,*

956 *England)* **380**, (2012).

957 5. Meyerhardt, J. A. *et al.* Impact of physical activity on cancer recurrence and
958 survival in patients with stage III colon cancer: findings from CALGB 89803. *J Clin*
959 *Oncol* **24**, 3535–3541 (2006).

960 6. Chastin, S. F. M. *et al.* Effects of Regular Physical Activity on the Immune System,
961 Vaccination and Risk of Community-Acquired Infectious Disease in the General
962 Population: Systematic Review and Meta-Analysis. *Sports Med* **51**, 1673–1686
963 (2021).

964 7. Hawley, J. A., Hargreaves, M., Joyner, M. J. & Zierath, J. R. Integrative biology of
965 exercise. *Cell* **159**, 738–749 (2014).

966 8. Booth, F. W., Roberts, C. K., Thyfault, J. P., Ruegsegger, G. N. & Toedebusch, R.
967 G. Role of Inactivity in Chronic Diseases: Evolutionary Insight and Pathophysiological
968 Mechanisms. *Physiol Rev* **97**, 1351–1402 (2017).

969 9. Pedersen, B. K. & Febbraio, M. A. Muscles, exercise and obesity: skeletal muscle
970 as a secretory organ. *Nat Rev Endocrinol* **8**, 457–465 (2012).

971 10. Sun, S. *et al.* A single-cell transcriptomic atlas of exercise-induced
972 anti-inflammatory and geroprotective effects across the body. *Innovation (Camb)* **4**,
973 100380 (2023).

974 11. K, C. *et al.* Molecular Choreography of Acute Exercise. *Cell* **181**, (2020).

975 12. Geng, L. *et al.* Systematic profiling reveals betaine as an exercise mimetic for
976 geroprotection. *Cell* **188**, 5403-5425.e33 (2025).

977 13. MoTrPAC Study Group, Lead Analysts, & MoTrPAC Study Group. Temporal
978 dynamics of the multi-omic response to endurance exercise training. *Nature* **629**,
979 174–183 (2024).

980 14. Mathis, D. & Shoelson, S. E. Immunometabolism: an emerging frontier. *Nat Rev*
981 *Immunol* **11**, 81 (2011).

982 15. Pedersen, B. K. & Saltin, B. Exercise as medicine - evidence for prescribing
983 exercise as therapy in 26 different chronic diseases. *Scand J Med Sci Sports* **25**
984 **Suppl 3**, 1–72 (2015).

985 16. Ki, S. *et al.* 12,13-diHOME: An Exercise-Induced Lipokine that Increases Skeletal
986 Muscle Fatty Acid Uptake. *Cell metabolism* **27**, (2018).

987 17. Shimazu, T. *et al.* Suppression of oxidative stress by β-hydroxybutyrate, an
988 endogenous histone deacetylase inhibitor. *Science* **339**, 211–214 (2013).

989 18. Jc, N. & E, V. β-Hydroxybutyrate: A Signaling Metabolite. *Annual review of*
990 *nutrition* **37**, (2017).

991 19. Pedersen, B. K. & Ullum, H. NK cell response to physical activity: possible
992 mechanisms of action. *Med Sci Sports Exerc* **26**, 140–146 (1994).

993 20. Pedersen, B. K. & Hoffman-Goetz, L. Exercise and the immune system:
994 regulation, integration, and adaptation. *Physiol Rev* **80**, 1055–1081 (2000).

995 21. Nieman, D. C. *et al.* The acute immune response to exhaustive resistance
996 exercise. *Int J Sports Med* **16**, 322–328 (1995).

997 22. Simpson, R. J., Kunz, H., Agha, N. & Graff, R. Exercise and the Regulation of
998 Immune Functions. *Prog Mol Biol Transl Sci* **135**, 355–380 (2015).

999 23. Yin, J., Zheng, Y., Huang, Z. & Zhou, W. Chinese Immune Multi-Omics Atlas.
1000 *Science* **391**, (2026).

1001 24. Bull, F. C. *et al.* World Health Organization 2020 guidelines on physical activity
1002 and sedentary behaviour. *Br J Sports Med* **54**, 1451–1462 (2020).

1003 25. Mann, S., Beedie, C. & Jimenez, A. Differential effects of aerobic exercise,
1004 resistance training and combined exercise modalities on cholesterol and the lipid
1005 profile: review, synthesis and recommendations. *Sports Med* **44**, 211–221 (2014).

1006 26. Aadahl, M. *et al.* Five-year change in physical activity is associated with changes
1007 in cardiovascular disease risk factors: the Inter99 study. *Prev Med* **48**, 326–331
1008 (2009).

1009 27. S, K. *et al.* Effect of aerobic exercise training on serum levels of high-density
1010 lipoprotein cholesterol: a meta-analysis. *Archives of internal medicine* **167**, (2007).

1011 28. Baxmann, A. C. *et al.* Influence of muscle mass and physical activity on serum
1012 and urinary creatinine and serum cystatin C. *Clin J Am Soc Nephrol* **3**, 348–354
1013 (2008).

1014 29. Waring, W. S. *et al.* Uric acid reduces exercise-induced oxidative stress in healthy
1015 adults. *Clin Sci (Lond)* **105**, 425–430 (2003).

1016 30. Stocker, R., Yamamoto, Y., McDonagh, A. F., Glazer, A. N. & Ames, B. N. Bilirubin
1017 is an antioxidant of possible physiological importance. *Science* **235**, 1043–1046
1018 (1987).

1019 31. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C.
1020 Differential abundance testing on single-cell data using k-nearest neighbor graphs.
1021 *Nat Biotechnol* **40**, 245–253 (2022).

1022 32. Schooneman, M. G., Vaz, F. M., Houten, S. M. & Soeters, M. R. Acylcarnitines:
1023 reflecting or inflicting insulin resistance? *Diabetes* **62**, 1–8 (2013).

1024 33. M, E., Ke, C. & B, E. Metabolism of ketone bodies during exercise and training:
1025 physiological basis for exogenous supplementation. *The Journal of physiology* **595**,
1026 (2017).

1027 34. Tan, J. *et al.* The role of short-chain fatty acids in health and disease. *Adv*
1028 *Immunol* **121**, 91–119 (2014).

1029 35. Sk, P. & Mj, J. Exercise-induced oxidative stress: cellular mechanisms and impact
1030 on muscle force production. *Physiological reviews* **88**, (2008).

1031 36. Christen, S., Thomas, S. R., Garner, B. & Stocker, R. Inhibition by
1032 interferon-gamma of human mononuclear cell-mediated low density lipoprotein
1033 oxidation. Participation of tryptophan metabolism along the kynurenine pathway. *J*
1034 *Clin Invest* **93**, 2149–2158 (1994).

1035 37. Gall, W. E. *et al.* alpha-hydroxybutyrate is an early biomarker of insulin resistance
1036 and glucose intolerance in a nondiabetic population. *PLoS One* **5**, e10883 (2010).

1037 38. Stockinger, B., Di Meglio, P., Gialitakis, M. & Duarte, J. H. The aryl hydrocarbon
1038 receptor: multitasking in the immune system. *Annu Rev Immunol* **32**, 403–432 (2014).

1039 39. Bhat, R. *et al.* Inhibitory role for GABA in autoimmune inflammation. *Proc Natl*
1040 *Acad Sci U S A* **107**, 2580–2585 (2010).

1041 40. Wlodarska, M. *et al.* Indoleacrylic Acid Produced by Commensal
1042 Peptostreptococcus Species Suppresses Inflammation. *Cell Host Microbe* **22**,
1043 25-37.e6 (2017).

1044 41. Hang, S. *et al.* Bile acid metabolites control TH17 and Treg cell differentiation.
1045 *Nature* **576**, 143–148 (2019).

1046 42. van Loon, L. J. C. & Goodpaster, B. H. Increased intramuscular lipid storage in
1047 the insulin-resistant and endurance-trained state. *Pflugers Arch* **451**, 606–616 (2006).

1048 43. O'Neill, L. A. J., Kishton, R. J. & Rathmell, J. A guide to immunometabolism for
1049 immunologists. *Nat Rev Immunol* **16**, 553–565 (2016).

1050 44. Hotamisligil, G. S. Inflammation and metabolic disorders. *Nature* **444**, 860–867
1051 (2006).

1052 45. Simpson, R. J., Kunz, H., Agha, N. & Graff, R. Exercise and the Regulation of
1053 Immune Functions. *Prog Mol Biol Transl Sci* **135**, 355–380 (2015).

1054 46. Storm van's Gravesande, K. *et al.* IFN regulatory factor-1 regulates
1055 IFN-gamma-dependent cathepsin S expression. *J Immunol* **168**, 4488–4494 (2002).

1056 47. van der Stoep, N., Quinten, E., Marcondes Rezende, M. & van den Elsen, P. J.
1057 E47, IRF-4, and PU.1 synergize to induce B-cell-specific activation of the class II
1058 transactivator promoter III (CIITA-PIII). *Blood* **104**, 2849–2857 (2004).

1059 48. Das, H. *et al.* Kruppel-like factor 2 (KLF2) regulates proinflammatory activation of
1060 monocytes. *Proc Natl Acad Sci U S A* **103**, 6653–6658 (2006).

1061 49. Liao, X. *et al.* Krüppel-like factor 4 regulates macrophage polarization. *J Clin*
1062 *Invest* **121**, 2736–2749 (2011).

1063 50. Shaulian, E. & Karin, M. AP-1 as a regulator of cell life and death. *Nat Cell Biol* **4**,
1064 E131-136 (2002).

1065 51. Igarashi, K. & Sun, J. The heme-Bach1 pathway in the regulation of oxidative
1066 stress response and erythroid differentiation. *Antioxid Redox Signal* **8**, 107–118
1067 (2006).

1068 52. Simpson, R. J., Kunz, H., Agha, N. & Graff, R. Exercise and the Regulation of
1069 Immune Functions. *Prog Mol Biol Transl Sci* **135**, 355–380 (2015).

1070 53. Intlekofer, A. M. *et al.* Effector and memory CD8+ T cell fate coupled by T-bet and
1071 eomesodermin. *Nat Immunol* **6**, 1236–1244 (2005).

1072 54. Pearce, E. L. *et al.* Control of effector CD8+ T cell function by the transcription
1073 factor Eomesodermin. *Science* **302**, 1041–1043 (2003).

1074 55. O'Shea, J. J. & Plenge, R. JAK and STAT signaling molecules in
1075 immunoregulation and immune-mediated disease. *Immunity* **36**, 542–550 (2012).

1076 56. Simpson, R. J., Kunz, H., Agha, N. & Graff, R. Exercise and the Regulation of
1077 Immune Functions. *Prog Mol Biol Transl Sci* **135**, 355–380 (2015).

1078 57. Vivier, E., Tomasello, E., Baratin, M., Walzer, T. & Ugolini, S. Functions of natural
1079 killer cells. *Nat Immunol* **9**, 503–510 (2008).

1080 58. Daussy, C. *et al.* T-bet and Eomes instruct the development of two distinct natural
1081 killer cell lineages in the liver and in the bone marrow. *J Exp Med* **211**, 563–577
1082 (2014).

1083 59. Ramirez, K. *et al.* Gene deregulation and chronic activation in natural killer cells
1084 deficient in the transcription factor ETS1. *Immunity* **36**, 921–932 (2012).

1085 60. Hayden, M. S. & Ghosh, S. Regulation of NF-κB by TNF family cytokines. *Semin*

1086 *Immunol* **26**, 253–266 (2014).

1087 61. Kobayashi, T. & Mattarollo, S. R. Natural killer cell metabolism. *Mol Immunol* **115**,
1088 3–11 (2019).

1089 62. A, C., Dr, L. & I, T. RUNX proteins in transcription factor networks that regulate
1090 T-cell lineage choice. *Nature reviews. Immunology* **9**, (2009).

1091 63. Kaplan, M. H., Schindler, U., Smiley, S. T. & Grusby, M. J. Stat6 is required for
1092 mediating responses to IL-4 and for development of Th2 cells. *Immunity* **4**, 313–319
1093 (1996).

1094 64. Ikawa, T. *et al.* An essential developmental checkpoint for production of the T cell
1095 lineage. *Science* **329**, 93–96 (2010).

1096 65. Ramirez, K. *et al.* Gene deregulation and chronic activation in natural killer cells
1097 deficient in the transcription factor ETS1. *Immunity* **36**, 921–932 (2012).

1098 66. Matsuyama, T. *et al.* Targeted disruption of IRF-1 or IRF-2 results in abnormal
1099 type I IFN gene induction and aberrant lymphocyte development. *Cell* **75**, 83–97
1100 (1993).

1101 67. Platanias, L. C. Mechanisms of type-I- and type-II-interferon-mediated signalling.
1102 *Nat Rev Immunol* **5**, 375–386 (2005).

1103 68. Wojakowski, W., Landmesser, U., Bachowski, R., Jadczyk, T. & Tendera, M.
1104 Mobilization of stem and progenitor cells in cardiovascular diseases. *Leukemia* **26**,
1105 23–33 (2012).

1106 69. Appay, V. & Rowland-Jones, S. L. RANTES: a versatile and controversial
1107 chemokine. *Trends Immunol* **22**, 83–87 (2001).

1108 70. Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. Requirement of transcription
1109 factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**,
1110 1573–1577 (1994).

1111 71. Lin, H. & Grosschedl, R. Failure of B-cell differentiation in mice lacking the
1112 transcription factor EBF. *Nature* **376**, 263–267 (1995).

1113 72. Gerondakis, S. & Siebenlist, U. Roles of the NF-kappaB pathway in lymphocyte
1114 development and function. *Cold Spring Harb Perspect Biol* **2**, a000182 (2010).

1115 73. R, S. *et al.* Graded expression of interferon regulatory factor-4 coordinates
1116 isotype switching with plasma cell differentiation. *Immunity* **25**, (2006).

1117 74. Rn, G., M, M.-S., A, N.-L. & Id, F. Systems biology in immunology: a

computational modeling perspective. *Annual review of immunology* **29**, (2011).

75. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, 1088 (2021).

76. M, F., M, H., S, G. & L, S. The role of resistin as a regulator of inflammation: Implications for various human pathologies. *Clinical immunology (Orlando, Fla.)* **133**, (2009).

77. Silswal, N. *et al.* Human resistin stimulates the pro-inflammatory cytokines TNF-alpha and IL-12 in macrophages by NF-kappaB-dependent pathway. *Biochem Biophys Res Commun* **334**, 1092–1101 (2005).

78. Leitinger, B. & Hohenester, E. Mammalian collagen receptors. *Matrix Biol* **26**, 146–155 (2007).

79. Griffith, J. W., Sokol, C. L. & Luster, A. D. Chemokines and chemokine receptors: positioning cells for host defense and immunity. *Annu Rev Immunol* **32**, 659–702 (2014).

80. O'Neill, L. A. J., Kishton, R. J. & Rathmell, J. A guide to immunometabolism for immunologists. *Nat Rev Immunol* **16**, 553–565 (2016).

81. Youm, Y.-H. *et al.* The ketone metabolite β-hydroxybutyrate blocks NLRP3 inflammasome-mediated inflammatory disease. *Nat Med* **21**, 263–269 (2015).

82. Yang, C. *et al.* Betaine Ameliorates Experimental Autoimmune Encephalomyelitis by Inhibiting Dendritic Cell-Derived IL-6 Production and Th17 Differentiation. *J Immunol* **200**, 1316–1324 (2018).

83. Bhandage, A. K. *et al.* GABA Regulates Release of Inflammatory Cytokines From Peripheral Blood Mononuclear Cells and CD4+ T Cells and Is Immunosuppressive in Type 1 Diabetes. *EBioMedicine* **30**, 283–294 (2018).

84. Pang, Z. *et al.* MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. *Nucleic Acids Res* **52**, W398–W406 (2024).

85. Pang, Z., Chong, J., Li, S. & Xia, J. MetaboAnalystR 3.0: Toward an Optimized Workflow for Global Metabolomics. *Metabolites* **10**, 186 (2020).

86. Fa, W., P, A. & Fj, T. SCANPY: large-scale single-cell gene expression data analysis. *Genome biology* **19**, (2018).

87. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst* **8**, 281-291.e9 (2019).

88. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289–1296 (2019).

89. Large-scale network dynamics underlying the first few hundred milliseconds after stimulus presentation: An investigation of visual recognition memory using iEEG - PubMed. https://pubmed.ncbi.nlm.nih.gov/37688546/.

90. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).

91. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545–15550 (2005).

92. S, A. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nature methods* **14**, (2017).

93. Moerman, T. *et al.* GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* **35**, 2159–2161 (2019).

94. Granja, J. M. *et al.* ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet* **53**, 403–411 (2021).

95. van Dijk, D. *et al.* Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**, 716-729.e27 (2018).

96. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

97. Yu, G., Wang, L.-G. & He, Q.-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).

98. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* **14**, 975–978 (2017).

99. Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **48**, D87–D92 (2020).

100. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576–589 (2010).

101. Tanigawa, Y., Dyer, E. S. & Bejerano, G. WhichTF is functionally important in your open chromatin data? *PLoS Comput Biol* **18**, e1010378 (2022).

102. Weirauch, M. T. *et al.* Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443 (2014).

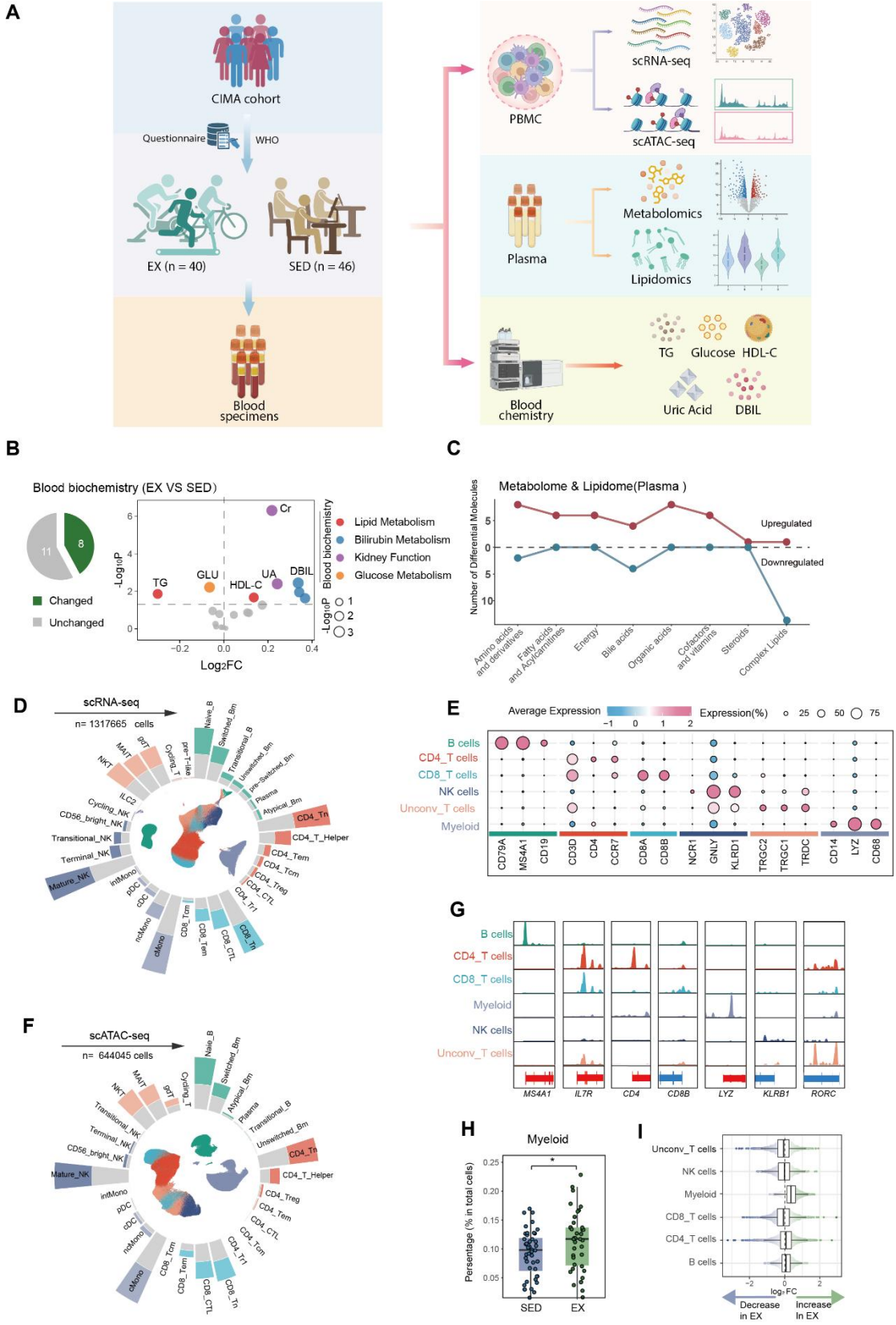


Figure 1. Study design, physiological characteristics, and multi-omics characterization of habitual exercisers versus sedentary controls.

(A) Overall framework of the present study design, about stratification of 86 participants from the Chinese Immune Multi-Omics Cohort (CIMA) into a regular active exercise group (EX, n=40) meeting WHO physical activity guidelines and a sedentary control group (SED, n=46), followed by integrated multi-omics profiling of plasma and peripheral blood mononuclear cell fractions and integrative analysis.

(B) Analysis of Differences in Blood Biochemical Indicators Between the regular active exercise group (EX) and sedentary control group (SED). A pie chart showing changes in blood biochemical indicators(left). A volcano plot displaying the log2 fold change (Log_2FC) and significance (-log10 p-value) of blood biochemical indicators across groups.

(C) A line chart illustrates the distribution of differentially expressed molecules across different categories of plasma metabolomics and lipidomics.

(D) Uniform Manifold Approximation and Projection (UMAP) visualization depicting 1,317,665 immune cells derived from single-cell RNA sequencing (scRNA-seq) data following dimensionality reduction, clustering, and cell annotation, colored by 34 immune lineages.

(E) A dotplot illustrating the average expression and expression percentage of marker genes across distinct immune cell subsets, showing both the magnitude and prevalence of gene expression.

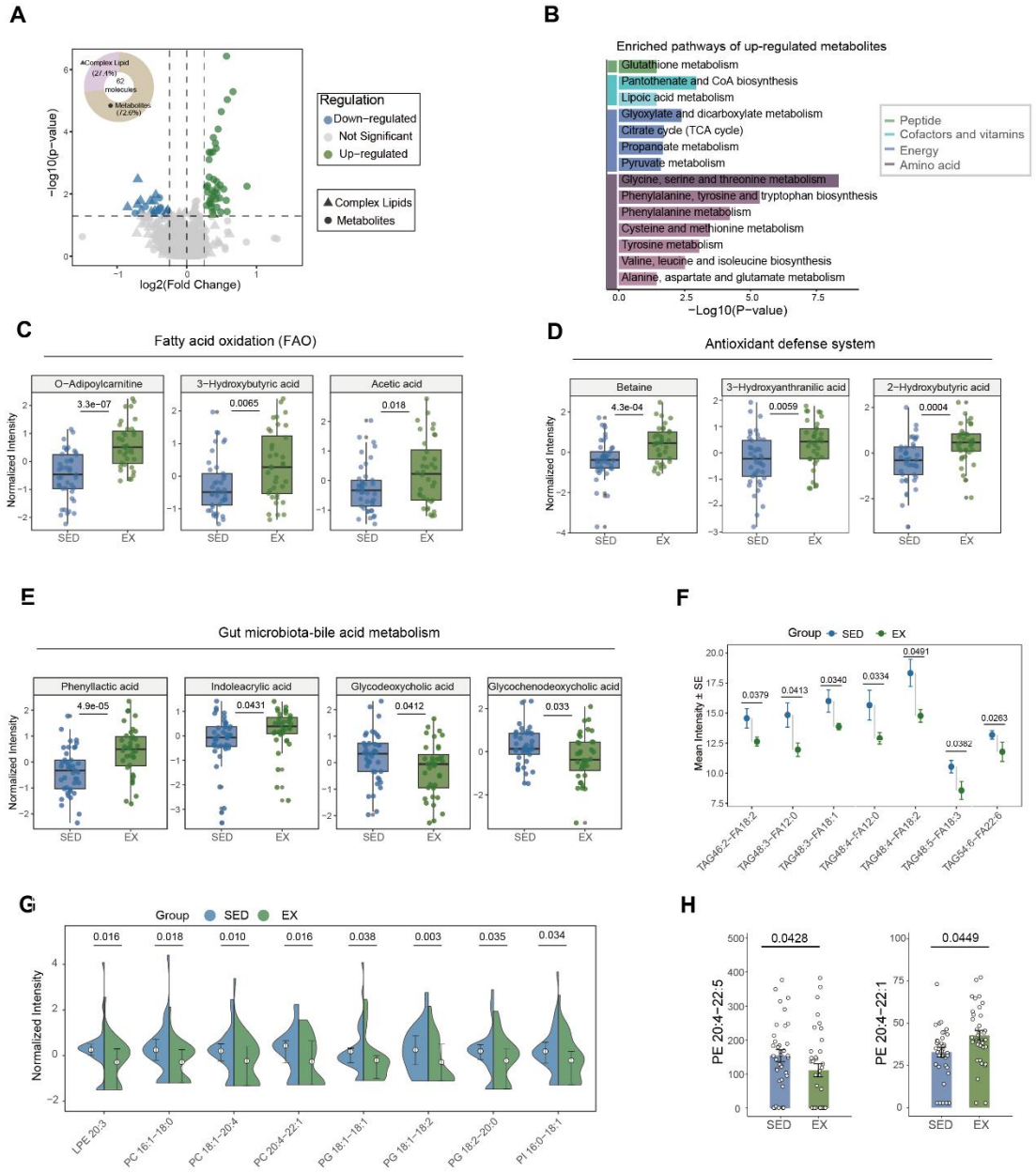
(F) UMAP plot depicting 644,045 immune cells from scATAC-seq datasets, where annotation labels were transferred from scRNA-seq data and each cell type was assigned a distinct color code. (match in D).

(G) Visualization of chromatin openness at key lineage gene promoter regions, including B cells (MS4A1), T cells (IL7R), CD4+ T cells (CD4), CD8+ T cells (CD8B), NK cells (KLRB1), and non-canonical T cells (RORC).

(H) Box plot showing the percentage of myeloid cells in the SED and EX groups.The distribution of myeloid cell percentages in the sedentary (SED, blue) and exercise (EX, green) groups is shown. The P value ($P < 0.05$) was derived from a linear model adjusted for age and sex, followed by Benjamini-Hochberg correction.

1217 (I) Differences in immune cell subsets abundance between EX and SED groups based

1218 on miloR analysis.



1219

1220 **Figure 2. Plasma metabolic and lipidomic changes induced by habitual physical**

1221 **activity.**

1222 (A) A volcano plot illustrating the differential abundance of plasma molecules

1223 (metabolites and complex lipids) between the EX and SED groups, the inset pie chart

1224 summarizes the composition of 62 differentially abundant molecules.

1225 (B) Pathway enrichment analysis of metabolites up-regulated in the EX group,
1226 highlighting significant enrichment in core energy metabolism pathways (e.g., TCA cycle),
1227 amino acid metabolism, and antioxidant defense pathways (e.g., glutathione metabolism).
1228 (C) Box plots showing the normalized intensity of key fatty acid oxidation (FAO)-related
1229 metabolites between the SED (blue) and EX (green) groups. P-values indicate the
1230 significance of group differences.
1231 (D) Regulatory analysis of antioxidants defense system-related metabolites (betaine,
1232 3-hydroxyindoleacetic acid, 2-hydroxybutyric acid) between the SED group (blue) and EX
1233 group (green). P-values indicate the significance of intergroup differences.
1234 (E) Changes in molecules associated with gut microbiota-bile acid metabolism between
1235 the SED group (blue) and EX group (green). P-values indicate significant differences
1236 between groups.
1237 (F) A box plot showing the mean intensity of distinct triglyceride (TAG) species between
1238 the SED (blue) and EX (green) groups. P-values indicate the significance of group
1239 differences.
1240 (G) Normalized Intensity Distribution of Membrane-Associated Complex Lipids in EX and
1241 SED.P-values indicate group difference significance, with violin plots showing intensity
1242 distribution.
1243 (H) Bar plots showing the intensity distribution of two PE species between the SED (blue)
1244 and EX (green) groups, with P-values denoting significance.

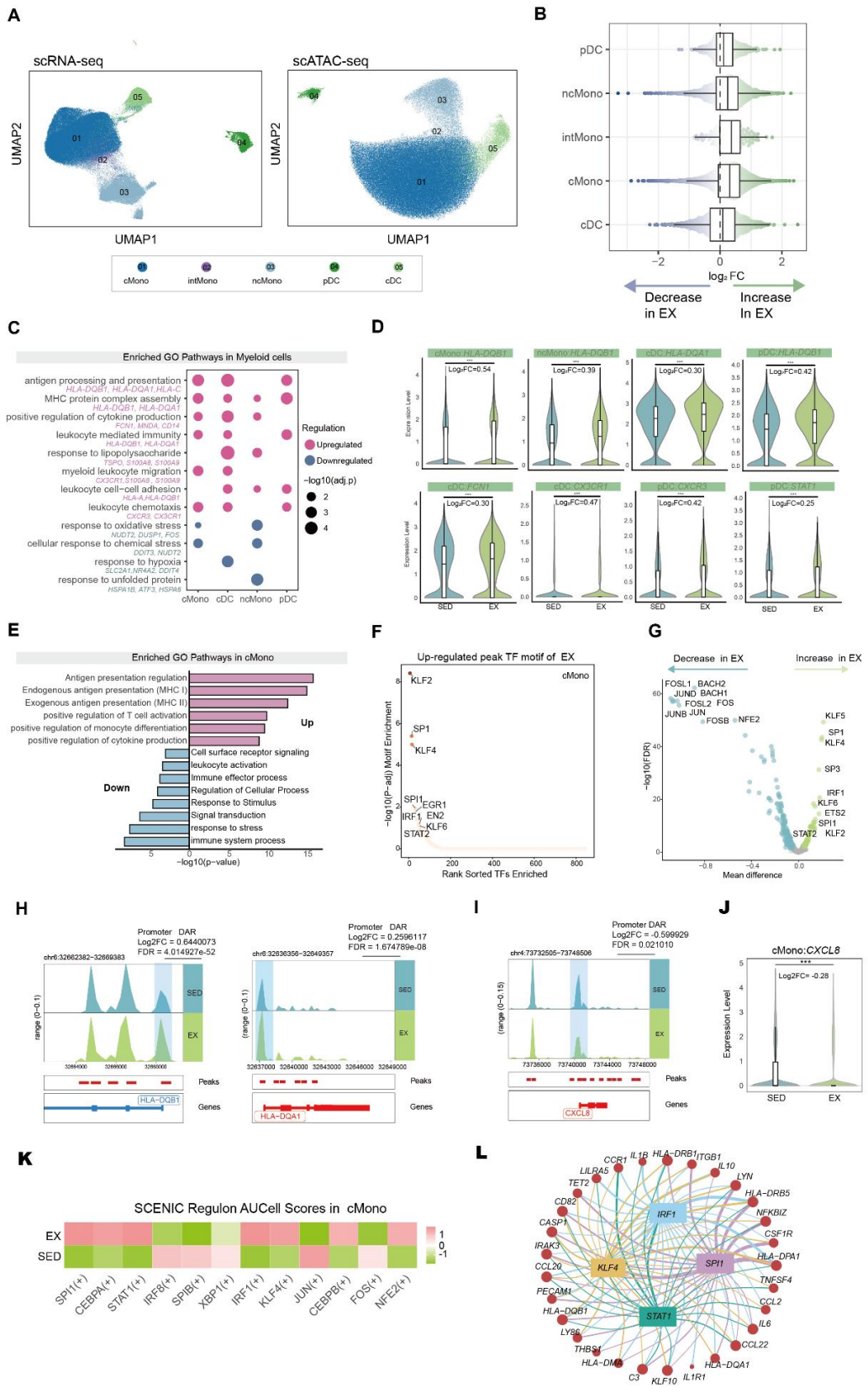


Figure 3. Multi-omics analysis of Myeloid cells heterogeneity in regular active exercise group (EX).

(A) UMAP plot displays scRNA-seq-derived Myeloid cells after subclustering and

1249 annotation, colored by five functional subsets(left). UMAP plot shows scATAC-seq-derived
1250 Myeloid cells, with annotation labels from corresponding scRNA-seq data and each cell
1251 type color-marked(right).
1252 (B) MiloR-based analysis was performed to assess differences in the abundance of
1253 Myeloid cell subsets between the EX and SED groups.
1254 (C) Enrichment analysis of Gene Ontology (GO) pathways in differentially expressed
1255 genes (DEGs) in each Myeloid cell subset, showing both upregulated and downregulated
1256 pathways in the EX group.
1257 (D) The violin plot showing changes in the expression levels of antigen
1258 presentation-related genes and cell migration-related genes within specific subsets,
1259 based on single-cell RNA data.
1260 (E) GO pathway enrichment analysis of differentially accessible regions (DARs) identified
1261 in classical monocytes (cMono) from scATAC-seq data.
1262 (F) Enrichment analysis of transcription factor (TF) binding motifs in DARs of cMono,
1263 revealing enriched TF motifs in the up-regulated chromatin-accessible peaks of the EX
1264 group.
1265 (G) Motif bias analysis showing the accessibility bias of key transcription factor binding
1266 motifs in cMono between the EX and SED groups. Positive values indicate increased
1267 accessibility in the EX group.
1268 (H) The chromatin accessibility at differential accessible regions (DARs) of HLA-DQB1
1269 (left) and HLA-DQA1 (right) between the SED and EX groups.
1270 (I) The chromatin accessibility at the differential accessible region (DAR) of CXCL8
1271 between the SED and EX groups.
1272 (J) Violin plot showing expression level of the pro-inflammatory gene CXCL8 in cMono of
1273 the EX group based on scRNA-seq data.
1274 (K) A heatmap showing SCENIC regulon AUCell scores (activity) of transcription factors
1275 (TFs) in cMono cells between the EX and SED groups.
1276 (L) A regulatory network plot showing the synergistic module of key TFs and their target
1277 genes in cMono cells.
1278

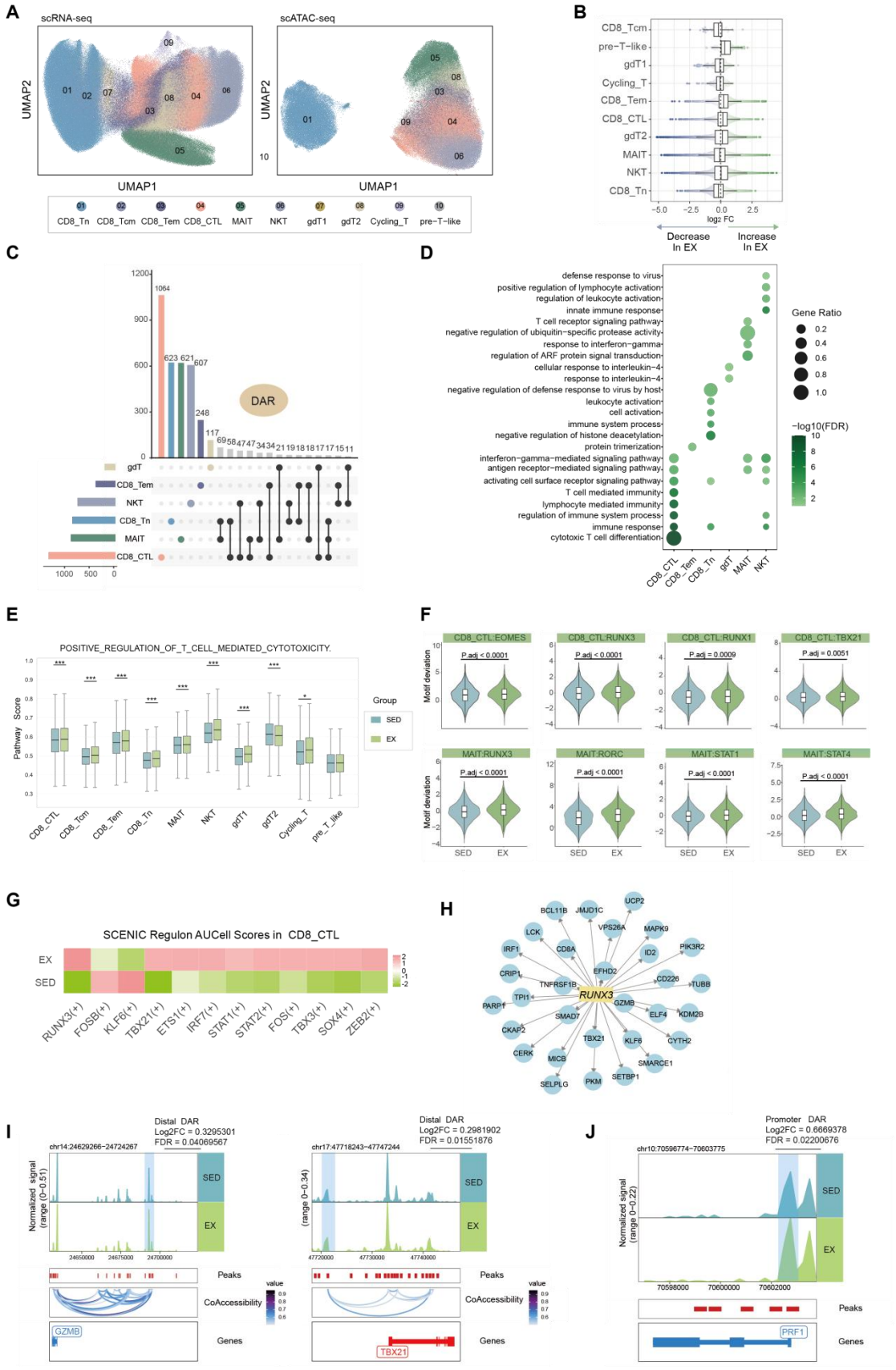


Figure 4. Epigenetic reprogramming enhances T cell functions in EX group.

(A) UMAP plot displays scRNA-seq-derived T cells after subclustering and annotation, colored by five functional subsets(left). UMAP plot shows scATAC-seq-derived T cells, with annotation labels from corresponding scRNA-seq data and each cell type color-marked(right).

(B) MiloR-based analysis was performed to assess differences in the abundance of T cell subsets between the EX and SED groups.

(C) Bar chart showing the number of upregulated differentially accessible regions (DARs) in each T cell subset from scATAC-seq data. The inset dot plot shows the proportion of DARs shared between two or more subsets.

(D) Enrichment analysis of Gene Ontology (GO) pathways in differentially expressed genes (DEGs) in each T cell subset, showing both upregulated and downregulated pathways in the EX group.

(E) Score distribution of characteristic gene sets in scRNA-seq data from T cell subpopulations differing between SED and EX groups.

(F) Violin plots showing the chromatin accessibility of core transcription factors (TFs) in CD8_CTL and MAIT cells between the SED and EX groups. P-adjusted (Padj) values indicate significance of group differences.

(G) A heatmap showing SCENIC regulon AUCell scores (activity) of transcription factors (TFs) in CD8_CTL cells between the EX and SED groups.

(H) Target gene regulatory network of RUNX3 in CD8_CTL cells.

(I) The chromatin accessibility at the differential accessible region (DAR) of GZMB(left) and TBX21(right) between the SED and EX groups in CD8_CTL cells,All significant peak–gene associations identified for each gene are presented alongside the corresponding track.

(J) The chromatin accessibility at the differential accessible region (DAR) of PRF1 between the SED and EX groups in MAIT cells.

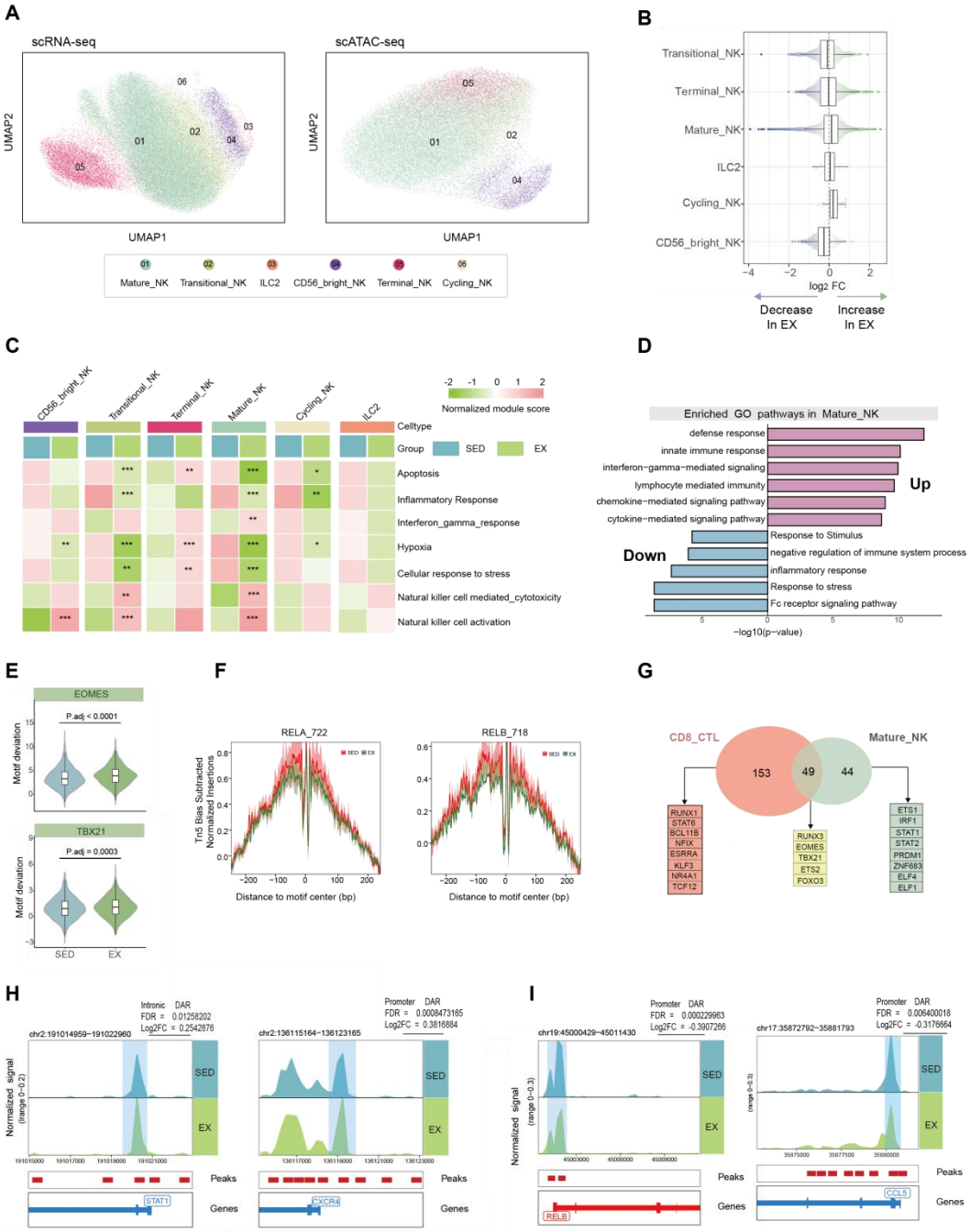


Figure 5. Multi-Omics characterization of NK cell heterogeneity in EX group.

(A) UMAP plot displays scRNA-seq-derived NK cells after subclustering and annotation, colored by five functional subsets(left). UMAP plot shows scATAC-seq-derived NK cells, with annotation labels from corresponding scRNA-seq data and each cell type color-marked(right).

(B) MiloR-based analysis was performed to assess differences in the abundance of NK cell subsets between the EX and SED groups.

(C) A heatmap showing normalized module scores of functional pathways across NK cell subsets between the SED and EX groups. Asterisks denote significance of group differences(*P<0.05,**P<0.01,***P<0.001).

(D) Enrichment analysis of Gene Ontology (GO) pathways in differentially expressed genes (DEGs) in each Mature_nk cell subset, showing both upregulated and downregulated pathways in the EX group.

(E) Violin plots showing the motif deviation of TFs (EOMES, TBX21) linked to NK cell effector function between the SED and EX groups. P-adjusted (Padj) values indicate significance of group differences.

(F) Footprint plots displaying the normalized Tn5 bias (y-axis) around the motif center for RELA and RELB in the SED and EX groups.

(G) Venn diagram showing the core transcription factors commonly upregulated in both CD8_CTL (adaptive immunity) and Mature_NK (innate immunity) cells of the EX group, as well as factors uniquely upregulated in each population.

(H) The chromatin accessibility at the differential accessible region (DAR) of STAT1(left) and CXCR4(right) between the SED and EX groups in Mature_NK cells.

(I) The chromatin accessibility at the differential accessible region (DAR) of RELB(left) and CCL5(right) between the SED and EX groups in Mature_NK cells.

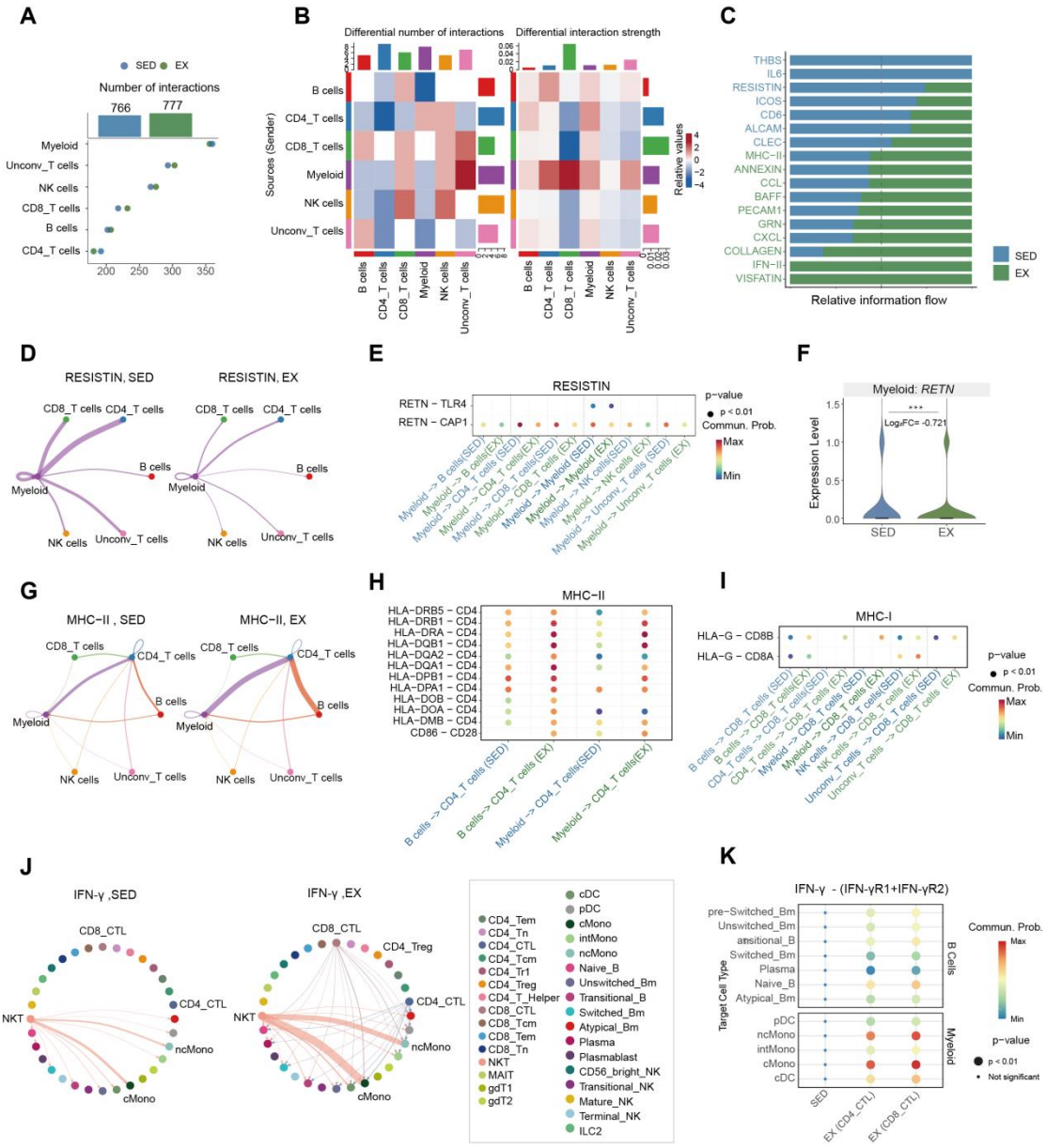


Figure 6. The impact of regular active exercise on the communication network among immune cells.

(A) A chart displaying the number of cell-cell interactions between immune cell subsets in the SED group and EX group. The top bar chart summarizes total interactions (766 in SED group, 777 in EX group), while the lower data points represent interaction counts for each subset.

(B) Heatmap showing the overall interaction strength between cell types in the SED (left) and EX (right) groups, computed using the CellChat algorithm.

(C) This diagram compares the information flow strength of selected signaling pathways (e.g., THBS, IL6, RESISTIN) between the SED (sedentary control, blue) and EX (exercise

intervention, green) groups.

(D) A regulatory network plot showing the cellular communication network for the RESISTIN signaling pathway in SED group(left) and EX group(right).

(E) Analysis of the key ligand-receptor axes in the RESISTIN pathway, comparing the interaction intensity for the RETN--CAP1 and RETN--TLR4 axes between the EX and SED groups.

(F) Violin plot comparing the expression level of the ligand gene RETN in myeloid cells between the two groups. Asterisks denote significance of group differences(***$P<0.001$), $Log_2FC=-0.721$.

(G) A regulatory network plot showing MHC-II-dependent signaling between immune cell subsets (nodes: cell types; edges: interaction strength) in the SED (left) and EX (right) groups.

(H) Cell-cell communication probability of MHC-II-related ligand-receptor pairs. Color depth reflects communication probability; larger bubbles mean more significant p - values.

(I) Cell-cell communication probability of MHC-I-related ligand-receptor pairs. Color depth reflects communication probability; larger bubbles mean more significant p - values.

(J) A regulatory network plot showing IFN-γ-dependent signaling (edges: interaction strength; nodes: cell types) in the SED (left) and EX (right) groups.

(K) A heatmap showing the communication probability (color gradient) and significance (markers) of the IFNG–(IFNGR1+IFNGR2) ligand-receptor axis: sources are cytotoxic T cells ($CD4^+$ CTL/$CD8^+$ CTL in EX group), and targets are myeloid/B cell subsets.

A

Participants' Characteristics at Baseline[a]

Characteristic	Physical Activity Pattern, No. (%) of Participants[b]			
	Total (n=86)	SED (n=46)	EX (n=40)	*p* value
Age, y, weighted mean (95% CI)	28.2 (27.2-29.2)	28.9 (27.5-30.3)	27.4 (26.0-28.8)	0.14
Sex, n (weighted %)				< 0.05
Male	28 (32.6)	9 (19.6)	19 (47.5)	
Female	58 (67.4)	37 (80.4)	21 (52.5)	
BMI, mean (SD)	21.3 (±3.0)	22.0 (±2.8)	20.7 (±3.0)	< 0.05
Lifestyle				
Smoker, n (weighted %)				0.26
Never	82 (95.3)	44 (95.7)	38 (95.0)	
Past	4 (4.7)	2 (4.3)	2 (5.0)	
Current drinker, n (weighted %)				0.15
Yes	16 (21.3)	5 (16.2)	11 (29.2)	
No	70 (78.7)	41 (83.8)	29 (70.8)	
Sleep duration, n (weighted %)				0.56
6-8h	72 (83.8)	40 (87.0)	32 (80.0)	
≥ 8h	14 (16.2)	6 (13.0)	8 (20.0)	
Dietary Habits, n (weighted %)				0.11
Balanced Diet	44 (51.1)	19 (41.3)	25 (62.5)	
Meat-Dominant Diet	12 (14.0)	9 (19.6)	3 (7.5)	
Other	30 (34.9)	18 (39.1)	12 (30.0)	
Medication Usage, n (weighted %)				0.08
No Medication	44(51.1)	28 (60.9)	16 (40.0)	
Used Medication	44(48.9)	18 (39.1)	24 (60.0)	

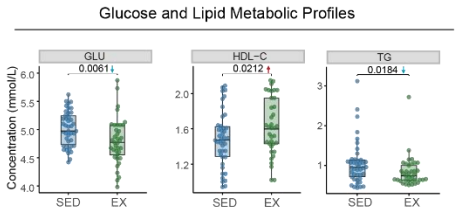
Abbreviations: BMI, body mass index(calculated as weight in kilogramsdivided by the height in meterssquared)

a Data were derived from the core cohort selected from the CIMA database, with a total of 86 participants included in this analysis.

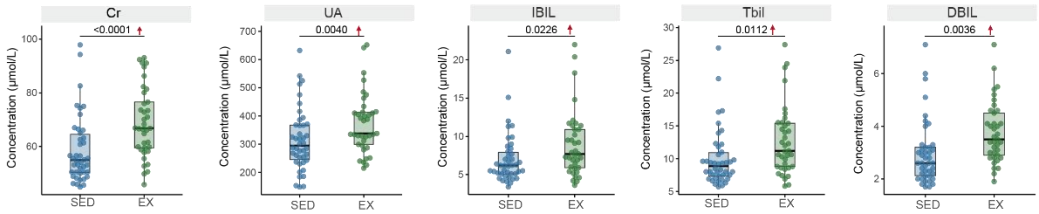
b Physical activity patterns were defined based on WHO criteria:
Sedentary (SED): Participants reported 0 minutes of moderate- or vigorous-intensity physical activity per week, and exhibited habitual sedentary behavior .
Regularly exercising (EX): Participants met the WHO recommended physical activity levels, i.e., at least 150 minutes per week of moderate-intensity activity, or at least 75 minutes per week of vigorous-intensity activity, distributed across ≥3 sessions.

c independent samples t-test within the same sex subgroup showed no significant difference in BMI between the two groups.

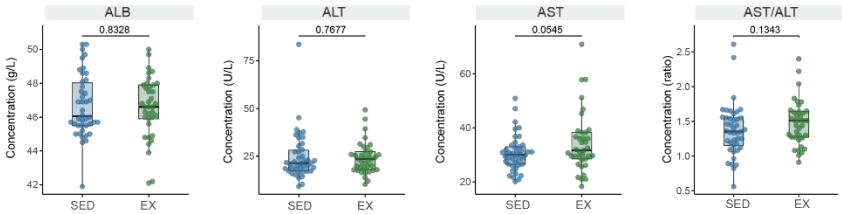
B



C



D



Supplementary Figure 1. Baseline characteristics and analysis of blood biochemical indices of the study population.

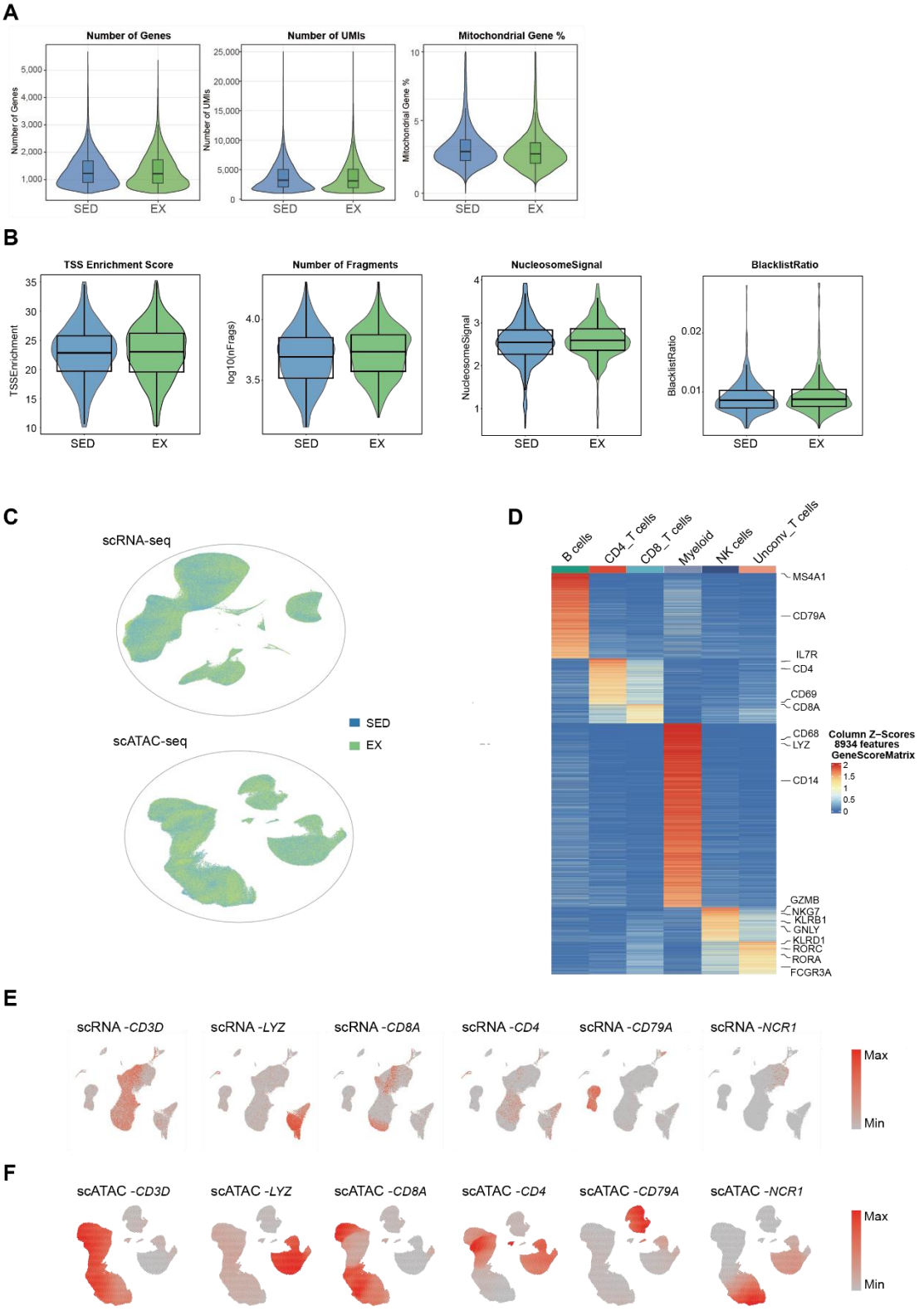
(A) A table summarizing the baseline characteristics of participants (n=86) stratified by physical activity pattern . The table includes metrics such as age, sex, BMI, lifestyle

factors (smoking, drinking, sleep), dietary habits, and medication usage, with weighted means/percentages and p-values indicating group differences.

(B) Box plots of blood metabolic indices between the EX and SED groups.Levels of fasting glucose (GLU), triglycerides (TG), and high-density lipoprotein cholesterol (HDL-C) are shown.

(C) Box plots of blood biochemical indices between the EX and SED groups. Levels of creatinine (Cr), uric acid (UA), total bilirubin (TBIL), direct bilirubin (DBIL), and indirect bilirubin (IBIL) are shown.

(D) Box plots showing the concentration of liver-related biochemical indicators (ALB: albumin; ALT/AST: hepatocellular injury markers; AST/ALT ratio) between the SED (blue) and EX (green) groups. The values above each plot denote p-values.



Supplementary Figure 2. Quality control, batch effect evaluation, and annotation validation for single-cell multi-omics data.

(A) Violin plots depicting key scRNA-seq quality control metrics (Gene Count, UMI Count, Mitochondrial Gene Proportion) between the SED (blue) and EX (green) cohorts;

consistent distribution of these metrics verified homogeneous data quality across groups.

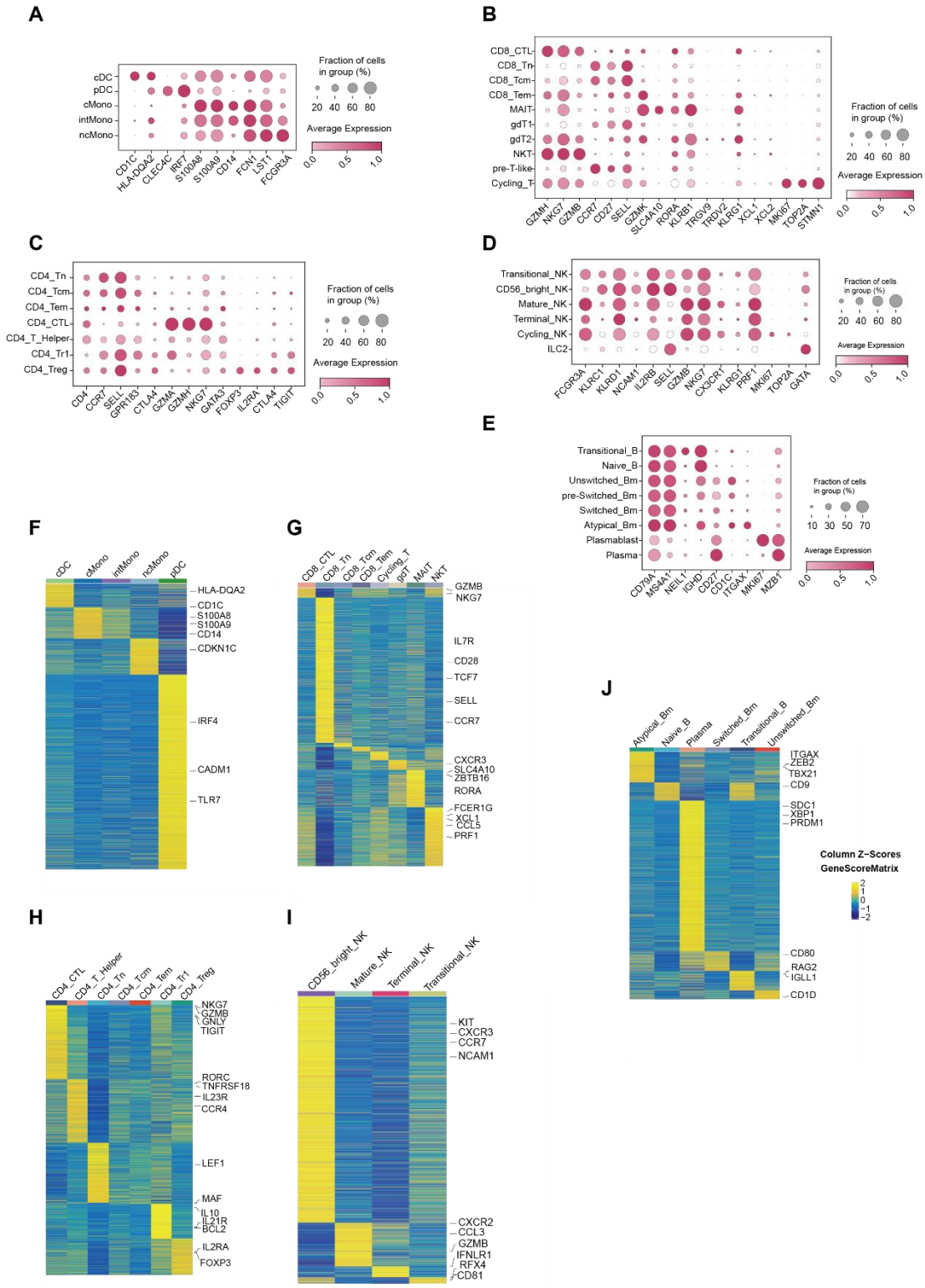
(B) Violin plots of scATAC-seq quality control indicators (TSS Enrichment Score, Number of Fragments, NucleosomeSignal, BlacklistRatio) comparing the SED and EX cohorts.

(C) UMAP plots displaying cell distribution in scRNA-seq (top) and scATAC-seq (bottom) datasets, with cells colored by group (SED/EX). Overlapping distribution of SED and EX cells indicates no prominent batch effects.

(D) Heatmap of main marker genes based on gene activity scores from scATAC-seq data, demonstrating high concordance between chromatin accessibility patterns and the annotated cellular identities across 30 cell subsets.

(E) UMAP plots showing the expression levels of key marker genes (e.g., CD3D, LYZ) in scRNA-seq data. The expression patterns align with the annotated cell subsets, supporting the cell typing results.

(F) UMAP plots showing chromatin accessibility at the promoter regions of key marker genes (e.g., CD3D, CD8A) in scATAC-seq data. The accessibility patterns match the corresponding cell subsets, validating the consistency between scATAC-seq and scRNA-seq annotations.



1403

1404 **Supplementary Figure 3. Visualization of marker genes: Dot plot from scRNA-seq**

1405 **and gene activity score heatmap from scATAC-seq.**

1406 (A) Expression dot plot of classical marker genes (e.g., CD14, FCGR3A, CD1C,

1407 LILRA4) for Myeloid cell subsets.

1408 (B) Expression dot plot of representative subset marker genes (e.g., CCR7, GZMB,

1409 KLRB1, TRDC) for $CD8^+$ T cell and non-conventional T cell subsets.

1410 (C) Expression dot plot of subset marker genes (e.g., FOXP3, IL7R, CXCR5, PRDM1) for

1411 the seven $CD4^+$ T cell subsets.

1412 (D) Expression dot plot of subset marker genes (e.g., NCAM1, KLRB1, FCER1A, IL1RL1)

1413 for NK cell subsets.

1414 (E) Expression dot plot of subset marker genes (e.g., IGHD, CD27, CD38, TNFRSF17)

1415 for B cell subsets.

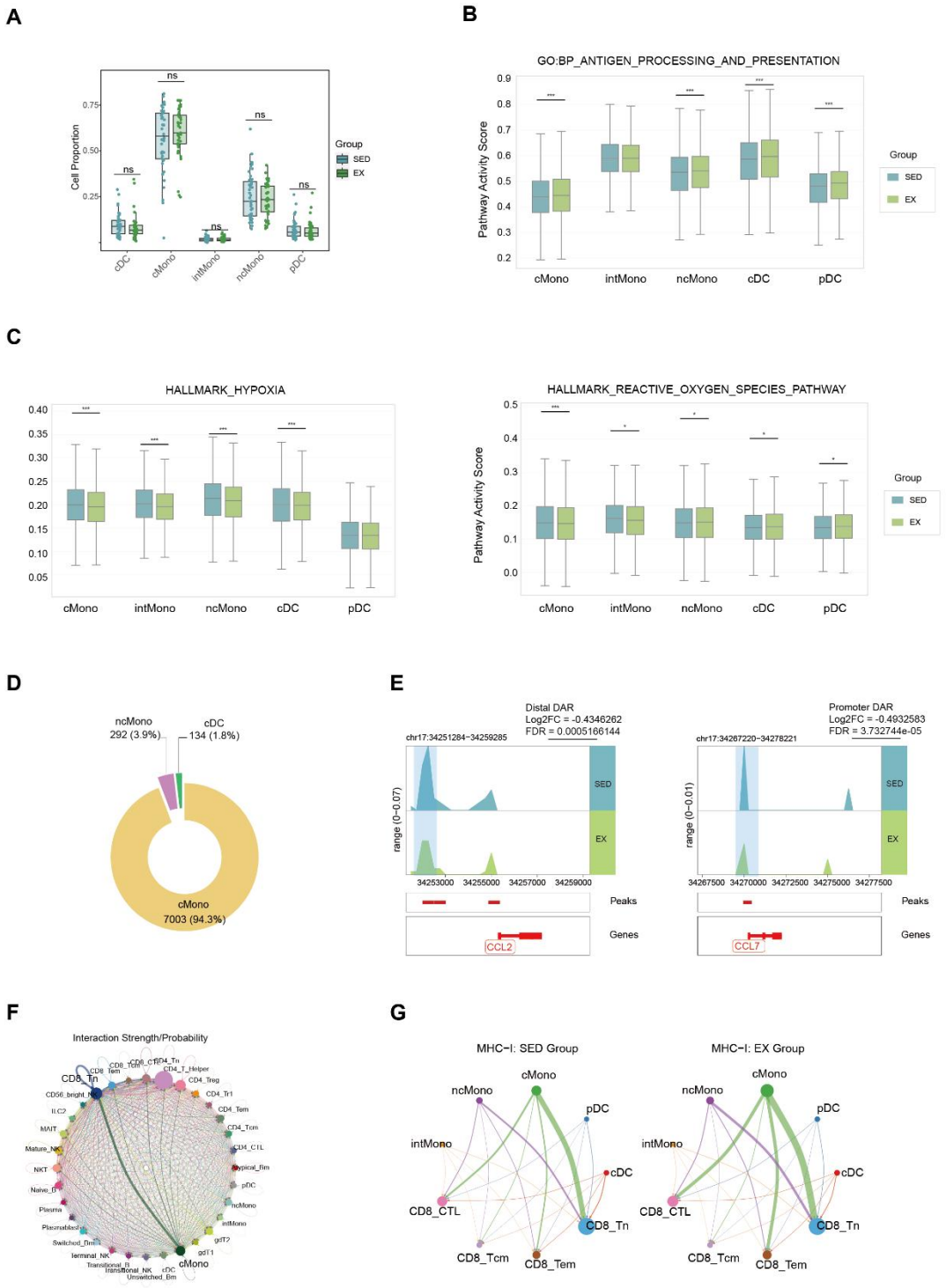
1416 (F) Gene activity score (GAS) heatmaps for Myeloid cell subsets.

1417 (G) GAS heatmaps for $CD8^+$ T&non-canonical T cells subsets.

1418 (H) GAS heatmaps for $CD4^+$ T cells subsets.

1419 (I) GAS heatmaps for NK cells subsets.

1420 (J) GAS heatmaps for B cells subsets

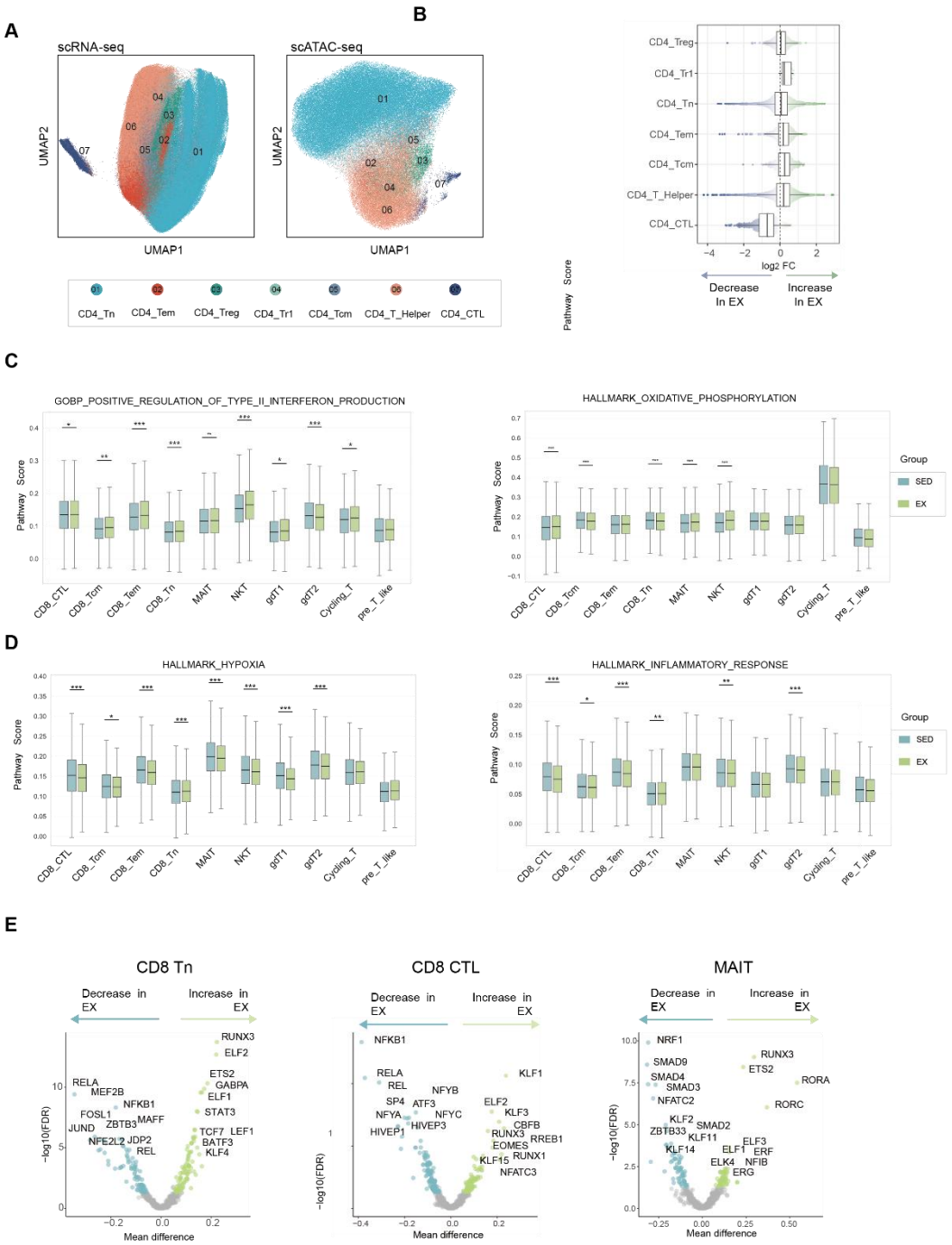


Supplementary Figure 4. Regulation of myeloid cell subset proportions, pathway activity, and chromatin accessibility by exercise.

(A) Box plots with overlaid scatter points showing the proportion of each myeloid cell subset (cDC, cMono, etc.) in the SED (blue) and EX (green) groups. “ns” indicates no significant difference between groups based on conventional testing.

(B) Bar chart showing the pathway activity scores for “GO:BP_ANTIGEN_PROCES

1428 SING_AND_PRESENTATION" in each myeloid cell subset for the SED and EX gro
1429 ups. The activity of this pathway was significantly upregulated in multiple subsets of
1430 the EX group.
1431 (C) Pathway activity scores for stress-related pathways. The left panel shows scores for
1432 "HALLMARK_HYPOXIA" and the right panel shows scores for
1433 "HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY". The activity of these two
1434 pathways was generally downregulated across myeloid subsets in the EX group.
1435 (D) Pie chart showing the distribution of differentially accessible regions (DARs) across
1436 myeloid cell subsets.
1437 (E) The chromatin accessibility at the differential accessible region (DAR) of
1438 inflammation-related genes between the SED and EX groups.
1439 (F) A network plot illustrating the interaction strength/probability between various immune
1440 cell subsets (nodes: cell types; edges: interaction connections).
1441 (G) A pair of cell-cell interaction network plots showing MHC-I-dependent signaling
1442 between myeloid cells (cMono, ncMono, etc.) and $CD8^+$ T cell subsets (CD8_CTL,
1443 CD8_Tn, etc.) in the SED group (left) and EX group (right).



Supplementary Figure 5. Regulation of CD4⁺ /CD8⁺ T cell subset proportions and functional pathways by exercise.

(A) UMAP plot displays scRNA-seq-derived CD4⁺ T cells after subclustering and annotation, colored by five functional subsets(left). UMAP plot shows scATAC-seq-derived CD4⁺ T cells, with annotation labels from corresponding scRNA-seq data and each cell type color-marked(right).

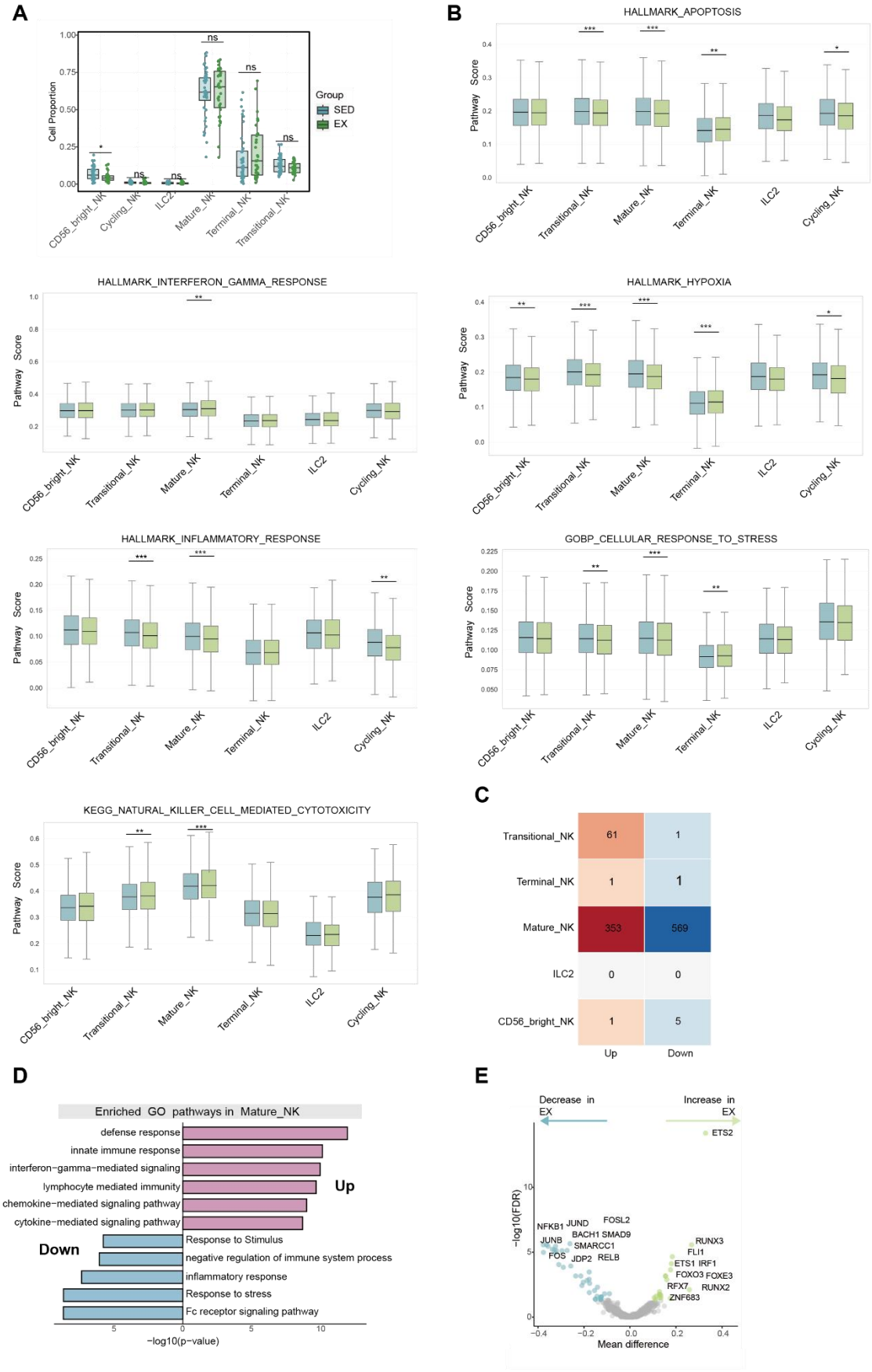
(B) MiloR-based analysis was performed to assess differences in the abundance of CD4⁺

T cell subsets between the EX and SED groups.

(C) Bar charts show the pathway activity scores for "GO:BP_POSITIVE_REGULATI ON_OF_TYPE_II_INTERFERON_PRODUCTION" and “HALLMARK_OXIDATIVE_PHO SPHORYLATION” in various T cell subsets (e.g., CD8_CTL, MAIT) for the SED an d EX groups.

(D) .Activity scores for stress-related pathways in T cells. The left panel shows sco res for "HALLMARK_HYPOXIA" and the right panel shows scores for "HALLMARK_ INFLAMMATORY_RESPONSE".

(E) Motif bias analysis of key transcription factors in CD8 Tn(left), CD8 CDL(middle), MAIT cells(right).


A
Cell Proportion
Group
SED
EX
CD56_bright_NK
Cycling_NK
ILC2
Mature_NK
Terminal_NK
Transitional_NK
B
HALLMARK_APOPTOSIS
HALLMARK_INTERFERON_GAMMA_RESPONSE
HALLMARK_HYPOXIA
HALLMARK_INFLAMMATORY_RESPONSE
GOBP_CELLULAR_RESPONSE_TO_STRESS
KEGG_NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY
Pathway Score
C
Up
Down
D
Enriched GO pathways in Mature_NK
defense response
innate immune response
interferon−gamma−mediated signaling
lymphocyte mediated immunity
chemokine−mediated signaling pathway
cytokine−mediated signaling pathway
Response to Stimulus
negative regulation of immune system process
inflammatory response
Response to stress
Fc receptor signaling pathway
−log10(p−value)
E
Decrease in EX
Increase in EX
−log10(FDR)
Mean difference

Supplementary Figure 6. Effects of exercise on NK cell subset proportions, functional pathways, and epigenetic regulation.

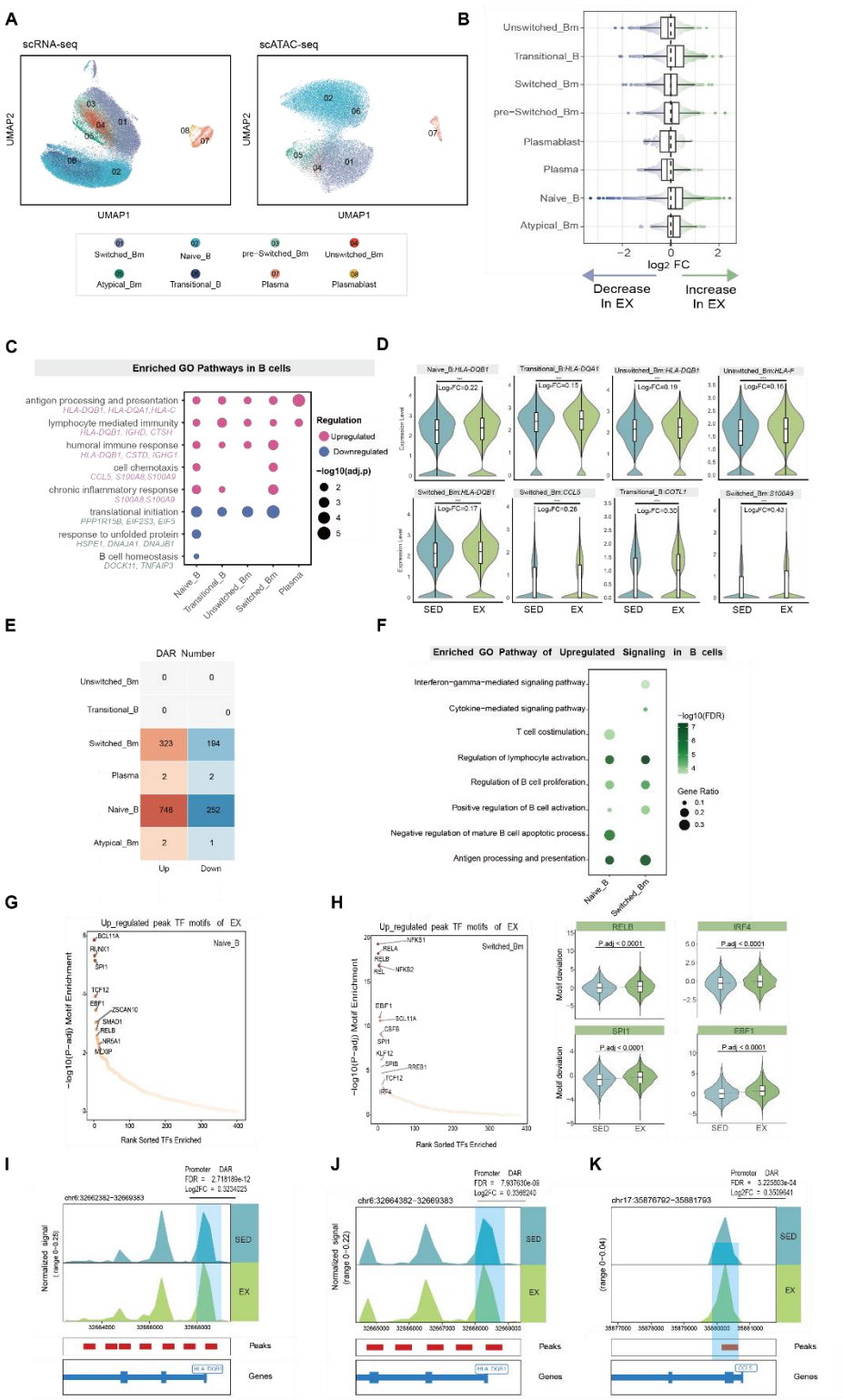
(A) Box plots with overlaid scatter points showing the proportion of each NK cell subset (CD56_bright_NK, Mature_NK, etc.) in the SED (blue) and EX (green) groups. "ns" indicates no significant difference. Asterisks denote significance of group differences(*P<0.05,**P<0.01,***P<0.001).

(B) Multiple bar charts showing pathway activity scores for various NK cell subsets in the SED and EX groups.

(C) Statistical table of differentially accessible regions (DARs) in NK cell subsets, showing the number of upregulated (Up) and downregulated (Down) DARs in each subset.

(D) DARs (e.g., immune defense, killing), and blue bars represent pathways associated with downreGene Ontology (GO) pathway enrichment analysis of DARs in Mature_NK cells. Red bars represent enriched pathways associated with upregulategulated DARs (e.g., stress response, inflammation).

(E) Motif bias analysis of key transcription factors in Mature_NK cells. The plot shows the distribution of $-\log_{10}$(p-value) versus the mean difference for TF motifs between the SED and EX groups.



Supplementary Figure 7. Effects of exercise on B cell subset proportions, functional pathways, and epigenetic regulation.

(A) UMAP plot displays scRNA-seq-derived B cells after subclustering and annotation, colored by five functional subsets(left). UMAP plot shows scATAC-seq-derived B cells, with annotation labels from corresponding scRNA-seq data and each cell type

color-marked(right).

(B) MiloR-based analysis was performed to assess differences in the abundance of B cell subsets between the EX and SED groups.

(C) Gene Ontology (GO) pathway enrichment analysis of DARs in B cells. Red bars represent enriched pathways associated with upregulated DARs (e.g., immune defense, killing), and blue bars represent pathways associated with downregulated DARs (e.g., stress response, inflammation).

(D) Violin plots showing the chromatin accessibility of antigen presentation-related genes (e.g., HLA-DQB1, CD86) in various B cell subsets between the SED (blue) and EX (green) groups; Log2FC values indicate fold changes in accessibility.

(E) Heatmap showing the number of upregulated (Up) and downregulated (Down) differentially accessible regions (DARs) in each B cell subset from scATAC-seq data.

(F) GO pathway enrichment analysis of upregulated signaling pathways in B cells.

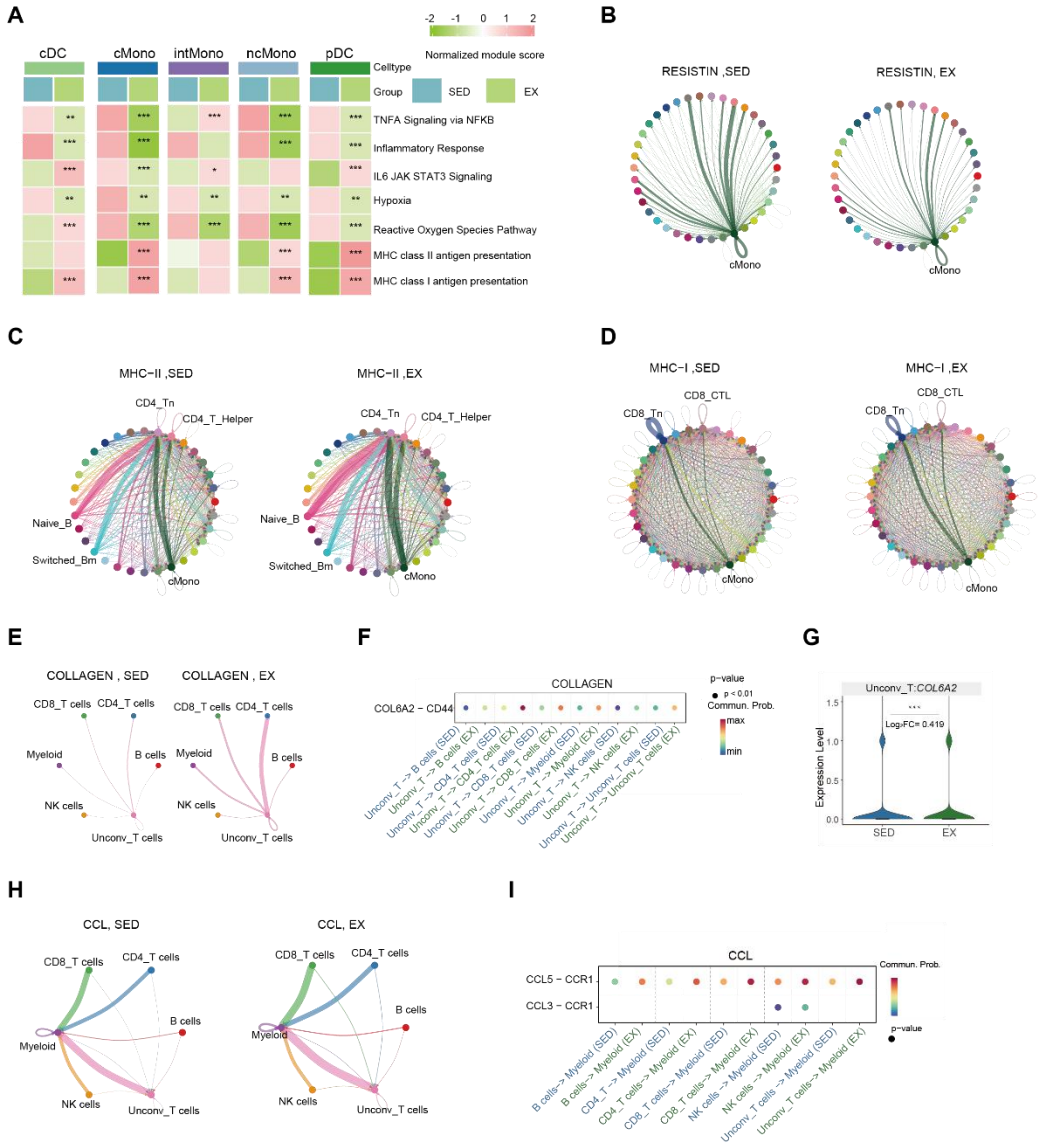
(G) Transcription factor (TF) binding motif enrichment analysis of upregulated DARs in Naive_B cells of the EX group; the y-axis denotes -log10(p-adj) (motif enrichment significance), and the x-axis represents rank-sorted enriched TFs.

(H) Left: TF binding motif enrichment analysis of upregulated DARs in Switched_Bm cells of the EX group. Right: Violin plots showing the motif deviation of key TFs in Switched_Bm cells between the SED and EX groups; Padj values indicate significance of group differences.

(I) The chromatin accessibility at the promoter DAR of HLA-DQB1 between the SED and EX groups in Switched_Bm cells; the track includes peak-gene links for the gene.

(J) The chromatin accessibility at the promoter DAR of HLA-DQA1 between the SED and EX groups in Switched_Bm cells.

(K) The chromatin accessibility at the promoter DAR of CCL2 between the SED and EX groups in Switched_Bm cells.



Supplementary Figure 8. Immune cell communication and pathway remodeling in the EX group

(A) Heatmap showing normalized module scores of functional pathways (e.g., TNFA Signaling via NFKB, MHC class II antigen presentation) across myeloid cell subsets (cDC, cMono, etc.) between the SED (blue) and EX (green) groups; asterisks denote significance of group differences (***P<0.001).

(B) Cell-cell interaction network plots of RESISTIN-mediated signaling in the SED (left) and EX (right) groups; nodes represent cell types, and edges represent interaction connections, centered on cMono.

(C) MHC-II-dependent cell communication network plots in the SED (left) and EX (right)

1524 groups; networks show interactions between B cells/Myeloid cells and $CD4^+$ T cell
1525 subsets, centered on cMono.
1526 (D) MHC-I-dependent cell communication network plots in the SED (left) and EX (right)
1527 groups; networks show interactions between various cell types and $CD8^+$ T cell subsets,
1528 centered on cMono.
1529 (E) COLLAGEN-mediated cell communication network plots in the SED (left) and EX
1530 (right) groups; networks show interactions between Myeloid/NK cells and B/T cells,
1531 centered on Unconv_T cells.
1532 (F) Heatmap showing communication probability of the COL6A2-CD44 ligand-receptor
1533 pair (COLAGEN pathway) across cell type interactions in the SED and EX groups; color
1534 intensity represents communication probability.
1535 (G) Violin plot showing the expression level of TCOL6A2 in Unconv_T cells between the
1536 SED and EX groups.
1537 (H) CCL-mediated cell communication network plots in the SED (left) and EX (right)
1538 groups; networks show interactions between Myeloid/B cells and T cells, centered on
1539 Unconv_T cells.
1540 (I) Heatmap showing communication probability of CCL5-CCR1 and CCL3-CCR1
1541 ligand-receptor pairs (CCL pathway) across cell type interactions in the SED and EX
1542 groups; color intensity represents communication probability, and black dots denote
1543 significance ($p<0.01$).