HAL
open science

Multiple metastatic clones assessed by an integrative multiomics strategy in clear cell renal carcinoma: a case study

Julien Dagher, Angelique Brunot, Bertrand Evrard, Solene-Florence Kammerer-Jacquet, Marion Beaumont, Laurence Cornevin, Fanny Derquin, Gregory Verhoest, Karim Bensalah, Alexandra Lespagnol, et al.

To cite this version:

Julien Dagher, Angelique Brunot, Bertrand Evrard, Solene-Florence Kammerer-Jacquet, Marion Beaumont, et al.. Multiple metastatic clones assessed by an integrative multiomics strategy in clear cell renal carcinoma: a case study. Journal of Clinical Pathology, 2022, 75, pp.426-430. 10.1136/jclinpath-2020-207326 . hal-03217234

HAL Id: hal-03217234

https://hal.science/hal-03217234v1

Submitted on 28 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **TITLE PAGE**

2 **Multiple metastatic clones assessed by an integrative multiomics strategy in clear cell**

3 **renal carcinoma: a case study.**

4

5 J Dagher[1,2], A Brunot[1,3], B Evrard[2], S-F Kammerer-Jacquet[1,2], M Beaumont[5], L Cornevin[5], F

6 Derquin[3], G Verhoest[4], K Bensalah[4], A Lespagnol[6], F Dugay[1,5], M-A Belaud-Rotureau[1,5], F

7 Chalmel[1,†,*], N Rioux-Leclercq[1,2,†].

8

9 **Affiliations** :

10 1 Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et

11 travail) - UMR_S1085, F-35000 Rennes, France

12 2 Pathology Department, University Hospital of Rennes, Rennes, France

13 3 Oncology Department, University Hospital of Rennes, Rennes, France

14 4 Urology Department, University Hospital of Rennes, Rennes, France

15 5 Department of Cytogenetics and Cell Biology, CHU de Rennes, Rennes, France

16 6 Department of Molecular Biology, CHU de Rennes, Rennes, France

17 *†These authors contributed equally to this work*

18

19 **Emails :**

20 Julien Dagher juliendagher@hotmail.com

21 Angélique Brunot angeliquebrunot@hotmail.com

22 Bertrand Evrard bertrand.evrard@inserm.fr

23 Solène-Florence Kammerer-Jacquet soleneflorence.kammerer-jacquet@chu-rennes.fr

24 Marion Beaumont marion.beaumont@chu-rennes.fr

Accepted manuscript

25 Laurence Cornevin laurence.cornevin@univ-rennes1.fr

26 Fanny Derquin fanny.derquin@hotmail.fr

27 Gregory Verhoest gregory.verhoest@chu-rennes.fr

28 Karim Bensalah karim.bensalah@chu-rennes.fr

29 Alexandra Lespagnol alexandra.lespagnol@chu-rennes.fr

30 Frédéric Dugay frederic.dugay@chu-rennes.fr

31 Marc-Antoine Belaud-Rotureau marc-antoine.belaud-rotureau@chu-rennes.fr

32 Frédéric Chalmel frederic.chalmel@inserm.fr

33 Nathalie Rioux-Leclercq nathalie.rioux-leclercq@chu-rennes.fr

34

35 *Corresponding author: IRSET, 9 Avenue du Professeur Léon Bernard, 35000 Rennes

36 Tel: +33 02 2323 5802; Fax: +33 02 2323 5055; Email: frederic.chalmel@inserm.fr

37

Accepted manuscript

ABSTRACT

The dynamics of metastatic evolution in clear cell renal cell carcinoma (ccRCC) is complex. We report a case study where tumor heterogeneity resulting from clonal evolution is a frequent feature and could play a role in metastatic dissemination.

We used an integrative multiomics strategy combining genomic and transcriptomic data to classify fourteen specimens from spatially different areas of a kidney tumor and three non-primary sites including a vein thrombus and two adrenal metastases.

All sites were heterogeneous and polyclonal, each tumor site containing two different aggressive subclonal populations, with differentially expressed genes implicated in distinct biological functions. These are rare primary-metastatic samples prior to any medical treatment, where we showed a multiple metastatic seeding of two subclonal populations.

Multiple interdependent lineages could be the source of metastatic heterogeneity in ccRCC. By sampling metastases, patients with resistance to therapies could benefit a combination of targeted therapies based on more than one aggressive clone.

Keywords

Kidney neoplasms; metastasis; tumor heterogeneity.

Accepted manuscript

56 **BACKGROUND**

57 Clear cell renal cell carcinoma (ccRCC) is the most frequent histological subtype of renal
58 cancer, frequently metastatic. Inter- and intra-tumor heterogeneity is a usual landscape in
59 ccRCC [1–3], and is the result of a continuous genetic diversity through clonal evolution [4,5].
60 The origins of metastatic heterogeneity through clonal evolution could be represented by drug-
61 resistant subclones and play a role in resistance to treatment.

62 To address this issue, we elaborated an experimental scheme in a case-study of a patient, by
63 performing an integrative multiomics clustering strategy combining cytogenetic and
64 transcriptomic technologies. Matched primary and multiple synchronous metastases from
65 different sites, before the interference of any medical treatment, revealed a complex
66 dissemination of multiple tumor clones, which could provide new insights into metastatic
67 interactions.

68

Accepted manuscript

69 **METHODS**

70

71 **Patient**

72 The four specimens were processed in sterile conditions, and multiple biopsies were taken from
73 each tumor (0.5 cm^3 each). Tumor cell content was assessed for each sample. The study was
74 approved by our local ethics committee.

75

76 **Array-CGH and GeneChip hybridation and data preprocessing**

77 *Array-CGH hybridization and raw data preprocessing.* DNA samples were hybridized to
78 Agilent Human Genome CGH microarray 180K (Agilent Technologies, Santa Clara, CA,
79 USA) as described in (Supplementary information, SI). Array-CGH data were normalized,
80 quality controlled (SI), Fig. S1A) and chromosomal imbalances (represented by SNOC regions,
81 for smallest non-overlapping chromosomal regions of deletions or amplifications) were
82 detected as described in SI

83 *GeneChip hybridization and data pre-processing.* Total RNA samples were hybridized to
84 Human Transcriptome Array 2.0 GeneChips as recommended by the manufacturer
85 (Affymetrix, Santa Clara, USA). The resulting CEL files were normalized, quality controlled
86 (SI, Fig. S1B) as described in SI.

87 Both ArrayCGH and GeneChip data were uploaded to the NCBI Gene Expression Omnibus
88 (GEO) repository under the accession numbers GSE113205 and GSE113204 [6].

89

90 **Multiple factor analysis and sample classification**

Accepted manuscript

91 We used the multiple factor analysis (MFA) function implemented in the *FactoMineR* package
92 [7] to classify samples based on both array-CGH and transcriptomic normalized data. This
93 method is an extension of the Principal Component Analysis (PCA) and allows studying tables
94 in which individuals (samples) are described by several groups of (quantitative and/or
95 qualitative) variables; i.e. SNOC regions for array-CGH data and gene expression for
96 transcriptomic data. Only SNOC regions showing at least a variation range of 10% across
97 samples and genes showing a minimal signal intensity fold change of 1.5 across all samples
98 were used for MFA. The first components explaining >90% of the information (variances)
99 were retained (Fig. 1, panel B). A hierarchical clustering based on the selected components
100 was used to estimate the degree of association between samples using the HCPC function
101 implemented in *FactoMineR* (Fig. 1, panel C). Samples within the resulting dendogram tree
102 were then automatically partitioned in classes based on both combined genomic and
103 transcriptomic data using the HCPC unsupervised method with default parameters (the nb.clust
104 parameter was set to -1 so that the tree is automatically cut at the level suggested by HCPC).

105

106 **Statistical filtration and cluster analysis**

107 *Array-CGH data analysis.* The statistical filtration of the SNOC regions displaying a
108 significant gain or loss between sample classes (C1-C3) (Fig. S2A) was performed using the
109 Annotation, Mapping, Expression and Network (AMEN) suite of tools [8]. Briefly, we first
110 filtered SNOC regions with a minimal gain or loss of at least 10% across classes. Finally, a
111 statistical test implemented in the LIMMA package (F-value adjusted with the false discovery
112 rate method: $P \leq 0.05$) was used to identify significantly differential regions across sample
113 classes [9]. The resulting SNOC regions were then clustered into groups using the k-means
114 algorithm implemented in AMEN. The ability of these clusters to discriminate between SNOC

Accepted manuscript

115 regions was verified using a silhouette plot. The resulting patterns were ordered according to
116 peak gain or loss levels in the distinct sample classes (Fig. S2, panel A).

117

118 *Transcriptomic data analysis.* Similar to Array-CGH data analysis, the statistical filtration of
119 the genes differentially expressed (DE) among the three sample classes (C1-C3) (Fig. S2B)
120 was performed using AMEN [8]. Briefly, we first filtered genes with at least one signal above
121 the background expression cutoff (BEC = 0.189, corresponding to the overall median intensity)
122 and with a minimal fold change of 1.5 across sample classes. Finally, a statistical test
123 implemented in the LIMMA package (*F*-value adjusted with the false discovery rate method:
124 $P \leq 0.05$) was used to identify significantly DE genes across sample classes [9]. The resulting
125 genes were then partitioned into gene expression clusters using the k-means method
126 implemented in R. The ability of these clusters to discriminate between genes was verified
127 using a silhouette plot. The resulting patterns were ordered according to peak expression levels
128 in the distinct sample classes (Table S1; Fig. S2, panel B).

129

130 **Functional analysis**

131 An enrichment analysis was used to measure the association between each gene expression
132 cluster (G1-G3) and gene ontology terms (biological process and cellular component) [10].
133 Briefly, enrichments were estimated by calculating the Fisher exact probability using the
134 Gaussian hypergeometric test implemented in the AMEN suite [7]. A given annotation term
135 was considered enriched in a group of genes when the adjusted *P*-value (adjusted with the false
136 discovery rate method) was ≤ 0.05 and the number of genes in this group bearing this
137 annotation term was ≥ 5 (Table S2).

Accepted manuscript

138 **RESULTS**

139

140 **Patient outcome and histology**

141 The patient developed these synchronous tumors in her late 60s. Further metastases occurred
142 6 months after surgery. No immunotherapy was conducted. Anti-angiogenic therapy
143 commenced in 2016, but was stopped due to severe adverse effects. Progressive disease was
144 confirmed radiologically and the patient died later that year.

145 We analyzed fourteen spatially separated samples (Fig. 1, panel A; Fig S3), originating from
146 the primary kidney tumor (five samples p1 to p5), a vena cava tumor thrombus (two samples
147 m1 and m2) and bilateral metastatic adrenal glands (four samples m3 to m6 from the right
148 adrenal and three samples m7 to m9 from the left adrenal). All samples were of overall high
149 grade (WHO/ISUP 3 or 4) even though primary samples were predominantly grade 2.
150 However, the morphological aspect of tumor cells was different, especially in sample *p2* and
151 some metastases, more frequently composed of rhabdoid and large pleomorphic cells, with
152 abundant cytoplasm and tumor necrosis (Fig S3). All samples showed the same inactivating
153 mutation of *VHL* and a variable loss of the short arm of chromosome 3 (SI).

154

155 **The integrative omics approach combining genomic and transcriptomic data enabled the**
156 **characterization of three sample classes.**

157 Given this uncommon collection with no drug interference, samples were assessed for
158 molecular heterogeneity. While clustering omic dataset separately can reveal patterns in the
159 data, integrative clustering combining several omics has the potential to expose more fine-
160 tuned structures that are not revealed by examining only a single data type [11]. The fourteen
161 samples were all analyzed by ArrayCGH and GeneChip microarrays (SI). The multiple factor

Accepted manuscript

162 analysis (MFA) based on both combined genomic and transcriptomic data followed by an
163 unsupervised clustering method automatically partitioned samples into three classes (termed
164 C1-C3). The resulting factor map and dendrogram tree are illustrated in Figure 1 (panels B and
165 C). As shown in panel C, C1 is only composed of primary samples (p1, p3 and p5), whereas
166 C2 (p4; m1, m3, m8 and m9) and C3 (p2; m2, m4, m5, m6 and m7) are composed of only one
167 primary sample, and all metastatic samples originating from the three metastatic sites.

168

169 **Significant chromosomal imbalances between sample classes contribute to genetic**
170 **intratumor heterogeneity**

171 The detailed genomic analysis of the fourteen samples identified 468 chromosomal imbalances
172 (CIs) (SI). The vast majority of CIs corresponded to region losses compared to the control
173 DNA. The mean number of chromosomal imbalances in metastases and primary tumors was
174 not significantly different (149 and 143 respectively). Further statistical analysis revealed that
175 among the 468 CIs, 91 showed significant imbalances between the three sample classes C1,
176 C2 and C3. These were partitioned into three chromosomal region patterns (R1-R3) (Fig. 2,
177 upper part). The R1 group was composed of 36 regions lost in C1, C2 and, to a lesser extent,
178 C3 (Fig S4).

179

180 **The three sample classes are associated with distinct transcriptional patterns**

181 We identified 1,008 genes showing a significant differential expression (DE) between the three
182 sample classes (SI). These DE genes were further clustered into three gene expression groups
183 (G1-3) (Fig. 2, lower part).

Accepted manuscript

184 The G1 group, containing 185 genes, showed a peak expression in class C1 comprised only of
185 primary samples. Functional analysis revealed that this group is significantly associated with
186 genes involved in immune response and T cell activation, but also in chemotaxis and plasma
187 membrane proteins. The G2 group, comprising 427 genes, had a high expression in C1 but
188 more importantly in class C2 compared to C3. This group is significantly enriched in genes
189 involved in B cell mediated immunity, and phagocytosis, fatty acid oxidation, cell adhesion,
190 kidney development and tryptophan metabolism. Finally, the G3 group, containing 396 genes
191 with a peak expression in C3 compared to the two other classes, is significantly associated with
192 various biological functions such as angiogenesis, extracellular matrix organization,
193 chemotaxis, cell adhesion, as well as kidney, neuron and bone development.

194

Accepted manuscript

195 **DISCUSSION**

196 This case study describes a rare collection of samples from *primary-thrombus-metastatic* pre-
197 therapeutic tumors, as most metastatic patients undergo medical treatment before surgical
198 excision. An integrative and unsupervised clustering strategy combining genomic and
199 transcriptomic data revealed three heterogeneous sample classes, suggesting a lineage
200 relationship between the "aggressive" subclones found in metastases. Although some
201 divergence exists between genomic and transcriptomic data, the combined analysis of the two
202 techniques separated two clones in the primary tumor that resemble metastatic samples, with
203 molecular heterogeneity in each site.

204 The spatial distribution analysis revealed that all metastatic sites were polyclonal, each
205 containing simultaneously both subclonal populations from C2 and C3. Further functional
206 analyses demonstrated that genes differentially expressed in the two subclones are associated
207 with distinct biological functions [12]. Two types of aggressive clones (multiple
208 dissemination), could be retraced back to the primary tumor. One hypothesis could be a
209 polyclonal seeding with interactions between each sub-population. [13,14] (Fig 3).

210 A metastasis at a certain time is a snapshot of a group of cells that have overcome the barriers
211 against metastatic dissemination. Dominant clones appear at an early phase in tumorigenesis,
212 and minor clones which could be phenotypically different could be observed after late
213 dissemination [15]. These clones could be mixed or separated, and the clonal architecture of a
214 given tumor varies with time [16]. However, in most models every lineage is separate and the
215 notion of complex multiple clones acting together in a same metastasis is rarely considered.

216 In a recent paper, genetic heterogeneity in primary samples of ccRCC was identified, with
217 genes like *PIK3CA* and *TP53* that could be specific of higher grades [17]. Some authors have
218 observed the transfer of multiple clones in melanoma, breast and prostate cancers [18], with
219 polyphyletic dissemination. In this study we confirm these observations in ccRCCs. In a similar

Accepted manuscript

220 setting, Ferronika et al. studied molecular profiles of metastases and matched primary samples
221 in ccRCC. A subgroup of three primary tumors with minor copy number changes was opposed
222 to a subgroup with a primary tumour, a thrombus, and lung metastases, all with a similar copy
223 number pattern and tetraploid-like characteristics [19].
224 Site-specific clones could be a limit to this work, as synchronous adrenal metastases which
225 represent early events, could share more common traits with primary clones than a
226 metachronous metastasis. However, the analysis of late metastases is limited as they are rarely
227 treated by surgical excision, especially before the initiation of any drug treatment.
228 A multi-patient analysis combining similar sampling strategies with single-cell transcriptomic
229 and genomic approaches would produce a clearer picture of the subclonal evolution in ccRCC
230 patients. This would be of great help in the transition towards a more personalized medicine.
231 In clinical routine, such experiments would be helpful and could be implemented in clinical
232 trials, by sampling metastases, and adapting the treatment to the specific aggressive profiles.
233 Patients with resistance to therapies could benefit a combination of targeted therapies based on
234 more than one aggressive clone.

Accepted manuscript

235 **Declarations**

236 ***Ethics approval and consent to participate***

237 Informed consent was signed from the patient and the study was approved by the Rennes
238 University Hospital ethics committee.

239 ***Availability of data and materials***

240 The datasets used and/or analysed during the current study are available from the
241 corresponding author on reasonable request.

242 ***Competing interests***

243 The authors declare that they have no competing interests.

244 ***Funding***

245 The authors would like to thank the *Ligue Régionale Contre Le Cancer, Institut national du*
246 *cancer (INCa)*. FC and BE were supported by the *Institut National de la Santé Et de la*
247 *Recherche Médicale* (*INSERM*), the *Université de Rennes 1* and the *Ecole des Hautes Etudes*
248 *en Santé Publique* (*EHESP* - School of Public Health).

249 ***Authors' contributions***

250 JD, AB, NRL and FC analysed and interpreted the patient data. JD, MB, LC, FD and MABR
251 performed genomic analyses. JD and BE performed DNA and RNA extraction. AL and JD
252 performed *VHL* gene analyses. FD, GV, and KB supplied clinical support. All authors read
253 and approved the final manuscript.

254

Accepted manuscript

255 **REFERENCES**

256 [1] Bianchi M, Sun M, Jeldres C, Shariat SF, Trinh Q-D, Briganti A, et al. Distribution of
257 metastatic sites in renal cell carcinoma: a population-based analysis. Ann Oncol Off J Eur
258 Soc Med Oncol 2012;23:973–80. https://doi.org/10.1093/annonc/mdr362.

259 [2] Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor
260 heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med
261 2012;366:883–92. https://doi.org/10.1056/NEJMoa1113205.

262 [3] Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic
263 architecture and evolution of clear cell renal cell carcinomas defined by multiregion
264 sequencing. Nat Genet 2014;46:225–33. https://doi.org/10.1038/ng.2891.

265 [4] Fidler IJ. Tumor heterogeneity and the biology of cancer invasion and metastasis. Cancer
266 Res 1978;38:2651–60.

267 [5] Turajlic S, Xu H, Litchfield K, Rowan A, Horswell S, Chambers T, et al. Deterministic
268 Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. Cell
269 2018;173:595-610.e11. https://doi.org/10.1016/j.cell.2018.03.043.

270 [6] NCBI Resource Coordinators. Database resources of the National Center for
271 Biotechnology Information. Nucleic Acids Res 2016;44:D7-19.
272 https://doi.org/10.1093/nar/gkv1290.

273 [7] Lê, S., Josse, J. & Husson, F. FactoMineR: An R Package for Multivariate Analysis. J
274 Stat Softw 251 Pp 1-18 2008.

275 [8] Chalmel F, Primig M. The Annotation, Mapping, Expression and Network (AMEN) suite
276 of tools for molecular systems biology. BMC Bioinformatics 2008;9:86.
277 https://doi.org/10.1186/1471-2105-9-86.

278 [9] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential
279 expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res
280 2015;43:e47. https://doi.org/10.1093/nar/gkv007.

281 [10] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology:
282 tool for the unification of biology. The Gene Ontology Consortium. Nat Genet
283 2000;25:25–9. https://doi.org/10.1038/75556.

284 [11] Rappoport N, Shamir R. Multi-omic and multi-view clustering algorithms: review and
285 cancer benchmark. Nucleic Acids Res 2018;46:10546–62.
286 https://doi.org/10.1093/nar/gky889.

287 [12] Aceto N, Bardia A, Miyamoto DT, Donaldson MC, Wittner BS, Spencer JA, et al.
288 Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. Cell
289 2014;158:1110–22. https://doi.org/10.1016/j.cell.2014.07.013.

290 [13] Massagué J, Obenauf AC. Metastatic Colonization. Nature 2016;529:298–306.
291 https://doi.org/10.1038/nature17038.

292 [14] Gundem G, Van Loo P, Kremeyer B, Alexandrov LB, Tubio JMC, Papaemmanuil E, et
293 al. The evolutionary history of lethal metastatic prostate cancer. Nature 2015;520:353–7.
294 https://doi.org/10.1038/nature14347.

Accepted manuscript

295 [15] Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for
296 cancer? Nat Rev Cancer 2012;12:323–34. https://doi.org/10.1038/nrc3261.

297 [16] Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of
298 genetic heterogeneity in cancer evolution. Nature 2013;501:338–45.
299 https://doi.org/10.1038/nature12625.

300 [17] Ferronika P, Kats-Ugurlu G, Haryana SM, Utoro T, Rinonce HT, Danarto R, et al.
301 Mutational heterogeneity between different regional tumour grades of clear cell renal cell
302 carcinoma. Exp Mol Pathol 2020;115:104431.
303 https://doi.org/10.1016/j.yexmp.2020.104431.

304 [18] Turajlic S, Swanton C. Metastasis as an evolutionary process. Science 2016;352:169–75.
305 https://doi.org/10.1126/science.aaf2784.

306 [19] Ferronika P, Hof J, Kats-Ugurlu G, Sijmons RH, Terpstra MM, de Lange K, et al.
307 Comprehensive Profiling of Primary and Metastatic ccRCC Reveals a High Homology of
308 the Metastases to a Subregion of the Primary Tumour. Cancers 2019;11.
309 https://doi.org/10.3390/cancers11060812.
310

Accepted manuscript

311 **Figure legends**

312 **Figure 1. Macroscopic distribution and sample classification using MFA.**

313 (A) Five samples were extracted from the primary tumor (p1 to p5) in different geographical
314 zones. Similar extractions were undertaken in the vein thrombus (two samples: m1 and m2)
315 and the right (four samples: m3 to m6) and left (three samples: m7 to m9) adrenal metastases.
316 Multiple factor analysis combining Array-CGH and transcriptomic data (panel B) followed
317 by hierarchical clustering (C) divided specimens into three sample classes: C1 composed of
318 three primary specimens; C2 and C3 composed each of one primary and all metastatic
319 specimens.

320

321 **Figure 2. Genomic and transcriptional heatmap of each cluster.**

322 The upper heatmap representation illustrates chromosomal gains and losses (green-losses,
323 purple-gains), which are partitioned into three groups R1 to R3 (green-loss, purple-gain). The
324 lower heatmap representation illustrates transcriptional patterns (blue-underexpression, red-
325 overexpression), which are also partitioned into three groups G1 to G3. The details of the
326 significant chromosomal regions (upper part) or significantly associated GO terms (lower part)
327 are noted at the right of each cluster. Color codes indicate gain (purple) or loss (green) for
328 Array-CGH array data and overrepresentation (red) and underrepresentation (blue) for
329 microarray data as indicated in the scale bars.

330

331 **Figure 3. Hypotheses explaining metastatic dissemination in this ccRCC.**

332 (A) Both aggressive clones are established in the primary tumor and colonize all metastatic
333 sites, as single cells or as clusters, following a functionally dependent and systematic scheme.

Accepted manuscript

334 (B) One clone in the primary tumor colonizes each metastatic site. In each metastasis, a
335 separate clonal evolution occurs with the emergence of two aggressive clones. These
336 aggressive clones recolonize each metastasis as well as the primary tumor.
337
338 **Supplementary Tables and Figures.**
339 **Table S1. Statistical filtration of differentially expressed genes.**
340
341 **Table S2. Functional analysis. Enriched biological process terms.**
342
343 **Figure S1. Quality control using box-plot diagram showing intensity signal distribution.**
344 The signal distribution of normalized Array-CGH (panel A) and transcriptomic (panel B)
345 data was controlled by box-plot diagrams.
346
347 **Figure S2. Statistical filtration and cluster analysis.**
348 Statistical tests were used to select SNOC regions (panel A) or genes displaying a significant
349 gain or loss of expression between sample classes (C1-C3). The resulting regions and genes
350 were then clustered into three groups R1 to R3 and G1 to G3 respectively.
351
352 **Figure S3. Illustration of histological aspects of extracted samples.**
353 Histological analyses showing similarities and differences between tumor samples. High
354 grade cells with pleomorphic nuclei and necrosis can be documented in p2, m2, m5 and m6.
355
356 **Figure S4. Detailed genomic analysis of the fourteen samples.**
357 Each chromosome is represented horizontally in regards to each cluster and specimen. Gains
358 are in red and losses in blue.

Accepted manuscript

359

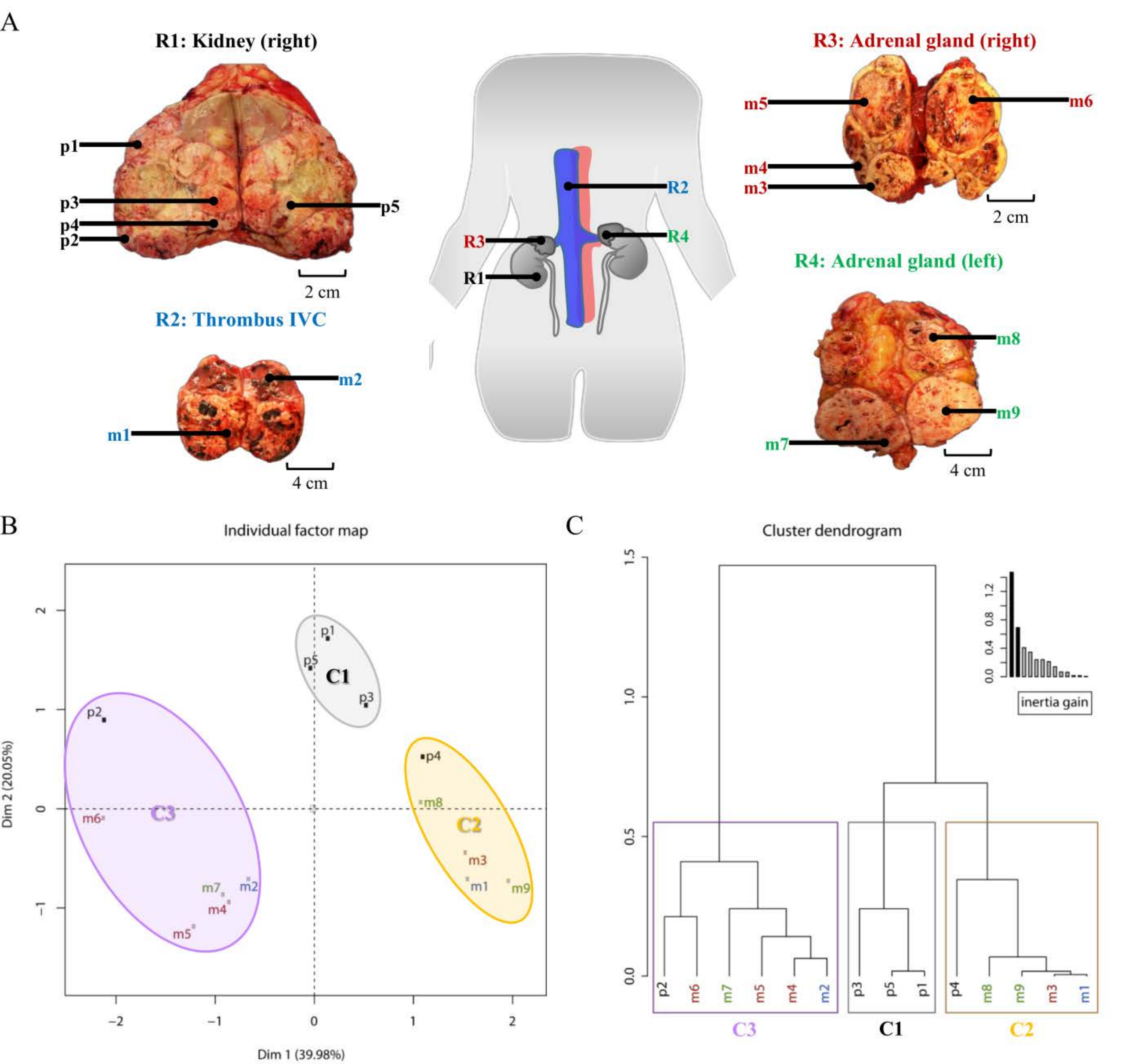
360 **Figure S5. Chromosome 7 ploidy in different samples.**

361 Compared to p3 and most metastatic samples where triploidy of chromosome 7 (3n) was

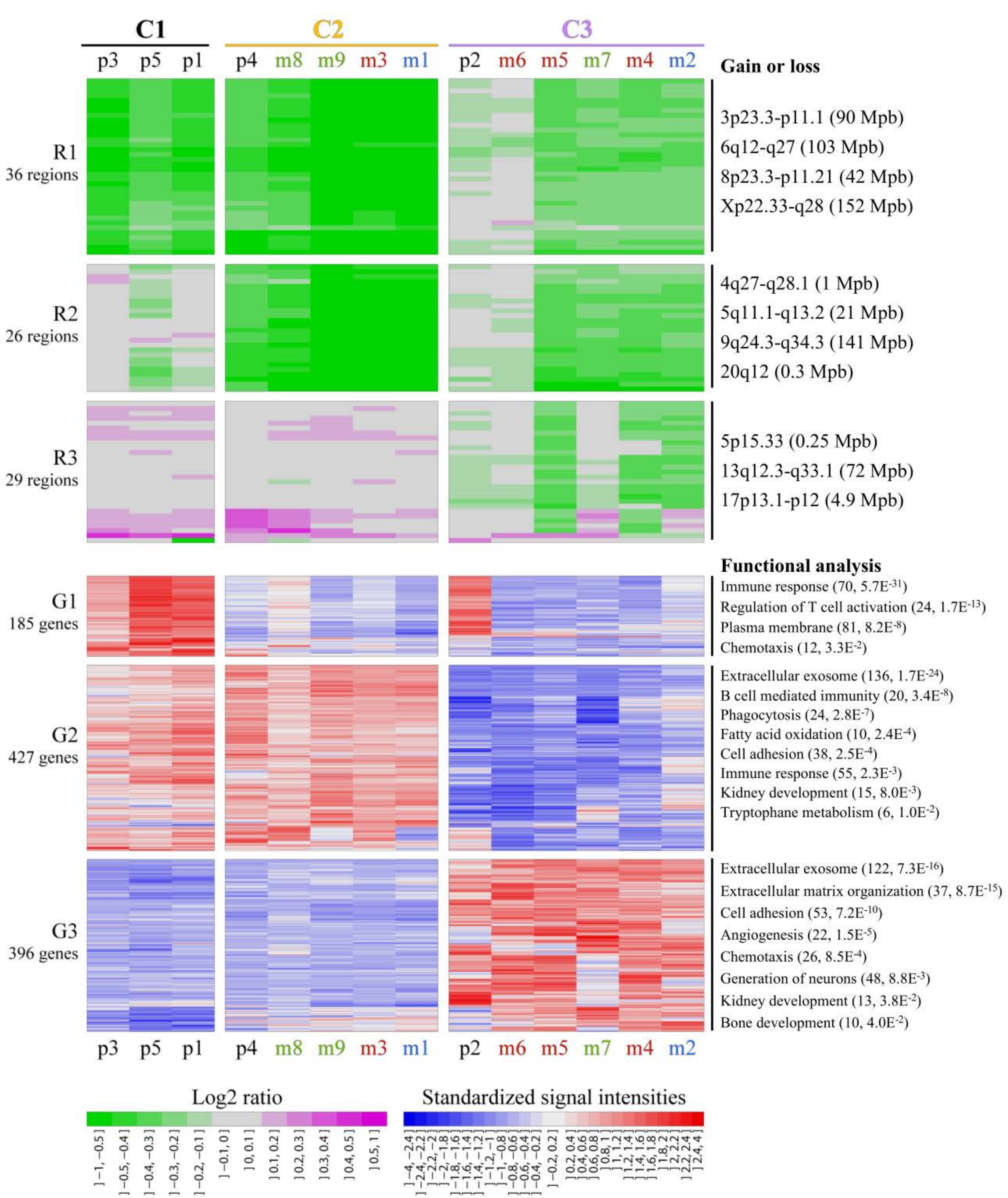
362 observed, most primary samples had two copies (2n) of chromosome 7.

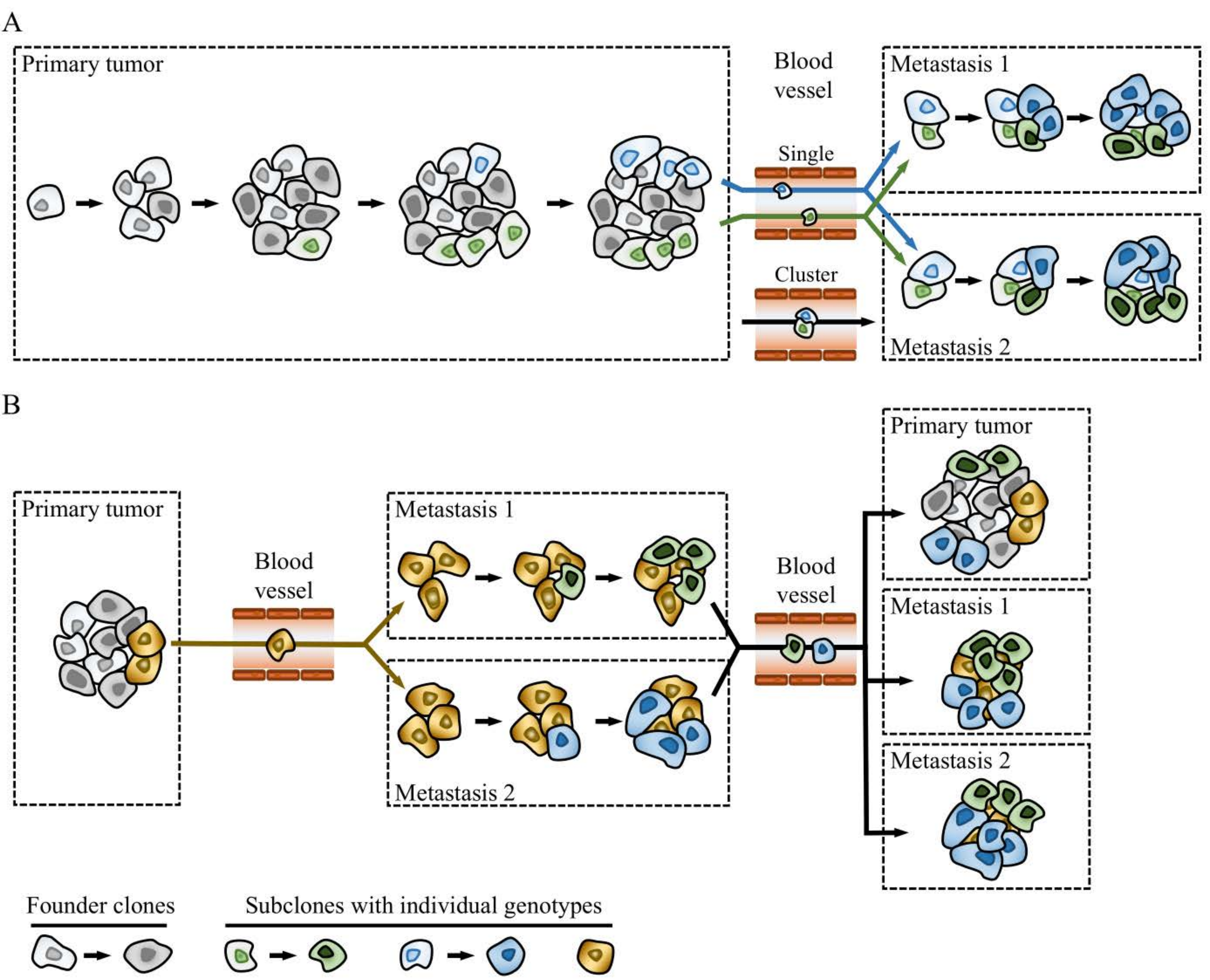
363

Accepted manuscript

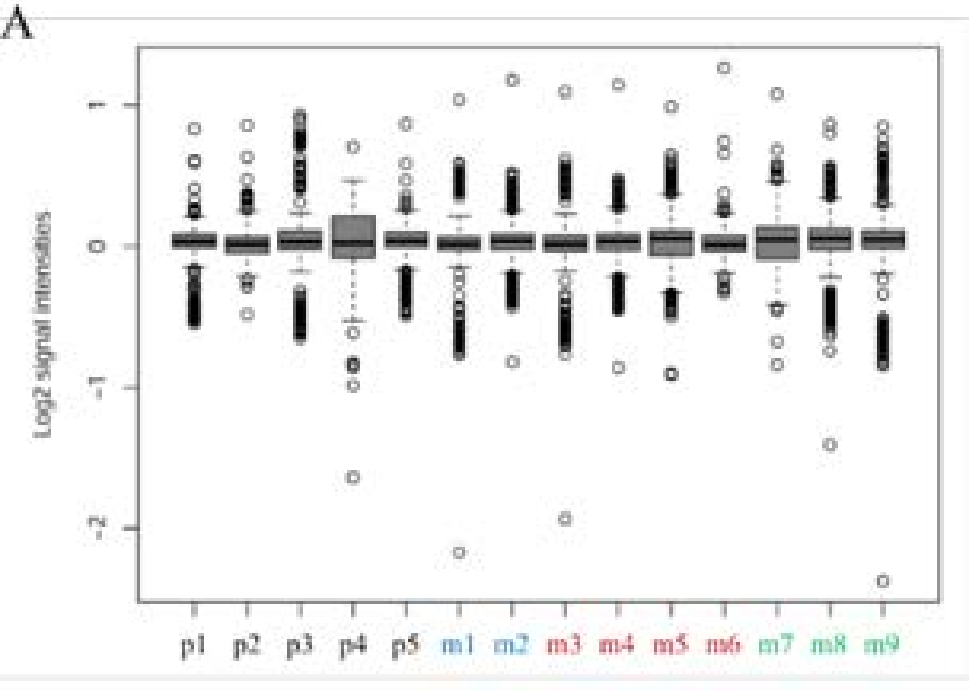


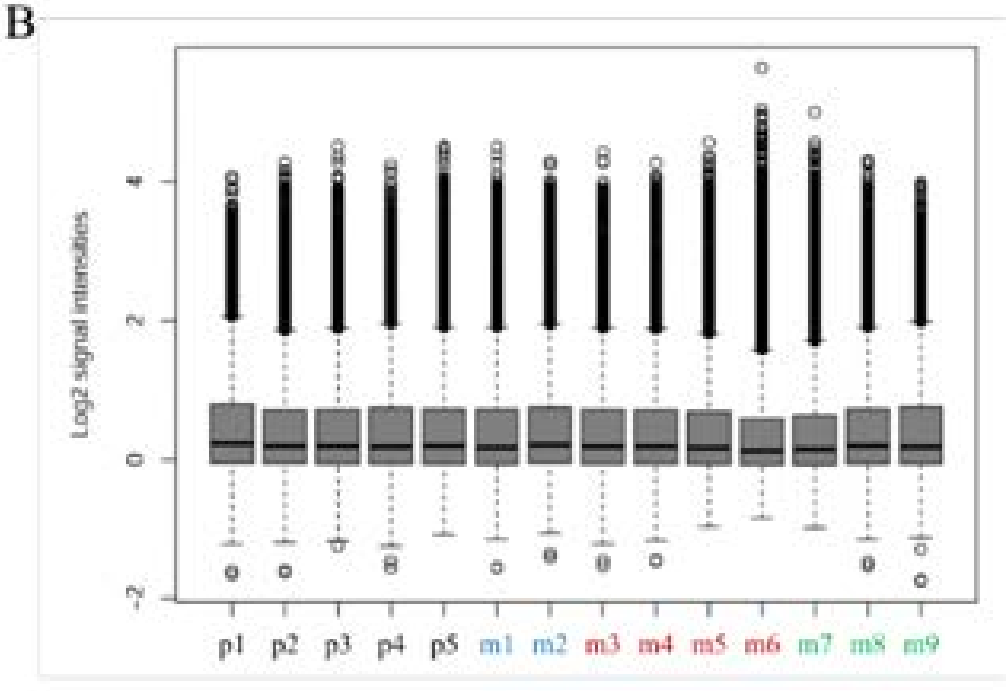
Accepted manuscript

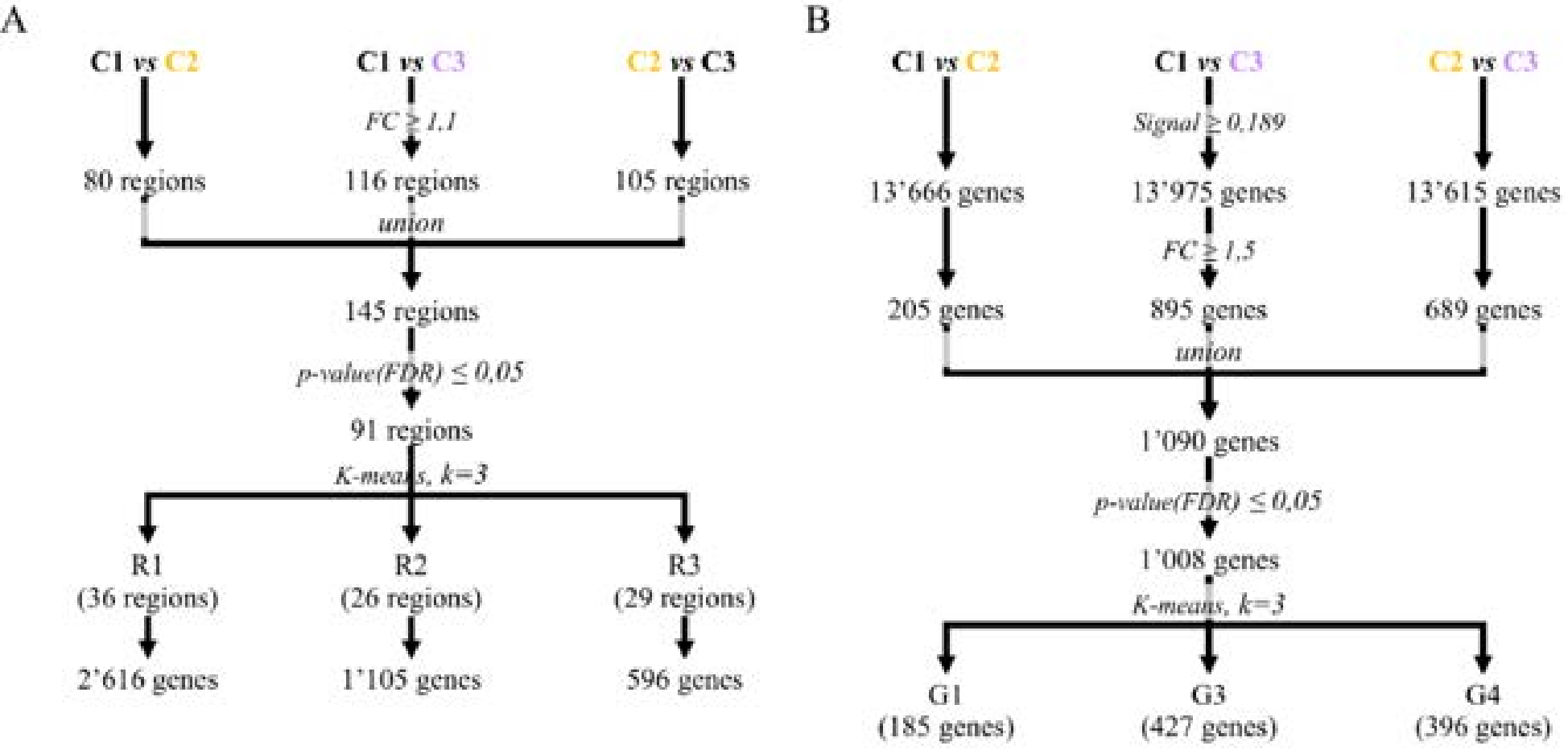

C1
C2
C3
p3 p5 p1
p4 m8 m9 m3 m1
p2 m6 m5 m7 m4 m2
Gain or loss
R1
36 regions
3p23.3-p11.1 (90 Mpb)
6q12-q27 (103 Mpb)
8p23.3-p11.21 (42 Mpb)
Xp22.33-q28 (152 Mpb)
R2
26 regions
4q27-q28.1 (1 Mpb)
5q11.1-q13.2 (21 Mpb)
9q24.3-q34.3 (141 Mpb)
20q12 (0.3 Mpb)
R3
29 regions
5p15.33 (0.25 Mpb)
13q12.3-q33.1 (72 Mpb)
17p13.1-p12 (4.9 Mpb)
Functional analysis
G1
185 genes
Immune response (70, 5.7E-31)
Regulation of T cell activation (24, 1.7E-13)
Plasma membrane (81, 8.2E-8)
Chemotaxis (12, 3.3E-2)
G2
427 genes
Extracellular exosome (136, 1.7E-24)
B cell mediated immunity (20, 3.4E-8)
Phagocytosis (24, 2.8E-7)
Fatty acid oxidation (10, 2.4E-4)
Cell adhesion (38, 2.5E-4)
Immune response (55, 2.3E-3)
Kidney development (15, 8.0E-3)
Tryptophane metabolism (6, 1.0E-2)
G3
396 genes
Extracellular exosome (122, 7.3E-16)
Extracellular matrix organization (37, 8.7E-15)
Cell adhesion (53, 7.2E-10)
Angiogenesis (22, 1.5E-5)
Chemotaxis (26, 8.5E-4)
Generation of neurons (48, 8.8E-3)
Kidney development (13, 3.8E-2)
Bone development (10, 4.0E-2)
p3 p5 p1 p4 m8 m9 m3 m1 p2 m6 m5 m7 m4 m2
Log2 ratio
]−1, −0.5]]−0.5, −0.4]]−0.4, −0.3]]−0.3, −0.2]]−0.2, −0.1]]−0.1, 0]]0, 0.1]]0.1, 0.2]]0.2, 0.3]]0.3, 0.4]]0.4, 0.5]]0.5, 1]
Standardized signal intensities
]−4, −2.4]]−2.4, −2.2]]−2.2, −2]]−2, −1.8]]−1.8, −1.6]]−1.6, −1.4]]−1.4, −1.2]]−1.2, −1]]−1, −0.8]]−0.8, −0.6]]−0.6, −0.4]]−0.4, −0.2]]−0.2, 0.2]]0.2, 0.4]]0.4, 0.6]]0.6, 0.8]]0.8, 1]]1, 1.2]]1.2, 1.4]]1.4, 1.6]]1.6, 1.8]]1.8, 2]]2, 2.2]]2.2, 2.4]]2.4, 4]

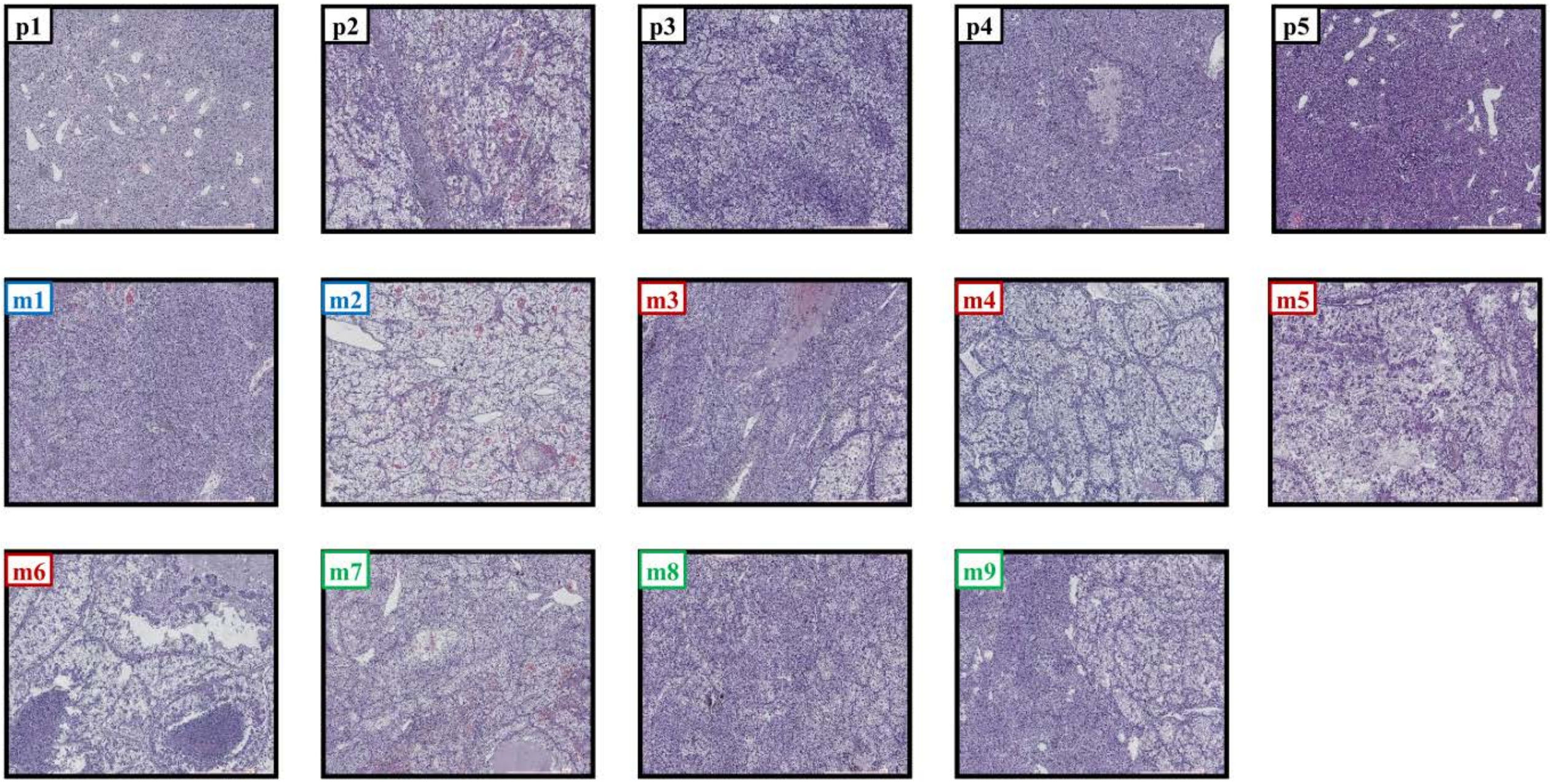


Accepted manuscript

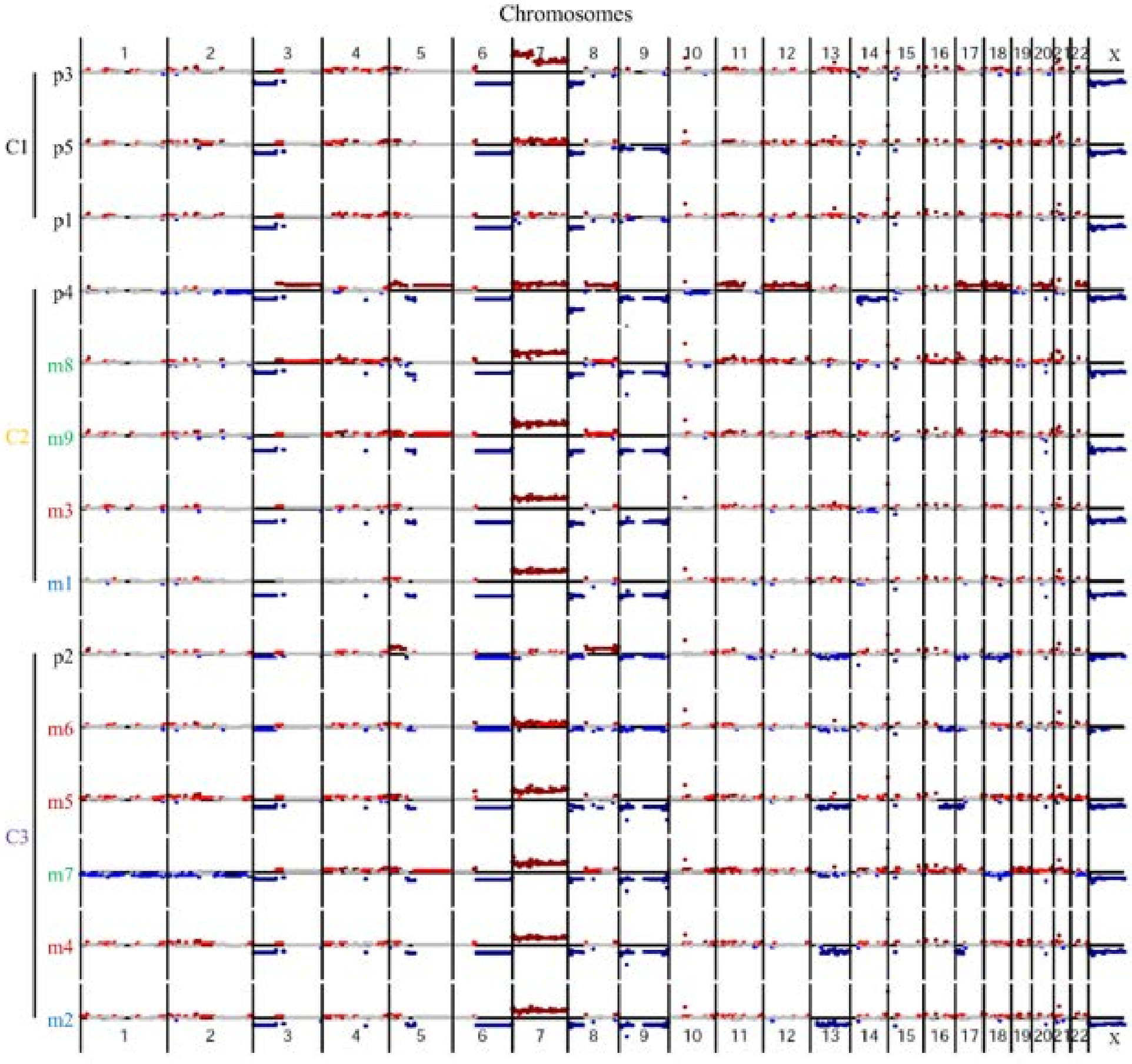

A
Log2 signal intensities
1
0
-1
-2
p1 p2 p3 p4 p5 m1 m2 m3 m4 m5 m6 m7 m8 m9

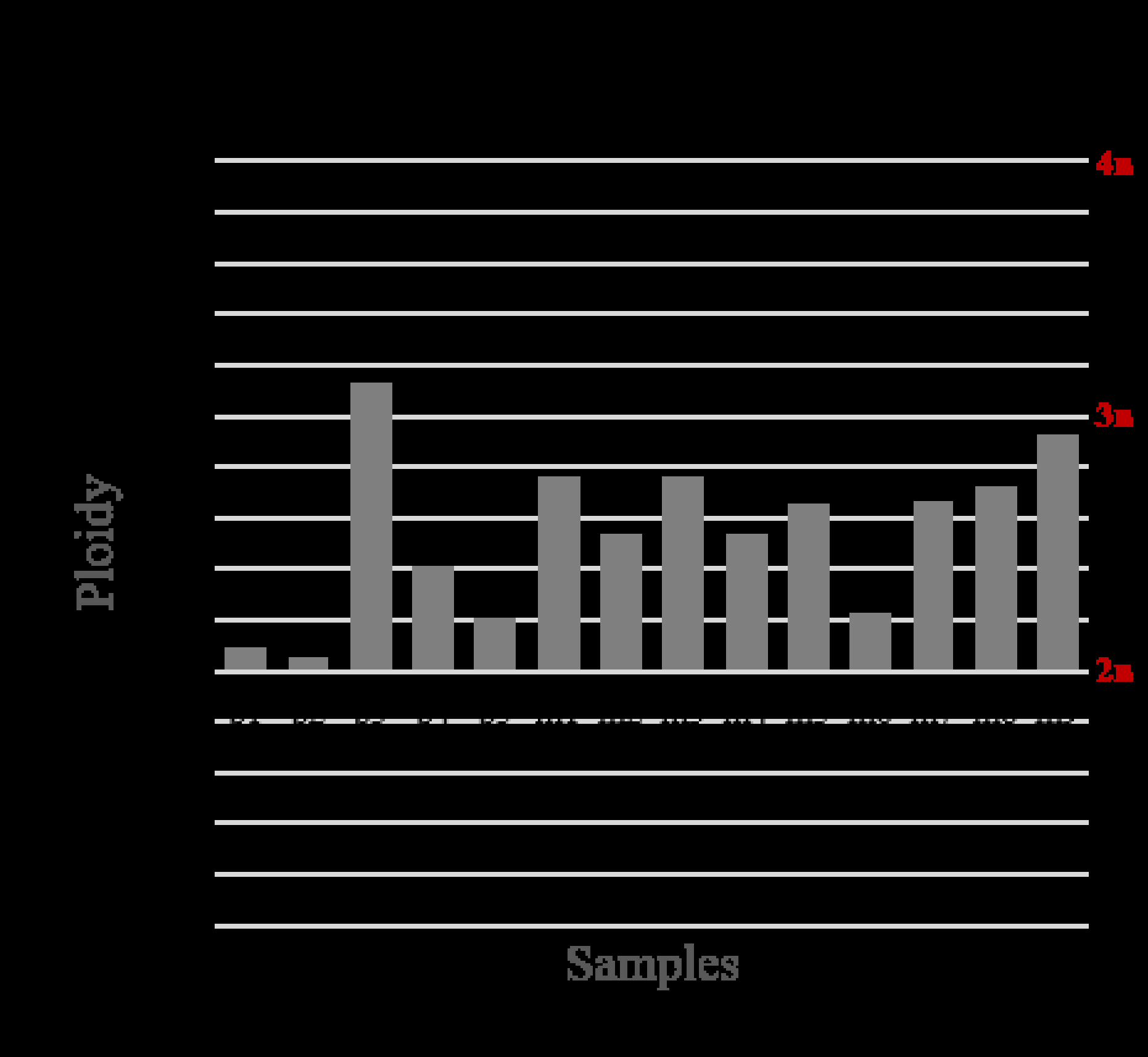

B
Log2 signal intensities
4
2
0
-2
p1 p2 p3 p4 p5 m1 m2 m3 m4 m5 m6 m7 m8 m9


A
C1 vs C2
C1 vs C3
C2 vs C3
FC ≥ 1.1
80 regions
116 regions
105 regions
union
145 regions
p-value(FDR) ≤ 0,05
91 regions
K-means, k=3
R1
(36 regions)
R2
(26 regions)
R3
(29 regions)
2'616 genes
1'105 genes
596 genes
B
C1 vs C2
C1 vs C3
C2 vs C3
Signal ≥ 0,189
13'666 genes
13'975 genes
13'615 genes
FC ≥ 1,5
205 genes
895 genes
689 genes
union
1'090 genes
p-value(FDR) ≤ 0,05
1'008 genes
K-means, k=3
G1
(185 genes)
G3
(427 genes)
G4
(396 genes)



Accepted manuscript


Chromosomes
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
C1
p3
p5
p1
C2
p4
m8
m9
m3
m1
C3
p2
m6
m5
m7
m4
m2


4n
3n
2n
Ploidy
Samples