

1 **Multioomic characterisation of high grade serous ovarian carcinoma enables high resolution patient**
2 **stratification**

3 Robert L Hollis^{1*}, Alison M Meynert², Caroline O Michie³, Tzyvia Rye¹, Michael Churchman¹, Amelia
4 Hallas-Potts¹, Ian Croy¹, W. Glenn McCluggage⁴, Alistair R W Williams⁵, Clare Bartos¹, Yasushi Iida^{1,6},
5 Aikou Okamoto⁶, Brian Dougherty⁷, J. Carl Barrett⁷, Ruth March⁸, Athena Matakidou⁹, Patricia
6 Roxburgh^{10,11}, Colin A Semple², D Paul Harkin^{12,13}, Richard Kennedy^{12,13}, C Simon Herrington¹ and
7 Charlie Gourley¹

8 *Correspondence: Dr Robb Hollis, Cancer Research UK Edinburgh Centre, MRC Institute of Genetics
9 and Cancer, University of Edinburgh, Crewe Road South, Edinburgh, EH4 2XU, Scotland, UK.
10 robb.hollis@ed.ac.uk

11

12 ¹Nicola Murray Centre for Ovarian Cancer Research, Cancer Research UK Edinburgh Centre, MRC
13 Institute of Genetics and Cancer, University of Edinburgh, UK

14 ²MRC Human Genetics Unit, MRC Institute of Genetics and Cancer, University of Edinburgh, UK

15 ³Edinburgh Cancer Centre, Western General Hospital, NHS Lothian, UK

16 ⁴Department of Pathology, Belfast Health and Social Care Trust, Belfast, UK

17 ⁵Division of Pathology, The Royal Infirmary of Edinburgh, Edinburgh, UK

18 ⁶The Jikei University School of Medicine, Tokyo, Japan

19 ⁷Translational Medicine, Oncology R&D, AstraZeneca, Waltham, MA, USA

20 ⁸Precision Medicine and Biosamples, Oncology R&D, AstraZeneca, Cambridge, UK

21 ⁹Centre for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca,
22 Cambridge, UK

23 ¹⁰Institute of Cancer Sciences, Wolfson Wohl Cancer Research Centre, University of Glasgow, UK

24 ¹¹Beatson West of Scotland Cancer Centre, Glasgow, UK

25 ¹²Almac Diagnostics, Craigavon, UK

26 ¹³Centre for Cancer Research and Cell Biology, Queen's University Belfast, UK

27 **Counts:** abstract 340 words, main text 3400 words, 5 figures, 1 table

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

28 **ABSTRACT**

29 **Background:** High grade serous ovarian carcinoma (HGSOC) is the most common type of ovarian
30 cancer; most patients experience disease recurrence which accumulates chemoresistance, leading to
31 treatment failure. Previous investigations have characterised HGSOC at the genomic and
32 transcriptomic level, identifying subtypes of patients with differential outcome and treatment
33 response. However, the relationship between molecular events identified at the gene sequence, gene
34 copy number and gene expression levels remains poorly defined.

35 **Methods:** We perform multi-layer molecular characterisation of a large retrospective HGSOC cohort
36 (n=362) with detailed clinical annotation to interrogate the relationship between patient groups
37 defined by gene mutation, copy number events, gene expression patterns and infiltrating immune cell
38 burden. We construct a high resolution picture of the molecular landscape in HGSOC and identify
39 features of tumours associated with distinct clinical behaviour in patients.

40 **Results:** *BRCA2*-mutant (*BRCA2m*) and *EMSY*-overexpressing cases demonstrated prolonged survival
41 (multivariable hazard ratio 0.40 and 0.53) and higher chemotherapy response rates at first- and
42 second-line treatment. *CCNE1*-gained (*CCNE1g*) cases demonstrated shorter survival (multivariable
43 hazard ratio 1.52, 95% CI 1.10-2.10), under-representation of FIGO stage IV cases (P=0.017) and no
44 significant difference in treatment response. We demonstrate marked overlap between the TCGA-
45 and derived subtypes: the TCGA DIF, IMR, PRO and MES subtypes correlated with the Tothill C4, C2,
46 C5 and C1 subtypes (P<0.001). IMR/C2 cases displayed higher *BRCA1/2m* frequency (25.5% and 32.5%)
47 and significantly greater infiltration of immune cells (P<0.001), while PRO/C5 cases had the highest
48 *CCNE1g* rate (23.9% and 22.2%) and were uniformly low in immune cell infiltration. The survival
49 benefit for cases with aberrations in homologous recombination repair (HRR) genes was apparent
50 across all transcriptomic subtypes (hazard ratio range 0.48-0.68). There was significant co-occurrence
51 of RB loss and HRR gene aberrations (P=0.005); RB loss was further associated with favourable survival
52 within cases harbouring HRR aberrations (multivariable hazard ratio 0.50, 95% CI 0.30-0.84).

53 **Conclusions:** These data paint a high resolution picture of the molecular landscape in HGSOC, better
54 defining patients who may benefit most from specific molecular therapeutics and highlighting those
55 for whom novel treatment strategies are needed to improve outcomes.

56 **Keywords:** high grade serous ovarian cancer, genomics, transcriptomics, survival, treatment response

57

58 **1. BACKGROUND**

59 High grade serous ovarian carcinoma (HGSOC) is the most common form of tubo-ovarian cancer. The
60 majority of HGSOC patients are diagnosed at advanced stage and experience poor prognosis, with a
61 five-year survival of approximately 30% in this population [1]. While the majority of HGSOC
62 demonstrate high levels of intrinsic sensitivity to platinum-based chemotherapy, most patients
63 experience disease recurrence which accumulates therapy resistance, leading to progressively shorter
64 treatment-free intervals until patients eventually succumb to disease [2, 3].

65 In the hope of identifying therapeutically exploitable disease biology, a wealth of data have been
66 produced over the last two decades characterising the genomic and transcriptomic landscape of
67 HGSOC [4-7]. At the gene sequence level, identification of mutational disruption in *BRCA1* and *BRCA2*
68 (*BRCA1/2m*) has ultimately paved the way for integration of poly-(ADP-ribose) polymerase (PARP)
69 inhibitor use into routine care for some patients [8-10]. Indeed, there continues to be an intense
70 research effort surrounding mechanisms and implications of homologous recombination DNA repair
71 (HRR) disruption beyond *BRCA1/2m* [11]; these include mutation of non-*BRCA1/2* HRR genes [12],
72 large-scale genomic variants disrupting *BRCA1/2* [13], epigenetic inactivation of HRR players such as
73 *BRCA1* and *RAD51C* [5, 14], and overexpression of the *BRCA2* regulator *EMSY* [15, 16].

74 At the gene expression level, numerous studies have characterised HGSOC samples, endeavouring to
75 identify clinically meaningful transcriptomic subtypes of disease or expression signatures predictive of
76 survival risk [4, 5, 7, 17, 18]. Most notably, Tothill et al. [4] and the TCGA investigators [5] each
77 identified multiple transcriptomic subtypes, associating these with differential survival profiles. These

78 analyses have identified favourable outcome in patients with tumours harbouring expression profiles
79 suggestive of active immune engagement (TCGA IMR subtype [5, 19], Tothill C2 subtype [4]) –
80 consistent with earlier reports of favourable outcome in cases with high levels of cytotoxic T cell
81 infiltration [20, 21]. However, transcriptomic subtyping is not currently used for clinical
82 prognostication or stratification of HGSOC patients, despite some investigators reporting differential
83 sensitivity of these groups to agents such as bevacizumab [22].

84 While multiple investigators have characterised either the genomic or transcriptomic landscape of
85 HGSOC, few have investigated the relationship between genomic and transcriptomic features.
86 Moreover, the relationship between these events and recently identified recurrent disruption of RB
87 and PTEN in HGSOC is poorly understood [6]. Integration of multiple layers of molecular
88 characterisation is required to paint a granular picture of the molecular landscape in HGSOC to better
89 inform rationally designed trials of novel treatment regimens or combination therapy strategies.
90 Indeed, some investigators have suggested that transcriptomic subtypes of HGSOC that appear to
91 derive greatest benefit from anti-angiogenic agents may be depleted for *BRCA1/2m* cases who benefit
92 most from PARP inhibition [5, 22]. This notion is consistent with mixed results observed from the
93 addition of anti-angiogenic agents to PARP inhibitors (PARPi) dependent on patient selection, therapy
94 line and agent combinations [23], exemplifying the need for comprehensive multi-layer molecular
95 characterisation to inform patient selection for future investigations of novel treatment approaches.
96 Here we perform matched genomic and transcriptomic characterisation of a large, well annotated
97 HGSOC cohort, dissecting the relationship between patient groups defined at the gene sequence, gene
98 copy number and gene expression level.

99

100 **2. METHODS**

101 *2.1 Patient cohort*

102 539 ovarian cancer patients treated at the Edinburgh Cancer Centre met the following study inclusion
103 criteria: primary ovarian, peritoneal or fallopian tube carcinoma (any histological type) diagnosed prior
104 to 2007; available formalin-fixed paraffin-embedded (FFPE) treatment-naïve surgical specimen; first-
105 line platinum-containing chemotherapy; minimum 3-year follow-up. Pathology review of H&E-stained
106 slides was undertaken by expert gynaecological pathologists (WGM, ARWW, CSH) to identify HGSOC
107 cases (figure S1); immunohistochemistry (IHC) for p53 and WT1 was used to clarify cases of uncertain
108 histological type (HGSOC: WT1 positive, p53 aberrant mutation-type expression pattern)
109 (Supplementary Methods Section 1) (figure S1). Ethical approval was obtained from South East
110 Scotland Human Annotated Bioresource (Lothian NRS Bioresource Ethics Committee reference
111 15/ES/0094-SR705 and SR752). The need for consent was waived by the ethics committee due to the
112 retrospective nature of the study.

113 *2.2 Genomic characterisation*

114 H&E stained slides were marked to identify tumour areas of high cellularity and used as a guide for
115 macrodissection of 10µm FFPE sections for DNA extraction. DNA extraction was performed using the
116 QIAamp FFPE DNA Kit and Qiagen Deparaffinisation Solution (Qiagen). Extracted DNA was quantified
117 by high sensitivity Qubit assay. *CCNE1* and *EMSY* copy number (CN) were quantified by TaqMan qPCR
118 (Supplementary Methods Section 2).

119 High throughput sequencing was performed using a custom Integrated DNA Technologies (IDT) gene
120 capture panel with unique molecular indices (UMIs) (Supplementary Methods Section 3). Whole
121 genome libraries were generated, pooled for target capture and sequenced using an Illumina NextSeq
122 550 at the Edinburgh Clinical Research Facility, Western General Hospital, Edinburgh, UK. The median
123 per-sample mean target coverage was 593X (range 205-3278X). Reads were processed using the bcbio
124 v1.0.6 high throughput sequence analysis pipeline (Supplementary Methods Section 4). Consensus
125 reads aligned to hg38 underwent variant calling using a majority vote system from three variant callers
126 (Freebayes, VarDict and Mutect2). Called variants were annotated using the Ensembl VEP v90.9

127 against Ensembl release 90 and filtered to retain only functional variation (Supplementary Methods
128 Section 5).

129 *2.3 Transcriptomic subtyping*

130 Transcriptomic data for the cohort were available from previous work identifying transcriptomic
131 subtypes of HGSOC [15, 17], including *EMSY* overexpression status (Supplementary Methods Section
132 6). TCGA (MES, PRO, IMR, DIF) and Tothill (C1, C2, C4, C5) transcriptomic subtyping calls were made
133 with the consensusOv R package using the consensusOv and Helland approaches [19] (Supplementary
134 Methods Section 6).

135 *2.4 Immune cell infiltration analysis*

136 Tumour infiltrating CD3-positive and CD8-positive immune cells were quantified by IHC of constructed
137 tumour tissue microarrays (TMAs) (Supplementary Methods Section 7); marker-positive cell burden
138 was quantified as percentage positive cells within tumour areas using QuPath version 0.1.2 [24].

139 *2.5 Detection of PTEN and RB loss by immunohistochemistry*

140 PTEN and RB protein loss was detected by IHC using sections of the HGSOC TMA (Supplementary
141 Methods Section 8). Loss was defined as complete absence of positive staining in tumour cells with
142 confirmed corresponding positive internal control stromal staining.

143 *2.6 Copy number analysis from off-target sequencing reads*

144 Copy number analysis was performed using CopywriteR [25] (Supplementary Methods Section 9): off
145 target reads were used to estimate the relative copy number of 50kB genome segments across each
146 chromosome, using the alignment bam files from the above sequencing analysis workflow. For
147 quantification of CN alteration burden, adjacent 50kB segments of gain/loss representing the same
148 large CN event were merged prior to quantification (Supplementary Methods Section 9).

149 *2.7 Clinical annotation*

150 Baseline clinicopathological features and outcome data were extracted from the Edinburgh Ovarian
151 Cancer Database [26], alongside chemotherapy response data (Supplementary Methods Section 10).
152 Overall survival (OS) and progression-free survival (PFS) were defined as the time from pathologically
153 confirmed diagnosis to patient death and disease progression or recurrence, respectively
154 (Supplementary Methods Section 9).

155 *2.8 Statistical analyses*

156 All statistical analyses were performed using R version 4.0.3 (R Foundation for Statistical Computing,
157 Vienna, Austria). Comparisons of frequency were performed using the Chi-squared test or Fisher's
158 exact test, as appropriate. Between-group comparisons of continuous variables was performed using
159 the Mann Whitney-U test. Survival analysis was performed using Cox proportional hazards regression
160 models and reported as hazard ratios (HR) with 95% confidence intervals (95% CI). For survival analysis
161 adjusted for other clinicopathological factors, multivariable hazard ratios (mHR) are reported. Median
162 follow-up time was calculated by the reverse Kaplan-Meir method. Adjustment for multiple testing
163 was applied using the Bonferroni method.

164

165 **3. RESULTS**

166 *3.1 Cohort characteristics*

167 Of 539 ovarian cancer cases that met eligibility criteria, 362 were classified as HGSOC following
168 pathology review and underwent molecular characterisation (n=27 insufficient tumor, n=131 non-
169 HGSOC, n=1 failed sequencing library preparation, n=8 failed quality control) (figure S1).
170 Clinicopathological features of the study cohort are summarised in table 1. The median follow-up time
171 was 15.0 years.

172 *3.2 Molecular landscape of high grade serous ovarian carcinoma*

173 The frequency of *TP53* mutation was 98.1% (355 of 362 cases) (figure 1, table S1). 12.7% and 6.6% of
174 cases harboured *BRCA1m* and *BRCA2m*. Eight cases (2.2%) demonstrated mutation of other HRR genes
175 (3 *BRIP1*, 2 *CHEK2*, 1 *RAD51C*, 1 *PALB2*, 1 concurrent *BAP1* and *NBN*). 14.9% of cases displayed CN gain
176 of *CCNE1* (*CCNE1g*) and 6.6% demonstrated amplification of *EMSY*. Tumours demonstrating *EMSY*
177 amplification were enriched for *EMSY* mRNA-overexpressing cases ($P < 0.001$) (Supplementary table 2);
178 however, *EMSY* CN was a poor predictor of *EMSY* overexpression status (positive predictive value 0.42,
179 95% CI 0.22-0.63; negative predictive value 0.88, 95% CI 0.84-0.91).

180 An HRR-centric stepwise taxonomy was constructed (figure 2A). Compared to the non-*CCNE1g* HRR-
181 wild-type reference population (HRRwt, 55.8% of cases), *BRCA2m* and *EMSY*-overexpressing (8.6%)
182 cases demonstrated favourable outcome (mHR for OS = 0.40, 95% CI 0.25-0.65 and 0.53, 95% CI 0.33-
183 0.84) (figure 2A, figure S2). Stage IV cases were under-represented in the *CCNE1g* group (8.2% vs 23.4%
184 in the HRRwt group, $P = 0.017$); *CCNE1g* cases demonstrated significantly shorter survival after
185 accounting for age, stage and debulking status (mHR for OS = 1.52, 95% CI 1.10-2.10). The *BRCA2m*
186 and *EMSY*-overexpressing subgroups demonstrated the highest rates of complete response to first
187 and second-line chemotherapy as determined by radiology or CA125 tumour marker (figure 2C).
188 Complete response rate was higher in the *BRCA2m*, *EMSY*-overexpressing and *BRCA1m* cases
189 compared to the HRRwt cases at first chemotherapy ($P < 0.001$, $P = 0.009$ and $P = 0.049$ for complete
190 GCIG CA125 response [confirmed normalisation from at least double upper limit of normal]); however,
191 only *BRCA2m* and *EMSY*-overexpressing cases had a significantly higher complete response rate after
192 adjusting for multiple testing ($P\text{-adj} = 0.001$, $P\text{-adj} = 0.027$, $P\text{-adj} = 0.148$). At relapse, *BRCA2m* and *EMSY*-
193 overexpressing cases retained a higher chemotherapy response rate ($P = 0.002$ and $P = 0.037$ for
194 complete CA125 response, respectively). Chemotherapy response rate was similar in the *CCNE1g* and
195 HRRwt groups at both primary treatment and relapse (figure 2C).

196 3.3 Relationship between transcriptomic subtypes

197 Two transcriptomic subtyping approaches were used (TCGA subtypes: DIF, IMR, PRO, MES; Tothill
198 subtypes: C1, C2, C4, C5). There was marked overlap between the subtyping approaches ($P < 0.0001$)
199 (figure 3A): PRO cases were overwhelmingly of the C5 subtype (91.0%, 61 of 67), while the vast
200 majority of MES cases were of the C1 subtype (88.9%, 88 of 99). The DIF group comprised mainly C4
201 tumours (69.6%, 71 of 102), while IMR cases were mostly of the C2 subtype (66.0%, 62 of 94).

202 *3.4 Genomic-transcriptomic correlates*

203 There was marked association between HRR-centric and transcriptomic subtypes (figure 3B and 3C).
204 Frequency of *BRCA1* and *BRCA2* mutation differed significantly between transcriptomic subtypes, with
205 the highest and lowest *BRCA1/2m* rates in the IMR/C2 and PRO/C5 subtypes, respectively (25.5% and
206 32.5% vs 6.0% and 8.6%, $P\text{-adj} = 0.009$ and 0.003) (figure 3B and 3C). Frequency of *CCNE1g* was highest
207 in PRO/C5 tumours (23.9% in PRO, 22.2% in C5), while the C2 subtype demonstrated the lowest
208 *CCNE1g* frequency (8.8%, $P = 0.002$) (figure 3B and 4C).

209 Prolonged survival for HRR deficient cases was apparent across all transcriptomic subtypes (HR range
210 0.48-0.68) (figure S3 and S4). We did not observe any significant differences in overall burden of CN
211 loss or gain events between transcriptional subtypes (figure S5).

212 *3.5 Immune cell infiltration burden*

213 The burden of tumour-infiltrating CD3+ and CD8+ cells was heterogeneous across samples, with higher
214 infiltration associated with prolonged survival (figure S6). *BRCA2m* cases demonstrated the highest
215 levels of CD3+ infiltration (figure 4A).

216 Subtypes defined by both transcriptomic subgrouping methodologies demonstrated marked
217 differences in infiltrating CD3+ ($P\text{-adj} < 0.0001$) (figure 4B and 4C) and CD8+ cells ($P\text{-adj} < 0.0001$) (figure
218 S7). The IMR/C2 subtypes demonstrated the highest infiltration levels, while the PRO/C5 subtypes
219 demonstrated uniformly low levels of infiltrating cells.

220 *3.6 RB and PTEN loss in HGSOc*

221 10.6% of cases (37 of 350 evaluable tumours) demonstrated PTEN protein loss (figure 1, figure 5A).
222 PTEN loss was a rare event in tumours of the PRO/C5 subtypes (3.0% in PRO, 2.5% in C5) (figure S9A
223 and S9B). Cases with loss of PTEN expression demonstrated significantly lower *PTEN* CN (P=0.0003)
224 (figure S9A).

225 16.8% of cases (59 of 352 evaluable tumours) demonstrated loss of RB protein (figure 1, figure S8C
226 and S8D). RB loss was ubiquitous among HGSOCs harbouring *RB1* mutation (11/11 cases
227 demonstrating loss; P<0.001 vs 48/341 in the absence of *RB1* mutation) (figure 1). Cases
228 demonstrating RB loss had a lower *RB1* CN (P=0.0258) (figure S9B) and there was significant co-
229 occurrence between RB loss and PTEN loss (22.8% PTEN loss in RB-lost cases, 13/57 vs 7.7%, 22/285
230 evaluable cases; P=0.001) (figure 1).

231 RB loss was significantly enriched among cases with HRR gene aberrations (*BRCA1m*, *BRCA2m*, *EMSY*-
232 overexpression or non-*BRCA*-HRR mutation) (26.0%, 27/104 evaluable cases vs 12.9%, 32/248;
233 P=0.005) (figure 5A) and was a rare event among *CCNE1g* cases (5.7%). In cases with HRR gene
234 aberrations, RB loss was associated with significantly longer survival (mHR for OS=0.50, 95% CI 0.30-
235 0.84) (figure 5B); conversely, RB loss was not associated with significant differences in survival within
236 the remaining population (mHR=0.71, 95% CI 0.53-1.06) (figure 5B).

237

238 4. DISCUSSION

239 Substantial advancements in our understanding of HGSOC biology have been made over the last two
240 decades, with many studies characterising HGSOC cases at the gene sequence and gene expression
241 level [4-6, 19]. These investigations have identified subgroups of patients with differential outcome
242 and therapy sensitivity, paving the way for molecular stratification of HGSOC patient care [8, 22, 27-
243 29]. However, the relationship between features described at the genomic and transcriptomic level is
244 poorly understood. We present matched genomic-transcriptomic characterisation – alongside

245 identification of other molecular features, including RB expression loss, PTEN expression loss, and
246 immune cell infiltration – in a large pathologically-confirmed HGSOC cohort with detailed clinical
247 annotation and extensive follow-up, revealing marked correlation across these levels of molecular
248 characterisation.

249 We utilised two transcriptomic subtyping approaches within our dataset: there was substantial
250 correlation between TCGA (PRO, MES, DIF, IMR) and Tothill (C1, C2, C4, C5) subtypes. The MES and
251 PRO TCGA subtypes demonstrated marker overlap with the C1 and C5 Tothill subtypes, while the
252 majority of DIF and IMR cases were of the C4 and C2 subtypes, respectively. This overlap is consistent
253 with previous reports of overlap between these subtyping approaches [19].

254 When comparing genomic features of these subtypes, the IMR/C2 groups demonstrated enrichment
255 for *BRCA1/2m*. These cases also demonstrated the highest immune cell infiltration burden, with
256 significantly greater levels of CD3+ and CD8+ cell infiltration. Together, the high *BRCA1/2m* rate and
257 high levels of immune engagement in IMR/C2 tumours likely underpin the favourable outcome
258 reported in these patient groups [19, 28]. In contrast, the vast majority of PRO/C5 cases were *BRCA1/2*
259 wild-type, and instead demonstrated the highest rates of *CCNE1g*; the PRO/C5 subtypes may therefore
260 represent the group least likely to benefit from PARP inhibition. Previous reports have suggested that
261 PRO cases may derive greatest benefit from anti-angiogenic agents such as bevacizumab [22]; the low
262 *BRCA1/2m* rate in this group may support the use of these agents over PARPi in this patient group.
263 Conversely, the IMR/C2 group harbour a large number of *BRCA1/2m* patients that are likely to benefit
264 from PARPi. Some investigators have suggested that anti-angiogenic therapies may not confer
265 greatest benefit in patients with HGSOC demonstrating immune-related gene expression signatures
266 [17], or may not benefit some groups of HRR-deficient patients [15, 30, 31]. Together, these data
267 suggest that further dissection of the relationship between transcriptional subtypes, HRR status and
268 relative benefit of single versus combined PARPi/anti-angiogenic strategies is required.

269 PRO/C5 cases were also uniformly low in tumour-infiltrating CD3+ and CD8+ cells, suggesting poor
270 engagement of the immune system against the tumour within this patient group; this may partially
271 account for the shorter survival time previously described in these cases [19, 28]. These data suggest
272 that immune checkpoint inhibitors, currently under investigation in ovarian cancer, are also unlikely
273 to represent viable therapeutic options for improving survival time in these patients [32].

274 Our multi-layer characterisation also sheds further light upon HRR pathway players and their
275 importance in HGSOc. We show that *EMSY*-overexpressing cases appear *BRCA2m*-like in their survival
276 profile and therapy sensitivity – consistent with *EMSY*'s role as a *BRCA2* regulator [16]. However, they
277 do not appear to be over-represented in the IMR/C2 transcriptomic subtypes, and do not demonstrate
278 a higher burden of tumour-infiltrating immune cells. We also demonstrate the importance of
279 aberrations in HRR genes regardless of the transcriptional subtype context; the hazard ratio for cases
280 with HRR gene aberrations ranged between 0.48-0.68 across all transcriptional subtypes. These data
281 confirm that the survival benefit among patients with tumours displaying HRR gene aberrations is not
282 due to differential distribution of transcriptional subtypes.

283 *CCNE1g* has been the focus of intense research interest since its identification as a recurrent event in
284 HGSOc [5, 33-36]. A number of studies have suggested that cases harbouring *CCNE1g* have poorer
285 survival, with some suggesting this is due to greater intrinsic chemoresistance [5, 33-35]. However,
286 these comparisons have typically been made against the wider non-*CCNE1g* population without
287 accounting for HRR-deficiency, which is associated with longer survival and increased platinum
288 sensitivity, confounding these comparisons. We compare the *CCNE1g* population directly to non-
289 *CCNE1g* HRRwt cases. *CCNE1g* was not associated with significantly poorer response rate to first-line
290 chemotherapy, or chemotherapy for relapsed disease, within our cohort. We show that, although the
291 most advanced stage cases are under-represented in the *CCNE1g* group, *CCNE1g* cases demonstrate
292 shorter survival time and that their survival is significantly poorer compared to non-*CCNE1g* HRRwt
293 patients upon multivariable analysis. *CCNE1g* tumours also demonstrated the lowest levels of

294 infiltrating immune cells compared to the other HRR-centric groups, which may contribute toward the
295 shorter patient survival time. Mutual exclusivity of *CCNE1g* and HRR gene events suggests that the
296 former are likely to represent a patient group who benefit least from PARP inhibition. Moreover, the
297 low immune infiltration levels demonstrated by these cases suggests that immune checkpoint
298 inhibitors are unlikely to be effective in these patients [32]. Given that *CCNE1g* is most frequent in the
299 PRO/C5 transcriptomic subtypes [22], and that the PRO subtype has been associated with greatest
300 benefit from bevacizumab in some reports, *CCNE1g* cases may represent those likely to derive benefit
301 from anti-angiogenic therapies. *CCNE1g* cases – alongside other HRR-proficient patient groups –
302 represent HGSOC patients with shorter survival time for which new treatment approaches are needed
303 to improve survival. However, the low frequency of other molecular events in *CCNE1g* cases (*BRCA1/2*
304 wild-type, RB intact, PTEN intact, low immune cell infiltration), represents a challenge toward
305 identifying further candidate biologically-targeted strategies within this patient group. Inhibition of
306 WEE1 kinase represents a potential strategy of interest for *CCNE1g* cases, with recent data
307 demonstrating objective responses to WEE1 inhibitors in treatment-refractory *CCNE1g* HGSOC [37].

308 Disruption of *PTEN* and *RB1* have only recently been identified as highly recurrent events in HGSOC
309 [6]. The relationship of these events to other molecular features and their impact on patient outcome
310 is poorly understood. We demonstrate that PTEN and RB protein loss are neither mutually exclusive
311 with one another, nor mutually exclusive with other recurrent genomic events in this tumour type.
312 Indeed, the frequency of RB loss was significantly higher in HRR-deficient cases and there was
313 significant co-occurrence between RB and PTEN loss. By contrast, RB loss was a rare event in *CCNE1g*
314 cases. Cases with RB or PTEN loss demonstrated reduced CN at their respective loci; however, not all
315 cases with loss demonstrated low CN, suggesting mechanisms of inactivation beyond CN loss,
316 consistent with reports of complex structural variants (SVs) affecting both *RB1* and *PTEN* [6]. Perhaps
317 most interestingly, RB status discriminated outcome within the cases showing HRR gene aberrations,
318 with the RB loss significantly associated with longer survival; RB-loss did not significantly impact
319 outcome in cases without identifiable events in HRR genes. It is unclear whether this phenotype is due

320 to differences in therapy sensitivity, or whether concurrent RB loss results in HRR-aberrant tumours
321 with more indolent behaviour. Mechanistic work investigating the phenotypic and signalling
322 consequences of RB loss in the context of HRR-deficiency is now warranted, including investigation of
323 the relative chemosensitivity of RB-lost and RB-intact HRR-deficient cells.

324 We present a large, pathologically-confirmed HGSOc patient cohort with extensive follow-up and
325 detailed clinical annotation, including chemotherapy response data. Together with the multiple layers
326 of molecular characterisation, these represent major strengths of this work. However, we were unable
327 to characterise genome-wide SVs due to a lack of whole genome sequencing, which is a limitation of
328 the study. Lack of *BRCA1/2* SV and *BRCA1* promoter methylation data will have likely resulted in a
329 more conservative HR estimate when comparing our HRR-aberrant and HRRwt populations. SVs such
330 as translocations and inversions are known to affect *NF1* in a proportion of HGSOc patients [6], and
331 we were unable to characterise this patient group in our study. Future work should seek to provide
332 even greater resolution within the HRR-proficient patient population, including characterisation of
333 cases with *NF1* loss.

334 5. CONCLUSION

335 Together, these data provide a high resolution picture of the molecular landscape in HGSOc,
336 integrating genomic sequencing with copy number data, transcriptomic profiling and immune cell
337 infiltration burden in a cohort of HGSOc with rich clinical annotation. Specific transcriptomic subtypes
338 are associated with marked differences in frequency of HRR gene aberrations, *CCNE1g* and infiltration
339 of immune cells; integration of these data highlights patient groups most likely to responding to
340 conventional chemotherapy and targeted biological therapeutics. Patients with *CCNE1g* and HRRwt
341 tumours represent those with greatest unmet clinical need; investigations of new treatment strategies
342 should focus on this patient group. RB and PTEN loss are common in HGSOc and frequently occur
343 alongside other molecular events, with RB loss affecting a large number of tumours with HRR gene
344 aberrations.

345 **Funding**

346 RLH was supported by an MRC-funded fellowship (I171113-1019) and received funding from Target
347 Ovarian Cancer, Tenovus Scotland (E19-11) and the Nicola Murray Foundation during the course of
348 this work. AMM and CAS received core funding from the UK Medical Research Council to the MRC
349 Human Genetics Unit (MC_UU_00007/16). COM received funding from The Melville Trust for Care and
350 Cure of Cancer. Sample collection was supported by Cancer Research UK Experimental Cancer
351 Medicine Centre funding. TR was supported by core funding from CRUK awarded to the CRUK
352 Edinburgh Centre. The genomic characterisation of this cohort was funded by a research grant from
353 AstraZeneca.

354 **Authors' contributions**

355 RLH: conceptualisation, formal analysis, investigation, methodology, visualisation, writing – original
356 draft. AMM: methodology, investigation, writing – review and editing. COM: data curation, resources,
357 writing – review and editing. TR: data curation, investigation, writing – review and editing. MC: data
358 curation, project administration, writing – review and editing. AHP: investigation, writing – review and
359 editing. IC: investigation, writing – review and editing. WGM: investigation, writing – review and
360 editing. ARWW: investigation, writing – review and editing. CB: data curation, writing – review and
361 editing. YI: investigation, writing – review and editing. AO: supervision, writing – review and editing.
362 BD: methodology, writing – review and editing. JCB: conceptualisation, supervision, writing – review
363 and editing. RM: conceptualisation, supervision, writing – review and editing. AM: conceptualisation,
364 supervision, writing – review and editing. PR: conceptualisation, supervision, writing – review and
365 editing. CAS: methodology, supervision, writing – review and editing. DPH: methodology, writing –
366 review and editing. RK: methodology, writing – review and editing. CSH: conceptualisation,
367 investigation, methodology, supervision, writing – review and editing. CG: conceptualisation,
368 methodology, funding acquisition, supervision, writing – review and editing.

369 **Declaration of competing interests**

370 RLH: consultancy fees from GSK outside the scope of this work. AMM: none. COM: none. TR: none.
371 MC: none. IC: none. AHP: none. WGM: none. ARWW: consultancy fees from Bayer and Mithra outside
372 the scope of this work. CB: none. YI: none. AO: none. BD, JCB, RM: employees and stockholders of
373 AstraZeneca. AM: employee and stockholder of AstraZeneca during the course of this work, and
374 employee of GSK. PR: grants from AstraZeneca during the conduct of this work, grants from
375 AstraZeneca outside of this work, personal fees from AstraZeneca and GlaxoSmithKline/Tesaro
376 outside of this work. CAS: none. DPH and RK: employees of Almac Diagnostics. CSH: none. CG: grants
377 from AstraZeneca during the conduct of this study, personal fees from MSD, GSK, Tesaro, Clovis,
378 Roche, Foundation One, Chugai, Takeda, Sierra Oncology and Cor2Ed outside the submitted work,
379 patents PCT/US2012/040805 issued, PCT/GB2013/053202 pending, 1409479.1 pending, 1409476.7
380 pending, and 1409478.3 pending.

381

382 **Acknowledgements**

383 We thank the patients who contributed to this study and the Edinburgh Ovarian Cancer Database from
384 which the clinical data reported here were retrieved. We are grateful to the NRS Lothian Human
385 Annotated Bioresource, NHS Lothian Department of Pathology and Edinburgh Experimental Cancer
386 Medicine Centre for their support. We thank the Edinburgh Clinical Research Facility, Western General
387 Hospital, Edinburgh, UK for their support with the high throughput sequencing described here.

388

389

390 References

- 391 1. Irodi, A., et al., *Patterns of clinicopathological features and outcome in epithelial ovarian*
392 *cancer patients: 35 years of prospectively collected data*. *BJOG: An International Journal of*
393 *Obstetrics & Gynaecology*. n/a(n/a).
- 394 2. Pignata, S., et al., *Treatment of recurrent ovarian cancer*. *Annals of Oncology*, 2017. **28**: p.
395 viii51-viii56.
- 396 3. Kuroki, L. and S.R. Guntupalli, *Treatment of epithelial ovarian cancer*. *BMJ*, 2020. **371**: p.
397 m3773.
- 398 4. Tothill, R.W., et al., *Novel molecular subtypes of serous and endometrioid ovarian cancer*
399 *linked to clinical outcome*. *Clin Cancer Res*, 2008. **14**(16): p. 5198-208.
- 400 5. Bell, D., et al., *Integrated genomic analyses of ovarian carcinoma*. *Nature*, 2011. **474**(7353): p.
401 609-615.
- 402 6. Patch, A.M., et al., *Whole-genome characterization of chemoresistant ovarian cancer*. *Nature*,
403 2015. **521**(7553): p. 489-94.
- 404 7. Waldron, L., et al., *Comparative meta-analysis of prognostic gene signatures for late-stage*
405 *ovarian cancer*. *J Natl Cancer Inst*, 2014. **106**(5).
- 406 8. Moore, K., et al., *Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian*
407 *Cancer*. *N Engl J Med*, 2018.
- 408 9. Mirza, M.R., et al., *Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian*
409 *Cancer*. *N Engl J Med*, 2016. **375**(22): p. 2154-2164.
- 410 10. Colombo, N. and J.A. Ledermann, *Updated treatment recommendations for newly diagnosed*
411 *epithelial ovarian carcinoma from the ESMO Clinical Practice Guidelines*. *Annals of Oncology*,
412 2021.
- 413 11. Gourley, C., et al., *Role of Poly (ADP-Ribose) Polymerase inhibitors beyond BReast CAncer*
414 *Gene-mutated ovarian tumours: definition of homologous recombination deficiency?* *Curr*
415 *Opin Oncol*, 2020. **32**(5): p. 442-450.
- 416 12. Walsh, T., et al., *Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal*
417 *carcinoma identified by massively parallel sequencing*. *Proceedings of the National Academy*
418 *of Sciences*, 2011. **108**(44): p. 18032.
- 419 13. Ewing, A., et al., *Structural Variants at the BRCA1/2 Loci are a Common*
420 *Source of Homologous Repair Deficiency in High-grade Serous Ovarian Carcinoma*. *Clinical*
421 *Cancer Research*, 2021.
- 422 14. Cunningham, J.M., et al., *Clinical characteristics of ovarian cancer classified by BRCA1, BRCA2,*
423 *and RAD51C status*. *Sci Rep*, 2014. **4**: p. 4026.
- 424 15. Hollis, R.L., et al., *High EMSY expression defines a BRCA-like subgroup of high-grade serous*
425 *ovarian carcinoma with prolonged survival and hypersensitivity to platinum*. *Cancer*, 2019.
426 **125**(16): p. 2772-2781.
- 427 16. Hughes-Davies, L., et al., *EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer*.
428 *Cell*, 2003. **115**(5): p. 523-35.
- 429 17. Gourley, C., et al., *Molecular subgroup of high-grade serous ovarian cancer (HGSOC) as a*
430 *predictor of outcome following bevacizumab*. *Journal of Clinical Oncology*, 2014. **32**(15_suppl):
431 p. 5502-5502.
- 432 18. Talhouk, A., et al., *Development and Validation of the Gene Expression Predictor of High-grade*
433 *Serous Ovarian Carcinoma Molecular SubTYPE (PROTYPE)*. *Clin Cancer Res*, 2020. **26**(20): p.
434 5411-5423.
- 435 19. Chen, G.M., et al., *Consensus on Molecular Subtypes of High-Grade Serous Ovarian Carcinoma*.
436 *Clin Cancer Res*, 2018. **24**(20): p. 5037-5047.
- 437 20. Zhang, L., et al., *Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer*. *N*
438 *Engl J Med*, 2003. **348**(3): p. 203-13.

- 439 21. Sato, E., et al., *Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory*
440 *T cell ratio are associated with favorable prognosis in ovarian cancer*. Proc Natl Acad Sci U S A,
441 2005. **102**(51): p. 18538-43.
- 442 22. Kommoss, S., et al., *Bevacizumab May Differentially Improve Ovarian Cancer Outcome in*
443 *Patients with Proliferative and Mesenchymal Molecular Subtypes*. Clin Cancer Res, 2017.
444 **23**(14): p. 3794-3801.
- 445 23. An, D., S. Banerjee, and J.M. Lee, *Recent advancements of antiangiogenic combination*
446 *therapies in ovarian cancer*. Cancer Treat Rev, 2021. **98**: p. 102224.
- 447 24. Bankhead, P., et al., *QuPath: Open source software for digital pathology image analysis*.
448 Scientific Reports, 2017. **7**(1): p. 16878.
- 449 25. Kuilman, T., et al., *CopywriteR: DNA copy number detection from off-target sequence data*.
450 Genome Biol, 2015. **16**(1): p. 49.
- 451 26. Irodi, A., et al., *Patterns of clinicopathological features and outcome in epithelial ovarian*
452 *cancer patients: 35 years of prospectively collected data*. BJOG, 2020. **127**(11): p. 1409-1420.
- 453 27. Ledermann, J., et al., *Olaparib maintenance therapy in platinum-sensitive relapsed ovarian*
454 *cancer*. N Engl J Med, 2012. **366**(15): p. 1382-92.
- 455 28. Gourley, C., et al., *Molecular subgroup of high-grade serous ovarian cancer (HGSOC) as a*
456 *predictor of outcome following bevacizumab*. Journal of Clinical Oncology, 2014. **32**(15_suppl):
457 p. 5502.
- 458 29. Tan, D.S., et al., *"BRCAness" syndrome in ovarian cancer: a case-control study describing the*
459 *clinical features and outcome of patients with epithelial ovarian cancer associated with BRCA1*
460 *and BRCA2 mutations*. J Clin Oncol, 2008. **26**(34): p. 5530-6.
- 461 30. Lorusso, D., et al., *Bevacizumab as maintenance treatment in BRCA mutated patients with*
462 *advanced ovarian cancer: A large, retrospective, multicenter case-control study*. Gynecol
463 Oncol, 2020. **159**(1): p. 95-100.
- 464 31. Liu, J.F., et al., *A phase III study comparing single-agent olaparib or the combination of*
465 *cediranib and olaparib to standard platinum-based chemotherapy in recurrent platinum-*
466 *sensitive ovarian cancer*. Journal of Clinical Oncology, 2020. **38**(15_suppl): p. 6003-6003.
- 467 32. Zamarin, D., et al., *Randomized Phase II Trial of Nivolumab Versus Nivolumab and Ipilimumab*
468 *for Recurrent or Persistent Ovarian Cancer: An NRG Oncology Study*. J Clin Oncol, 2020. **38**(16):
469 p. 1814-1823.
- 470 33. Nakayama, N., et al., *Gene amplification CCNE1 is related to poor survival and potential*
471 *therapeutic target in ovarian cancer*. Cancer, 2010. **116**(11): p. 2621-34.
- 472 34. Etemadmoghadam, D., et al., *Synthetic lethality between CCNE1 amplification and loss of*
473 *BRCA1*. Proc Natl Acad Sci U S A, 2013. **110**(48): p. 19489-94.
- 474 35. Etemadmoghadam, D., et al., *Integrated genome-wide DNA copy number and expression*
475 *analysis identifies distinct mechanisms of primary chemoresistance in ovarian carcinomas*. Clin
476 Cancer Res, 2009. **15**(4): p. 1417-27.
- 477 36. Sapoznik, S., et al., *CCNE1 expression in high grade serous carcinoma does not correlate with*
478 *chemoresistance*. Oncotarget, 2017. **8**(37): p. 62240-62247.
- 479 37. Fu, S., et al., *Abstract 974: Phase II trial of the Wee1 inhibitor adavosertib in advanced*
480 *refractory solid tumors with CCNE1 amplification*. Cancer Research, 2021. **81**(13 Supplement):
481 p. 974.

482

483

484 Tables

485 Table 1. Clinicopathological features of high grade serous carcinoma

		N	%
Total cases	N	362	
Age at diagnosis	Median years	61	Range 33-86
FIGO stage at diagnosis	I	15	4.3
	II	31	8.8
	III	237	67.5
	IV	68	19.4
	NA	11	-
RD following surgical debulking	No visible RD (0cm)	65	19.4
	Macro RD (0.1-2cm)	66	19.7
	Gross macro RD (\geq 2cm)	187	55.8
	Macro RD of unknown size	17	5.1
	Unknown	27	-
First line chemotherapy	Single-agent platinum	217	59.9
	Platinum-taxane combination	135	37.3
	Other platinum-containing regimes	10	2.8
Vital status at last follow-up	Alive	36	9.9
	Deceased - died of OC	307	84.8
	Deceased - other causes	19	5.2
Median follow-up time	Years	15.0 (95% CI 12.8-19.9)	
Median PFS	Years	1.17 (95% CI 1.09-1.28)	
Median OS	Years	2.60 (95% CI 2.40-3.07)	

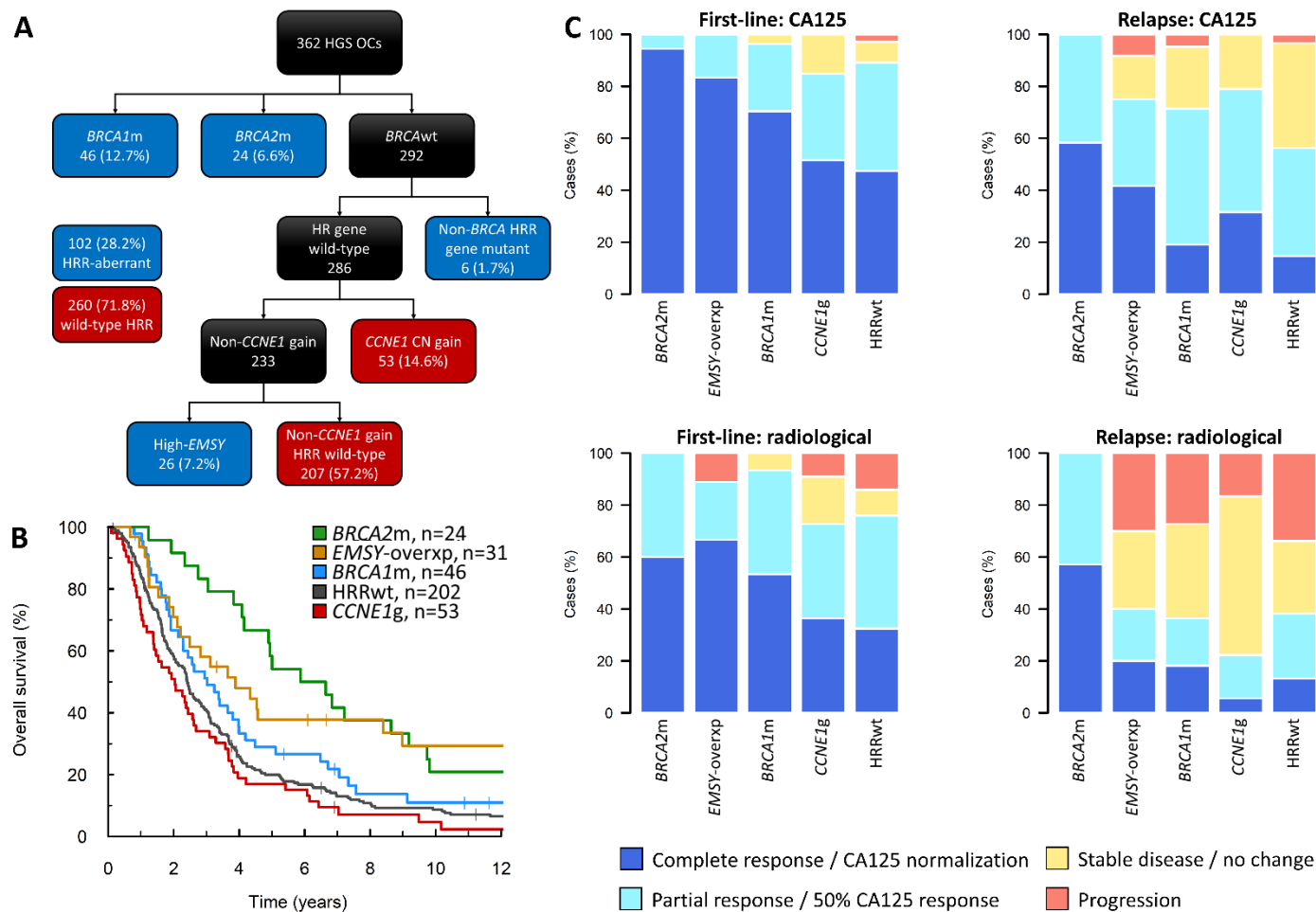
486 FIGO, International Federation of Gynecology and Obstetrics; NA, not available; Macro, macroscopic;

487 RD, maximal residual disease diameter; OC, ovarian carcinoma



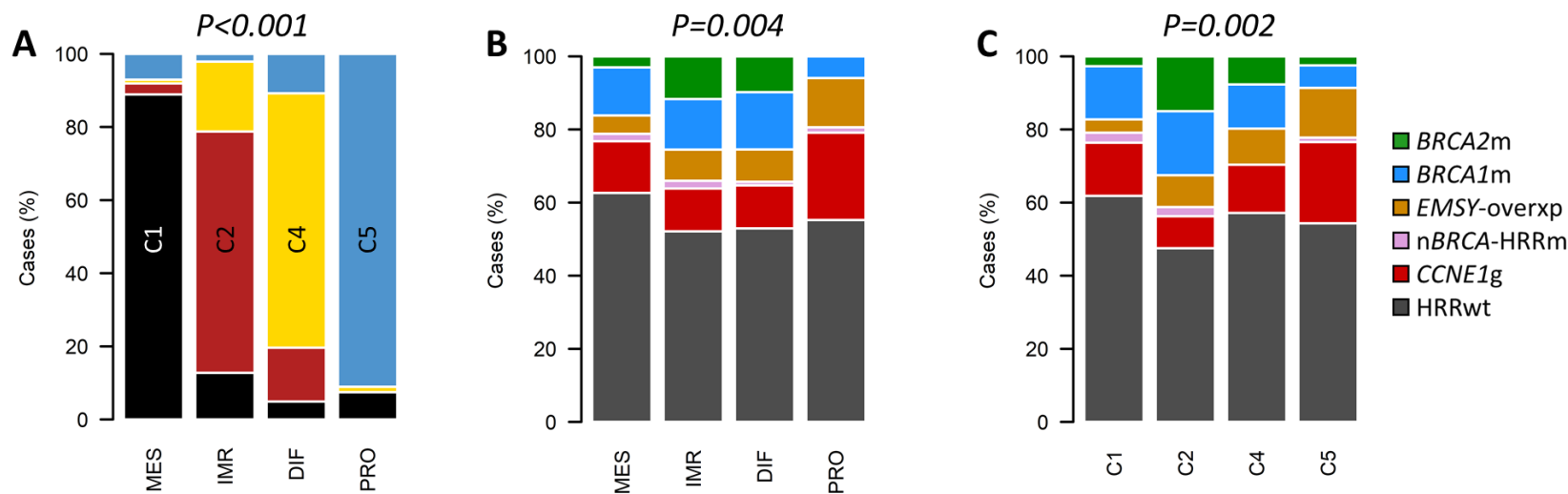
489

490 Figure 1. Molecular landscape of high grade serous ovarian carcinoma. ¹Mutation in non-*BRCA1/2* homologous recombination repair (HRR) genes: 3 *BRIP1*, 2
 491 *CHEK2*, 1 *RAD51C*, 1 *PALB2*, 1 concurrent *BAP1* and *NBN*. *CCNE1* copy number (CN) gain, ≥ 4 copies by Taqman CN assay. *EMSY* CN amplification, ≥ 6 copies by
 492 Taqman CN assay.



493

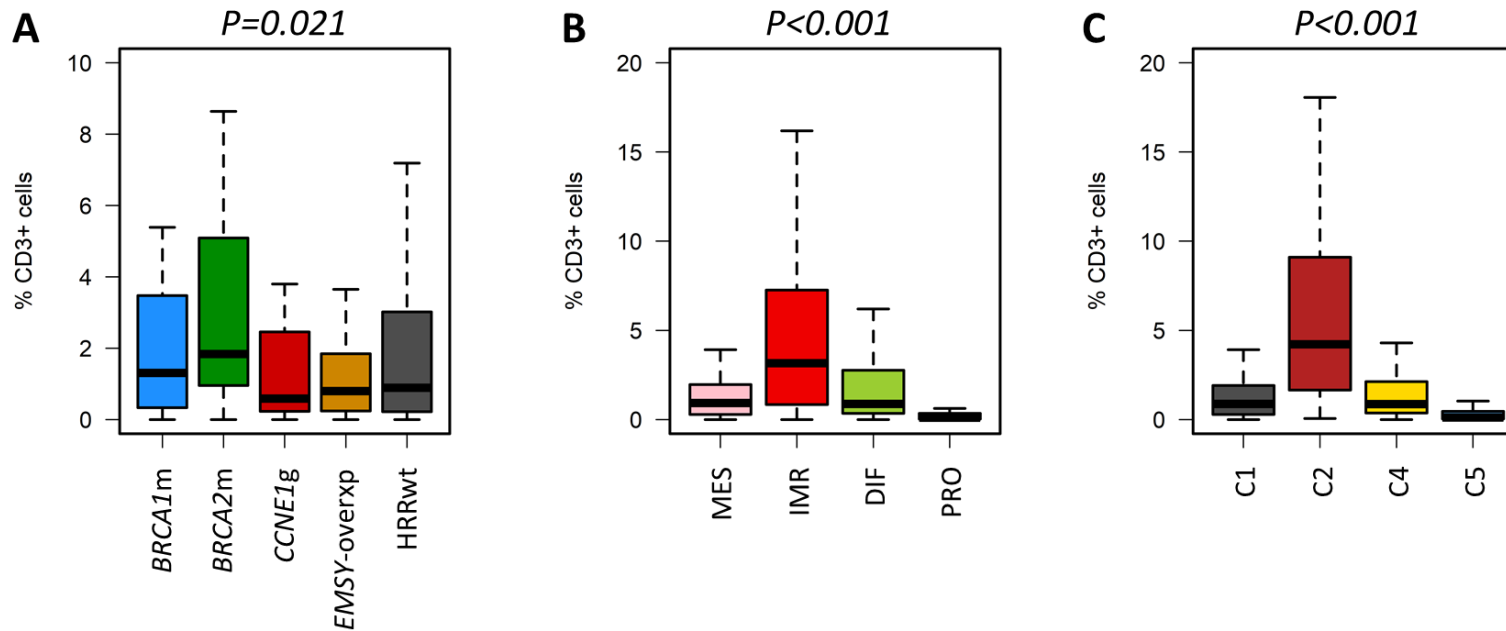
494 Figure 2. Homologous recombination repair pathway (HRR)-centric subtyping of high grade serous carcinoma. (A) HRR-centric classification taxonomy. (B)
 495 Overall survival profile of HRR-centric subtypes. (C) Chemosensitivity of HRR-centric subtypes at first-line treatment (left hand panels) and treatment for
 496 disease relapse (right hand panels) as determined by CA125 tumour marker (top panels) and radiology (bottom panels). *BRCA2m*, *BRCA2* mutant; *BRCA1m*,
 497 *BRCA1* mutant; *EMSY-overxp*; overexpression of *EMSY*; *CCNE1g*, gain of *CCNE1*; HRRwt, non-*CCNE1g* homologous recombination proficient.



499

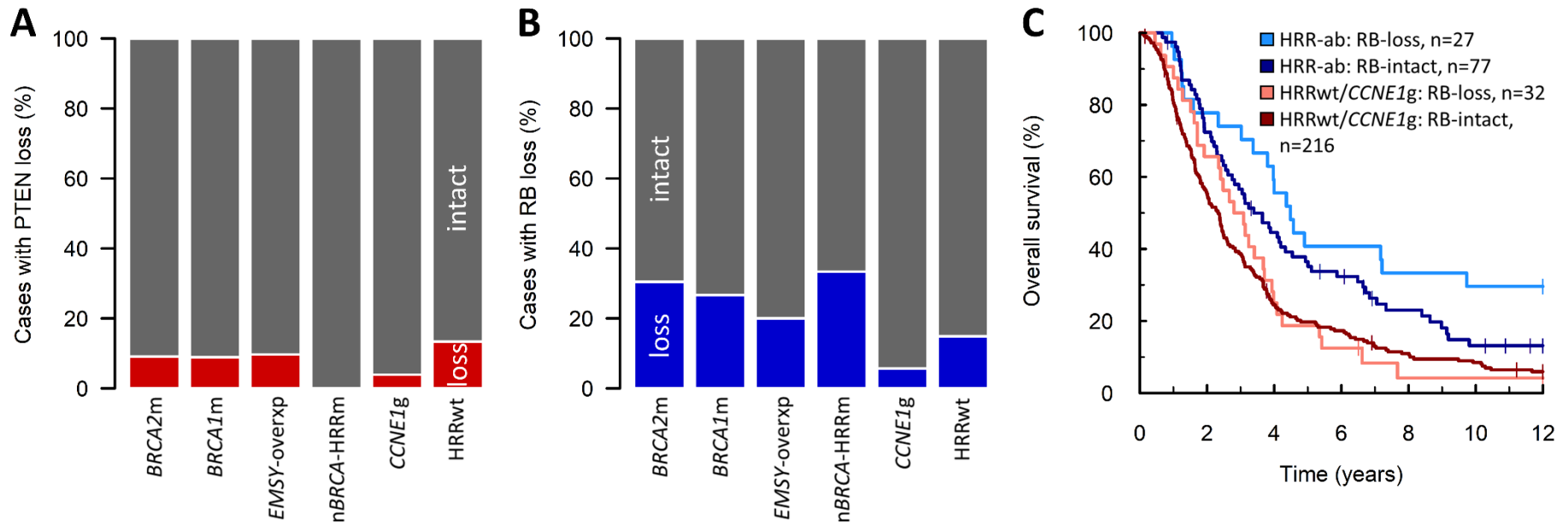
500 Figure 3. Relationship between subgroups methodologies. (A) Comparison of transcriptomic subgrouping approaches: composition of Tothill subtypes across
 501 each of the TCGA subtypes; labelled P value represents comparison of Tothill subtype frequency across all TCGA subtypes by Chi-squared test. (B) Distribution
 502 of homologous recombination repair (HRR)-centric subtypes across each of the TCGA transcriptomic subtypes; labelled P value represents comparison of
 503 $BRCA1/2m$ frequency across all groups by Chi-squared test; P-adj=0.009. (C) Distribution of homologous recombination repair (HRR)-centric subtypes across
 504 each of the Tothill transcriptomic subtypes; labelled P value represents comparison of $BRCA1/2m$ frequency across all groups by Chi-squared test; P-adj=0.003.
 505 $BRCA2m$, $BRCA2$ mutant; $BRCA1m$, $BRCA1$ mutant; $EMSY$ -overxp, overexpression of $EMSY$; nBRCA-HRRm, non- $BRCA1/2$ HRR gene mutation; $CCNE1g$, gain of
 506 $CCNE1$; HRRwt, non- $CCNE1g$ homologous recombination proficient.

507



508

509 Figure 4. Tumour-infiltrating immune cells across high grade serous carcinoma subtypes. (A) CD3+ infiltration across HRR-centric subtypes; labelled P value
 510 represents comparison of *BRCA2m* and *CCNE1g* groups using the Mann Whitney-U test. (B) CD3+ infiltration across TCGA transcriptomic subtypes; labelled P
 511 value represents comparison of IMR and PRO groups using the Mann Whitney-U test. (C) CD3+ infiltration across Tothill transcriptomic subtypes; labelled P
 512 value represents comparison of C2 and C5 groups using the Mann Whitney-U test. *BRCA2m*, *BRCA2* mutant; *BRCA1m*, *BRCA1* mutant; *EMSY-overxp*;
 513 overexpression of *EMSY*; *CCNE1g*, gain of *CCNE1*; HRRwt, non-*CCNE1g* homologous recombination proficient.



514

515 Figure 5. PTEN and RB loss in high grade serous ovarian carcinoma. (A) Frequency of loss of PTEN protein expression across homologous recombination repair
 516 (HRR)-centric subtypes. (B) Frequency of loss of RB protein expression across homologous recombination repair (HRR)-centric subtypes. (C) Impact of RB loss
 517 on survival in patients based on HRR status. Multivariable hazard ratio (mHR) for HRR-ab: RB-loss vs HRR-ab RB-intact=0.50, 95% CI 0.30-0.84; mHR for
 518 HRRwt/CCNE1g: RB-loss vs HRRwt/CCNE1g: RB-intact=0.71, 95% CI 0.48-1.06. BRCA2m, BRCA2 mutant; BRCA1m, BRCA1 mutant; EMSY-overxp;
 519 overexpression of EMSY; nBRCA-HRRm, non-BRCA1/2 HRR gene mutation; CCNE1g, gain of CCNE1; HRRwt, non-CCNE1g HRR wild-type. HRR-ab, HRR-aberrant:
 520 BRCA1m, BRCA2m, EMSY-overxp or nBRCA-HRRm.