A novel sorting method reveals superior metabolic activity of mononucleated over binucleated tetraploid hepatocytes

Yusuke Watanabe[1], Jonathan H. Sussman[2], Tomoya Anami[3], Moena Hattori[3], Saaya Yamane[3], Masaki Nishikawa[3], Yasuyuki Sakai[1,3], Takeshi Katsuda[3,*]

[1]Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

[2]Medical Scientist Training Program, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

[3]Department of Chemical System Engineering, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

*Correspondence: Takeshi Katsuda, email: tkatsuda@g.ecc.u-tokyo.ac.jp

Author contributions:

Y.W., M.N., and T.K. conceived the study. Y.W., T.A., and T.K. performed the methodology. Y.W. and T.K. designed the experiments. Y.W., M.H., and S.Y. conducted the experiments. Y.W., J.H.S., and T.K. analyzed the bioinformatics and wrote the manuscript. M.N., Y.S., and T.K. acquired the resources and funding. T.K. supervised the study.

ABSTRACT

Hepatocytes display notable ploidy diversity, varying both in the number of genomic copies and in the number of nuclei. Most mouse hepatocytes are tetraploid (4n), which can exist as either mononucleated (1x4n) or binucleated (2x2n) cells. Despite this distinction, these two cell types have traditionally been grouped and studied as a single population. One likely reason for this is that conventional ploidy-sorting methods classify hepatocytes based solely on total DNA content, without distinguishing between mononucleated and binucleated states. In this study, we developed a novel flow cytometry strategy to distinguish and isolate 1x4n and 2x2n hepatocytes. Our approach leverages Hoechst-Area to assess total ploidy and Hoechst-Height to differentiate mononucleated and binucleated hepatocytes. Transcriptome analysis comparing these two populations revealed that 1x4n hepatocytes exhibit a broader and more metabolically active gene expression profile. Importantly, this enhanced metabolic activity was independent of liver zonation, a well-known driver of metabolic heterogeneity in hepatocytes. Our findings uncover a previously underappreciated layer of functional diversity in the liver and provide a new framework for studying the physiological and pathological roles of nuclear configuration in hepatocytes.

MAIN

Unlike most other mammalian cell types, hepatocytes exhibit variable numbers of genomic copies (ploidy). The most common state in mouse hepatocytes is a binucleated tetraploid cell with each nucleus containing two genomic copies (2x2n), while other cells are mononucleated with four genomic copies (1x4n). A smaller percentage of hepatocytes contain two tetraploid nuclei (2x4n) or one nucleus with 8 genome copies (1x8n). While tetraploid hepatocytes have traditionally been studied as a single entity, several recent studies reported that binucleated hepatocytes decrease in pathological conditions such as after a partial hepatectomy[1] and carbon tetrachloride (CCl_4)-induced injury[2], suggesting that there are functional differences between mononucleated and binucleated hepatocytes.

A major limitation in studying the differences between 1x4n and 2x2n hepatocytes has been the lack of a method to isolate them. A flow cytometry study in plants reported that DNA-stained pollen populations could be distinguished based on their fluorescence and intensity parameters when using propidium iodide (PI), allowing for selective sorting of mononucleated and binucleated populations[3]. We hypothesized that a similar approach could be applied to hepatocytes, using Hoechst 33342, a DNA dye compatible with live cells. First, we confirmed that the conventional gating strategy enabled the sorting of freshly isolated mouse hepatocytes with different ploidy states (**Figure 1A, Supplementary Figure S1**), yielding mononucleated or binucleated cells whose nuclear sizes increased in accordance with their expected ploidy states (**Figure 1B**). We then analyzed hepatocytes using Hoechst-Area, representing the total integrated fluorescence signal, and Hoechst-

60 Height, representing the maximum intensity of the signal pulse. As expected, 4n hepatocytes
61 separated into two distinct populations, one with a higher Hoechst-Height signal (termed 4n-
62 top) and one with a lower signal (termed 4n-bottom) (**Figure 1C**). To examine whether these
63 populations correspond to 1x4n and 2x2n hepatocytes, we sorted the whole 4n (4n-whole),
64 4n-top and 4n-bottom hepatocytes. The proportions of mononucleated and binucleated cells
65 were determined by microscopy after plating, confirming that 4n-top and 4n-bottom
66 populations were enriched for mononucleated and binucleated cells, respectively (**Figure**
67 **1D-E**). Specifically, the 4n-whole population consisted of 48.0% of mononucleated and 52.0
68 ± 3.7% of binucleated cells (mean ± SEM). In contrast, the mononucleated fraction in the 4n-
69 top population increased to 59.6 ± 1.6% while the binucleated fraction in the 4n-bottom
70 population reached 87.5 ± 2.1% of 2x2n cells (n = 9) (**Figure 1E**). To further test our
71 hypothesis that the 4n-top population is enriched for 1x4n hepatocytes, we performed three
72 rounds of sequential sorting. This gradually increased the proportion of 1x4n cells to 75.5 ±
73 1.2% (n = 3) (**Figure 1F-G**). In conclusion, our new gating strategy enables the selective
74 sorting of the 1x4n and 2x2n hepatocytes.

75 Using this sorting method, we next investigated differences of gene expression
76 profiles among hepatocytes with different ploidy. We sorted 2n-whole, 4n-whole, 4n-top and
77 4n-bottom populations and performed RNA sequencing (RNA-seq). For 4n-top hepatocytes,
78 we used cells after a single round of sorting, rather than repeated sorting, to ensure enough
79 cells for RNA-seq. We first performed a global principal component analysis (PCA) across
80 all the samples, which revealed a clear separation between the 2n-whole population and the
81 three 4n populations along the first principal component (PC1) (**Figure 2A**). A binary

comparison of the 4n-whole versus the 2n-whole population identified 1363 differentially expressed genes (DEGs, p-adj<0.05, |log2FC|>0.5) (**Figure 2B**) and 934 differentially enriched pathways (FDR<0.05) from the Gene Ontology Biological Process (GOBP) category using a Gene Set Enrichment Analysis (GSEA) (top enriched pathways shown in **Figure 2C**). These results indicated that our RNA-seq dataset reliably captures the difference between 2n and 4n cells, while representing the overall similarity among the 4n populations.

We next investigated differences between 1x4n and 2x2n hepatocytes. Notably, the 4n-bottom population was positioned closer to the 2n-whole population along PC1 than the other 4n populations were. This observation is consistent with a previous report suggesting that the gene expression profile of 2x2n nuclei is more similar to that of 1x2n nuclei than 1x4n nuclei[4]. Despite the overall similarity, a direct comparison between 4n-top and 4n-bottom populations identified 54 DEGs (**Figure 2D**, representative genes shown in **Figure 2E**). Moreover, GSEA revealed differential enrichment across multiple pathway categories, including 99 in the GOBP (top enriched pathways shown in **Figure 2F**), 14 in KEGG, 122 in Reactome and 10 in Hallmark (**Supplementary Figure S2A**). Notably, many pathways enriched in 4n-top were involved in metabolism, including *branched chain amino acid catabolism* and *porphyrin-containing compound metabolism*. Given that the liver is spatially portioned into three metabolic zones (known as zonation), we considered whether these differences were driven by zonation rather than ploidy. To assess the potential influence of zonation, we curated zone-specific gene lists based on published RNA-seq datasets[5,6] GSEA using these lists revealed that Zone 1 genes were enriched in 4n-bottom, while Zone 2 and 3 genes were enriched in 4n-top (**Figure 2G**). This result suggests that 2x2n

104 hepatocytes are more prevalent in Zone 1, whereas 1x4n hepatocytes are enriched in Zone 2
105 and 3, consistent with previous reports[4,7].

106 To further explore the relationship between ploidy and zonation, we examined
107 enrichment of zonation-associated metabolic pathways[8,9]. Unexpectedly, Zone 1-related
108 *fatty acid β-oxidation* and Zone 3-related *xenobiotic metabolism* were both enriched in 4n-
109 top hepatocytes (**Supplementary Figure S2B**). Moreover, 4n-top showed enrichment of
110 many other metabolism-related pathways that were not previously linked to zonation,
111 including *nucleoside bisphosphate metabolism* and *tetrapyrrole metabolism*
112 (**Supplementary Figure S2C**). In contrast, we found only two metabolic pathways enriched
113 in the 4n-bottom population: *formate metabolism* and *glutamine family amino acid*
114 *metabolism* (**Supplementary Figure S2D**). Taken together, these analyses suggest that
115 zonation alone cannot fully explain the observed differences between 1x4n and 2x2n
116 hepatocytes. Therefore, it is likely that these two populations have inherently distinct
117 metabolic programs, with 1x4n hepatocytes exhibiting broader and more active metabolic
118 profiles.

119 In summary, our novel sorting strategy enabled the selective isolation of 1x4n and
120 2x2n hepatocytes, allowing us to uncover transcriptomic differences between these two
121 populations. Although the greatest differences existed between 4n and 2n cells, 1x4n and
122 2x2n hepatocytes exhibited clear transcriptional differences, including more active metabolic
123 profiles in 1x4n hepatocytes compared to 2x2n hepatocytes. These findings motivate future
124 work on the roles of different nucleation states on cellular function in the liver. Recent
125 developments in transcriptional profiling, such as spatial transcriptomics will facilitate a

126 more granular dissection of the phenotypes contained within each ploidy state.

127

128 # METHODS

129 **Mice**

130 C57BL/6JJcl mice were purchased from CLEA Japan. Mice were fed with a standard
131 diet CE-2 (10 kGy) (CLEA Japan) and kept in a 12-hour light–dark cycle. 7- to 15-week-old
132 male mice were used for experiments.

133 **Hepatocyte isolation**

134 Hepatocyte isolation from the mouse liver was performed as previously
135 described[10]. Briefly, livers were perfused with HBSS (Wako), Liver Perfusion Medium
136 (Gibco), and HBSS with 5 mM CaCl2 (Wako) and 500 μg/ml collagenase (Wako) in order.
137 Following perfusion, livers were mechanically dispersed with tweezers, resuspended in wash
138 medium (DMEM (Wako) with 5% FBS (BioWest)) supplemented with 40 μg/ml DNase
139 (Millipore), and filtrated with a 70 μm cell strainer. The cells were centrifuged at 4 °C at 48
140 × g for 5 min. Then, the cells were resuspended in complete percoll solution (10.8 ml percoll
141 (Cytiva), 12.5 ml wash medium, and 1.2 ml 10× HBSS per liver) and centrifuged at 4 °C at
142 48 × g for 10 min. After a single wash with wash medium, hepatocytes and were use d for
143 the downstream experiments.

144 **Fluorescent staining**

145 Hepatocytes were suspended with DMEM supplemented with 10% FBS and 1 x
146 Penicillin-Streptomycin-Amphotericin B Suspension (P.S.A.) (Wako) at a concentration of
147 2e6 /ml. After 15 μg/ml Hoechst 33342 (Wako), 5 μM reserpine (Wako) and 40 μg/ml DNase
148 were added, hepatocytes were incubated at 37 °C for 30 min, with inverted every 10 minutes.
149 Following the centrifuge at 4 °C at 48 × g for 5 min, hepatocytes were resuspended with flow

150 buffer (HBSS supplemented with 1 M HEPES (Sigma), 1 M $MgCl_2$(Wako), 200 mg/ml
151 glucose (Sigma), 1 x P.S.A., 1 x MEM Non-essential Amino Acids Solution (Wako), 1 X
152 GlutaMAX Supplement (Wako), 1 mM sodium pyruvate (Wako), 40 μg/ml DNase, 15 μg/ml
153 Hoechst 33342 and 5 % FBS and pH was adjusted to 7.4 with NaOH (Wako)) at a
154 concentration of 1e7 /ml. Then, hepatocytes were filtrated with a 40 μm cell strainer, and 100
155 nm/ml TO-PRO-3 Iodide (Invitrogen) was added for staining dead cells.

156 **Flow cytometry**

157 Flow cytometry was performed using SH800S (Sony) with a 130 μm nozzle. Events
158 other than hepatocytes and doublets were removed by gating on BSC-A vs. FSC-A and FSC-
159 H vs. FSC-A, respectively. Dead cells were removed according to TO-PRO-3 fluorescence
160 by gating on FSC-H vs. APC-A. Then, hepatocytes were plotted by FSC-W vs. Hoechst
161 33342-A or Hoechst 33342-H vs. Hoechst 33342-A. Sorted cells were collected into 1:1
162 mixture of DMEM and FBS. After sorting, cells were immediately centrifuged at 4 °C at 48
163 × g for 5 min for 4n and 8n cells, or 800 × g for 10 min for 2n cells, respectively. Cells were
164 then resuspended with DMEM supplemented with 10% FBS and 1 x P.S.A. and kept on ice
165 before the downstream experiments.

166 **Counting of mononucleated and binucleated hepatocytes**

167 A portion of sorted hepatocytes were plated onto a 96-well plate with DMEM
168 supplemented with 10% FBS and 1 x P.S.A. Cells were incubated at 37 °C to for a few hours,
169 stained with 1 μg/ml Hoechst 33342 again and observed by fluorescence microscope BZ-
170 X810 (Keyence). The numbers of mononucleated and binucleated cells were counted,
171 respectively, and the proportions of these cells in sorted samples were calculated.

172 **RNA extraction**

173 80 ~ 300 K of hepatocytes were sorted by flowcytometry. RNA was extracted with
174 FastGene RNA basic kit including DNase I treatment. The concentration, A260/A280 and
175 A260/A230 of extracted RNA were measured with NanoDrop One (Thermo). RNA with low
176 purity was cleaned up with Monarch Spin RNA Cleanup Kit (10 μg) (New England Biolabs).

177 **RNA-seq**

178 Library preparation and sequencing were performed by Azenta (Tokyo, Japan) using
179 NovaSeq (Illumina).

180 **Bioinformatics for RNA-seq**

181 Reads were aligned to the mouse genome (GRCm39) using STAR aligner with
182 default parameters[11]. The gene-count matrix was produced by featureCounts[12]. Batch
183 effects due to individual differences among mice were adjusted using ComBat. Then, to
184 compare gene expression between samples, expression levels were normalized on the
185 "median of ratios" method using DESeq2[12]. Binary comparisons were performed using
186 the results() function in DEseq2. Differentially expressed genes were defined following the
187 criteria: adjusted p value < 0.05 and $|\log_2(\text{fold-change})| > 0.5$. Differentially expressed
188 pathways were identified by gene sets enrichment analysis with the cut-off values of false
189 discovery rate < 0.05.

190 **Preparation of lists of zonated genes**

191 For Zone1- and 3-related genes, the counts per million matrix from
192 PRJNA556572[5] was used. Eight layers along hepatic zones were merged in pairs to create

193 the expression values for four layers, and Kruskal-Walis test was performed for every gene.
194 188 of Zone1- and 259 of Zone3-related genes were identified by two criteria: FDR < 0.05
195 and the median values of the layers exhibit a monotonic increase or decrease. For Zone2-
196 related genes, we used 46 of previously identified genes by Halpern *et al*[6].

197

198 **ACKNOWLEDGEMENTS**

199 This work was supported by grants from the Uehara Memorial Foundation and the
200 Nakatomi Foundation.

201

202 **REFERENCES**

203 1. Sladky VC et al. *Developmental Cell* 2020; **52**: 335-349.e7.

204 2. Lin YH et al. *Gastroenterology* 2020; **158**: 1698-1712.e14.

205 3. Kron P et al. *Cytometry Part A* 2015; **87**: 943-957.

206 4. Yang L et al. *Developmental Cell* 2023; **58**: 1996-2010.e6.

207 5. Ben-Moshe S et al. *Nature Metabolism* 2019; **1**: 899-911.

208 6. Halpern KB et al. *Nature* 2017; **542**.

209 7. Richter ML et al. *Nature Communications* 2021; **12**.

210 8. Paris J et al. *Hepatology* 2022; **76**: 1219-1230.

211 9. Martini T et al. *Annu. Rev. Pathol. Mech. Dis* 2025; **46**: 439-466.

212 10. Katsuda T et al. *Nature Communications* 2024; **15**.

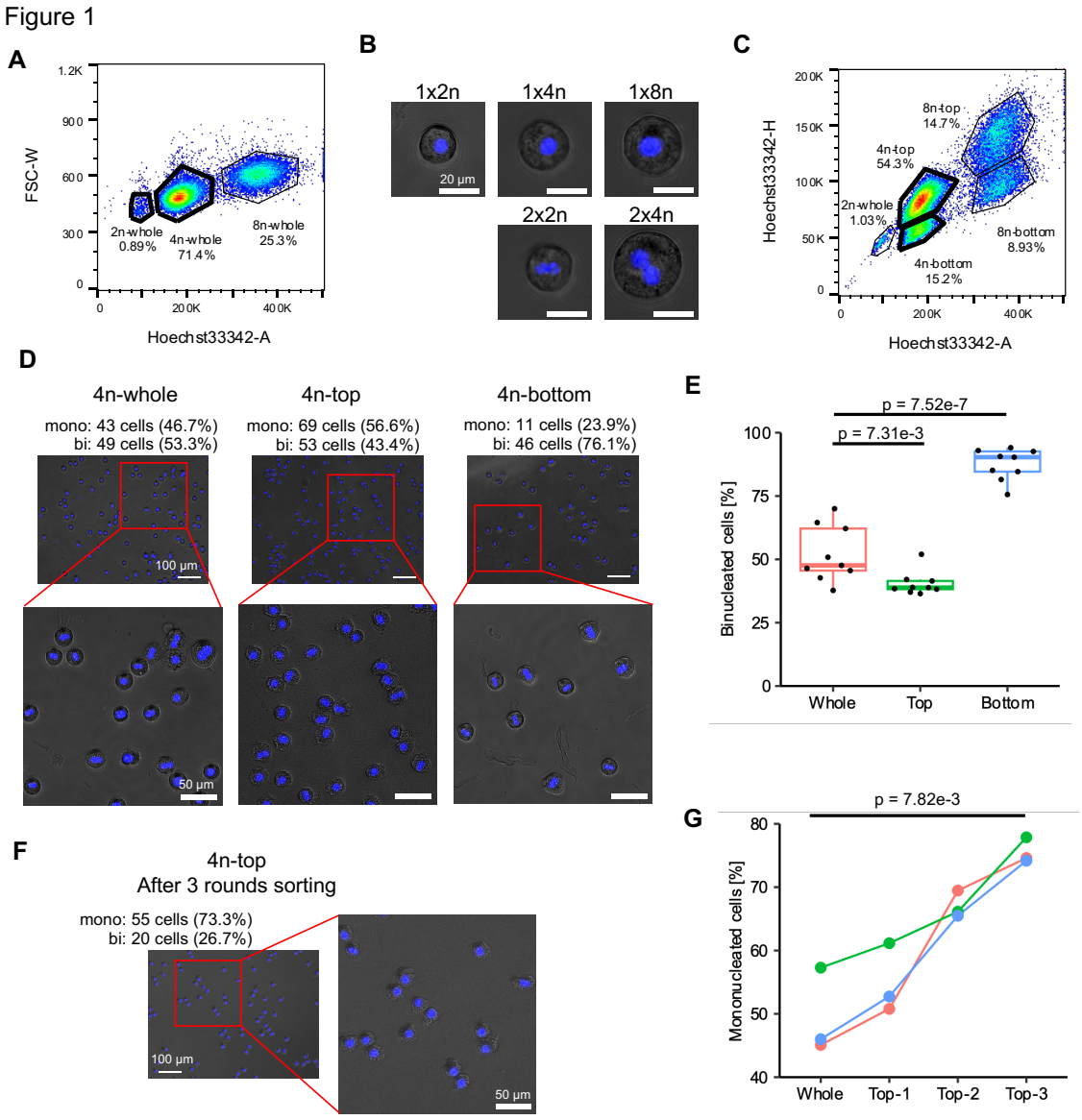
213 11. Dobin A et al. *Bioinformatics* 2013; **29**: 15-21.

214 12. Love MI et al. *Genome Biology* 2014; **15**.

215

216

217 **FIGURES**



218

219 **Figure 1. Development of a novel strategy for isolating 1x4n and 2x2n hepatocytes**

220 (A) The conventional gating strategy for sorting whole 2n, 4n and 8n hepatocytes.

221 (B) Representative images of sorted hepatocytes with distinct ploidy states.

222 (C) The newly developed gating strategy for presumable 1x4n and 2x2n hepatocyte

223 populations.

224 (D) Representative images of plated 4n hepatocytes sorted with the conventional (**A**) or new

225 (**C**) gating strategy.

226 (E) Proportions of binucleated cells in each of the sorted 4n populations (n = 9). Significance

227 was assessed using a one-sided Welch's *t*-test.

228 (F) A representative image of plated cells after three rounds of sorting of 4n-top hepatocytes.

229 (G) Proportions of mononucleated cells after three rounds of sorting of 4n-top hepatocytes

230 (n = 3). Significance was assessed using a one-sided Welch's t-test.

231

232

Figure 2

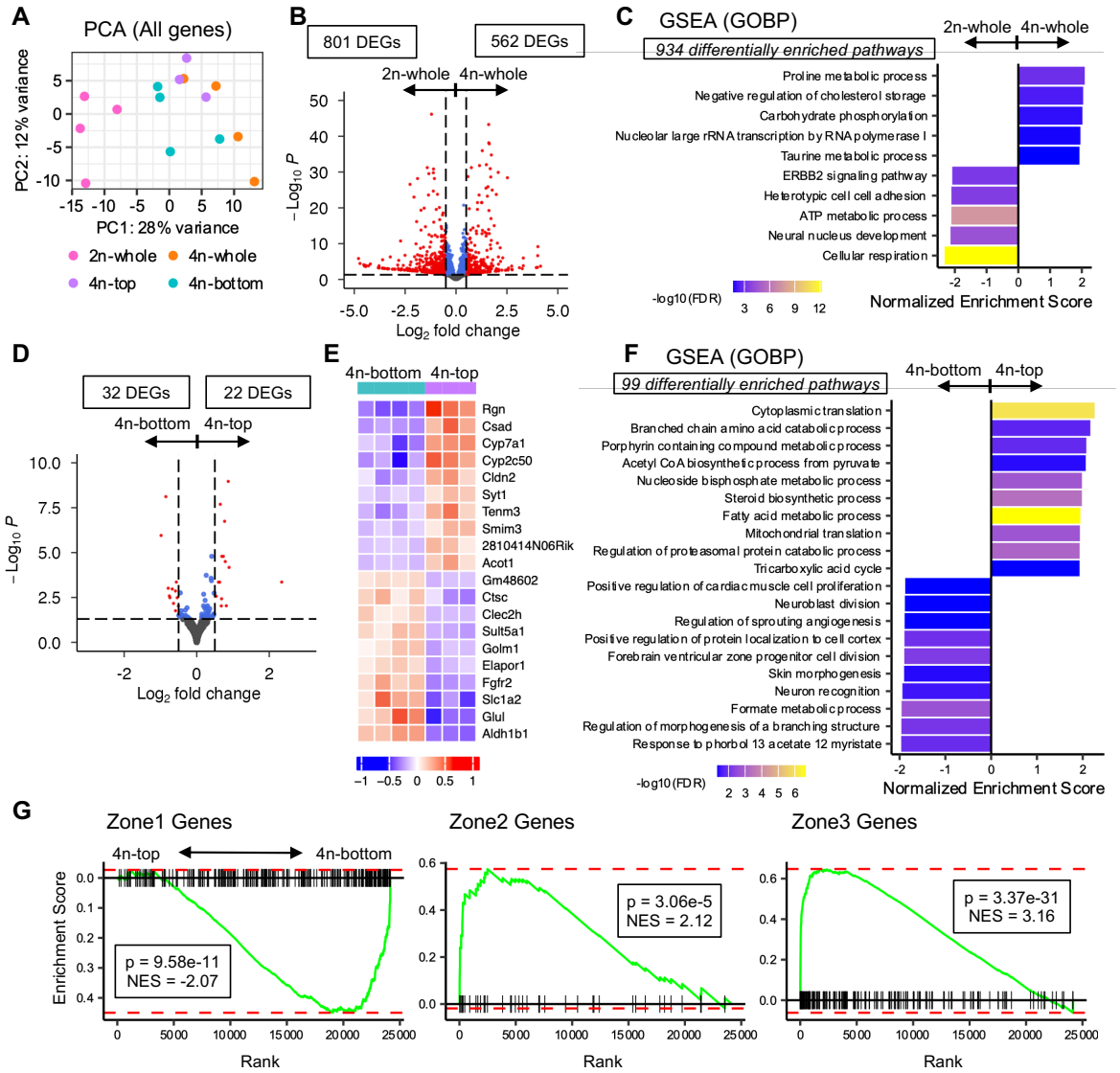


Figure 2. Transcriptomic analysis of hepatocytes with different ploidy states.

(A) Principal component analysis (PCA) across all samples.

(B) A volcano plot comparing gene expression between 4n-whole and 2n-whole hepatocytes. The log fold change shrunken data was plotted. Significant genes defined as p-adj < 0.05 and |Log2FC| > 0.5.

(C) Representative pathways differentially enriched between 4n-whole and 2n-whole

240 hepatocytes. The Gene Ontology Biological Process (GOBP) database was used. FDR, false
241 discovery rate.
242 (D) A volcano plot comparing 4n-top and 4n-bottom hepatocytes. The log fold change
243 shrunken data was plotted. Significant genes defined as p-adj < 0.05 and |Log2FC| > 0.5.
244 (E) A heatmap comparing 4n-top and 4n-bottom hepatocytes. Among the 54 DEGs, the top
245 10 genes enriched in 4n-top and 4n-bottom are shown, respectively. Genes are colored by z-
246 score normalized expression.
247 (F) Representative pathways differentially enriched between 4n-top and 4n-bottom
248 hepatocytes.
249 (G) GSEA using manually curated gene lists characteristic to each zone. The analysis was
250 conducted by comparing 4n-top (left) and 4n-bottom (right) hepatocytes. FDR, false
251 discovery rate. NES, normalized enrichment score.
252
253

SUPPLEMENTARY FIGURES

Supplementary Figure S1. Flow cytometry gating strategy.

Supplementary Figure S2. Additional GSEA results.

(A) Top enriched pathways by GSEA using KEGG, Hallmark, and Reactome categories.

(B-D) GSEA plots comparing 4n-top (left) and 4n-bottom (right). Zonated pathways (B). pathways enriched in 4n-top (C), and pathways enriched in 4n-bottom (D).