Gene networks and expression quantitative trait loci associated with platinum-based chemotherapy response in high-grade serous ovarian cancer

Jihoon Choi[1], Anastasiya Tarnouskaya[2], Sean Nesdoly[2], Danai G. Topouza[1], Madhuri Koti[1], Qing Ling Duan[1,2]

[1]Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario, Canada; [2]School of Computing, Queen's University, Kingston, Ontario, Canada;

Correspondence: QL Duan (qingling.duan@queensu.ca)

Abstract

Introduction: A major impediment in the treatment of ovarian cancer is the relapse of platinum-resistant tumors, which occurs in approximately 25% of patients. A better understanding of the genetic mechanisms underlying response to platinum-based chemotherapy will improve treatment efficacy.

Objectives: We used a systems biology approach to identify novel gene networks and regulatory sequence variants associated with platinum-based chemotherapy response.

Methods: From the Cancer Genome Atlas (TCGA), we classified high-grade serous ovarian carcinoma (HGSOC) patients who remained cancer-free 12 months following completion of platinum-based chemotherapy as “chemo-sensitive” (N=160) and those who had cancer recurrence within six months of chemotherapy as “chemo-resistant” (N=110). Both univariate and multivariate analysis of gene expression microarrays identified differentially expressed genes and co-expression networks associated with chemotherapy response. Moreover, we integrated genomics data to determine expression quantitative trait loci (eQTL).

Results: One differentially expressed gene encoding Valosin-containing protein (VCP) and five co-expression network modules were associated with chemotherapy response in HGSOC patients. These genes contribute to protein processing in the endoplasmic reticulum, which has been previously implicated in variable chemotherapy response. In addition, 192 eQTLs were associated with co-expressed gene modules as well as genes regulating cholesterol levels, also previously described to affect chemotherapy response.

Conclusion: This study implicates known and novel gene networks underlying response to platinum-based chemotherapy among HGSOC patients. A better understanding of the genetic

30 determinants of chemotherapy response will facilitate genetic testing for predicting drug
31 response, which will increase treatment efficacy and identify patients for alternative therapies.
32

Introduction

Ovarian cancer is the most lethal gynecological malignancy and the 8th leading cause of cancer death in women around the world.[1] According to the Global Cancer Observatory report in 2012, ovarian cancer accounts for 3.6% of all cancer cases and 4.3% of all cancer related deaths worldwide.[2] High-grade serous ovarian carcinoma (HGSOC) is the most malignant form of ovarian cancer that accounts for up to 70% of all cases.[3] Routine diagnosis is often difficult due to the lack of mass screening methods and heterogenous manifestations of the cancer symptoms, which result in approximately 75% of patients diagnosed with advanced stages.[4] The average 5-year survival rates are 39% for Stage 3 and 17% for Stage 4 cancers.[1]

The current standard of care for ovarian cancer is aggressive cytoreductive surgery followed by platinum-based chemotherapy.[4] However, this standard of care is not effective for all patients, with approximately 25% experiencing relapse within six months following platinum-based therapy, likely due to the development of antineoplastic resistance.[5] The median survival time for recurrent ovarian cancers range from 12-24 months.[6,7] Treatment options for patients with recurrent ovarian cancer include non-platinum-based chemotherapy regimens, immunotherapy, and molecular targeted therapy.[7,8]

Ovarian cancer has a multifactorial etiology that includes genetic and non-genetic risk factors. An estimated 23% of cases are hereditary, but the majority are sporadic with multiple reported risk factors such as history of gravidity, infertility, and late age menopause.[9,10] A better understanding of the etiology of ovarian cancer, as well as the genetic mechanisms underlying variable response to platinum-based chemotherapy is needed for improved diagnosis and treatment. For example, previous studies reported that *BRCA1* and

BRCA2 genes, associated with increased risk of ovarian cancer, harbor mutations associated with platinum drug sensitivity and survival time.[11] Similarly, tumor suppressor genes such as *RB1*, *NF1*, *RAD51B*, *PTEN* have been associated with acquired chemotherapy resistance.[12] Earlier studies have also highlighted the importance of the immune system in the treatment of ovarian cancer. For example, loss of chemokines and disruptions to the IFN-γ pathway have been associated with poor treatment outcomes in HGSOC paients[13] whereas the NFκB signaling pathway and elevated expression of *STAT1* were associated with increased response to platinum therapy.[14,15,16] However, these known genetic variations do not account for all of the variability in chemotherapy response among HGSOC patients and there is currently no screening method to accurately predict prognosis prior to start of chemotherapy. Thus, further studies are necessary to determine additional modulators of chemotherapy response, which can be used as biomarkers for genetic testing.

In this study, we classified HGSOC patients as sensitive to platinum-based chemotherapy if they did not experience a recurrence one year after completing chemotherapy. Our definition differs from previous studies which classified patients as sensitive if they did not relapse after a platinum-free interval of six months or greater (The Cancer Genome Atlas Research Network, 2011). In addition, we classified patients as resistant if they experienced tumor progression during chemotherapy or had a recurrence within the six months following the end of therapy. Our classification of sensitive and resistant patients are in accordance with the standard Gynecologic Oncology Group (GOG) criteria.[5] Consequently, we excluded patients who had a recurrence between six to 12 months after platinum therapy. Our strategy for dichotomizing chemotherapy response using these "extremes" of response instead of a single threshold aims to enrich for genetic differences between resistant and sensitive patients.

In addition to our novel classification of chemotherapy response in ovarian cancer patients, we applied a multivariate, systems biology analysis approach to identify and characterize the gene networks associated with this variable therapeutic outcome. Earlier studies largely focused on univariate analysis of gene expression data known as differential gene expression (DGE) analysis[17,18]. This analysis assumes that each gene functions in isolation within the genome and is limited in statistical power due to multiple testing corrections to reduce likelihood of false positives. The multivariate approach used in this study identifies novel connections among genes, which represent potential interactions of genes within biological pathways or networks. Specifically, we applied Weighted Gene Co-expression Network Analysis[19] (WGCNA), which uses an unsupervised machine-learning algorithm to identify clusters of highly correlated or co-expressed genes. Moreover, we correlated sequence variations with co-expressed gene networks known as expression Quantitative trait loci (eQTLs). A better understanding of the biological mechanisms regulating chemotherapy response will enable more effective treatments either by improving the accuracy of genetic testing or through the identification of novel therapies for cancer patients.

Methods

Patient Classification:

Using a cohort of 608 high-grade grade serous ovarian carcinoma (HGSOC) patients from The Cancer Genome Atlas Project (TCGA), we retrieved clinical data from the Genomic Data Commons Portal via the TCGAbiolinks R/Bioconductor package[20]. Among the HGSOC cases, 587 had available clinical data; these were filtered and subsequently classified into two distinct groups based on their response to platinum-based chemotherapy. Specifically, we focused on the time interval between a patients' last primary platinum treatment and the onset of a recurrent tumor, or progression of an existing one for those patients whose initial tumor did not undergo complete remission[17]. We excluded patients with insufficient clinical information, were administered non-platinum drugs, and those who had died for any reason within a year of initial diagnosis, as they cannot be defined conclusively as sensitive or resistant to their adjuvant chemotherapy. Of those remaining, a total of 270 patients were classified as either resistant or sensitive to platinum-based chemotherapy. Patients who developed a new tumor in less than six months following their last platinum treatment were defined as resistant (N=110). In contrast, those who did not have a recurrent tumor event more than a year after their last primary platinum treatment were defined as sensitive (N=160). Those who had a recurrent tumour event between six months to one year following therapy were excluded from the study. This strategy was used to enrich for the genetic differences between the two drug response groups.

Transcriptomics Data Processing and Analysis:

Expression Microarrays: Expression data from the Affymetrix ht_hg_u133a microarray chip was used for univariate and multivariate (co-expression) analysis. Of the 270 subjects

120 classified as sensitive or resistant to chemotherapy, 238 had expression array data (N=138
121 sensitive, 100 resistant). Array measures were adjusted with robust multi-array average
122 (RMA)[21], followed by quantile normalization, and log-transformation using the *affy* package
123 from Bioconductor[22]. Two potential outliers and two duplicated samples were removed from
124 the study using the *arrayQualityMetrics* package for quality control (see **Supplemental Data**
125 **1** for normalization, transformation and quality control assessments), resulting in 135
126 sensitive and 99 resistant subjects remaining. Probe sets were sorted by median absolute
127 deviation (MAD), and the top 50% of the probes with highest variation (n=11,107) were
128 included for analysis. This non-specific filtering step was used to remove non-variable probes,
129 therefore, reducing multiple testing and likelihood of false positives.

130 *Differential Gene Expression Analysis:* The *Limma*[23] package in Bioconductor[24] was used to
131 identify differentially expressed genes between chemo-sensitive and resistant groups, using
132 linear models. We applied the false discovery rate (FDR) method for multiple testing
133 correction to reduce the likelihood of false-positive results. Age at diagnosis and stage of
134 tumor were included as covariates in this analysis (**Supplemental Data 2**).

135 *Weighed Gene Co-expression Network Analysis (WGCNA):* We performed hierarchical
136 clustering of genes with the R package *WGCNA*[19], grouping genes based on their similarity in
137 expression. This was achieved by first creating a similarity matrix using Pearson correlations
138 of expression among all genes. The resulting matrix was raised to a soft-thresholding power
139 of 9, as suggested by the soft-thresholding power estimation plot (**Supplemental Figure 1**).
140 This process amplifies the connection strength between genes and reduces noise in the
141 adjacency matrix. To account for topological similarity of genes, the number of shared
142 neighbors among genes was used to give weight to gene-gene connectivity. To avoid
143 excessive splitting of genes into smaller modules, minimum module size was set to 30, split

sensitivity (deep split) to 4, and modules with similar expression profiles were merged at a height of 0.5 (**Supplemental Figure 2**). Using principal component analysis, we calculated module eigengene for each co-expression cluster to summarize module expression into a single measure. Each module eigengene was tested for association with platinum chemotherapy response using generalized linear models, while adjusting for age at diagnosis and stage of cancer as covariates. Finally, we used *Cytoscape*[25], an open source bioinformatics platform, to visualize gene co-expression networks.

Gene function and pathway annotations: We used the Database for Annotation, Visualization and Integrated Discovery (DAVID) [26] to identify biological pathways and functions that were enriched in each significant gene co-expression module. We also screened significant genes in the GeneMANIA[27] database to identify for functional connections reported in published literature. Next, we searched the UCSC TFBS conservation sites track with DAVID to identify enriched motifs of transcription factors that may co-regulate genes within the associated clusters. Finally, we used the drug–gene interaction database (DGIdb)[28], a public database with curation of data describing relationships between genes, chemicals, drugs, and pathological phenotype, to identify genes with prior reported associations with chemotherapeutic agents.

Genomics Data Processing and Analysis:

Genomics Data: Single nucleotide polymorphisms (SNPs) data from the germline tissues (blood derived normal or solid tissue normal, based on availability) were obtained from the TCGA legacy database. The Affymetrix Genome-Wide Human SNP Array 6.0 was used to capture genetic variations, which detected 906,600 SNPs. Of the 270 subjects classified as resistant or sensitive to platinum-based chemotherapy, 266 (157 sensitive and 109 resistant)

had genotype data available.

Imputations: The imputation of autosomal chromosomes was performed using the Michigan imputation server pipeline[29]. We used the 1000 Genome Project phase 3 sequencing data (version 5)[30] reference panel for the imputation of missing genotypes. We then used Eagle v.2.3[31] for phasing of the genotypes to their respective chromosomes. For the imputation of variants on the X chromosome, SHAPEIT[32] was used for phasing in combination with the 1000 genomes project phase 3 (version 5) reference panel (**Supplemental Data 3**).

Quality control:

Subject level: Two pairs of individuals had a relatedness coefficient (pi-hat) > 0.9, which are likely duplicated samples. One subject from each pair was randomly removed from the dataset. Next, inbreeding coefficients (F) were computed for each subject using PLINK[33]. A total of 18 subjects with high homozygosity (F>0.05) or heterozygosity (F<-0.05) rates were excluded. Moreover, sex was computed based on heterozygosity rates (F) of the X chromosome. Four subjects whose predicted gender was undefined (F>0.2) were removed from the study.

SNP level: SNPs with minor allele frequencies (MAF) less than 1% or with genotyping call rate less than 90% were removed. This step removed 38,430,595 SNPs with MAF < 0.01, resulting in 9,528,963 SNPs to be used for further analysis.

Genome-wide Association Study: After imputations and quality control filtering, 240 subjects (N= 142 sensitive, 98 resistant) and a total of 9,528,963 SNPs (MAF > 0.1) remained for analysis (**Supplemental Data 4**). Plink (v.1.90) was applied with association analysis of platinum-based chemotherapy response using a logistic regression method. Age at diagnosis and tumor stages were included as covariates.

Expression Quantitative Trait Loci (eQTL) Analysis:

Common SNPs (MAF > 0.01) were tested for association with gene expression of individual genes and module eigengenes using the *matrixeQTL* R package[34]. Correlated loci indicated potential regulatory function of the SNPs on the expression of the nearby corresponding gene, known as cis-expression Quantitative Trait Loci (cis-eQTL). Cis-eQTL were defined as correlated SNPs within 1 Mb from the gene transcriptional start site (TSS).

Results

Differential gene expression (DGE) analysis returned one significant probe mapping to Valosin Containing Protein (*VCP*) with FDR adjusted p-value < 0.05 (**Figure 1**) as well as 606 probes mapping to 521 unique genes with nominal significance (unadjusted p-value < 0.05; **Supplemental Table 1**). Functional annotation analysis showed that the 521 genes were enriched for many oncogenic processes, such as cellular response to DNA damage stimulus (GO:0006974), DNA repair (GO:0006281), positive regulation of cell growth (GO:0030307), and positive regulation of apoptotic process (GO:0043065). Pathway analysis identified many immune related pathways, such as mitogen-activated protein kinase (MAPK) pathway and B-Cell antigen receptor (BCR) signaling pathway. A detailed list of functional annotations and identified pathways of differentially expressed genes can be found in **Supplemental Data 5**.

The hierarchical clustering of genes using *WGCNA* resulted in 86 unique modules of co-expressed genes (**Supplemental Table 2**). Each module was assessed for association with chemotherapy response (**Supplemental Figure 3**). Five gene clusters (*honeydew1, lightcyan1, lightpink3, orangered4, and skyblue3*) were identified to be correlated with sensitivity to chemotherapy ($p < 0.05$); one gene cluster remained significant after Bonferroni correction

(*honeydew1*, p-value=7e-05) (**Figure 2**). The significant modules were annotated using DAVID, which identified gene enrichment for biological pathways including apoptosis, negative regulation of the Wnt signaling pathway, transcription, immune responses, and DNA double-strand break processing involved in repair via single-strand annealing. GeneMANIA analysis showed that genes in these modules were previously reported in 49 publications, some of which documented associations with oncogenic pathways and chemotherapeutic outcomes. We searched in these genes in the gene-drug interaction database (DGIdb) and found that 35 were associated with chemotherapeutic agents such as carboplatin and paclitaxel. Finally, we identified over-represented conserved transcription factor binding sites located in genes from each module. A detailed list of functional annotations, transcription factors and pathways related to gene modules can be found in **Supplemental Data 6.**

Our genome-wide association study (GWAS) of SNPs did not identify significantly correlated variants for platinum chemotherapy response. The Manhattan plot (**Figure 3**), shows that all SNPs fell below the genome-wide significance threshold ($p<5x10-8$), as indicated by the red horizontal line. As this may due to insufficient statistical power, we performed a targeted association analysis on two well-known genes associated with chemotherapeutic outcomes in ovarian cancer: *BRCA1* and *BRCA2*. Of the 237 SNPs in *BRCA1* and 256 in *BRCA2*, we identified 34 SNPs in *BRCA2* and 1 SNP in *BRCA1* that were significantly associated with chemotherapy response (**Supplemental Table 3**).

Next, SNPs were tested for correlation with expression of the 606 differentially expressed probes and 5 co-expression modules. This identified 5,242 cis-eQTL associated with the expression of differentially expressed probes, and 192 cis-eQTL associated with co-expression networks (**Supplemental Data 7**).

Discussion

In this manuscript, we identified novel genes and gene networks correlated with variable response to platinum-based chemotherapy in HGSOC patients. Using a univariate analysis approach, we identified a differentially expressed locus encoding the valosin-containing protein (*VCP*) associated with sensitivity to platinum-based chemotherapy. In addition, we applied a multivariate, co-expression analysis method for identifying groups of interconnected genes that could contribute to common biological pathways. This method identified 5 clusters of co-expressed genes nominally correlated with chemo-response, one of which remained significant after correction of multiple testing. Genes in these modules have been annotated to be associated with biological pathways such as apoptosis, transcription, immune response, negative regulation of the Wnt signaling pathway and DNA double-strand break processing involved in repair via single-strand annealing.

The most significantly associated probe identified in the DGE analysis was for a gene encoding Valosin-containing protein (*VCP*, p = 3.91E-06). *VCP* plays a critical role in disintegrating large polypeptide cellular structures for further degradation by proteolytic enzymes. It functions to regulate important pathways of DNA repair, replication and cell cycle progression by removing faulty polypeptide structures from chromatin material, ribosomes, endoplasmic reticulum and mitochondria. This gene has been previously identified as a potential biomarker for predicting the success of platinum-based chemotherapy in lung cancer patients.[35] Drugs that inhibit *VCP*, such as the novel inhibitor CB-5083[36], are currently being evaluated for their prospective anticancer properties.

Many of the nominally correlated genes from our DGE analysis are enriched for

pathways known to be critical for oncogenesis and chemotherapeutic drug resistance. For example, probe 213532_at mapping to the A Disintegrin and Metalloproteinase-17 (*ADAM17*) gene was upregulated in the chemo-resistant group (p = 0.017). This is consistent with prior studies, which reported that overexpression of *ADAM17* results in reduced cisplatin-induced apoptosis in hepatocellular carcinoma cells and may contribute to cisplatin resistance via the EGFR pathway[37]. Moreover, we identified another up-regulated probe 205239_at (p = 0.003) for the gene encoding Amphiregulin (AREG), a protein found in the EGFR signaling pathway, which has been reported to promote ovarian cancer stemness and drug resistance to anti-cancer therapy.[38] Taken together, our findings support a potential role of the AREG-EGFR signaling pathway in the development of platinum-resistance in ovarian carcinoma.

It is also important to note that one of the most enriched pathways from our DGE analysis was the protein processing in endoplasmic reticulum (hsa04141) pathway. Using DAVID, we have identified 18 differentially expressed probes mapping to 15 unique genes within this pathway. Prior studies have highlighted that endoplasmic reticulum (ER) stress may cause cisplatin resistance in ovarian carcinoma by inducing autophagy in cancer cells, allowing them to escape apoptosis.[39,40] Findings from this pathway analysis suggest that future studies of the contribution of ER stress in cisplatin sensitivity is needed to improve our understanding of platinum-resistance in ovarian cancer.

In our co-expression network analysis, the gene module "*honeydew1*" showed the most significant correlation with chemotherapy response (p = 6.53e-05). This module includes two probes that map to *VCP*, a gene that was associated with platinum-based chemotherapy response in our DGE analysis. Additional genes in this module were associated with positive regulation of mitochondrial membrane potential, protein ubiquitination, mitosis, alternative splicing, and apoptotic processes. Pathway analysis showed that this module is

283 involved in protein processing in the endoplasmic reticulum. A prior study has found that
284 *VCP* is critical for extraction and degradation of unfolded proteins in endoplasmic reticulum
285 is crucial for cancer cell survival and lower expression of *VCP* is associated with poor
286 response to platinum□based chemotherapy.[41] In alignment with this finding, genes co-
287 expressed in the *honeydew1* module were co-downregulated in chemo-resistant patients.

288 The *honeydew1* module is composed of 76 probes mapping to 54 unique genes, and
289 of these, 44 genes are found to be in chromosome 9, demonstrating the importance of
290 chromosome 9 in the regulation of chemo-resistance in ovarian cancer. These findings
291 support previous studies, where genetic imbalance and alterations in chromosome 9 have
292 been associated with progression of ovarian cancer and increased cisplatin resistance.[42]
293 Analysis of overrepresented transcription factor binding sites demonstrated that genes in this
294 module may be co-regulated by a common transcription factor known as organic cation
295 transporter 1 (OCT1). We found that over 96% of genes in this module (49/54 genes) contain
296 a nucleotide motif bound by OCT1. Prior studies have reported that silencing OCT1 impaired
297 cisplatin-induced apoptosis in esophageal cancer cells, and that cisplatin-resistant cells were
298 already expressing significantly reduced levels of OCT1.[43] Taken together, these findings
299 characterize a network of co-expressed genes that is associated with platinum-resistance in
300 ovarian cancer. Genes within this module may be co-regulated by the OCT1 transcription
301 factor, which may be used as a novel potential target for ovarian cancer therapies

302 The other four co-expression modules associated with platinum-based response also
303 include genes known to be involved in oncogenic process and drug response outcomes. For
304 example, the *orangered4* module is composed of genes that are critical for regulation of
305 immune response, which have been shown to play a role in chemotherapy response in
306 HGSOC.[1415] Genes in this module are associated with functional annotation terms including

immunoglobulin receptor binding, antigen binding, B cell receptor signaling pathway, and phagocytosis (DAVID). In addition, 10 of the 58 genes in this module are enriched for a common transcription factor binding site: acute myeloid leukemia 1 protein (AML1). This transcription factor is involved in haematopoiesis process and immune functions such as thymic T-cell development. Studies have reported that the AML1 transcription factor is overexpressed in ovarian cancer patients, and may contribute to cancer cell proliferation, migration and invasion.[44] In addition, we found that the *lightpink3* module is strongly associated with transcription regulation process, which plays a pivotal role in cancer progression. Finally, genes in modules *lightcyan1* and *skyblue3* are target regions of chemotherapeutic agents such as tyrosine kinase inhibitors (**Supplemental data 6**).

The analysis of *BRCA1* and *BRCA2* SNPs demonstrated that 28 out of 35 variants associated with chemotherapy response were also cis-acting eQTLs, correlated with the expression of *BRCA2* as well as neighboring genes *N4BP2L1*, *N4BP2L2*, *FRY*, and *STARD13* (nominal p-value <0.05). Both *BRCA2* and *STARD13* are well known tumor-suppressors, and upregulation of *N4BP2L1* and *N4BP2L2* is associated with positive prognosis in ovarian cancer cases.[45] This finding shows the potential regulatory effect of the variants in *BRCA2*.

In addition, our results show that 47% of variants identified in *BRCA2* are associated with LDL/HDL cholesterol levels (**Supplemental Table 3**). Prior studies of lung and ovarian cancers consistently reported that cholesterol levels may affect the efficacy of platinum-based chemotherapeutic agents.[46,47] Our findings indicate a new link between genetic variants in *BRCA2* and platinum-based chemotherapy response through cholesterol level regulation.

Conclusion

In this study, we identified genes and gene networks correlated with platinum-based

chemotherapy response in high-grade serous ovarian cancer patients, which implicate known and novel biological mechanisms. Specifically, we identified that reduced expression of *VCP* is associated with platinum-resistance. This gene is critical for removing unfolded proteins from the endoplasmic reticulum and previously correlated with cancer cell survival and platinum-based chemotherapy response. In addition, we report potentially regulatory variants in the *BRCA2* gene correlated with chemotherapy response and the expression of genes that determine cholesterol levels. Moreover, we identified a novel group of potentially co-regulated genes on chromosome 9 that are correlated with platinum resistance using a machine-learning algorithm. In addition, genes from this module including *VCP* are also involved in protein processing in the endoplasmic reticulum. This manuscript supports earlier studies which implicated *VCP* and *BRCA2* genes in chemotherapy response. We also identified regulatory variants in *BRCA2* and additional genes co-regulated with *VCP* on chromosome 9 that also contribute to protein removal in the ER. Findings from our study could facilitate genetic testing through the identification of gene signatures that may predict chemotherapy response as well as lead to novel drug targets, given a better understanding of the biological mechanisms underlying chemotherapy response.

Figure Legends

Figure 1. Gene co-expression modules correlated with platinum-based chemotherapy response. Network plot showing the five significant gene co-expression modules from *WGCNA*: honeydew1 - centre, lightcyan1 - left, lightpink3 - top, orangered4 - bottom, and skyblue3 - right. Nodes represent probes and edges are connections among the probes. Co-expressed probes (i.e. belonging to a single module) are indicated in the same colour.

353 **Figure 2. Differential expression analysis of platinum-based chemotherapy response in**
354 **HGSOC patients.** Volcano plot showing univariate analysis results. One probe, 208648_at,
355 which maps to the Valosin-Containing Protein (*VCP*) gene is significantly correlated with
356 chemotherapy outcome as indicated by the red line (FDR-corrected p value < 0.05). A total of
357 606 probes mapping to 521 unique genes with nominal correlate as indicated by the green
358 line (p = 0.05).

359 **Figure 3. GWAS of chemotherapy response.** Manhattan plot shows no SNPs are
360 significantly associated (i.e. surpasses the Red line indicating a p-value threshold of 5x10^-8).

361

362 **Acknowledgements**
363 Computations in this manuscript were performed on resources and with support provided by
364 the Centre for Advanced Computing (CAC) at Queen's University in Kingston, Ontario. The
365 CAC is funded by: the Canada Foundation for Innovation, the Government of Ontario, and
366 Queen's University. JC is funded by Queen's University, Faculty of Health Sciences Dean's
367 Doctoral Award. QLD receives funding from the Canadian Institutes of Health Research and
368 Queen’s University.

369 **Conflict of interest**

370 The authors declare no conflicts of interest.

371 **Author contribution**

372 J.C. performed the data analyses and drafted the manuscript. D.G.T., S.N. and A.T. assisted in
373 the data analyses. Q.L.D. designed the research project, supervised data analyses and assisted
374 in the writing of the manuscript. M.K. assisted in the study design and editing of the

manuscript.

References

1. Reid, F. World Ovarian Cancer Coalition 2018. *World Ovarian Cancer Coalit.* (2018).
2. International Agency for Research on Cancer Website. Globocan 2012 - Home. *Webpage* (2012). doi:NO:11
3. Brett M., R. *et al.* Epidemiology of ovarian cancer: a review. *Cancer Biol. Med.* (2017). doi:10.20892/j.issn.2095-3941.2016.0084
4. Cannistra, S. A. Cancer of the Ovary. *N. Engl. J. Med.* **351**, 2519–2529 (2004).
5. Miller, D. S. *et al.* Phase II Evaluation of Pemetrexed in the Treatment of Recurrent or Persistent Platinum-Resistant Ovarian or Primary Peritoneal Carcinoma: A Study of the Gynecologic Oncology Group. *J. Clin. Oncol.* **27**, 2686–2691 (2009).
6. Armstrong, D. K. Relapsed ovarian cancer: challenges and management strategies for a chronic disease. *Oncologist* **7 Suppl 5**, 20–28 (2002).
7. Foley, O. W., Rauh-Hain, J. A. & del Carmen, M. G. Recurrent epithelial ovarian cancer: an update on treatment. *Oncology (Williston Park).* **27**, 288–94, 298 (2013).
8. Ozols, R. F. Recurrent ovarian cancer: Evidence-based treatment. *Journal of Clinical Oncology* **20**, 1161–1163 (2002).
9. Walsh, T. *et al.* Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proc. Natl. Acad. Sci.* (2011). doi:10.1073/pnas.1115052108
10. Booth, M., Beral, V. & Smith, P. Risk factors for ovarian cancer: A case-control study. *Br. J. Cancer* (1989). doi:10.1038/bjc.1989.320
11. Vencken, P. M. L. H. *et al.* Chemosensitivity and outcome of BRCA1- and BRCA2-associated ovarian cancer patients after first-line chemotherapy compared with sporadic ovarian cancer patients. *Ann. Oncol.* (2011). doi:10.1093/annonc/mdq628
12. Patch, A. M. *et al.* Whole-genome characterization of chemoresistant ovarian cancer. *Nature* (2015). doi:10.1038/nature14410
13. Hao, D. *et al.* Immunogenomic analyses of advanced serous ovarian cancer reveal immune score is a strong prognostic factor and an indicator of chemosensitivity. *Clin. Cancer Res.* (2018). doi:10.1158/1078-0432.CCR-17-3862

14. Koti, M. *et al.* A distinct pre-existing inflammatory tumour microenvironment is associated with chemotherapy resistance in high-grade serous epithelial ovarian cancer. *Br. J. Cancer* (2015). doi:10.1038/bjc.2015.81

15. Au, K. K. *et al.* STAT1-associated intratumoural TH1 immunity predicts chemotherapy resistance in high-grade serous ovarian cancer. *J. Pathol. Clin. Res.* (2016). doi:10.1002/cjp2.55

16. Koti, M. *et al.* Identification of the IGF1/PI3K/NF κB/ERK gene signalling networks associated with chemotherapy resistance and treatment response in high-grade serous epithelial ovarian cancer. *BMC Cancer* **13**, 549 (2013).

17. Cancer, T. & Atlas, G. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–15 (2011).

18. Li, J., Wood, W. H., Becker, K. G., Weeraratna, A. T. & Morin, P. J. Gene expression response to cisplatin treatment in drug-sensitive and drug-resistant ovarian cancer cells. *Oncogene* **26**, 2860–2872 (2007).

19. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

20. Colaprico, A. *et al.* TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkv1507

21. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).

22. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. Affy - Analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).

23. Smyth, G. limma: Linear Models for Microarray Data. in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* 397–420 (2005). doi:citeulike-article-id:5722720

24. Gentleman, R. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).

25. Shannon, P. *et al.* Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

26. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

27. Warde-Farley, D. *et al.* The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **38**, (2010).

28. Cotto, K. C. *et al.* DGIdb 3.0: a redesign and expansion of the drug–gene interaction database. *Nucleic Acids Res.* (2017). doi:10.1093/nar/gkx1143

29. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* (2016). doi:10.1038/ng.3656

30. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

31. Loh, P. R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* (2016). doi:10.1038/ng.3679

32. Delaneau, O. *et al.* Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nat. Commun.* (2014). doi:10.1038/ncomms4934

33. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

34. Shabalin, A. A. Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**, 1353–1358 (2012).

35. Peng, J. *et al.* VCP gene variation predicts outcome of advanced non-small-cell lung cancer platinum-based chemotherapy. *Tumor Biol.* **34**, 953–961 (2013).

36. Lan, B., Chai, S., Wang, P. & Wang, K. VCP/p97/Cdc48, A Linking of Protein Homeostasis and Cancer Therapy. *Curr. Mol. Med.* (2018). doi:10.2174/1566524018666180308111238

37. Wang, X. J., Feng, C. W. & Li, M. ADAM17 mediates hypoxia-induced drug resistance in hepatocellular carcinoma cells through activation of EGFR/PI3K/Akt pathway. *Mol. Cell. Biochem.* (2013). doi:10.1007/s11010-013-1657-z

38. Tung, S. L. *et al.* miRNA-34c-5p inhibits amphiregulin-induced ovarian cancer stemness and drug resistance via downregulation of the AREG-EGFR-ERK pathway. *Oncogenesis* (2017). doi:10.1038/oncsis.2017.25

39. Tian, J., Liu, R. & Qu, Q. Role of endoplasmic reticulum stress on cisplatin resistance in ovarian carcinoma. *Oncol. Lett.* (2017). doi:10.3892/ol.2017.5580

40. Xu, Y. *et al.* Tolerance to endoplasmic reticulum stress mediates cisplatin resistance in human ovarian cancer cells by maintaining endoplasmic reticulum and mitochondrial homeostasis. *Oncol. Rep.* (2015). doi:10.3892/or.2015.4283

41. Bastola, P., Neums, L., Schoenen, F. J. & Chien, J. VCP inhibitors induce endoplasmic reticulum stress, cause cell cycle arrest, trigger caspase-mediated cell death and synergistically kill ovarian cancer cells in combination with Salubrinal. *Mol. Oncol.* (2016). doi:10.1016/j.molonc.2016.09.005

42. Devlin, J., Elder, P. A., Gabra, H., Steel, C. M. & Knowles, M. A. High frequency of chromosome 9 deletion in ovarian cancer: Evidence for three tumour-suppressor loci. *Br. J. Cancer* (1996). doi:10.1038/bjc.1996.75

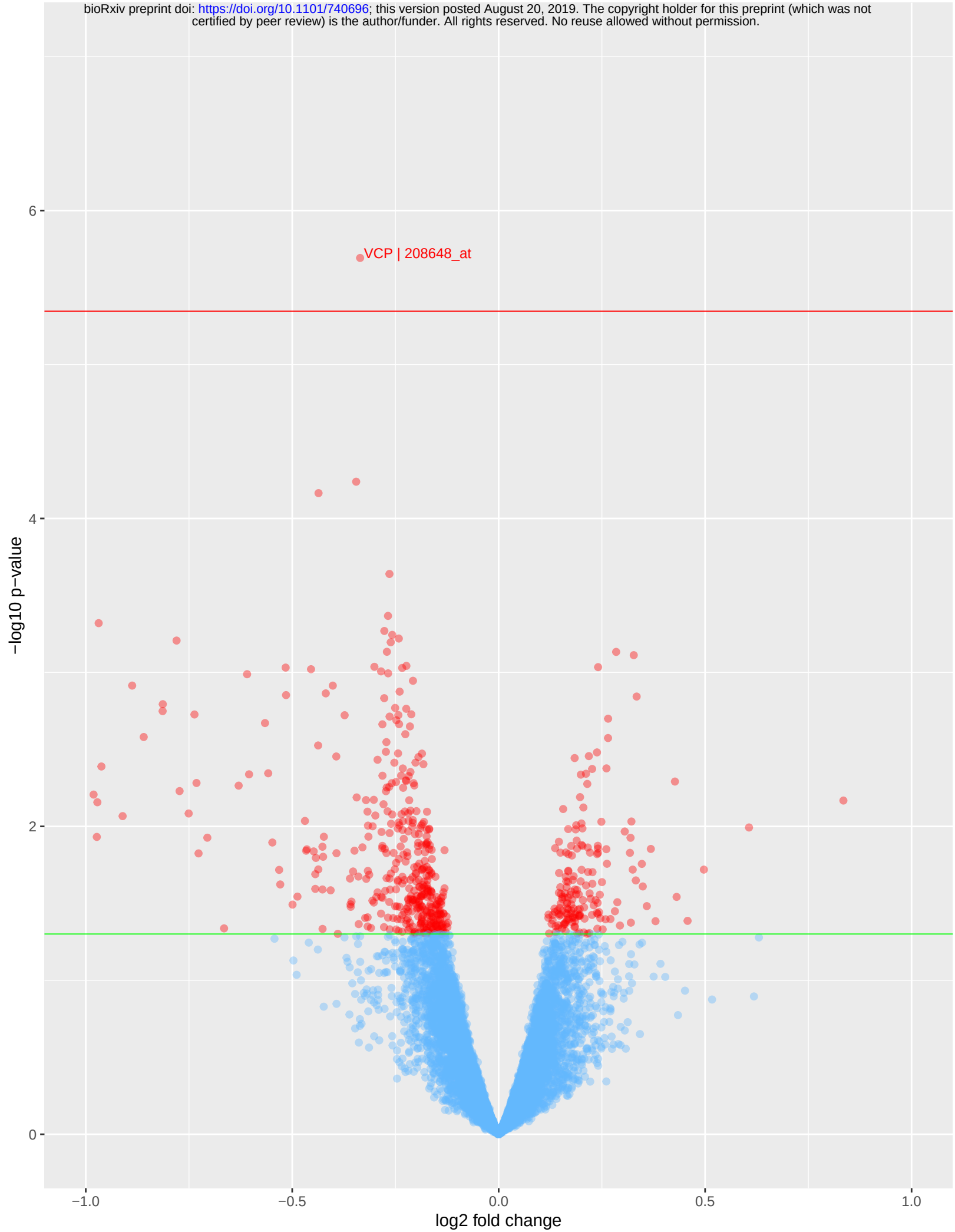
43. Lin, R. *et al.* Long-term cisplatin exposure promotes methylation of the OCT1 gene in human esophageal cancer cells. *Dig. Dis. Sci.* (2013). doi:10.1007/s10620-012-2424-9

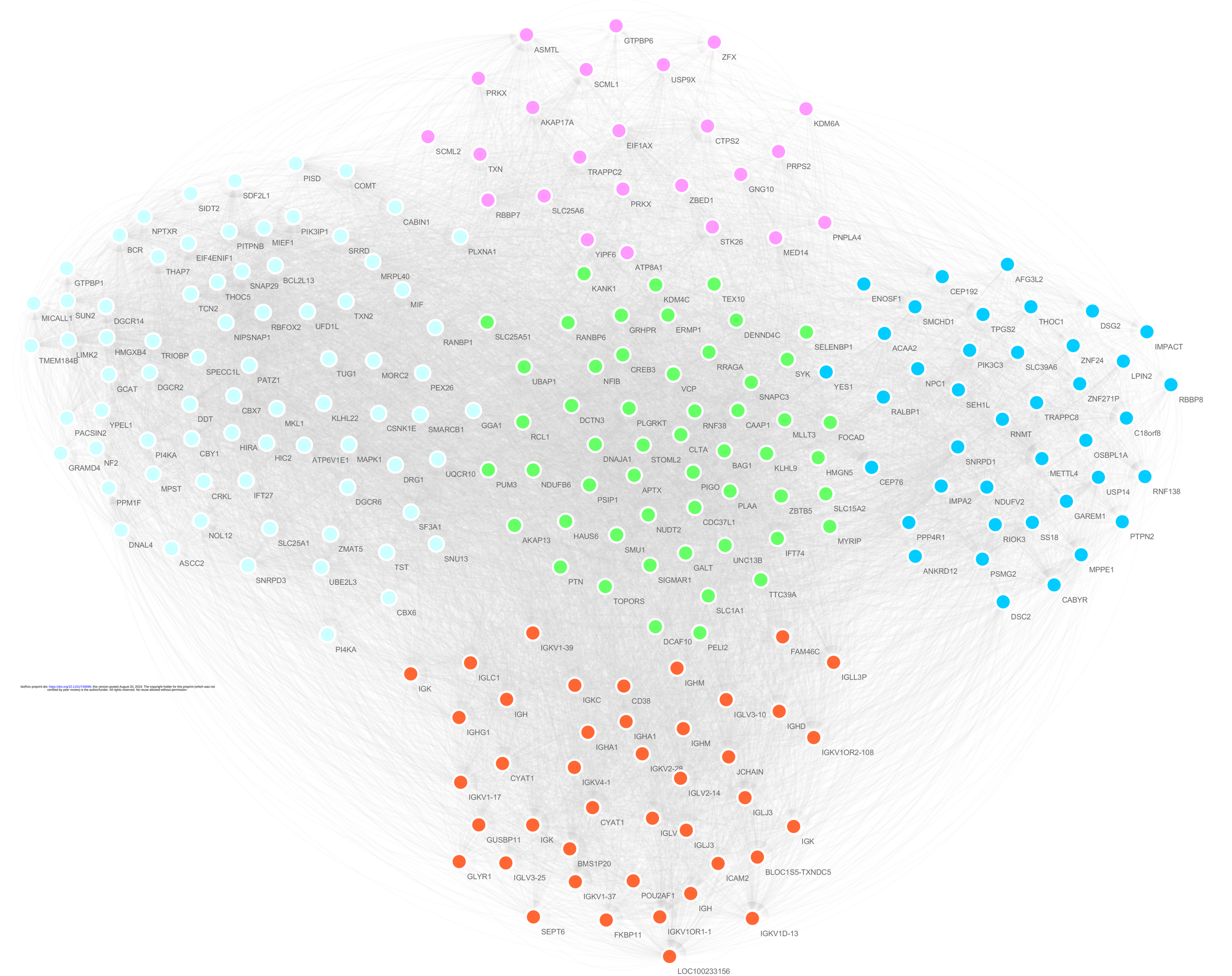
44. Keita, M. *et al.* The RUNX1 transcription factor is expressed in serous epithelial ovarian carcinoma and contributes to cell proliferation, migration and invasion. *Cell Cycle* (2013). doi:10.4161/cc.23963

45. Koussounadis, A., Langdon, S. P., Harrison, D. J. & Smith, V. A. Chemotherapy-induced dynamic gene expression changes in vivo are prognostic in ovarian cancer. *Br. J. Cancer* **110**, 2975–2984 (2014).

46. Wu, Y. *et al.* Cholesterol reduces the sensitivity to platinum-based chemotherapy via upregulating ABCG2 in lung adenocarcinoma. *Biochem. Biophys. Res. Commun.* (2015). doi:10.1016/j.bbrc.2015.01.035

47. Kim, S., Lee, M., Dhanasekaran, D. N. & Song, Y. S. Activation of LXR□/$β$ by cholesterol in malignant ascites promotes chemoresistance in ovarian cancer. *BMC Cancer* **18**, 1232 (2018).





Manhattan Plot