Genomic and Epigenomic study of Stage I Epithelial Ovarian Carcinomas to Implement Novel Prognostic Molecular Biomarkers

Thesis

Istituto di Ricerche Farmacologiche Mario Negri IRCCS

The Open University

Doctor of Philosophy Program in

School of Life, Health and Chemical Science

Genomic and Epigenomic study of Stage I Epithelial Ovarian Carcinomas to implement novel prognostic molecular biomarkers

PhD Dissertation of

Chiara Pesenti

Personal Identifier:

H8412174

Director of Studies Dr. Maurizio D'Incalci

Supervisor Professor James Brenton

DECEMBER 2021

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University or Institute. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements.

Dr Luca Beltrame was responsible for the bioinformatic analyses; Dr Angelo Velle and Prof. Chiara Romualdi were responsible for the statistical analyses, Dr Nicolò Panini was responsible for cytofluorimetry experiments and Dr. Milena Cribiù, Dr. Fulvio Borrella and Dr. Marta Jaconi were responsible for HE-stained tumour slides evaluation.

This work was performed at the Istituto di Ricerche Farmacologiche Mario Negri IRCCS in Milan, Italy, from 2019 to 2021, under the supervision of Dr Maurizio D'Incalci (director of studies), Professor James Brenton (external supervisor) and Dr Roberta Frapolli (third party monitor).

ABSTRACT

Stage I epithelial ovarian carcinomas represent the earliest stage of ovarian carcinoma, when the disease is confined to the ovaries, and is amenable to efficient treatment by cytoreductive surgery followed by adjuvant chemotherapy in cases considered at high-risk of relapse. It is generally characterized by a favourable outcome; with just about 20% of relapse. However, the staging of the tumour is often suboptimal causing on one hand over-treatment of many who receive adjuvant chemotherapy to prevent recurrence, and on the other hand under-treatment of those improperly defined as "low risk", who experience relapses with a much poorer prognosis. The histologically heterogeneity of stage I epithelial ovarian tumours further complicates the development of efficient prognostic markers. Nevertheless, previous results on the transcriptomic signatures of stage I tumours obtained in the hosting Lab, prompted us to extend the current knowledge about the molecular landscape of stage I tumours to potentially identify novel parameters able to stratify patients based on the prognosis.

In my PhD project, I exploited high-throughput Next Generation Sequencing approaches to evaluate the most recurrent Single Nucleotide Variants in a subset of frequently altered genes in ovarian cancer and the Somatic Copy Number Alterations distribution across the genome on a unique cohort of 205 stage I epithelial ovarian cancer patients. These analyses revealed the existence of three different genomic instability patterns namely stable, unstable and highly unstable, based on the on the percentage of genome affected by copy number alterations and their length. These patterns are strictly related to distinct etiopathogenetic processes that drive tumour evolutionary routes. In an effort to define potential mechanisms involved in the generation of genomic instability, five copy number signatures related to different mutational pathways were defined and global DNA methylation was assessed through LINE1 promoter methylation status. The significant association of the genome instability with a reduction in global methylation levels by LINE-1 retrotransposons supports a close relationship between the genomic landscape and tumour epigenetic regulation.

Finally, the three SCNA patterns were correlated to patients' survival and resulted strongly predictive of patients' prognosis also in multivariate models with the currently used clinical variables.

These results show that genomic and epigenomic analyses offer novel possibilities to identify markers useful for the management of the disease, offering an improved patient prognosis prediction.

Acknowledgments

First, I would like to thank Maurizio for having entrusted me with this PhD project and for mentoring me, continuously offering me new ideas for discussion and for the esteem he has shown me, which is absolutely reciprocal, working with you and your team has been an incredible opportunity.

I would also like to thank Professor James Brenton because our meetings have brought a new critical point of view on the work, always allowing me to ask new questions and test new ideas. It was an honour to have him as a supervisor.

A huge thank you goes to Sergio Marchini's group in which I had the pleasure to work in, in particular to Luca, who not only translated my ideas into bioinformatic language, but also became a very good friend that it is always a pleasure to chat with. I can not help but thank Laura, Ilaria, Silvana, Sara and Nicolò, current and old lab mates with whom I spent wonderful moments both working in the lab and cheering at the pub after work (this second part fortunately does not end yet).

Thanks to Deborah, Tommaso and Lara, the three thesis students who helped me in this project, I hope I made you love this work because you are really good.

There are so many collaborators that I have met during this journey and that I would like to thank because they have always taught me something and that has always been special to me; special mentions go to Chiara and Angelo, whom I often bothered for the statistical analyses, Milena who put up with me in reading the hundreds of slides and the group of gynaecologists in Monza, who helped me to recover all the cases analysed.

Finally, a huge thank you to all my family, my parents, my grandmother, my two little brothers and Roberto, they have always been there and that is not to be taken for granted. I hope that this further piece of my study and work career will make them even more proud of me.

Publications by the PhD Candidate

Conforti F, Pala L, Pagan E, Rocco EG, Bagnardi V, Montagna E, Peruzzotti G, De Pas T, Fumagalli C, Pileggi S, **Pesenti C**, Marchini S, Corso G, Marchio' C, Sapino A, Graffeo R, Collet L, Aftimos P, Sotiriou C, Piccart M, Gelber RD, Viale G, Colleoni M, Goldhirsch A. Biological and clinical features of triple negative Invasive Lobular Carcinomas of the breast. Clinical outcome and actionable molecular alterations. Breast. 2021 Oct;59:94-101.

Paracchini L, **Pesenti C**, Delle Marchette M; Beltrame L; Bianchi T; Grassi T; Buda A; Fabio Landoni F; Ceppi L; Bosetti C; Paderno M; Adorni M; Vicini D; Perego P; Leone BE; D'Incalci M, Marchini S; Fruscio R. Detection of TP53 Clonal Variants in Papanicolaou Test Samples Collected up to 6 Years Prior to High-Grade Serous Epithelial Ovarian Cancer Diagnosis. JAMA Network Open. 2020;3(7):e207566.

Tabano S., Azzollini J., **Pesenti C**., Lovati S., Costanza J., Fontana L., Peissel B., Miozzo M., Manoukian S. Analysis of BRCA1 and RAD51C Promoter Methylation in Italian Families at High-risk of Breast and Ovarian Cancer. Cancers (Basel). 2020 Apr 8;12(4):910.

Lopez G, Noale M, Corti C, Gaudioso G, Sajjadi E, Venetis K, Gambini D, Runza L, Costanza J, **Pesenti C**, Grossi F, Maggi S, Ferrero S, Bosari S, Fusco N. PTEN Expression as a Complementary Biomarker for Mismatch Repair Testing in Breast Cancer. Int J Mol Sci. 2020 Feb 21;21(4). pii: E1461.

Pesenti C, Navone SE, Guarnaccia L, Terrasi A, Costanza J, Silipigni R, Guarneri S, Fusco N, Fontana L, Locatelli M, Rampini P, Campanella R, Tabano S, Miozzo M, Marfia G. The Genetic Landscape of Human Glioblastoma and Matched Primary Cancer Stem Cells Reveals Intratumour Similarity and Intertumour Heterogeneity. Stem Cells Int. 2019 Mar 7;2019:2617030.

Azzollini J, **Pesenti C**, Pizzamiglio S, Fontana L, Guarino C, Peissel B, Plebani M, Tabano S, Sirchia SM, Colapietro P, Villa R, Paolini B, Verderio P, Miozzo M, Manoukian S.

Constitutive BRCA1 Promoter Hypermethylation Can Be a Predisposing Event in Isolated Early-Onset Breast Cancer. Cancers (Basel). 2019 Jan 9;11(1).

Colombo C, Muzza M, Proverbio MC, Tosi D, Soranna D, **Pesenti C**, Rossi S, Cirello V, De Leo S, Fusco N, Miozzo M, Bulfamante G, Vicentini L, Ferrero S, Zambon A, Tabano S, Fugazzola L. Impact of Mutation Density and Heterogeneity on Papillary Thyroid Cancer Clinical Features and Remission Probability. Thyroid. 2019 Feb;29(2):237-251.

Fusco N, Lopez G, Corti C, **Pesenti C**, Colapietro P, Ercoli G, Gaudioso G, Faversani A, Gambini D, Michelotti A, Despini L, Blundo C, Vaira V, Miozzo M, Ferrero S, Bosari S. Mismatch Repair Protein Loss as a Prognostic and Predictive Biomarker in Breast Cancers Regardless of Microsatellite Instability. JNCI Cancer Spectr. 2018 Dec 13;2(4):pky056.

Paganini L., **Pesenti C.**, Milani D., Fontana L., Motta S., Sirchia S.M., Scuvera G., Marchisio P., Esposito S., Cinnante C.M., Tabano S., Miozzo M. A novel splice site variant in ITPR1 gene underlying recessive Gillespie Syndrome. American Journal of Medical Genetics (Part A). 2018 Apr.

Bonaparte E., **Pesenti C.**, Fontana L., Falcone R., Paganini L., Marzorati A., Ferrero S., Nosotti M., Mendogni P., Sirchia S., Tabano S., Bosari S., Miozzo M., Molecular profiling of lung cancer specimens and liquid biopsies using MALDI-TOF mass spectrometry. Diagnostic Pathology. 2018 Jan 12;13(1):4.

Del Gobbo A.; Morotti A, Colombo A.E., Vaira V., Ercoli G., **Pesenti C.**, Bonaparte E., Guerini-Rocco E., Di Cristofori A., Locatelli M., Palleschi A., Ferrero S. IMP3 expression in nsclc brain metastases reveals its prognostic role in non-neuroendocrine phenotypes. Medical Oncology. 2017 Dec 1;35(1):2.

Pesenti C., Muzza M., Colombo C., Proverbio M.C., Ferrero S., Miozzo M., Fugazzola L., Tabano S. MassARRAY-based simultaneous detection of hotspot somatic mutations and recurrent fusion genes in papillary thyroid carcinoma: the PTC-MA assay. Endocrine. 2017 Dec.

Pesenti C., Paganini L., Fontana L., Veniani E., Runza L., Ferrero S., Bosari S., Menghi M., Marfia G., Caroli M., Silipigni R., Guerneri S., Tabano S., Miozzo M. Mass spectrometry-based assay for the molecular diagnosis of glioma: concomitant detection of chromosome 1p/19q codeletion, and IDH1, IDH2, and TERT mutation status. Oncotarget. 2017 Jul 8;8(34):57134-57148.

Azzollini J., **Pesenti C**., Ferrari L., Fontana L., Calvello M., Peissel B., Portera G., Tabano S., Carcangiu ML., Riva P., Miozzo M., Manoukian S., Revertant Mosaicism for Family Mutations Is not Observed in BRCA1/2 Phenocopies.PloS One. 2017;12(2).

Rawal CC., Riccardo S., **Pesenti C**., Ferrari M., Marini F., Pellicioli A. Reduced kinase activity of polo kinase Cdc5 affects chromosome stability and DNA damage response in S. cerevisiae. Cell Cycle. 2016 Nov;15(21):2906-2919.

Marfia G., Navone SE., Fanizzi C., Tabano S., **Pesenti C**., Abdel Hadi L., Franzini A., Caroli M., Miozzo M., Riboni L., Rampini P., Campanella R. Prognostic value of preoperative von Willebrand factor plasma levels in patients with Glioblastoma. Cancer Med. 2016 Aug;5(8):1783-90.

Fontana L., Tabano S., Bonaparte E., Marfia G., **Pesenti C**., Falcone R., Augello C., Carlessi N., Silipigni R., Guerneri S., Campanella R., Caroli M., Maria Sirchia S., Bosari S., Miozzo M. MGMT-Methylated Alleles Are Distributed Heterogeneously Within Glioma Samples Irrespective of IDH Status and Chromosome 10q Deletion. J Neuropathol Exp Neurol. 2016 Jun 26

Pesenti C., Gusella M., Sirchia S.,Miozzo M. Germline Oncopharmacogenetics, a promising field in cancer therapy. Cell Oncol (Dordr). 2015 Feb;38(1):65-89.

List of Abbreviations

EOC	Epithelial Ovarian Carcinoma
FIGO	International Federation of Obstetrics and Gynaecology
HGSOC	High Grade Serous Ovarian Carcinoma
WHO	World Health Organization
EC	Endometrioid Carcinoma
OCCC	Clear Cell Carcinoma
MOC	Mucinous Carcinoma
LGSOC	Low-Grade Serous Carcinoma
PARPi	Poly (ADP-Ribose) Polymerase inhibitors
HR	Homologous Recombination
HRD	Homologous Recombination Deficient
NGS	Next Generation Sequencing
MMR	Mismatch Repair
MMRd	Mismatch Repair deficient
MRI	Magnetic Resonance Imaging
CT	Chemotherapy
OS	Overall Survival
PFS	Progression Free Survival
SET	Solid, pseudo-Endometrioid and Transitional

TIL	Tumour-Infiltrating Lymphocyte
ER	Estrogen
PR	Progesterone
STIC	Serous Tubal Intraepithelial Carcinoma
SEE-FIM	Sectioning and Extensively Examining the Fimbriated End
STIL	Serous Tubal Intraepithelial Lesion
SCOUT	Secretory Cell OUTgrowth
NCCN	National Comprehensive Cancer Network
WT1	Wilms Tumor 1
SNV	Single Nucleotide Variant
SCNA	Somatic Copy Number Alterations
NSMP	No Specific Molecular Profile
LINE1	Long Interspersed Nuclear Element 1
TCGA	The Cancer Genome Atlas
ISC	Integrated Signature Classifier
sWGS	shallow Whole Genome Sequencing
FFPE	Formalin Fixed Paraffin Embedded
HE	Haematoxylin and Eosin
BER	Base Excision Repair
NER	Nucleotide Excision Repair

NHEJ	Non Homologous End Joining
CNB	Copy Number Burden
MAF	Median Allelic Fraction
LoF	Loss of Function
IQR	InterQuantile Range
HU	Highly Unstable
U	Unstable
S	Stable
DBS	Double Strand Break

Table of Contents

INTRODUCTION

Stage I Epithelial Ovarian Carcinoma (EOC)

Stage I Epithelial Ovarian Carcinomas (EOCs) represent the earliest stage of EOCs (following the International Federation of Obstetrics and Gynaecology staging, FIGO), when the disease is confined to the ovaries and still usually amenable to efficient treatment by cytoreductive surgery, as described below. Although most stage I EOC patients experience a favourable outcome, almost 20% of them experience relapse developing platinum-resistant disease. The randomized trial "ICON1/ACTION"[1] proffered definitive criteria to be applied to current clinical practice. These criteria suggest the use of adjuvant platinum-based chemotherapy in patients considered at high risk of relapse but the withholding of such treatment in patients with low-grade differentiated tumour limited to the ovaries according to optimal surgical staging. As explained in the "Therapeutic Treatments" chapter, a comprehensive surgical procedure allows to stage EOC excluding, in case of Stage I, the spread of the disease outside the ovaries. However, the current clinical classification lacks sensitivity and specificity, failing to optimally predict the risk of relapse. As further detailed in the next chapters, this often prevents the correct stratification of the patients according to the relapse risk. Therefore, common practise is to treat most of stage I EOC patients with adjuvant chemotherapy. There are thus two undesirable consequences for patients: on one hand over-treatment of many who receive adjuvant chemotherapy to prevent recurrence, and on the other hand under-treatment of those improperly defined at low relapse risk.

This is further complicated by the fact that malignant EOCs were historically considered as a single disease with treatments approaches essentially based on the behaviour of the most frequent subtype, High Grade Serous Ovarian Carcinoma (HGSOC). However, to date, EOC is regarded as a group of different tumours with distinct etiopathogenetic mechanisms, precursor lesions, molecular genetics, risk factors, clinical course, response to chemotherapy, and prognosis. Following the World Health Organization (WHO) classification of gynaecological cancers, EOC is divided into five major histological subtypes: HGSOC, endometrioid carcinoma (EC), clear cell carcinoma (OCCC), low-grade

serous carcinoma (LGSOC), and mucinous carcinoma (MOC)[2]. Recent studies have allowed a deeper comprehension of the biology and the molecular alterations of each histotype, allowing to uncover novel opportunities of a more personalized therapeutic approach and treatments with targeted drugs (e.g., poly (ADP-ribose) polymerase inhibitors, PARPi)[3,4].

Therefore, the development of molecular classifiers more tailored to predict relapse risk and to open novel potential treatment windows is essential to improve the management of stage I EOC patients. In this dissertation, I will summarize current knowledge about stage I EOC and present novel data, produced during my PhD project to uncover the genomic landscape of stage I EOC and define novel potential prognostic markers.

Epidemiology

Ovarian tumours comprise a heterogeneous group of diseases that affect women from the age of 20; approximately 80% of these tumours are benign, the remaining are malignant and 90% of them are diagnosed in women older than 40 years. EOC represents almost 90% of the malignant ovarian tumours and, unfortunately, the median 5-year survival rate for patients diagnosed with EOC is 46% (https://www.cancer.net. Accessed in September 2021). EOC represents the third lethal tumour among feminine cancers and affects women both in reproductive and post-menopausal age. In Italy, it has been estimated that in 2020 there would be about 5200 new diagnoses, with about 3000 deaths[5]. The high mortality rate is mainly related to the fact that the pathology is asymptomatic until the spread outside of the ovaries, when its metastasization to the peritoneum makes a complete removal challenging. Moreover, no efficient screening strategies have been developed yet, and, accordingly, about 75-80% of the EOCs are diagnosed at advanced stages (FIGO III or IV), with only about 10% of patients being diagnosed with stage I disease. FIGO staging system is a conventional system to describe cancers enabling clinicians to internationally cooperate to define better treatments options (Table 1).

Table 1 reports the current criteria for FIGO staging of EOCs (adapted from[6]).

STAGE I		Tumour confined to ovaries or fallopian tube(s)
○ IA		Tumour limited to one ovary (capsule intact) or fallopian tube. No tumour on ovarian or fallopian tube surface; no malignant cells in the ascites or peritoneal washings.
○ IB		Tumour limited to both ovaries (capsules intact) or fallopian tubes. No tumour on ovarian or fallopian tube surface; no malignant cells in the ascites or peritoneal washings.
○ IC		Tumour limited to one or both ovaries or fallopian tubes
	• IC1	Surgical spill
	• IC2	Capsule rupture before surgery or tumour on ovarian or fallopian tube surface.
	• IC3	Malignant cells in the ascites or peritoneal washings.
STAGE II		Tumour involves one or both ovaries or fallopian tubes with pelvic extension (below the pelvic brim) or primary peritoneal cancer
○ IIA		Extension and/or implant on uterus and/or fallopian tubes and/or ovaries
○ IIB		Extension to other pelvic intraperitoneal tissues
STAGE III		Tumour involves one or both ovaries or fallopian tubes, or primary peritoneal cancer, with cytologically or histologically confirmed spread to the peritoneum outside the pelvis and/or metastasis to the retroperitoneal lymph nodes
○ IIIA		Positive retroperitoneal lymph nodes and/or microscopic metastasis beyond the pelvis
	• IIIA1	Positive retroperitoneal lymph nodes only (cytologically or histologically proven) i) Metastasis ≤ 10 mm in greatest diameter ii) Metastasis > 10 mm in greatest diameter
	• IIIA2	Microscopic, extrapelvic (above the pelvic brim) peritoneal involvement ± positive retroperitoneal lymph nodes
○ IIIB		Macroscopic, peritoneal metastasis beyond the pelvis up to 2 cm in greatest dimension ± metastasis to the retroperitoneal lymph nodes.
○ IIIC		Macroscopic, peritoneal metastasis beyond the pelvis more than 2 cm in greatest dimension ± metastasis to the retroperitoneal lymph nodes. Includes extension of tumour to capsule of liver/spleen without parenchymal involvement of either organ.
STAGE IV		Distant metastasis excluding peritoneal metastasis
○ IVA		Pleural effusion with positive cytology
○ IVB		Parenchymal metastasis, metastasis to extra- abdominal organs (including inguinal lymph nodes and lymph nodes outside of the abdominal cavity)

For EOCs, the tumour staging represents the most important prognostic factor. Advanced tumours unfortunately have a 5-years survival of approximately 30%, in contrast with the 90% of stage I disease[7]. Nevertheless, as mentioned, the asymptomatic nature of EOCs causes the late presentation of the disease, making the diagnosis of Stage I EOCs extremely rare.

Staging and prognostic factors

The optimal staging is performed during surgery that includes total abdominal hysterectomy, bilateral salpingo-oophorectomy, omentectomy, inspection of peritoneal surfaces with biopsy or removal of any suspicious areas, and para-aortic and pelvic lymph node dissection[6]. Surgery should remove all the disease without leaving any residual. Once optimal staging is completed, histopathological diagnosis on tumour tissue must be performed to define the histological subtypes and thus better stratify the patients' prognosis and potential different treatment approaches.

Current clinic-pathological prognostic factors for stage I EOC include:

- Histological grade, this is considered the most relevant prognostic element
- FIGO substage (detailed above)
- Age at diagnosis
- Histotype (detailed below)
- Capsule rupture
- Presence of ascites

Following these criteria Stage I EOC are divided in three classes with distinct 5-years survival rates[7]:

- Patients at low risk of relapse (5-years survival rate >90%) → Stage IA and IB grade 1.
- Patients at intermediate risk of relapse (5-years survival rate 70-80%) → Stage IA and IB grade 2 and stage IC grade 1.

- Patients at high risk of relapse (5-years survival rate 50-60%) → Stage IA and IB grade 3 and stage IC grade 2 and 3, all OCCCs.

In the past years other minor prognostic features were suggested, such as DNA ploidy and tumour size, but they are not currently used due to insufficient evidence about their efficacy.

Genomic predisposition and risk factors

Genomic predisposition to ovarian cancer is now recognized in about 15% of tumours. The majority of hereditary EOCs is determined by deleterious mutations in *BRCA1* and *BRCA2* genes, responsible for more than 50% of cases belonging to the hereditary Breast and Ovarian cancer syndrome. These genes encode for two proteins involved in a DNA repair system called Homologous Recombination (HR). Beside the hereditary high predisposition to gynaecological cancers due to germline mutations in these genes, the presence of such mutations in the tumour cells makes them defective in HR (HRD) and thus responsive to targeted treatment with PARP inhibitors. Along with germline mutations in *BRCA1* and *BRCA2*, other less penetrant predisposing genes involved in HR are responsible for about 5% of cases, such as *RAD51C*, *CHEK2*, *PALB2*[8–10]. These genes are now included in multigene testing panels for Next Generation Sequencing (NGS) platforms to identify those women affected by this syndrome and, thus, to involve the families in proper surveillance programs and to set the optimal treatment option.

Other hereditary syndromes that increase the lifetime risk to develop EOC include the Lynch syndrome, that accounts for 15% of hereditary EOC and is caused by mutations in genes (*MLH1*, *PMS2*, *MSH6*, *MSH2*) involved in the DNA repair system named Mismatch Repair (MMR), Li-Fraumeni syndrome, cause by mutation in *TP53* and other rare conditions such as Peutz-Jegher and Gorlin syndromes[11–13].

Beside the presence of germline mutations in these genes, nowadays the massive molecular characterization of tumours has uncovered the presence of somatic mutations in the same genes that could be targetable by specific drugs. Particularly, PARPi in case

of mutations in the HR genes especially in the HGSOC subtype or immunotherapy in case of MMR deficiency (MMRd) especially in ECs, OCCCs and MOCs.

Other endocrine and environmental risk factors for EOCs include the number of lifetime ovulations (absence of pregnancy, early age of menarche, and late age at menopause), benign gynaecological conditions (such as endometriosis and pelvic inflammatory disease), use of oral contraceptives, smoking and alcohol consumption[14].

Screening and diagnosis

Besides the huge efforts to define precise biomarkers to early diagnose EOCs in the general population, no approved screening strategies are currently available.

Several randomised trials, such as the UKCTOCS trial (NCT00058032) or the PLCO Cancer Screening Randomized Controlled Trial, tested the CA125 antigen measurement alone or in combination with other tests such as transvaginal ultrasound, but failed to identify a reduction in mortality rate exploiting these screening procedures[15,16]. Other biomarkers have been tested with CA125, such as human epididymis protein 4, leptin, prolactin, osteopontin, insulin-like growth factor II (IGF-II) and macrophage inhibitory factor (MIF), but further studies are still necessary to define their performances[17]. The only recommendations for women belonging to high-risk families is to perform risk-reducing salpingo- oophorectomy. EOC symptoms are not specific signs and include abdominal and back pain, satiety, nausea, weight loss and fatigue.

The diagnosis of EOC usually starts with the measurement of CA125 followed by pelvic ultrasound. Once determined the presence of a potential EOC, further imaging investigations include chest and abdomen or pelvis Computed Tomography for staging, and potentially a pelvic Magnetic Resonance Imaging (MRI). During the surgery the optimal staging is performed to remove any suspicious neoplastic areas[6]. Finally, to complete the diagnosis, histopathological evaluation of tumour tissue must be performed to define the histological subtype.

Therapeutic Treatments

The therapeutic and surgical treatments for EOCs have constantly improved over the past decades, however, recurrent disease is still almost fatal. The treatment guidelines have largely been based on HGSOCs, since this is the most frequent and aggressive subtype in advanced tumours. This is not true for example in stage I EOC where the other subtypes are more frequent. However, randomized clinical trials on early-stage disease are challenging to perform due to the rarity of stage I tumours.

Primary treatment

Primary debulking surgery followed by chemotherapy represents the standard of care for advanced EOC. Otherwise, for a subset of cases, such as older women, patients with large disease burden or cases with several comorbidities, neoadjuvant chemotherapy followed by interval debulking surgery could be an alternative approach. The recommendations are based on the results of the EORTC 55971 randomised trial[18], patients with stage IIIC disease and metastases smaller than 5 cm have better Overall Survival (OS) performing primary debulking surgery, while patients with stage IV disease have better survival with neoadjuvant chemotherapy.

The treatment plan for early stages is determined by the comprehensive surgical staging, that, as recommended by FIGO, should include hysterectomy, bilateral salpingo-oophorectomy, cytology of ascites or peritoneal washings, peritoneal biopsies, infracolic omentectomy, and pelvic- and para-aortic lymphadenectomy. Comprehensive staging is essential in early stages, when the tumour is considered confined to the ovaries or the pelvis, since the analysis of multiple tissues is essential to exclude the spread of the tumour and to set the best treatment. Indeed, adjuvant chemotherapy is considered beneficial only in clinically early-stage EOC patients with possible unidentified residual disease due to a suboptimal staging[1,19,20]. The ICON and ACTION trials recommended the use of adjuvant chemotherapy in early-stages disease, with carboplatin or cisplatin and paclitaxel, considering the histology subtype and the tumour grade (all stage IB and IC grade 2 and 3 tumours, as well as OCCCs should receive chemotherapy)[21]. Fertility

conservative surgery could be discussed with young patients with early stages disease, after the histopathological diagnosis. Fertility preservation could be considered in cases of HGSOC stage IA or IC1 or LGSOC, EC, expansile MOC and OCCC stage IA-IC. The decision must be examined with the patient.

The volume of residual tumour after the debulking surgery is the strongest prognostic factor both for OS and Progression Free Survival (PFS).

Standard chemotherapy includes combination of carboplatin and paclitaxel. The only exception to this combinatory approach is applied to tumours belonging to the mucinous subtype, for which more benefit could be derived from treatments similar to gastrointestinal-type chemotherapy regimens; due to the different biology of this histotype, as better described below[22,23].

Two phase III randomised trials, GOG218 and ICON7, showed a significantly increased PFS, but not OS with the addition of the anti-angiogenesis inhibitor bevacizumab (directed against vascular endothelial growth factor)[24,25]. These findings led to the approval in 2018 of the addition of bevacizumab to paclitaxel and carboplatin for front-line use in stage III or IV ovarian cancers.

In 2006, the GOG 172 trial proposed intraperitoneal chemotherapy as a standard of care[26]. However, the subsequent trial GOG 252 reported no significant advantages in PFS or OS of the intraperitoneal chemotherapy regimen compared to the intravenous approach in optimally resected patients with stage III disease. Moreover, intraperitoneal chemotherapy had higher incidence of toxicity[27].

Recently, two randomised studies from Dutch and Korean groups supported the use of hyperthermic intraperitoneal chemotherapy for women with primary ovarian cancer[28,29]. However, both studies were conducted on a limited cohort, and further research on larger sample sizes is necessary to correctly estimate the advantages of hyperthermic intraperitoneal chemotherapy as a standard of care.

Recurrence

Recurrence remains incurable in about 75% of women who present with advanced disease. Surveillance frequently includes a combination of clinical evaluation strategies, such as pelvic examinations, imaging and CA 125 monitoring. A functional algorithm uses the platinum-free interval to infer the prognosis and select subsequent therapy at the onset of relapse. Patients are defined platinum-sensitive if recurrence appeared at least 6 months after platinum treatment; otherwise, patients are defined as platinum-resistant. Accordingly, the treatment of recurrences is different for these two types of patients.

Second debulking surgery could be proposed to platinum sensitive patients with isolated or small-volume disease and minimal ascites, followed by chemotherapy with carboplatin in combination with bevacizumab, liposomal doxorubicin, gemcitabine, or a taxane. In 2017, three randomized phase III trials showed that maintenance therapy after initial response to platinum-based treatment with the three PARPi, olaparib, rucaparib and niraparib, was associated with improved PFS in patients with platinum-sensitive recurrent ovarian cancers[30–32]. Therefore, they were approved by the FDA for maintenance treatment of women with platinum-sensitive recurrent ovarian cancers. PARPi are a novel class of drug that offer the possibility to precisely treat tumour cells with defects in the DNA repair pathway HR. BRCA1 and BRCA2 mutated tumours are defective in the HR pathway, but all the other tumours with alterations in other components of this pathway are HR-deficient (HRD) as well. Four recent trials, SOLO1, PRIMA, PAOLA1 and VELIA, showed that the PFS of all advanced tumours with HRD improves if they are treated with PARPi as maintenance therapy in the first-line treatment[33–36]. As a result, there are now three FDA approvals for PARPi as maintenance therapy in the front-line settings. Moreover, these trials exploited novel genomic tests to assess the HRD condition, thus making the development of an approved test for the assessment of HRD an urgent clinical need.

Platinum-resistant recurrent tumours are on the other hand treated with variety of single-agent cytotoxic combinations; single agent approach led to less severe toxicity and

multiple possible combinations of subsequent single-agents could be applied as reported by the National Comprehensive Cancer Network guidelines[37].

Despite the evolution in EOC therapeutic approaches, many questions remain to be addressed to improve the treatment of platinum resistant patients and especially patients with advanced tumour of rare histotypes, such as OCCCs, that recently showed promising results when treated with immunotherapy, in contrast with HGSOCs[38–40].

EOC histotypes

EOCs is a heterogeneous group of tumours that differ from site of origin to clinical course. The novel reported molecular data allowed to establish the main characteristics of the different subtypes. According to the last World Health Organization classification, the main types of ovarian carcinomas are: HGSOC, LGSOC, EC, OCCC and MOC[2]. They can be classified according to histopathology, immunoprofile, and molecular analyses. Table 2 summarizes the main features of each histotype. Basing on the etiopathogenetic mechanism, the molecular and clinic-pathologic features, EOCs are divided into Type 1 tumours, which include LGSOCs, ECs, OCCCs, and MOCs, and Type 2 tumours, which include HGSOCs[41,42].

Table 2: Major pathologic and molecular features of EOC histotypes (adapted from[43])

	HGSOC	LGSOC	EC	OCCC	MOC
Site of origin	Fallopian tube	Endosalpingiosis/ Fallopian tube	Endometriosis	Endometriosis	Teratoma/ Unknown
Precursor lesion	Serous tubal intraepithelial carcinoma (STIC)	Serous borderline tumour	Atypical endometriosis endometrioid borderline tumour	Atypical endometriosis; clear cell borderline tumour	Mucinous borderline tumour
Hereditary Cancer Syndrome	BRCA1/BRCA2-associated hereditary breast and ovarian cancer syndrome (HBOC)	-	Lynch syndrome	Lynch syndrome	-
Molecular alterations	*TP53* *BRCA1/BRCA2* HRD Chromosomal instability Copy-number alterations	*KRAS* *NRAS* *BRAF* HER2	*CTNNB1* *PIK3CA* *PTEN* *KRAS* *ARID1A* MSI *POLE* *TP53*	*ARID1A* *PIK3CA* *PTEN* MSI	*CDKN2A* copy-number loss *KRAS* HER2 amplification *TP53*
Potential targeted therapies	PARPi	MEK inhibitor	mTOR inhibitors; Immune checkpoint inhibitors	Tyrosine kinase inhibitor; Immune checkpoint inhibitors	Trastuzumab

HGSOC

HGSOC is the most frequent EOC subtype, accounting for the 70% of advanced stages tumours. Unfortunately, about 80% of HGSOC are diagnosed at advanced stages, only 5-10% of HGSOC are diagnosed at stage I.

From a morphological point of view, two types of HGSOC have been described: the classic type and SET (Solid, pseudo-Endometrioid and Transitional) variant[2,44]. Classic HGSOC shows variable architectural features including papillary, micropapillary and solid growth patterns. The tumour cells typically exhibit prominent nucleoli in the nucleus and high mitotic activity. SET variant is characterized by solid sheets of cells like those typical of endometrioid and/or transitional cell carcinomas. Necrosis and tumour-infiltrating lymphocytes (TILs) are typical in these tumours.

Immunohistochemistry of HGSOCs shows positivity for Wilms tumor 1 (WT1), p16, abnormal-type pattern of p53, and variable expression of Estrogen and Progesterone receptors (ER and PR)[2]. The SET pattern is more commonly associated to *BRCA1* and *BRCA2* germline mutations[44].

HGSOCs derives from cells belonging to the epithelium of the fimbria of the fallopian tube. This evidence comes from molecular and morphological data, indeed, in the fallopian tubes of prophylactic salpingo-oophorectomies executed on women with *BRCA1* or *BRCA2* germline mutations it was identified the Serous Tubal Intraepithelial Carcinoma (STIC) that shows the same characteristic of HGSOC, such as *TP53* alteration and high Ki-67 index. This is not the precursor lesion of HGSOC, it is the early histologic manifestation of HGSOC. Precise protocol to dissect the fimbriae during risk reducing salpingo-oophorectomies, the Sectioning and Extensively Examining the Fimbriated End (SEE-FIM protocol)[45], allowed to identify three types of precursor lesions, defined as: p53-signature, normal-appearing tubal epithelium that overexpresses p53, Serous Tubal Intraepithelial Lesion (STIL), a lesion composed of at least 12 consecutive secretory cells with alterations of *TP53* and SCOUT (Secretory Cell OUTgrowth), defined as secretory cell outgrowth to more than 30 cells with alterations in *TP53*[46]. Several studies showed that the p53 signature, STILs and STICs harboured the same *TP53* alteration of matched tumour biopsy, corroborating that they are the initial event of HGSOCs[47–49].

LGSOC

LGSOCs is a rare subtype of EOC, it accounts for about 3% of total EOCs. Only 2–5% of women are diagnosed with stage I disease. No definitive treatment guidelines are available for early stage disease, as previously mentioned; the National Comprehensive Cancer Network (NCCN) suggests observation for women with stage IA or IB LGSOC, while there is no treatment standard for those with stage IC disease, in this case observation, chemotherapy or endocrine therapy could all be considered suitable treatment options[37].

Advanced stage diseases are often associated with a worse prognosis as they are poorly responsive to conventional platinum-based chemotherapy[50].

Morphologically, the tumour is usually solid and cystic involving the ovarian surface and the parenchyma. It is usually characterized by cuboidal, low columnar, and rarely flattened cells with an amphophilic or lightly eosinophilic cytoplasm. Nuclei seldom show atypia, and the mitotic index is usually low. Psammoma bodies are frequent. LGSOCs are positive to WT1, CK7, PAX8, ER and PR. The Ki-67 proliferation index is low (less than 3%) and p53 shows a wild-type expression[51].

LGSOCs can arise as an evolution of Serous Borderline Tumours (SBTs) or as *de novo* malignancy from the ovary or the peritoneum. According to the dualistic tumorigenesis model of EOCs[52], LGSOC evolves in a stepwise fashion from a benign serous neoplasm, to SBT, then to non-invasive LGSOC (micropapillary/cribriform serous borderline tumour) and ultimately to invasive LGSOC. Indeed, LGSOCs differentiate from SBTs by the presence of gross stromal invasion. Regarding the precursor lesion of LGSC, two hypotheses have been formulated, but both remain controversial for some aspects. One hypothesis is that the epithelial progenitor cells migrate from the fallopian tube to the ovaries during ovulation, giving origin to serous inclusion cysts and then to serous cystadenomas. It has also been postulated that papillary tubal hyperplasia represents the precursor of serous borderline tumours[52,53].

EC

EC represents the second most diffused EOC histotype, diagnosed in 10-15% of patients; most of the ECs are stage I-II disease and are associated to a synchronous endometrial cancer in 20% of the patients.

Most ECs are low-grade carcinomas with glandular, cribriform, and/or villoglandular patterns. The glands are usually composed of tall, stratified columnar cells with sparse eosinophilic cytoplasm. High grade ECs are usually poorly-differentiated tumours with multiple patterns of atypia and high mitotic index. These tumours are similar to the SET variant of HGSOC from which they should be distinguished exploiting endometroid confirmatory features, such as: metaplastic features (squamous, moral, hobnail, or mucinous) or other alterations in cellular phenotype (eosinophilic or secretory change);

association with endometriosis, ovarian endometrioid adenofibroma or endometrioid borderline tumour; or the presence of a synchronous uterine endometrioid neoplasm[54].

ECs are positive to PAX8, Vimentin, ER, PR and β-catenin in some cases; TP53 alteration can be observed in high grade tumours[51].

ECs are frequently associated with endometriosis and contain areas of endometrioid adenofibroma and endometrioid borderline tumour. Thus, the precursor lesion of ECs is represented by endometriosis in 40% of cases. The hypothesis that ECs evolve from these precursors event is corroborated by the common molecular alterations between tumour and adjacent endometriosis[54–56].

OCCC

OCCCs represent 10% of EOC cases, are usually diagnosed at early stages, about 60-80% of OCCCs tumours belong to stage I-II EOC and they account for approximately 20-25% of all stage I EOC. Early-stage disease usually has a good prognosis with a five-year survival rate close to 90%, while advanced stages have a worse prognosis compared to that of HGSOCs at the same stage and this is ascribable to resistance to platinum-based chemotherapy[57–60]. This histotype is more diffused among Asian (11% of EOCs) compared to people of white and black ancestry (3-4% of EOCs)[61,62].

Morphological features include a combination of papillary, tubulocystic, and solid patterns and stromal hyalinization, tumour cells have clear and eosinophilic cytoplasm. OCCC immunophenotype is characterized by positivity for the hepatocyte nuclear factor 1-beta (HNF-1β), negative staining for WT1, ER, PR, and wild-type pattern of p53[2,51].

Similarly to ECs, OCCCs are strongly associated with endometriosis, and many tumours contain adenofibromatous and borderline areas, suggesting the close relationship between these two histotypes.

MOC

MOCs currently accounts for only 3% of EOCs, the incidence of MOCs has decreased since the introduction of current identification criteria that allowed the distinction between benign and malignant MOCs, but especially between primary mucinous carcinoma of the ovary and metastatic carcinoma of the ovary, which does not have an ovarian origin[63]. Features suggestive of primary ovarian MOC include large size (>10 cm), unilaterality, and absence of ovarian surface involvement. Approximately 80% of MOCs are diagnosed as stage I, when the prognosis is generally good. As for the other subtypes, tumour spread outside the tissue of origin sensibly worsen the patients' prognosis. MOCs usually present as a unilocular cystic mass containing mucinous fluid. According to the growth and invasion pattern, MOCs can be divided into expansile and infiltrative subtypes[6]. The expansile subtype has no destructive stromal invasion, but exhibits back-to-back or complex malignant glands with minimal or absent stroma. The infiltrative type is associated with stromal invasion and with a desmoplastic stromal reaction. This last subtype is generally associated to higher risk of relapse. MOC immunohistochemistry is characterized by positivity for CK7 with variable, but not diffuse expression of CK20. All the primary tumours not associated with cystic teratomas are negative for SATB2. CDX2 is usually expressed, while WT1, ER, and PR expression is absent[64].

From an etiopathogenetic point of view there are multiple theories about the development of MOC, a clear origin of ovarian MOC is not established yet. One of the hypotheses is that it evolves in a stepwise fashion from benign epithelium to borderline tumour to invasive carcinoma. The existence of mucinous cystadenoma and mucinous borderline areas in MOC supports this. Another possible origin is from germ cell, this hypothesis is supported by the association with mature teratoma in 5% of cases. However, most MOCs do not have any teratomatous components. They have been suggested to be mucinous metaplasia of the ovarian surface epithelium or within the lining of cortical inclusion cysts, while the similarity to endocervical-like or Mullerian mucinous tumours suggests an association with endometriosis. Finally, the mucinous epithelium frequently presents with Brenner tumours, thus MOCs, mainly belonging to the intestinal type, may evolve from transitional cells or metaplasia at the fallopian tube-peritoneal junction[42,57,65–67].

Single Nucleotide Variants (SNVs)

Single Nucleotide Variants (SNVs) in oncosuppressors and oncogenes represent one of the molecular fingerprints, along with Somatic Copy Number Alterations (SCNAs), that drives tumour evolution. Recurrent and clonal SNVs are responsible for tumour insurgence and characterize all tumour cells, passenger SNVs stochastically occur in neoplastic cells creating different malignant subpopulations that shape tumour behaviour. Many SNVs represent targetable markers for precise therapies and could also be responsible for resistance to treatments.

EOC subtypes are characterized by almost always distinct, but sometimes common, recurrent SNVs that can explain their behaviour.

HGSOC: The common characteristic of almost all HGSOCs is the presence of a clonal mutation in *TP53*, several molecular studies suggested that HGSOC carcinogenesis is initiated by early p53 loss that favours the consequent disruption of DNA repair pathway, followed by increasing levels of chromosomal instability[68,69]. Indeed, a huge number of somatic copy-number alterations (SCNAs) represents the major determinant of progression of HGSOC[70,71]. No other recurrent altered genes were identified apart from *BRCA1* and *BRCA2*, which are mutated in approximately 20% of HGSOCs and other genes of HR pathway that globally cause HRD in about 50% of HGSOCs[72].

LGSOC: The paucity of cases prevented large studies on the molecular landscape of LGSOCs. The predominant alterations of this histotype are activating mutations in *KRAS*, *BRAF*, *ERBB2*, *NRAS* and *NF1*, underlining the important role of the MAP-Kinase signalling pathway in the tumorigenesis of this cancer. *BRAF* mutations are more frequently associated to early disease stages. These alterations are found also in SBT, suggesting their etiopathogenetic role. Other driver mutations have been described in *PIK3CA*, *FFAR1*, *USP9X* and *EIF1AX*. USP9X and EIF1AX proteins are both linked to regulation of mTOR, which is a downstream effector of the MAPK pathway[73–75]. To date, data regarding potential treatments with MEK inhibitors, exploiting the alteration of MAP-Kinase pathway, as well as the implementation of endocrine therapy are still insufficient to

formulate better targeted agents that could overcome the resistance to chemotherapy of LGSOCs.

EC: EC is a heterogeneous group of tumours with several molecular defects. Among the most common molecular alterations are mutations of the β-catenin gene *CTNNB1*, which are usually associated to stage I -II ECs (89% of the cases) and favourable prognosis, mutations in *ARID1A*, and consequent loss of expression of the encoded oncosuppressor BAF250a, involved in the chromatin remodelling SWI/SNF pathway, alterations of the apoptosis controlling PI3K pathway through loss of function mutations in *PTEN* or activating mutations in *PI3KCA* and defects in the MAP-Kinase pathway through mutations in its components such as *KRAS* and *NRAS*[76]. These events are present also in endometriosis adjacent to the EC and are thus considered etiopathogenetic events. Beside these recurrent mutations, ECs are also characterized by mutations in MMR genes, responsible for the Lynch syndrome, in 10-20% of cases. The presence of MMR deficiency characterizes one of the four molecular classes of ECs recently recognized by The Cancer Genome Atlas (TCGA)[77]; precisely, the molecular characterization of ECs includes ultramutated tumours due to *POLE* exonuclease domain mutations (~5%), hypermutated tumours due to mismatch repair deficiency (MMRd; ~13%), *TP53*-mutated (9–13%) cases, and tumours with no specific molecular profile (NSMP; 69–73%). This molecular classification strongly correlates to patients 'outcomes.

OCCC: Mutations in *ARID1A* are present in about half of the OCCCs, followed by alterations in *PIK3CA* in about 40% of cases. These mutations are detected also in endometriosis adjacent to the tumour mass, confirming their driver role. Beside these alterations, there are other less frequent defects, such as MMR deficiency, found in less than 10% of cases and alterations in *PTEN*, occurring in about 20% of cases. These alterations open novel treatment opportunities that should be further investigated[78,79]. Mutations in *TERT* are rare in early stages and most diffused in advanced disease where they correlated with a worse prognosis, these mutations are absent in endometriosis, supporting the idea that they should be a late event[80–82].

MOC**:** The most frequent molecular alterations are *CDKN2A* copy number loss, *KRAS* activating mutations and *TP53* loss of function variants. *CDKN2A* and *KRAS* alterations occur early during the tumorigenic process, indeed they are present also in borderline tumours, while *TP53* mutations occur later as they are exclusively detected in mucinous carcinoma. Other less frequent mutations are amplification of *HER2* and variants in *RNF43*, *BRAF*, *PIK3CA*, and *ARID1A*[65]. MOCs are usually less responsive to conventional platinum-based chemotherapy; moreover, due to the similarity to gastro-intestinal tumours, therapy regimens like those for this type of tumours have been proposed, as well as treatments against *HER2* amplification. These alternative treatments should be deeply explored to assess their efficacy[22].

Genomic Instability

SCNAs are responsible for genomic instability; with SNVs, they represent an additional tool to distinguish tumours and understand their evolution. The close relation between tumour behaviour, patient prognosis and genomic aberrations emerged several years ago, when multiple studies showed that DNA ploidy could be an independent prognostic factor to predict patients' risk of relapse and to select patients at high risk of recurrence eligible for adjuvant chemotherapy[83]. Polyploid and aneuploid stage I EOC tumours, independently from the histotype, should be considered at high risk of recurrence and thus shall be treated with adjuvant chemotherapy. This association between DNA ploidy and patients' prognosis remained significant also in multivariate analysis with the other clinical variables exploited to stratify patients according to the risk of recurrence[84].

Subsequent studies mainly focused on the most frequent subtype HGSOC and found genomic instability to be an intrinsic characteristic of EOCs. Zall and colleagues showed that the genomic aberrations of stage I EOC subtypes are mostly similar to those of the matched tumours subtypes at advanced stages, confirming that stage I EOC and advanced stages EOC are not distinct diseases with different metastatic potential, but represented the same disease diagnosed at different phases of tumour evolution[85]. Moreover, they suggested a correlation between increased amounts of genomic aberrations and worse prognosis. Few years later, investigating the genomic landscape

of advanced stage EOCs, Wang et al. defined structural aberrations signatures, strictly related to point mutations, that were able to stratify EOCs into 7 types reflecting different genetic backgrounds responsible for diverse biological behaviours[80]. SCNA patterns in human cancers could be considered a sort of genomic scar indicative of the mutational processes that causes it and can be thus exploited to assist molecular stratification of tumours for precision medicine. Macintyre and colleagues have recently identified and validated seven SCNAs signatures caused by multiple mutational processes in HGSOCs[70]. These SCNA signatures were found to be associated with patient outcome and platinum resistance relapse. Even if these SCNA signatures were identified in HGSOC tumour type, they were also evident in other tumours characterized by high degree of genomic instability, such as head and neck squamous cell carcinoma, confirming their prognostic value[86]. Recently, Graf and colleagues, showed that in HGSOCs the presence of SCNA in specific 13 genomic regions could be used to calculate a CNV risk score that predicts OS of patients with the current clinic-pathological variables[71]. Moreover, SCNAs profiles are peculiar of each tumour type and thus can be exploited to identify cancer type and the organ of origin[87], corroborating the theory that SCNAs landscape represent a specific tool to discriminate tumours and analyse their behaviour.

Long Interspersed Nuclear Element 1 (LINE1) methylation

DNA methylation is one of the epigenetic mechanisms that regulate chromatin organization and several other processes, such as gene expression. Alterations of DNA methylation is one of the features currently emerging as a potential diagnostic and prognostic marker. Indeed, global DNA hypomethylation, the genome-wide decrease of methylated CpGs sites, is an aspect typical of all cancer types and is associated with tumour behaviour and prognosis. Indeed, alterations of methylation in specific genomic loci can sustain the expression of oncogenes and, concomitantly, the silencing of oncosuppressors. Moreover, genome-wide hypomethylation contributes through several mechanisms to the increase of chromosome instability. One of these mechanisms is the loss of DNA methylation in mobile elements such as Long Interspersed Nuclear Element 1 (LINE1). LINE1 retrotransposons are highly repeated and widely interspersed human

retrotransposon sequences. LINE1s constitute about 17% of the human genome; with up to 600,000 copies of LINE1s, approximately 2000 of which are still full length and capable of transposition. DNA methylation occurs in these sites to maintain them silenced. Several studies observed that hypomethylation of LINE1 was correlated with histological subtypes, higher FIGO and advanced tumour grade. Moreover, a decrease in LINE1 methylation is also associated to worse OS[88–90]. A loss of methylation of LINE1 seems to be an early event in carcinogenesis, indeed several studies observed reduction in LINE1 methylation also in precancerous lesions[91–93]. These results suggested that LINE1 could be a prognostic marker also in stage I EOCs.

Previous data on potential prognostic molecular classifiers, the use of transcriptional signatures

Several studies already reported the utility of transcriptional markers not only to understand tumour biology and behaviour, but also to predict patients' outcome. For example, large scale genomic studies, conducted by The Cancer Genome Atlas (TCGA) Research Network, determined four classes of HGSOCs, "immunoreactive," "differentiated," "proliferative," and "mesenchymal", based on gene expression profiles[68]. These transcriptional classes are associated with distinct clinical outcomes, particularly, the immunoreactive subtype resulted associated with *BRCA1* alterations and a better prognosis. It was proposed that these subtypes may reflect the patterns of the mutational pathways responsible for the tumorigenesis.

The laboratory of Maurizio D'Incalci, whom I worked with for my PhD course, has a robust expertise in the study of genetic and epigenetic markers useful for the diagnosis and the prognosis of EOCs. In 2008, they defined a subset of genes involved in cell cycle regulation and chromosome maintenance whose expression was significantly associated to risk of relapse[94]. Among these genes, cyclin E and MCM5, already identified as predictive markers in advanced stages EOC, resulted up-regulated in tumours of patients who relapsed, and their expression levels correlated with OS and PFS. In a subsequent study, they investigated association of miRNAs expression with stage I EOC prognosis on a

multicentric cohort of stage I EOCs[95]. miRNAs are 21–23 nucleotide long non-coding RNAs involved in the regulation of gene expression. Alterations of miRNAs could be genetic, potentially modifying their functionality, or epigenetic, in proteins involved in their synthesis and expression, such as DICER and DROSHA. Since miRNAs are a regulatory system of gene expression, alterations in miRNA expression and specific miRNA signatures can be used to identify cancerous tissues, as well as to predict tumour prognosis. In this study, my research group identified mir200-c as the first miRNA with prognostic role in stage I EOCs by microarrays analyses[95]. Particularly, the loss of miR-200c correlated with poor prognosis in stage I EOCs, causing increased expression of VEGFA, which is a putative miR-200c downstream target that correlated with reduced PFS and OS in multivariate analysis. Moreover, they defined robust miRNAs histotype markers, particularly miR-30a as a marker of OCCCs and miR-192/194 as marker of MOCs[96]. miR-30a is highly expressed in OCCCs compared to other EOC subtypes and its expression levels were negatively correlated to those of E2F3, a transcription factor involved in control of cell proliferation. miR192/194 is instead highly expressed in MOCs samples, compared to other EOCs, also in mucinous borderline tumours, even if at low levels. These miRNAs are probably involved in TP53 regulation. Deepening the study of potential prognostic transcriptional signatures, the Micrographite, a new pathway-based integrative analysis tool developed by our group[97], was exploited to reconstruct a circuit composed of 26 elements, 16 miRNAs and 10 genes, associated to both OS and PFS[98]. The components of the circuit span different cell pathways: cell cycle regulation, Activins/Inhibins and Hedgehog signalling. Their expression levels were integrated to predict the risk of recurrence; precisely, their expression levels were translated in an activation state index, called Integrated Signature Classifier (ISC), able to stratify patients into classes of risk. This ISC was able to predict patients' prognosis (PFS) with 80% of sensitivity and 95% of specificity, independently from tumour histotype. Beside miRNAs, other regulatory RNAs are indicators of patients' risk of relapse and outcome, particularly, increased transcription of lnc-SER-TAD2-3, lnc-SOX4-1, lnc-HRCT1-1, and PVT1 resulted independent prognostic markers of relapse and poor prognosis[99]. The increased levels of

lnc-SOX4-1 and PVT1 probably cause increased activation of the AKT/PIK3CA pathway involved in regulation of cell proliferation.

These results prompted us to extend the knowledge about the molecular landscape of stage I EOC to define novel potential prognostic markers feasible for clinical practice. This was the basis for my PhD project.

AIM AND SCOPE

Stage I EOC encompasses five different histological subtypes characterized by distinct etiopathogenetic, morphological and clinicopathological features. Nevertheless, despite the five histotypes represent distinct diseases, common mechanisms that drive tumour evolution have been depicted, corroborating the hypothesis that, albeit they originate from different etiopathogenetic processes, all EOCs share common molecular markers that could indicate similar tumour behaviour and that could be exploited to predict patients' outcome. As a consequence, as highlighted by previous studies conducted in my hosting laboratory[94–96,98,99], the development of molecular classifiers to predict relapse risk of stage I EOCs is an unmet clinical need that could undoubtedly improve the clinical management of stage I EOC patients.

My PhD project aims at increasing the knowledge of the molecular features of stage I EOCs both from a genomic and epigenomic point of view, to better explain tumour behaviour and to finally improve clinical practice.

To achieve this aim, a large retrospective multi-center cohort composed of 205 cases of stage I EOC was selected as explained in the Results Chapter "Sample Cohort Description".

The study was based on the use of high throughput NGS technologies available at Hosting Institution and was organized as follows:

- definition of recurrent SNVs in 139 frequently altered in cancers through a targeted-amplicons NGS approach;
- analysis of SCNAs exploiting a low coverage whole genome sequencing called shallow Whole Genome Sequencing (sWGS);
- integration of SNVs and SCNAs to better understand the mechanisms that drive tumour behaviour;
- investigation of the global DNA methylation status by the evaluation of LINE1 promoter methylation;

- integration of genomics (SCNAs and SNVs) and epigenomics (LINE1) data to explore possible relationships between the genomic landscape and tumour cells epigenetics;
- correlation of genomics and epigenomics data to patients' follow-up information to identify novel potential prognostic markers.

This study was carried out on DNA purified from two types of tumour specimens, snap-frozen biopsies and Formalin Fixed Paraffin Embedded (FFPE) samples, to define novel potential classifiers on all the biological materials commonly available in the Hospitals' Pathology Units.

Figure 1 graphically summarizes the study workflow.

All the data were analysed to expand the knowledge of the earliest and rarest stage of EOCs and to define novel potential prognostic markers useful in clinical practice.

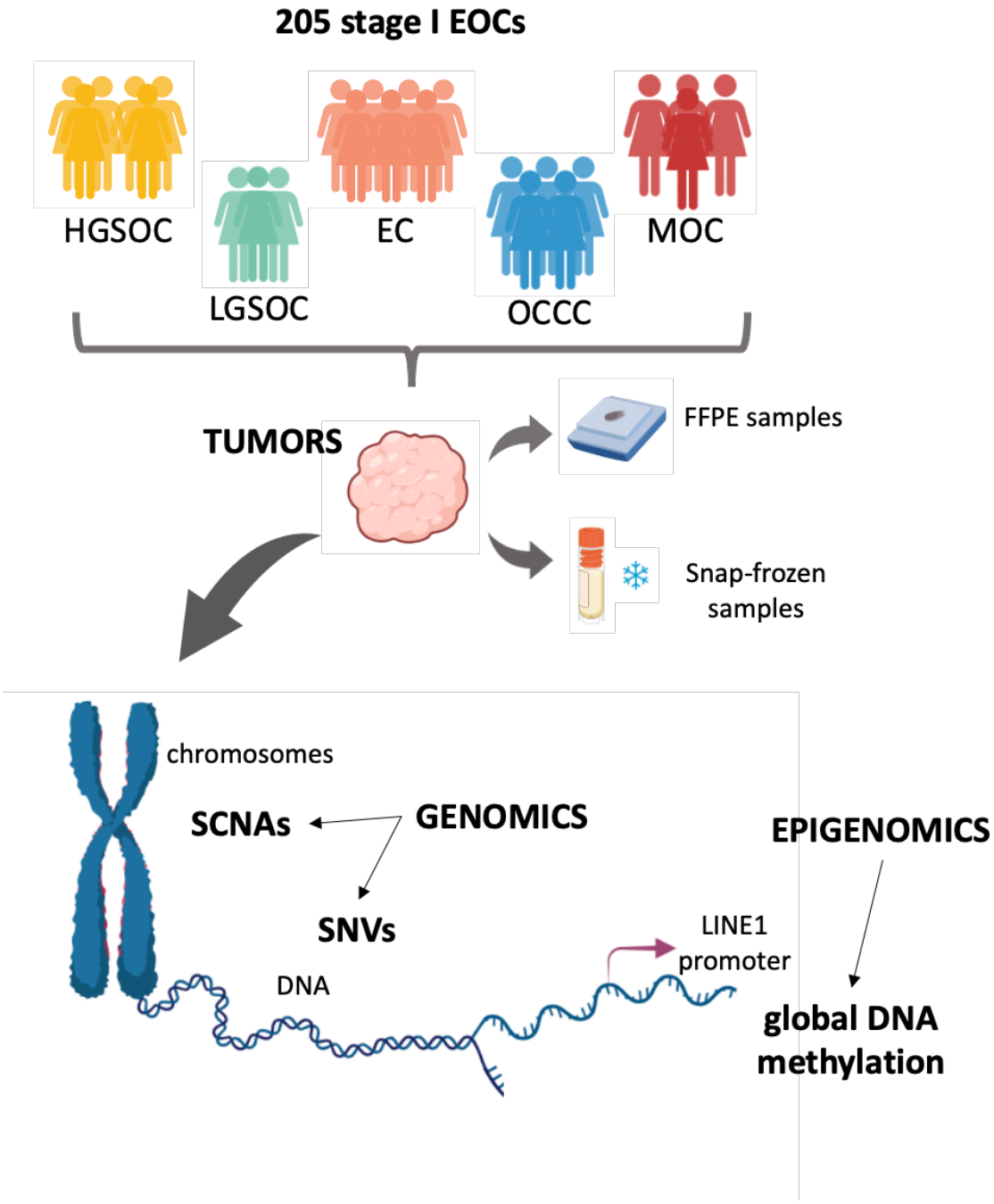


Figure 1: Schematic representations of the study workflow.

MATERIALS AND METHODS

My PhD project involved several collaborators with precise expertise in order to achieve all the results presented in this thesis. Particularly, Dr Luca Beltrame from the Istituto di Ricerche Farmacologiche Mario Negri (Milan, Italy) was responsible for the bioinformatic analyses; Dr. Angelo Velle and Prof. Chiara Romualdi from Università degli Studi di Padova (Padoa, Italy) were responsible for the survival analyses, Dr. Nicolò Panini from the Istituto di Ricerche Farmacologiche Mario Negri (Milan, Italy) was responsible for cytofluorimetry experiments and Dr. Milena Cribiù from Ospedale Maggiore Policlinico (Milan, Italy), Dr. Fulvio Borrella from Sant'Anna Hospital in Torino (Italy) and Dr. Marta Jaconi from San Gerardo Hospital in Monza (Italy) were responsible for HE-stained tumour slides evaluation and histopathological revision. The collaboration with these different professional figures was necessary to complete all the analyses I conceived. I personally took care of the interpretation of the results of each analysis.

Samples collection

The study cohort was selected starting from a collection of 225 stage I EOC cases stored in the Pandora tumour tissue biobank at the Istituto di Ricerche Farmacologiche Mario Negri, IRCCS Milano Italy. Cases were collected from two Italian hospitals from 1989 to 2018; the cohort A was composed by 172 cases from San Gerardo Hospital in Monza (Italy), while the cohort B included 53 patients surgically managed at the Sant'Anna Hospital in Torino (Italy). As depicted in the REMARK diagram (Figure 2), the first level of cohort revision was the exclusion of undifferentiated and mixed histological subtypes to gather only those cases belonging to the five main histotypes (HGSOC, LGSOC, OCCC, MOC, EC). The second level of revision was applied on MOC cases to prevent the inclusion of tumours suspected of not having a primary ovarian origin[63]. Mucinous subtype cases with: i) bilateral tumours, ii) tumours with size smaller than 10 cm, and iii) tumours with overall survival (OS) shorter than 12 months were excluded.

Since the guidelines for histopathological diagnosis have changed during the recent years (e.g. serous subtype has been divided into low and high grade instead of grade 1, 2 or 3),

the samples were recently revised by two pathologists, Dr. Marta Jaconi from San Gerardo Hospital in Monza (Italy) and Dr. Fulvio Borella from Sant'Anna Hospital in Torino (Italy), following the current guidelines of the World Health Organization for EOC [2].

The final study population was composed by 205 Stage I EOCs, 160 cases from San Gerardo Hospital and 45 cases from Sant'Anna Hospital. One sample was excluded from the analyses due to poor DNA quality.

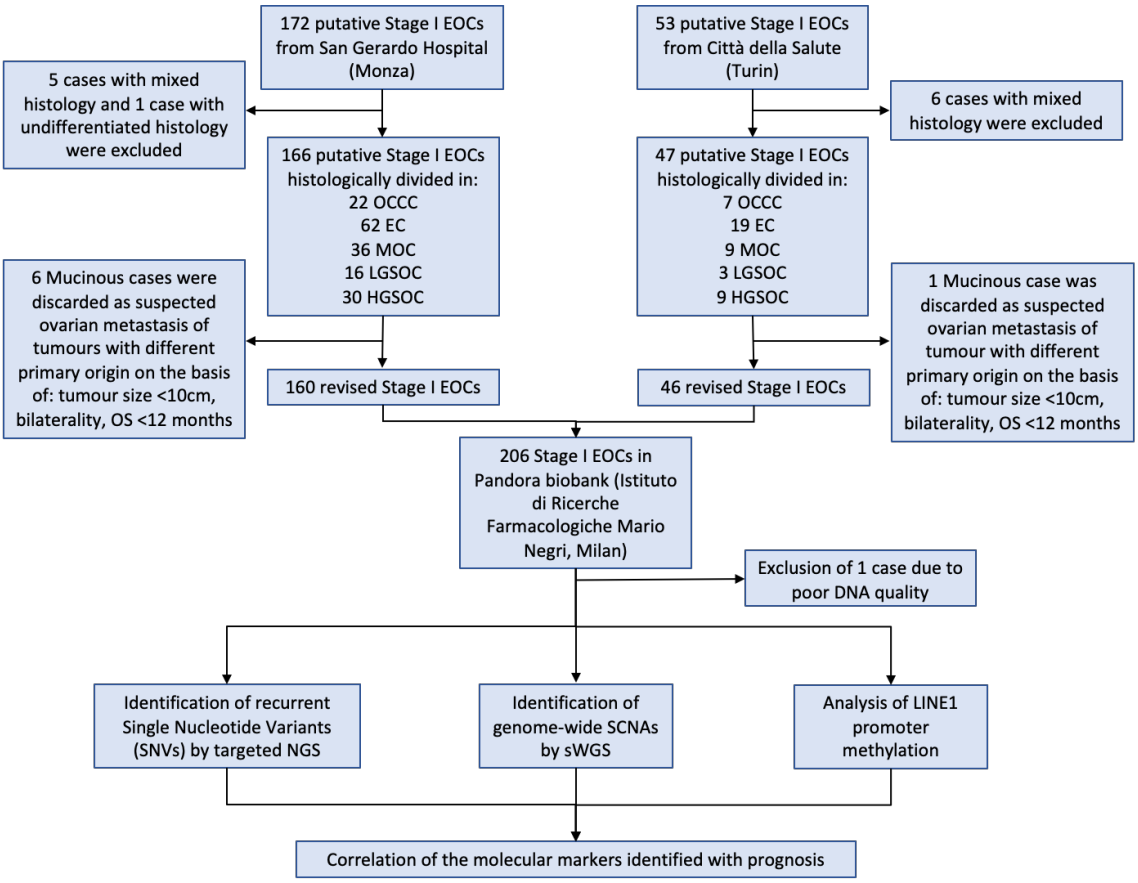


Figure 2: REMARK diagram describing the sample cohort selection

Snap-frozen tumour tissue specimens obtained directly during surgery (Table 3) were collected for 194 cases, while for eleven cases only FFPE tumour samples were gathered from the Pathology Department of the hospital. Both type of samples, FFPE and snap-frozen, were available and analysed for 20 cases. The tumour content of all the FFPE samples was evaluated by pathologists and it was at least 60% for all the cases. For 8 bilateral tumours, tumour samples from both the ovaries were recovered (HGSOC n = 5,

LGSOC n = 2, and OCCC n = 1; Table 3). In addition, snap-frozen samples from a second surgery were obtained from four patients who relapsed (HGSOC n = 1, EC n=2, OCCC n = 1).

Table 3: Histotype distribution of sample type (FFPE, snap-frozen, or both), bilateral and relapsing samples in the patient cohort.

	HGSOC	LGSOC	EC	OCCC	MOC
Only snap-frozen	37	16	65	21	35
Snap-frozen + FFPE	2	2	11	3	2
FFPE	0	1	4	5	1
Samples from both the ovaries	5	2	0	1	0
Samples from the relapse	1	0	2	1	0

The study was performed following the principles of the Declaration of Helsinki; the scientific ethical committee "Brianza" approved collection and usage of tumour, blood and plasma samples (N° 1065, on November 10th, 2015, emended on February 22nd, 2018). Written informed consent was obtained from all patients enrolled in the study.

DNA extraction

The QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) was used to purify DNA from about 25mg of snap-frozen tumour biopsies and 200 μl of blood samples on the automatic nucleic acid purification platform Qiacube (Qiagen, Hilden, Gemany). The automatic nucleic acid purification system Maxwell® (Promega, Madison, Wisconsis, USA) with the RSC DNA FFPE Kit (Promega, Madison, Wisconsis, USA) was used to separate DNA from all the FFPE tumour specimens, starting from 2 to 4 μm tissue sections where the tumour area was delineated by pathologists to reach the maximum tumour purity.

The quantity and the quality of the purified DNA were evaluated with Qubit® dsDNA High Sensitivity Assay Kit (Invitrogen, Carlsbad, California, USA) and 4200 Tapestation (Agilent Technologies, Santa Clara, California, USA), respectively.

Haematoxylin and eosin (HE) staining

In order to evaluate the tumour cell content of the snap-frozen tumour biopsies, HE-staining was performed on 4 μm sections obtained using a cryo-microtome from the snap-frozen tissue samples adjacent to the one used to purify DNA and to the one used to flow cytometry analysis. These sections were dehydrated with absolute ethanol, stained with HE (Bio-Optica, Milan, Italy) and washed with xylene. The estimation of the tumour cell content was performed by a specialized pathologist.

Ploidy analysis through flow cytometry

Flowcytometry analysis was performed on about 25mg of snap-frozen samples. Specimens were fixed in 70% ethanol and stored at 4°C before DNA staining with 2 ml of 25ug/ml propidium iodide solution and 25 **µ**l of RNAsi 0.5mg/ml.

Prepared samples were run on a FACSCalibur system (Becton Dickinson, Sunnyvale, CA, USA) at the acquisition rate of 500cells/s and excluding doublets.

DNA index (R) was used to indicate the ploidy as the ratio between the G1, peak of tumour sample and the G0/G1 peak of peripheral blood standards. The samples were defined diploid if the DNA-index was 1 ± 0.2.

The overlapping of the G1 peak of tumour diploid cells with the G1 peak of normal cells present in the snap-frozen biopsies prevented the calculation of the percentage of cells in each cell cycle phase. Nevertheless, when present, the peak representing the G2 phase was considered mainly composed by neoplastic cells, since these should actively proliferate.

Microsatellite Instability Analysis

In order to assess Microsatellite instability (MSI), and thus MMRd, MSI analysis was performed in all the ECs and in the OCCC cases harbouring a SNV in the MMR genes (*MLH1, MSH2, PMS2* and *MSH6*) and with a matched-blood sample. The analysis was carried out on DNA purified from both tumour and blood samples using a panel composed by five mononucleotide repeat microsatellites (BAT-25, BAT-26, NR-21, NR-24 and NR- 27) amplified with a PCR-based system (Promega, Madison, Wisconsis, USA), following the manufacturer's instructions. Amplified PCR products were diluted with formamide and run on a 3130xl automated capillary electrophoresis DNA sequencer (Applied Biosystems, Foster City, CA). The allelic sizes of each microsatellite was estimated using GeneMapper 3.1 software (Applied Biosystems, Foster City, CA).

A tumour was considered MSI and MMRd if at least one of the microsatellite profiles in the tumour sample was different from the corresponding profile in the matched blood sample.

Targeted Amplicon Sequencing

In order to assess the SNVs landscape of stage I EOC I decided to focus on genes frequently altered in cancer and with known pathogenic significance. Thus, I exploited a an in-house designed Qiaseq Targeted DNA Panel kit (Qiagen, Hilden, Germany) for targeted amplicon NGS that covers the exonic regions of 139 genes involved in DNA damage signalling and repair pathways, cell cycle regulation and signal transduction (the complete list of genes is available in Table 4). Library were prepared using an automatic liquid handling station "BRAVO B" (Agilent Technologies, Santa Clara, California, USA). To achieve the coverage necessary to detect somatic variant (2000X), 33 libraries were pooled together and sequenced on NextSeq 500 in pair-end 2x150 bp mode (Illumina, San Diego, CA).

Table 4: Targeted amplicon sequencing panel of 139 genes. BER Base Excision Repair, HR Homologous Recombination, MMR Mismatch Repair, NER Nucleotide Excision Repair, NHEJ Non Homologous End Joining.

DNA Repair						Cell Signalling			Recurrent Cancer Genes	Cell Cycle
DNA checkpoint	**BER**	**HR**	**MMR**	**NER**	**NHEJ**					
ATR	*ALKBH2*	*BRCA1*	*MLH1*	*DDB1*	*POLM*	*AKT1*	*KRAS*	*NOTCH1*	*APC*	*RB1*
ATM	*ALKBH3*	*BRCA2*	*MLH3*	*DDB2*	*NHEJ1*	*AKT2*	*BRAF*	*NOTCH2*	*ARID1A*	*CDH1*
CHEK1	*MGMT*	*BRIP1*	*MSH2*	*ERCC1*	*XRCC4*	*MTOR*	*NRAS*	*NOTCH3*	*ARID2*	*CDK12*
CHEK2	*PARP1*	*C17orf70*	*MSH3*	*ERCC2*	*XRCC5*	*TP53*		*MYC*	*DROSHA*	*CDK4*
H2AFX	*PARP2*	*c11orf30*	*MSH6*	*ERCC3*	*XRCC6*	*TP53BP1*	*ERBB2*	*NF1*	*DICER1*	*CDKN2A*
MDC1	*XRCC1*	*FANCA*	*PMS1*	*ERCC4*	*POLD1*		*ERBB3*		*EP300*	*CYLD*
RAD1		*FANCB*	*PMS2*	*ERCC5*	*POLE*	*PI3KCA*	*ERBB4*		*B2M*	*CCND1*
RAD17		*FANCC*	*DNMT3A*	*ERCC6*	*POLE4*	*PIK3R1*	*EGFR*		*FOXL2*	*CCNE1*
RAD9A		*FANCD2*		*ERCC8*	*PRKDC*	*PPP2R1A*	*IGF1R*			*CDKN1A*
RNF8		*FANCE*		*RAD23A*		*PTEN*	*MET*		*WNT1*	*CDKN1B*
		FANCF		*XPA*					*CTNNB1*	*CDK16*
		FANCG		*XPC*			*THBS2*		*SMAD7*	*CDK17*
		FANCI					*PDGFRB*		*SMAD2*	
		MUS81					*FGFR2*		*Zeb1*	
		NBN					*FGFR3*		*APOBEC3B*	
		PALB2					*FGFR4*		*TOP1*	
		RAD50					*FLT1*		*TOP2B*	
		RAD51					*KDR*		*TOP2A*	
		RAD51C					*FLT4*		*TOPBP1*	
		RAD51D							*ABCB1*	
		RAD54L							*TUBB3*	
		XRCC2							*BARD1*	
		XRCC3							*TAP1*	
		SHFM1							*TAP2*	
									TAPBP	
									CREBBP	
									RNF213	
									STK11	

Variant Calling

BWA was used to align raw reads to the reference genome (hg38)[100]. Reads were then pre-processed to identify regions where variant calling would be possible ("callable" regions). Two different programs, MuTect 2[101] and VarDict Java[102], were applied to call variants.

Variants were annotated with the Variant Effect Predictor (VEP, version 90[103]) and then loaded in a database GEMINI[104] for further analysis.

Variants were extensively filtered to remove artifacts and false positives, moreover several public databases were exploited to deeply interpret the biological significance of each filtered variants and remove those without biological relevance.

Annotation and interpretation of the variants was performed using dbSNP, version 151 (https://www.ncbi.nlm.nih.gov/snp/), COSMIC, version 90 (https://cancer.sanger.ac.uk/cosmic), ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), BRCA Exchange (https://brcaexchange.org) and the IARC TP53 database[6], version 19.

The variants filtering and interpretation were conducted as follows:

- Discard variants with less than 200X coverage and less than 10% allelic fraction;
- Discard variants caused by sequencing artifacts (e.g., poly-X regions, present in all samples regardless of histotype);
- Discard variants appearing in at least two samples out of a pool of 70 unrelated normal samples used as reference;
- Discard all variants with at least 1% frequency in at least one of three large population sequencing data sets (ExAC, ESP, gnomAD);
- Discard non-benign variants (defined by the presence of a COSMIC record or a non-benign ClinVar status) with high population frequency (> 1%) in at least two out of three data sets (ExAC, ESP, gnomAD);
- Discard variants whose frequency in the general population was greater than 1%;

- Discard variants with no effect on the protein (synonymous, start retained, stop retained);
- Discard variants known to be polymorphisms by checking both COSMIC and dbSNP records;
- Discard variants with low probability of impact on the protein (e.g., intronic, intergenic);
- Discard variants with ClinVar records indicating “Benign” or “Likely Benign”;

For *BRCA1* and *BRCA2* genes, somatic variants with mutated allele fractions higher than 40% were considered potentially germline.

In case of FFPE samples, an additional step was performed to correct for bias caused by DNA damage during formalin treatment, using DKFZBiasFilter (https://github.com/DKFZ-ODCF/DKFZBiasFilter).

The resulting variant call set was then loaded into the Cancer Genome Interpreter (CGI)7, either using the HGVSp notation for missense, frameshift, stop gained variants or coordinates for those which were not (splice acceptor / donor variants). After analysis, only variants with status “known driver” or “predicted driver” (either Tier I or Tier II) were kept for subsequent analyses.

Shallow Whole Genome Sequencing (sWGS)

sWGS was applied on all tumour DNA samples to identify SCNAs across the whole genome. This technique allows to detect gains and losses through a low coverage sequencing with a resolution of 60kb.

For DNA purified from snap-frozen samples, libraries were prepared using the KAPA HyperPrep Kit (Roche, Basel, Switzerland) starting from 200ng DNA fragmented on Bioruptor (Diagenode, Liége, Belgium) for 7th cycles, 30’’ ON and 90’’ OFF with the high potency level and then purified with AMPure XP beads (Beckman Coulter, Brea, California, USA).

For FFPE DNA samples, libraries were prepared using the KAPA HyperPLUS Prep Kit (Roche, Basel, Switzerland), with 200ng DNA enzymatically digested to reach fragments of the proper size to proceed with library preparation. Quantity and quality of the libraries were assessed by Qubit® dsDNA High Sensitivity Assay Kit (Invitrogen, Carlsbad, California, USA) and 4200 Tapestation (Agilent Technologies, Santa Clara, California, USA), respectively.

In each experiment 24 single index barcoded libraries were pooled together in equimolar amount to reach a coverage of 0.5X with 1Mbp of resolution. Prior to sequencing the required sequencing depth was estimated to ensure a power of at least 0.9 when detecting subclonal events (50% of the cell population) of 1Mbp with a minimum tumour purity of 50%, and clonal CNAs with a minimum tumour purity of 25% (https://gMacintyre.shinyapps.io/sWGS_power/)[105], a pessimistic scenario.

The library pool was diluted at 1.45 pM and sequenced on NextSeq 500 platform (Illumina, San Diego, California, US) in 2x150bp mode.

Copy Number Alteration Analysis

Reads were aligned to the reference genome (hg38) using BWA[100].

Aligned reads were pre-processed, removing unreliable or badly mapping reads, and excluding PCR artifacts. QDNAseq[106] was used to generate log2 ratios from sequencing data by grouping reads mapped across the genome in bins of 30kbp. Error rates and data dispersion were calculated by comparing the results on each tumour sample versus a pooled reference of 54 individuals from the 1000 Genomes project, removing outlier bins from the subsequent computations. Bins were then segmented with circular binary segmentation (CBS) as implemented in DNAcopy[107].

Due to poor alignment in the chromosome X for all samples (an indication of a library artifact), chromosome X was excluded from the copy number analysis.

Absolute Copy Number adjustment for purity and ploidy

To accurate predict SCNAs is necessary measured copy number changes on an absolute scale (copies per cancer cell). The main challenges in determining absolute copy number are: tumour content, ploidy and tumour cells populations heterogeneity, especially for those alterations defined as arm-level. In this study, ACE[108] was used to calculate an estimate of tumour purity ("cellularity") and of tumour ploidy. The best fit of absolute copy number, cellularity and ploidy was calculated for each sample using as a guide the MAF of specific variants, previously identified by targeted NGS and with likely clonal status. In case no candidate variant was available to perform the estimation, the combination of purity and ploidy with the highest purity and the lowest error between the ACE estimation and the observed data was selected. The procedure used was as follows:

- Run ACE on the QDNAseq output for all samples for four different ploidy states (2n, 3n, 4n, 5n);
- Collect all the solutions found by ACE for each ploidy;
- For each sample, identify a potential clonal variant from variant calling data and its associated allelic fraction;
- Iterate through the purity/ploidy combinations and extract the log2 ratio and absolute copy number for each calculation;
- Perform two different steps in case the variant is on an oncogene, or on a tumour suppressor gene: in the former case, if the fraction is lower than 60% and there are no copy number alteration events on the gene itself, consider the event to be monoallelic and thus, halve the predicted purity for the subsequent step, if not, consider the change to be biallelic and use the purity value directly; in the latter case, consider the purity value directly;
- Use the predicted purity and the cellularity (adjusted for the type of gene: see step 5) to determine a hypothetical allelic fraction for the variant;
- For all the calculated fractions, select the purity/ploidy combination where the difference with the observed allelic fraction is minimal;

- In case of no clonal variant exploitable, choose the ACE solution for ploidy and purity with the highest cellularity and the lowest error.
- In case of inconsistent results, such has high variance across bins, manually evaluate the ACE plot for a correct solution;
- Repeat the process for all samples.

Then, segments containing corrected copy numbers and log2 ratios were obtained from ACE using the `postanalysisloop` function with the selected purity/ploidy combinations for all samples. Log2 ratios were calculated using the following formula:

$$ratio = log_2\left(\frac{CN}{ploidy}\right)$$

where CN is the calculated copy number from ACE, and ploidy the estimated ploidy of the sample.

GISTIC analysis

Recurrent SCNAs were calculated using GISTIC 2.23[109] from the segmentations produced by QDNAseq and adjusted by ACE. GISTIC allows the determination of focal, broad and arm-level SCNAs and to define the most recurrent ones. Focal SCNAs were considered as regions spanning less than 25% of the chromosome arm. SCNAs greater than this were termed “broad”, if they covered more than 25% of a chromosome arm, but less than the entire chromosome arm or “arm-level” if they involved the entire chromosome arm.

The following parameters were used:

- minimum ratio threshold, 0.3;
- confidence, 99%;
- exclude chromosome X from the computation;
- allow arm-level analysis;
- maximum q-value cutoff; 0.05;
- maximum allowed copy number, 9

Length distribution of SCNAs

SCNA segments (where called copy number was different than the detected ploidy) were first normalized versus the length of the chromosome arm they occurred in, then frequency histograms were made for all sample groups being considered. As for GISTIC analysis, the SCNA was classified as focal, broad or arm-level.

Tumour copy number burden (CNB) calculation

CNB was calculated as the percentage of the genome with altered copy number multiplied by 100. Copy number was considered altered when the absolute copy number called by ACE was different from the sample's reported ploidy, that is, if the following condition was true:

$$CN_{call} - ploidy \neq 0$$

Somatic Copy Number Alteration (SCNA) Signature definition

SCNA signatures were calculated following the procedure by Macintyre et al.[70] using the software published as part of the "cnsignatures" repository (https://bitbucket.org/britroc/cnsignatures).

ACE copy number calls were used to extract copy number features and Gaussian mixture models were fit to the data. Using the function `predictSignatures`. The optimal number of signatures was determined by comparing results from the data versus randomly generated data (resampled 10000 times).

Signatures were then generated using negative matrix factorization (NMF) following the procedure in the "cnsignatures" repository. Exposure to each signature was calculated with the YAPSA R package (https://www.bioconductor.org/packages/release/bioc/html/YAPSA.html) and the `quantifySignatures` function from the cnsignatures repository. Assignment of each sample to a signature was done with the `predict` function from the NMF R package[110].

Differential exposure analysis of copy number signatures

To perform differential exposure, a selection of 149 unique genes was built by merging two gene sets. The first included 80 genes from the targeted gene panel whose SNVs emerged from the variant analysis, and two genes commonly affected by copy number amplifications (*CCNE1* and *MYC*).

The second set was a collection of 81 genes whose amplification or deletion was regarded as a potential driver, and was selected as follows:

- From the GISTIC data on the whole cohort, select all genes with strong amplifications or deletions (absolute value of discretized change > 1);
- Narrow the initial gene set by performing an intersection with the 765 genes present in the Cancer Genome Interpreter, correcting for differences in gene symbol naming.
- Analyse the remaining genes (separately for amplification and deletion) in the CGI and include only those marked as “known driver” or “predicted driver” status.

Then, the SNV data and the copy number data were merged, and the result discretized for each sample. As genes from the variant and copy number data partially overlapped, the 161 genes set obtained from the merge was reduced to 149 unique genes.

The presence of a variant in a gene classified the sample as “altered”. With regards to the copy number related genes, they were marked as “altered” only in presence of strong amplifications (copy number > 5 from the calculated ploidy value) or deletions (copy number <= 2 from the calculated ploidy value). The remaining samples were classified as “non altered”.

Once the assignments were complete, the Mann-Whitney U test was used for each gene in each signature, comparing the exposure of the “altered” samples versus the “non altered” samples. Raw p-values were corrected with the False Discovery Rate[111] method. Only genes with a corrected p-value < 0.05 and a median difference in exposure between “altered” and “non altered” > 0.08 were included.

The Fisher's exact test was used to compare altered genes in samples belonging to each signature to the same genes in the S samples group. A p-value of 0.1 was used as threshold of significance.

Regression tree classification

A classification tree algorithm as implemented in *rpart* R package was applied to a manual selection of 130 samples with clear SCNA pattern (named training set). *rpart* R package was run with the default parameters to grow the tree, the Gini impurity and a 10-fold cross-validation procedure. Average error rate was calculated across the 10-fold classification trees. Then, based on the classification criteria obtained by the classification tree we predicted the class of the remaining 75 samples. Then a manual revision was performed on all the classified samples.

LINE1 methylation analysis

The DNA methylation status of LINE1 promoter was evaluated by using bisulfite treatment followed by pyrosequencing in 187 out of 205 stage I EOC cases, those for which enough DNA was available. Around 500 ng of DNA were subjected to bisulfite modification using EZ DNA methylation-Gold™ kit (Zymo Research, Irvine, CA, USA).

For each sample, LINE1 promoter methylation levels were defined as the mean percentage of methylation of four consecutive CpG sites detected by the PyroMark Q96 CpG LINE1 kit (Qiagen, Hilden, Gemany). Commercial fully methylated and unmethylated DNA (Zymo Research, Irvine, CA, USA) were used as positive and negative controls in each experiment.

Statistical Analysis

Comparisons of CNB between sample groups were carried out with the Mann-Whitney test. Multivariate Cox proportional hazard models were used to test differences in survival measures adjusted for grade, substage, histotype, chemotherapy and *TP53* mutational status.

RESULTS

Samples cohort description

To achieve the aims of my PhD project, a retrospective cohort of 205 stage I EOC biopsies was selected from a tumour collection of 225 tumour cases assembled at two independent Italian clinical centers, as described in Materials and Methods (Figure 2).

The final study cohort was histologically distributed as follows: 18% of MOC, 14 % of OCCC, 39% of EC, 20% of HGSOC and 9% of LGSOC (Table 4). FIGO substage proportion as well as grade composition for MOC and EC cases are detailed in Table 4. The majority of cases belonged to FIGO substage c (n = 116, 57%), 33% to FIGO substage a (n = 68), and 17 cases (10%) to FIGO substage b. MOCs were mainly divided in grade 1 (n = 26, 68.4%) and grade 2 (n= 10, 26.3%), only 1 case was diagnosed as grade 3. Instead, ECs were prevalently classified as grade 2 (n= 38, 47.5%) and grade 3 (n = 28; 35%); 14 out of 80 cases were diagnosed as grade 1 (18%).

Table 4: Summary of the patients' clinicopathological features. G grade; y years; FU Follow-up; n.d. non defined

Clinical Annotations	Number of patients	% of patients	Number of patients with FU	Median FU [IQR 1-3]
	Histology and Grade			
MOC				
G1	26	13		
G2	10	5	38	11.2y [7.2y-14.3y]
G3	1	0.5		
n.d	1	0.5		
OCCC	29	14	28	5.7y [2.5y-14.3y]
EC				
G1	14	7	79	10.1y [6.6y-14.7y]
G2	38	18		
G3	28	14		
LGOSC	19	9	19	14.6y [10.3y-17.2y]
HGSOC	39	19	39	7.3y [5.2y-15.7y]
	FIGO substages			
A	68	33		
B	17	8		
C	116	57		
n.d	4	2		
Median age at diagnosis [min-max];		54.9 y [16.5y – 89.3y]		
Total number of patients		205		

The median age at diagnosis was 54.9 years (IQR 16.5– 89.3), the median follow-up was 8.7 years (IQR 5.6-14.1). Most of the cases were completely cured (n = 145, 71%); however, 21% of patients (n = 43) relapsed after platinum-based chemotherapy. For 17 (8%) cases no information regarding relapse were available. Fifty-one patients were never treated with chemotherapy (25%), while 148 women received platinum-based

chemotherapeutic treatments (72%). The median Overall Survival (OS) was 10.1 years (IQR 6 -15.5), and the median Progression Free Survival (PFS) was 8.7 years (IQR 4.8 - 14.4).

To check if the study cohort was representative of stage I EOC, the clinical variables were associated to survival. The Cox proportional hazards model indicates that grade '3' and sub-stage 'c' were significantly associated with both PFS and OS, while chemotherapy and age were significantly associated with PFS and OS, respectively (Table 5). There was no difference in PFS among the five histotypes, while there was difference in OS among OCCC and HGSOC versus EC, LGSOC and MOC (Figure 3). The study cohort thus reflect the typical parameters of stage I EOCs, and it is suitable to deepen the molecular knowledge of stage I EOCs.

Table 5: Univariate survival models (OS and PFS) for the clinical variables Grade, FIGO2, CT, and AGE (each tested independently). Hazard Ratio (HR) with Confidence Interval at 95% of confidence (CI95%) and p-value of the log-rank test are reported. Highly ($p<0.05$) significant variables are highlighted in bold.

	PFS			OS		
	HR	CI (95%)	P-value	HR	CI (95%)	P-value
Grade						
Low/1(reference)	-	-	-	-	-	-
High/2/3	2.81	1.47-5.37	**0.002**	2.37	1.26-4.48	**0.008**
FIGO2						
a (reference)	-	-	-	-	-	-
b	1.08	0.23-4.99	0.925	2.96	0.86-10.14	0.084
c	2.34	1.11-4.94	**0.026**	3.08	1.35-7.02	**0.008**
CT						
NO (reference)	-	-	-	-	-	-
YES	8.04	1.94-33.32	**0.004**	1.32	0.63-2.78	0.459
Age						
	1.01	0.98-1.04	0.458	1.06	1.03-1.09	**0.00001**
Histotype						
OCCC (reference)	-	-	-	-	-	-
EC	0.87	0.32-2.40	0.793	0.30	0.13-0.70	**0.005**
MOC	0.32	0.08-1.36	0.123	0.32	0.12-0.88	**0.028**
HGSOC	2.59	0.95-7.07	0.064	0.73	0.31-1.68	0.454
LGSOC	0.41	0.08-2.13	0.292	0.17	0.04-0.78	**0.023**

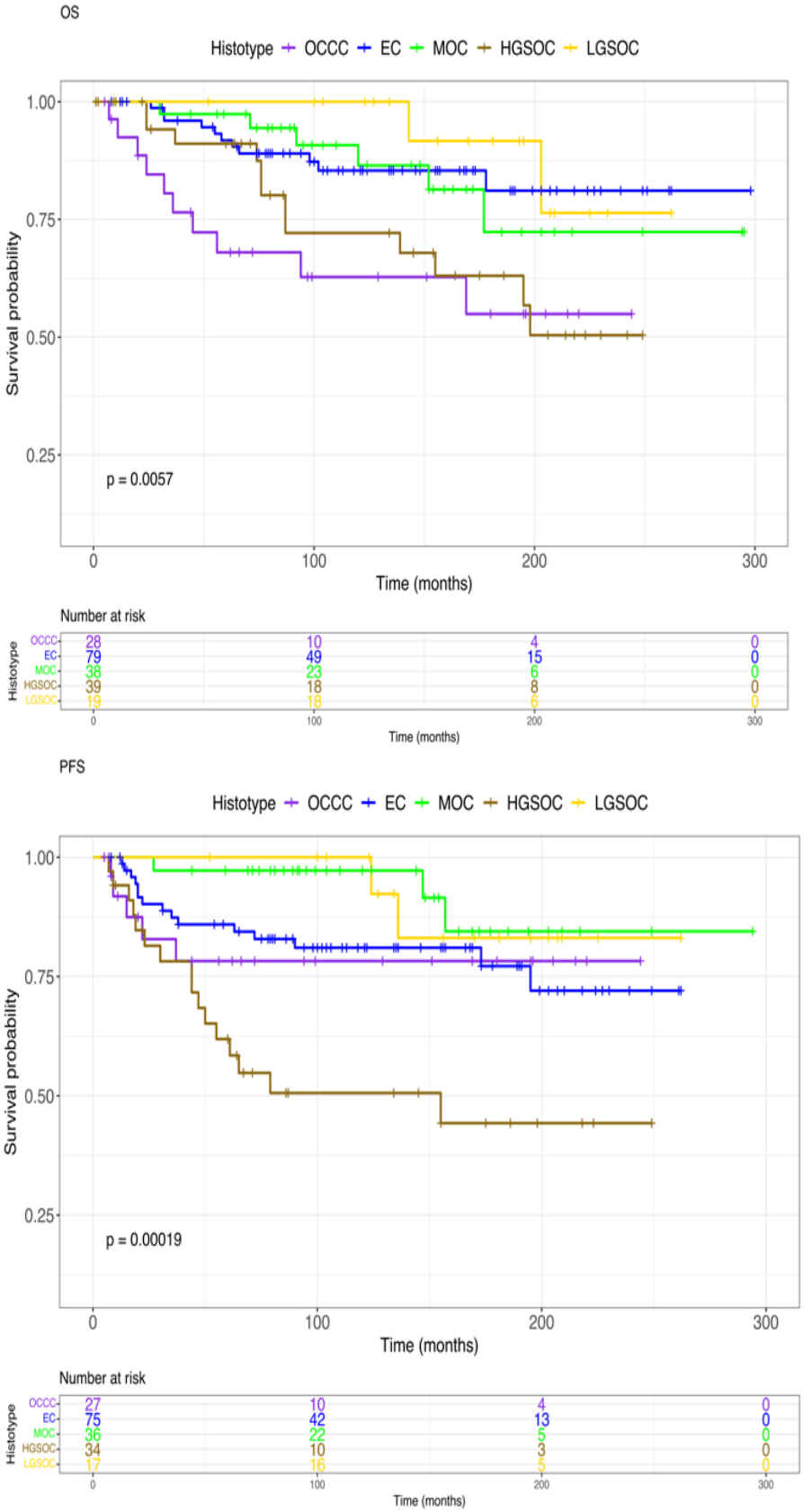


Figure 3: Survival analyses among the five histotypes in the study cohort. The p-value of the log-rank test is reported in the plot along with the risk table.

Analysis of Single Nucleotide Variants in the different histotypes of stage I EOC

As a first step in the molecular characterization of stage I EOC, the single nucleotide variants (SNVs) in cancer-related genes were investigated to highlight the presence of recurrent putative driver mutations across the five histotypes. The analysis was focused on a subset of genes (n=139) known to be frequently altered in cancer and involved in DNA repair, cell cycle regulation and DNA damage signalling. The genes included in the in house-designed targeted amplicon-based panel were divided in nine main categories: DNA checkpoint, Base Excision Repair (BER), HR, MMR, Nucleotide Excision Repair (NER), Non-Homologous End Joining (NHEJ), Cell Signalling and Cell Cycle and finally a miscellaneous of cancer frequently altered genes (see Materials and Methods). The choice to analyse only these genes was taken considering that EOC is characterised by low tumour mutational burden[112].

In the whole sample cohort of stage I EOC 1919 somatic SNVs were identified (with at least 10% allelic fraction). The description of the SNVs repertoire is divided into the five histotypes.

The oncoprint (Figure 4) reports the complete mutational landscape of the altered genes in Stage 1 EOC histotypes. The type of variants identified is indicated by the colours.

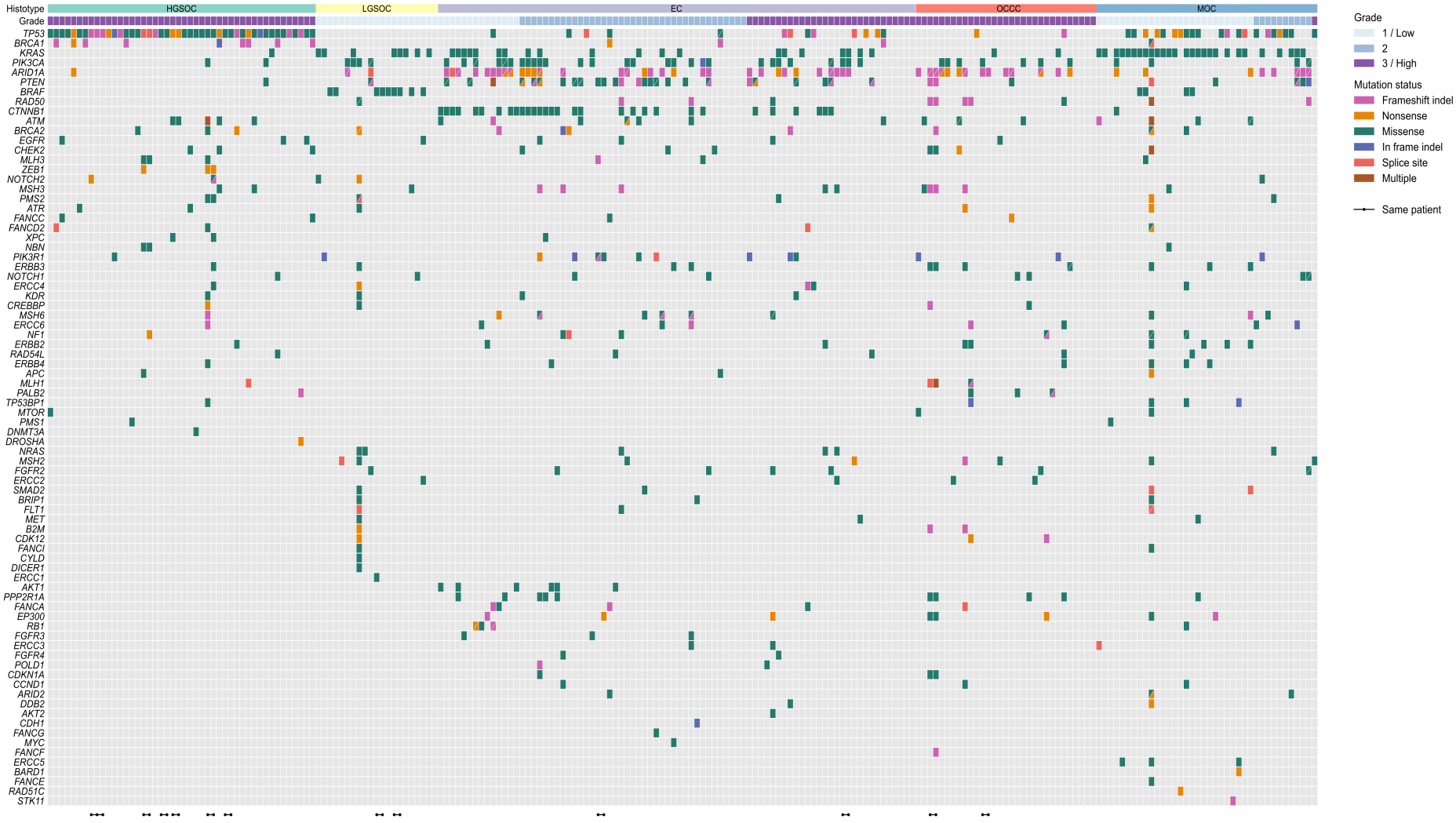

Histotype
HGSOC
LGSOC
EC
OCCC
MOC
Grade
TP53
BRCA1
KRAS
PIK3CA
ARID1A
PTEN
BRAF
RAD50
CTNNB1
ATM
BRCA2
EGFR
CHEK2
MLH3
ZEB1
NOTCH2
MSH3
PMS2
ATR
FANCC
FANCD2
XPC
NBN
PIK3R1
ERBB3
NOTCH1
ERCC4
KDR
CREBBP
MSH6
ERCC6
NF1
ERBB2
RAD54L
ERBB4
APC
MLH1
PALB2
TP53BP1
MTOR
PMS1
DNMT3A
DROSHA
NRAS
MSH2
FGFR2
ERCC2
SMAD2
BRIP1
FLT1
MET
B2M
CDK12
FANCI
CYLD
DICER1
ERCC1
AKT1
PPP2R1A
FANCA
EP300
RB1
FGFR3
ERCC3
FGFR4
POLD1
CDKN1A
CCND1
ARID2
DDB2
AKT2
CDH1
FANCG
MYC
FANCF
ERCC5
BARD1
FANCE
RAD51C
STK11
Grade
1 / Low
2
3 / High
Mutation status
Frameshift indel
Nonsense
Missense
In frame indel
Splice site
Multiple
Same patient

Figure 4: Repertoire of SNVs in Stage 1 EOC histotypes. Samples are organized in rows, grouped by histotype and grade. Dots on the right-hand side indicate different biopsies belonging to the same patient. Columns represent genes (ordered by frequency). Only genes with at least one mutation passing the selection filters (see Materials and Methods) are shown.

The plots (Figure 5) summarize the frequency, the Median Allelic Fraction (MAF) of altered alleles and the putative role of the variants in genes mutated in at least 15% of samples for each histotype, the numerical values of frequency and MAF are specified in Table 6.

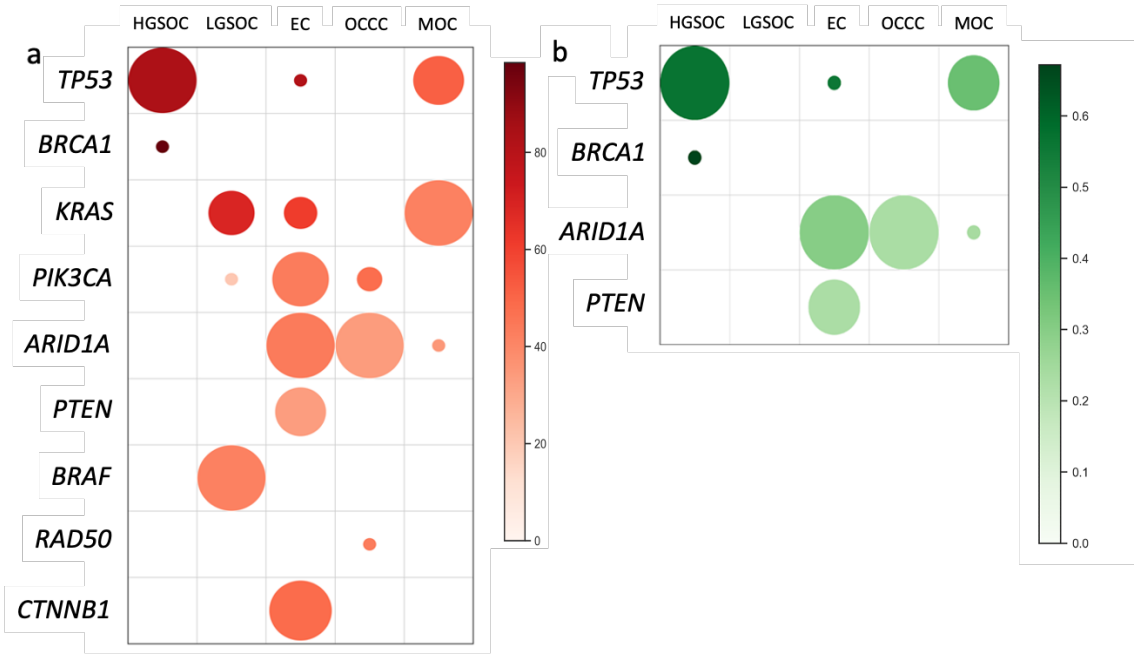


Figure 5: Summary of the Median allelic fraction, frequency and putative role of variants identified in the most frequently altered genes in Stage I EOC samples.

a) *Median allelic fraction and frequency of the variants identified in these genes. Each row is a gene with variants in at least 15% of samples per histotype. The radius of the circle indicates the frequency: the larger the radius, the higher the frequency of variants in that histotype. The colour of the circles shows the median allelic fraction of the variants in a specific gene and histotype: the darker the colour, the higher the frequency.*
b) *Frequency and fraction of putative loss of function (LoF) variants in the most recurrently altered genes in Stage I EOC histotypes. The larger the circle, the more frequent the*

variant is in samples belonging to the same histotype. The darker the colour of each circle, the higher the mutated allele fraction (MAF).

Table 6. Frequency and mutated allelic fraction (MAF) of main genes for each histotype, sorted by decreasing frequency. Only variants present in at least 15% of the sample population in each histotype are included. IQR, interquantile range.

HGSOC			
Gene	Frequency	Median MAF	MAF IQR
TP53	100%	56.98%	39.26%-71.07%
BRCA1	19.15%	67.16%	62%-75.84%
LGSOC			
Gene	Frequency	Median MAF	MAF IQR
BRAF	45%	28.67%	20.49%-40.78%
KRAS	40.0%	46.9%	31.42%-56.31%
PIK3CA	20.0%	14.39%	13.85%-22.31%
EC			
Gene	Frequency	Median MAF	MAF IQR
ARID1A	42.66%	30.14%	19.42%-40.19%
CTNNB1	36.59%	33.22%	24.85%-41.43%
PIK3CA	34.15%	31.68%	22.02%-42.72%
PTEN	31.71%	23.76%	19.13%-34.42%
KRAS	24.39%	43.76%	38.10%-54.17%
TP53	18.66%	65.02%	29.57%-82.00%
OCCC			
Gene	Frequency	Median MAF	MAF IQR
ARID1A	45.16%	23.23%	17.92%-30.45%
PIK3CA	22.58%	32.60%	22.85%-36.37%
RAD50	16.13%	29.66%	20.37%-42.77%
MOC			
Gene	Frequency	Median MAF	MAF IQR
KRAS	71.05%	28.64%	22.67%-36.11%
TP53	55.26%	36.01%	28.65%-38.96%
ARID1A	21.05%	25.83%	22.50%-39.80%

HGSOC

The clonal pathogenic somatic variant in the *TP53* gene was the main feature of HGSOC, all of which harboured a *TP53* variant with a MAF of 57% (39.26%-71.07%; Figure 4-5 and Table 6). Noteworthy, the possibility to perform targeted sequencing in HGSOC cases with matched bilateral biopsy or with a paired second biopsy at time of relapse (Table 7),

allowed the identification of the same *TP53* variant in all the biopsies from the same patient, confirming the previously reported clonal and etiopathogenetic nature of *TP53* in the HGSOC histotype[113].

BRCA1 (23%) was the second most frequently mutated gene in HGSOC (MAF of 67.2%, 62%-75.84%, Figure 4-5 and Table 6). *BRCA2* was instead altered only in 3 out of 39 cases (Figure 4). Considering the allele fractions, in ten out of 12 cases (83,3%) the *BRCA1* or *BRCA2* variants could be considered potentially germline, with MAF higher than 60%.

Both *TP53* and *BRCA1* variants were classified in the Cancer Genome Interpreter (CGI) as loss of function mutations, as expected given the tumour suppressor role of these two genes (Figure 5).

No other recurrent mutated genes were identified in HGSOC, corroborating the idea that also in stage I, as in stage III-IV, the genomic landscape of HGSOC is characterised by few recurrent altered genes and in general few SNVs.

Table 7: summary of the main common potentially clonal SNVs between bilateral HGSOC samples and primary and relapse HGSOC tumours. MAF, mutant allele frequency.

Sample	Type	Gene	Codon change	Amino acid change	MAF (%)
31	primary relapse 1 relapse 2	*TP53*	c.988del	p.Leu330PhefsTer15	68.75 77.20 85.96
131	left Right	*TP53*	c.949C>T	p.Gln317Ter	57.28 70.28
202	left right	*TP53*	c.844C>T	p.Arg282Trp	61.79 55.22
123	left right	*TP53*	c.711G>C	p.Met237Ile	25.30 29.46
113	left right	*TP53*	c.376-1G>A	NA	88.19 77.42
183	left right	*TP53*	c.332T>A	p.Leu111Gln	54.71 26.03

LGSOC

The LGSOCs SNVs landscape was completely different from that of HGSOCs, confirming the different etiopathogenesis of the two serous subtypes.

LGSOC cases are indeed mainly characterized by mutations in the *BRAF* gene (45%; MAF 28.67%, Figure 4-5 and Table 6). *KRAS* resulted as the second most frequently mutated gene in LGSOC, with hotspot pathogenic mutations in codon 12 and 13 (40%; MAF 46.9%), leading to the production of constitutively active proteins. Analysis in matched samples from the same patient (Table 8) revealed the clonal origin of the BRAF V600E (case 161) and KRAS G12D (case 187) variants. Hotspot variants in *PIK3CA* were identified in a small proportion of cases (n = 5; 26%).

Table 8: summary of the main common potentially clonal SNVs between bilateral LGSOC samples and primary and relapse LGSOC tumours. MAF, mutant allele frequency.

Sample	Type	Gene	Codon change	Amino acid change	MAF (%)
161	left	*BRAF*	c.1799T>A	p.Val600Glu	16.93
	right				62.56
187	left	*KRAS*	c.35G>A	p.Gly12Asp	42.02
	right				53.40

Surprisingly, one LGSOC case (5%) harboured a POLE hotspot variant in the exonucleasic domain (POLE V411L). POLE hotspot variants (POLE V411L and P286R) were also identified in five out of 80 EC tumours (6%). The presence of these mutations caused an hypermutator phenotype, as suggested by the fact that the median number of SNVs in these samples is 53 (IQR 32.25-79.5) compared to all the other tumours without alteration in this locus (5; IQR 4-8) (Figure 6). Due to this intrinsic higher incidence of SNVs, these samples were excluded from the analysis of the SNVs repertoire of LGSOC and EC cases, to avoid overestimation of mutation frequency of in the other analysed genes.

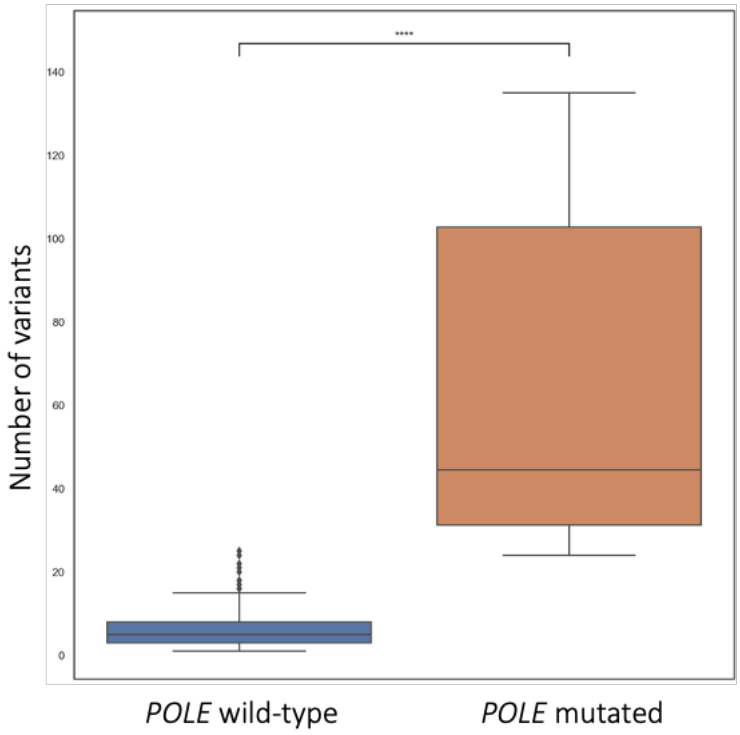


*Figure 6: Number of SNVs distribution between POLE mutated tumours and POLE wild-type tumours. Each group is represented by a boxplot. ****, Mann-Whitney test p-value < 0.001.*

EC

EC cases showed the greatest heterogeneous mutational landscape (Figure 4). The most recurrent gene was *ARID1A*, altered in 32 out of 75 cases (42.66%; MAF 30.14%; IQR 19.42%-40.19%; Figure 4-5 and Table 6). Endometrioid cases showed a plethora of many other altered genes, some of which are being investigated as therapeutic targets in several clinical trials, such as: *PIK3CA* in 34.15% of cases (MAF 31.68%), *PTEN* in 31.71% (MAF 23.76%), hotspot exon 3 of *CTNNB1* in 36.59% (MAF 33.32%), hotspot codons 12, 13, 59 and 61 of *KRAS* in 24.39% (MAF 43.76%), *PIK3R1* in 12.5% of cases (MAF 29.75%, IQR 19.87-33.68%), and AKT1 E17K in 7.5% (MAF 76.5%, IQR 73.24-87.51%). These data corroborate the heterogeneous nature of EC tumour subtype. As shown in Table 9, potentially clonal SNVs were confirmed in both samples of bilateral cases or in paired second biopsy taken at relapse even in EC cases.

Table 9: summary of the main common potentially clonal SNVs between bilateral EC samples and primary and relapse EC tumours. MAF, mutant allele frequency.

Sample	Type	Gene	Codon change	Amino acid change	MAF (%)
188	primary	*ARID1A*	c.3454_3484del	p.Ser1152HisfsTer18	16.49
	relapse 1				51.46
188	primary	*PIK3CA*	c.1633G>A	p.Glu545Lys	37.97
	relapse 1				42.68,
97	primary	*PTEN*	c.389G>C	p.Arg130Pro	11.10
	relapse 1				51.82
188	primary	*KRAS*	c.34G>T	p.Gly12Cys	60.54
	relapse 1				54.23

Following the molecular classification proposed for endometrial carcinoma[76], 14 out of 75 ECs (excluding the mentioned five samples bearing a mutation in a hotspot locus of *POLE*) resulted mutated in *TP53* with a MAF of 65.02%. Interestingly, the presence of *TP53* variants was mutually exclusive with SNVs in other recurrent EC altered genes, *ARID1A*, *CTNNB1*, *PIK3CA* and *PTEN* (pairwise Fisher's Test $p<0.05$; Table 10). On the contrary, there is a significant association between mutations in *ARID1A* and *PIK3CA* (18/75 cases harboured both alterations, pairwise Fisher's Test $p = 0.007$), supporting the already reported tumorigenic role of the coexistence of these molecular events [114,115]. Thirteen out of 75 cases harboured at least one pathogenic variant in one of the four main genes involved in MMR pathway (*MLH1, MSH2, MSH6, PMS2*). Unfortunately, it was possible to confirm the MSI phenotype through MSI analysis only in four of these cases for which matched blood samples were available. The absence of immunohistochemistry and DNA methylation analyses prevented the accurate evaluation of the MSI status of all the other endometrioid samples. The samples altered in MSI did not exhibit a higher mutational rate, but this phenotype could be masked by the fact that MMR deficiency cause instability especially in repetitive sequences, but the targeted panel exploited in this work covered only exonic regions of 139 genes.

Table 10: Summary of the mutually exclusive pair of genes EC cases. Ratio indicates the number of samples carrying a variant for both genes listed in the "Gene pair" column. p-value, result from the pairwise Fisher's test.

Gene pair	Event	Ratio	p-value
CTNNB1, TP53	Mutually Exclusive	0/43	0.0005
ARID1A, TP53	Mutually Exclusive	2/41	0.008
PIK3CA, TP53	Mutually Exclusive	1/37	0.012
PTEN, TP53	Mutually Exclusive	1/37	0.024

OCCC

OCCC were prevalently mutated in *ARID1A* gene (45.16% of cases, MAF 23.23%, IQR 17.92%-30.45%; Figure 4-5 and Table 6). *PIK3CA* was less frequently altered in OCCC compared to EC (22.58%; MAF 32.6, Figure 4-5 and Table 6). The last recurrent altered gene (at least 15% of cases) was *RAD50* (16.13%; Figure 4-5 and Table 6). Hotspot codon 12 of *KRAS* was altered in two samples (6.45%) as well as *TP53*; *PTEN* in 9.7% of cases. Noteworthy, for three patients with two available tumour samples, the same *ARID1A* variant was detected in bilateral tumours obtained at the first diagnosis and in paired tumour samples collected at the relapse. In the last bilateral tumour, a variant of *RAD50* resulted potentially clonal being detected in both biopsies of the patient (Table 11).

Regarding the MSI status, four out of 29 cases (13.8%) harboured at least one pathogenic mutation in one of the four main MMR genes; MSI analysis on the three eligible samples available confirmed MMR deficiency. As for the EC cases, the MSI analysis was performed only on a small subgroup of cases for which matched blood samples were available.

As for ECs, also in OCCCs mutations in the SWI/SNF pathway component ARID1A and in the PI3K/AKT pathway seemed to correlate, but did not reach a statistical significance, probably for the paucity of cases. In particular, six out of seven *PI3KCA* altered cases harboured also *ARID1A* mutations, moreover, two OCCC cases altered in *PIK3R1* were also mutated in *ARID1A*, suggesting the strong associations between the two altered pathways.

Table 11: summary of the main common potentially clonal SNVs between bilateral OCCC samples and primary and relapse OCCC tumours. MAF, mutant allele frequency.

Sample	Type	Gene	Codon change	Amino acid change	MAF (%)
115	primary	*ARID1A*	c.5203dup	p.Glu1735GlyfsTer4	67.90
	relapse 1				18.43
23	left	*ARID1A*	c.5548del	p.Asp1850ThrfsTer33	12.90
	right				20.19
23	left	*ARID1A*	c.6420del	p.Phe2141SerfsTer59	17.39
	right				23.09
23	left	*RAD50*	c.2167dupG	p.E723Gfs*5	13.01
	right				20.37

MOC

MOC were found mainly mutated in hotspot codons 12 and 61 of *KRAS* (71.05% of cases; MAF 28.64%; Figure 4-5 and Table 6) and *TP53* gene (55.26% of cases; MAF 36.01%, Figure 4-5 and Table 6). Interestingly, the median *TP53* MAF in MOC cases is significantly lower than that of HGSOC cases (36.01 vs 56.04, respectively; Mann-Whitney test, p = 0.01), suggesting the subclonal nature of this mutation in MOC cases, and corroborating the different role of this molecular event in the genesis of these two distinct histological subtypes. Other less recurrently mutated genes were: *ARID1A* (21.05%), *PTEN* (13.16%), *PIK3CA* and *BRAF* (10.53%). Variants located on the hotspot exon 3 of *CTNNB1* were identified in two cases (2.63%).

Lastly, also in MOC cases, as for ECs, variants in *TP53* were mutually exclusive with alterations in *PIK3CA* and *PTEN* (Table 12).

Table 12: Summary of the mutually exclusive pair of genes MOC cases. Ratio indicates the number of samples carrying a variant for both genes listed in the "Gene pair" column. p-value, result from the pairwise Fisher's test.

Gene pair	Event	Ratio	p-value
PIK3CA, TP53	Mutually Exclusive	0/26	0.012
PTEN, TP53	Mutually Exclusive	0/25	0.032

Whole genome SCNAs analysis shows the existence of different recurrent SCNAs and distinct SCNAs distribution across the five histotypes

The genome of EOCs is mainly characterized by structural variants, as already reported for the advanced stages[85]. Thus, the next step in the genomic characterization of Stage I EOC was the analysis of SCNAs across the whole genome. To reach this goal, a low-coverage whole genome sequencing (called shallow Whole Genome Sequencing, sWGS) was performed on all 205 cases (218 biopsies, including bilateral samples and paired relapsed tumours), as described in Materials and Methods. The somatic variants data were exploited to accurately estimate tumour purity and ploidy of each case basing on potential clonal mutations and then the sWGS copy number data were adjusted accordingly (see Materials and Methods). GISTIC algorithm was applied to observe the SCNAs distribution across the five histological subtypes (Figure 7) and to determine the presence of recurrent SCNAs; strongly recurrent events were considered those found in more than 50% of cases. Table 13 reports SCNA events spanning less than 25% of the chromosome arm, defined as “focal” and those involving more than 25% of the chromosome arm but less than its entirety, called “broad”. Table 14 showed all the SCNAs that affected the entire chromosome arm, thus “arm-level”.

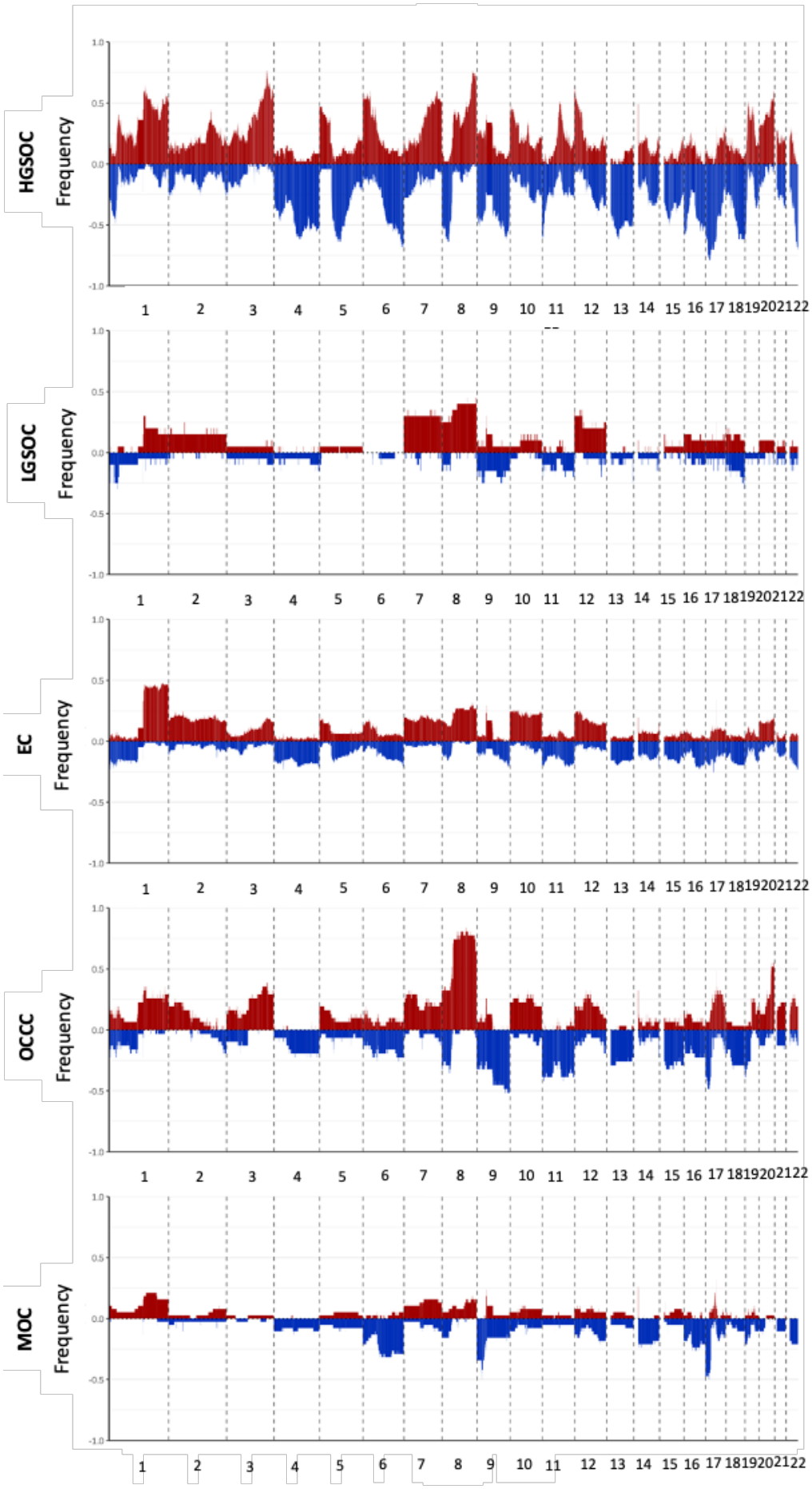

HGSOC
LGSOC
EC
OCCC
MOC
Frequency
1.0
0.5
0.0
-0.5
-1.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

Figure 7: SCNAs distribution across Stage 1 EOC histotypes. Only alterations with an absolute value of log2 ratio higher than 0.3 are reported. Blue indicates losses, while red indicates gains.

Table 13: List of focal and broad SCNA events (peaks) identified by the GISTIC algorithm in stage I EOC histotypes with q-value < 0.05. Amp, Amplification; Del; Deletion. Genomic coordinates refer to the hg38 genome assembly.

HGSOC							
Chromosome	Cytoband	Start	End	Length (Mb)	Type	Type	Frequency
chr1	1p36.11	6610001	33310000	26.70	Del	focal	53.33
chr1	1q21.1	121570002	145650000	24.08	Del	focal	24.44
chr2	2q37.3	241850002	242193529	0.34	Del	focal	40.00
chr4	4q25	73740001	142940000	69.20	Del	broad	64.44
chr4	4q35.2	148070001	190214555	42.14	Del	focal	55.56
chr5	5q11.2	45690001	108730000	63.04	Del	broad	53.33
chr5	5q14.3	61540001	106820000	45.28	Del	focal	66.67
chr6	6q26	159390001	170805979	11.42	Del	focal	68.89
chr8	8q24.21	127220001	129010000	1.79	Amp	focal	82.22
chr9	9p24.2	1	31380000	31.38	Del	broad	46.67
chr9	9q34.11	89710001	138394717	48.68	Del	broad	60.00
chr11	11p15.5	1	7520000	7.52	Del	focal	64.44
chr12	12q23.1	57190001	133275309	76.09	Del	broad	44.44
chr13	13q14.11	30660001	62010000	31.35	Del	broad	60.00
chr14	14q11.2	1	19720000	19.72	Amp	focal	48.89
chr15	15q14	22220001	60010000	37.79	Del	broad	46.67
chr16	16p13.3	1	15440000	15.44	Del	broad	64.44
chr16	16q24.3	48620001	90338345	41.72	Del	broad	55.56
chr18	18q23	36890001	80373285	43.48	Del	broad	62.22
chr19	19q12	29660001	30010000	0.35	Amp	focal	44.44
chr19	19p13.3	1	6590000	6.59	Del	broad	62.22
chr22	22q13.33	46730001	50818468	4.09	Del	focal	68.89
LGSOC							
Chromosome	Cytoband	Start	End	Length (Mb)	Type	Type	Frequency
chr1	1p35.1	23780001	42390000	18.61	Del	focal	30
chr1	1q21.1	121570001	145650000	24.08	Del	focal	20
chr9	9p11.1	39560001	42700000	3.14	Amp	focal	35
chr13	13q34	111340002	112700000	1.36	Del	focal	20
chr15	15q14	34350001	34760000	0.41	Del	focal	5
chr16	16q22.1	66580001	90338345	23.76	Del	broad	15
chr17	17q21.31	46100001	46720000	0.62	Amp	focal	25

EC

Chromosome	Cytoband	Start	End	Length (Mb)	Type	Type	Frequency
chr1	1p36.32	3840001	5999999	2.16	Del	focal	23.46
chr1	1q21.1	121570002	145650000	24.08	Del	focal	14.81
chr2	2q33.1	202030001	203710000	1.68	Del	focal	14.81
chr2	2q37.3	241850001	242193529	0.34	Del	focal	19.75
chr4	4q13.2	69180001	69490000	0.31	Del	focal	20.99
chr5	5q11.2	50430001	133429999	83.00	Del	broad	17.28
chr11	11p15.4	1	32390000	32.39	Del	broad	19.75
chr14	14q32.33	105530001	107043718	1.51	Del	focal	19.75
chr15	15q14	34350001	34760000	0.41	Del	focal	24.69
chr17	17q21.31	46130001	46720000	0.59	Amp	focal	34.57
chr17	17q12	32490001	34930000	2.44	Del	focal	24.69
chr22	22q13.32	43170001	50818468	7.65	Del	focal	25.93

OCCC

Chromosome	Cytoband	Start	End	Length (Mb)	Type	Type	Frequency
chr1	1q21.1	121570002	145689999	24.12	Del	focal	23.33
chr2	2q37.3	207760001	242193529	34.43	Del	focal	23.33
chr6	6p11.2	57220002	61679999	4.46	Del	focal	23.33
chr7	7q11.23	76370001	77320000	0.95	Del	focal	16.67
chr8	8p11.21	41720001	46570000	4.85	Amp	focal	53.33
chr9	9q34.2	66810001	138394717	71.58	Del	broad	56.67
chr11	11q22.1	79440001	118230000	38.79	Del	broad	46.67
chr14	14q11.2	1	19720000	19.72	Amp	focal	30.00
chr14	14q32.33	105530001	107043718	1.51	Del	focal	26.67
chr15	15q11.2	23680001	24680000	1.00	Del	focal	33.33
chr15	15q14	34350001	34760000	0.41	Del	focal	40.00
chr17	17q21.31	46100001	46720000	0.62	Amp	focal	53.33
chr17	17q23.2	61370001	62620000	1.25	Amp	focal	36.67
chr19	19p13.3	1	13100000	13.10	Del	broad	46.67
chr20	20q13.2	53480001	54010000	0.53	Amp	focal	60.00

MOC

Chromosome	Cytoband	Start	End	Length (Mb)	Type	Type	Frequency
chr1	1q21.1	121570002	145680000	24.11	Del	focal	13.51
chr9	9p21.3	21510001	22449999	0.94	Del	focal	48.65
chr15	15q14	34350001	34760000	0.41	Del	focal	13.51
chr17	17q12	39560001	39820000	0.26	Amp	focal	21.62

Table 14: List of arm levels SCNA events identified by the GISTIC algorithm in stage I EOC samples with q-value < 0.05. Amp, Amplification; Del; Deletion.

HGSOC			
Arm	Kind	Frequency	q-value
1q	Amp	35%	0.006
4q	Del	48%	0.0006
6q	Del	41%	0.008
8p	Del	48%	0.0009
9p	Del	48%	0.0004
9q	Del	39%	0.011
13q	Del	48%	0.004
17p	Del	72%	< 0.0001
17q	Del	57%	< 0.0001
20q	Amp	48%	0.006
22q	Del	41%	0.008
LGSOC			
Arm	Kind	Frequency	q-value
7p	Amp	29%	0.004
7q	Amp	29%	0.002
8p	Amp	24%	0.02
8q	Amp	38%	< 0.0001
12p	Amp	29%	0.004
EC			
Arm	Kind	Frequency	q-value
1q	Amp	37%	< 0.0001
2p	Amp	20%	0.01
2q	Amp	18%	0.02
8p	Amp	17%	0.04
8q	Amp	22%	0.002
10p	Amp	23%	< 0.0001
10q	Amp	20%	0.01
12p	Amp	23%	< 0.0001
12q	Amp	18%	0.02
17p	Del	21%	0.03
20q	Amp	18%	0.03
22q	Del	20%	0.03

OCCC			
Arm	Kind	Frequency	q-value
8p	Amp	42%	0.0005
8q	Amp	71%	< 0.0001
9p	Del	35%	0.05
9q	Del	45%	0.0023
11p	Del	45%	0.0023
11q	Del	35%	0.002
17p	Del	45%	0.0002
20q	Amp	35%	0.008
MOC			
Arm	Kind	Frequency	q-value
1q	Amp	24%	0.02
6q	Del	29%	< 0.0001
9p	Del	34%	< 0.0001
14q	Del	21%	0.01
16p	Del	18%	0.03
16q	Del	21%	0.01
17p	Del	47%	< 0.0001
17q	Del	21%	0.01
19p	Del	18%	0.026
22q	Del	21%	0.01

HGSOC showed the highest level of genomic instability, with many SCNAs prevalently involving less than an entire chromosome arm across the whole genome. As depicted in Figure 7, all chromosomes are affected by SCNAs, especially at focal events. Moreover, multiple regions were recurrently altered in more than 50% of cases (Table 13-14). HGSOC was the only histotype with this SCNAs pattern with many strongly recurrent altered regions. Among the arm-level SCNAs in HGSOC, two highly recurrent events involved the entire chromosome arm: the deletion of chromosome 17p, where the *TP53* gene is located, observed in in 72% of cases and the deletion of chromosome 17q, where the *BRCA1* gene is located, found in 57% of cases (Table 14), while *BRCA2* gene was involved in a broad deletion of chromosome 13q in about 60% of HGSOCs (Table 13). Among the main focal and broad SCNAs amplifications, there is the amplification of chromosome 8q24.21, containing the *MYC* oncogene, detected in 82% of HGSOCs; this

SCNA was already reported as the one of the most recurrently altered regions in HGSOCs [116], along with the amplification of chromosome 3q26.2, involving *MECOM* gene, but this second alteration was not identified by GISTIC in stage I HGSOCs samples, suggesting that it could be associated with advanced disease stages. Interestingly, *CCNE1*, located at chromosome 19q12, was focally amplified in 44% of cases [117–119]. No other focal or broad recurrent amplifications were detected. Among the recurrent deletions there were several regions already reported to be frequently deleted also in other cancer types, such as multiple deletions of chromosome arms 4q and 5q [120,121]. Several putative tumour suppressor genes were located in recurrent deleted regions, such as the SWI/SNF chromatin remodelling complex member *ARID1A* in the focal deletion of 1p36.11 in 53% of HGSOCs[122]; the *STK11* gene in the broad deletion of 19p13.3 in 62% of cases[123], the *CDKN2A* gene in the broad deletion of chromosome 9p24.2 in 47% of cases [117,124], and *ZNF516*, already reported as chromosome instability suppressors, in the deletion of 18q23 in 62% of HGSOCs[125].

Regarding the other histotypes, recurrent SCNAs with a frequency higher than 50% were almost absent (Figure 7, Table 13-14). This is not ascribable to an insufficient number of cases in each subtype to define recurrent SCNAs, as no recurrent SCNAs were identified among endometrioid samples, the most abundant group of tumours in the cohort. Rather, it seemed to reflect tumour behaviour, for example the fact that ECs showed the highest variability also in terms of SNVs. OCCC was the only histotype with a limited number of recurrent SCNAs. Among them there is one arm-level event: the amplification of the entire chromosome 8q in 71% of tumours, not only of the focal region encompassing *MYC*, further supporting the different mechanism of instability (Table 14). Few strongly recurrent focal SCNAs are present, among these the amplification of 20q13.2 including the well-known *ZNF217* oncogene is observed in 62% of clear cells samples[81] (Table 13).

The MOC histotype presents only one strongly recurrent focal SCNA in approximately 53% of the samples, a 960 kbp deletion of 9p21.3 involving precisely the oncosuppressor gene *CDKN2A* (Table 13). Interestingly, the focal amplification of *ERBB2* on chromosome

17q12 was found in about 22% of MOC (Table 13), the presence of this alteration could represent a potential therapeutic target [126,127].

These findings indicate that higher levels of genomic instability cause higher numbers of recurring SCNAs, which may also drive further instability and shape tumour evolution, as suggested by the presence of SCNAs affecting oncogenes and tumour suppressors.

Identification of three distinct genomic instability patterns by sWGS analysis

Despite the heterogeneity among DNA regions involved in SCNAs between the five main histotypes, it is possible to distinguish three common different genomic instability patterns based on the size of SCNAs and the amount of genome involved in SCNAs (Figure 8). This last variable was expressed as the percentage of genome expressing copy number alterations (CNB, Copy Number Burden). The three patterns were intuitively defined as stable (S), unstable (U) and highly unstable (HU) indicating gradually higher number of SCNAs and CNB. Particularly, the S pattern is characterized by the absence of SCNAs at the resolution of sWGS, the U pattern by large SCNAs, especially arm-level, on generally few chromosomes and the HU pattern by a huge number of SCNAs, also focal, on about all chromosomes.

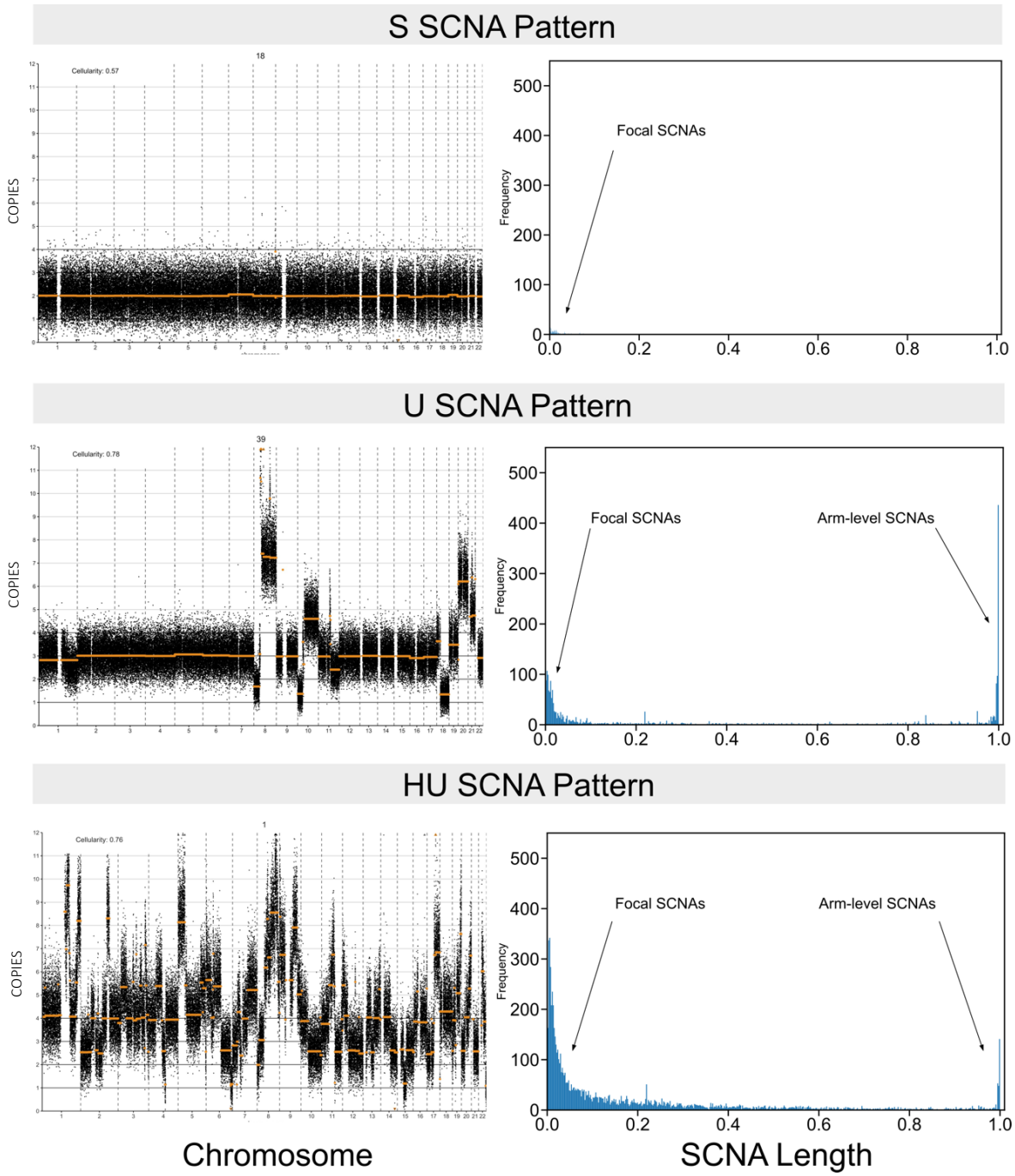


Figure 8: Raw sWGS plots indicates three distinct SCNA patterns in stage I EOCs. The left-hand panel reports log2ratios (y axis) for each 30kbp bin (black points) across the genome (x axis, from chromosome 1 to chromosome 22) for the three representative cases of stable (S), unstable (U), and highly unstable (HU) SCNA pattern (cases n. 18, 39 and 1, respectively). The right-hand panels show the frequency (y axis) of the length of DNA regions affected by SCNAs normalized per chromosome arm (ranging from 0, indicating focal SCNAs, to 1, entire chromosomal arms) for each SCNA pattern.

The genomic instability pattern was initially assigned manually to 130 cases for which it was clearly ascribable as S (n=20), U (n=70) and HU (n=40). On this manually selected set a classification tree algorithm was trained. Regression trees method refers to a decision tree algorithm that can be used for data classification. It can be considered as a supervised machine learning algorithm based on training and validation steps. In the training step the input data will produce a binary tree able to create a set of if-else conditions that allow an accurate prediction or classification of cases (in our case the SCNA pattern of each sample). In the validation step the trained algorithm is tested in terms of specificity and sensitivity. The classification tree algorithm was trained on the 130 cases for which the SCNA pattern was manually assigned. Among a set of 10 different genomic variables, the algorithm identified SCNA lengths normalized by chromosome arm and the CNB as the two continuous variables derived from sWGS experiments (see Materials and Methods) as the two variables able to classify samples with the lowest error rate. The leaf nodes of the binary tree in Figure 9 represent the cut-offs of the two variables identified by the algorithm to create the if-else conditions. The distribution of the values of the SCNA lengths normalized by chromosome arm and the CNB is reported in Figure 10. In the validation step, the classification tree was tested on the remaining 75 cases to automatically predict the genomic instability pattern of each of them.

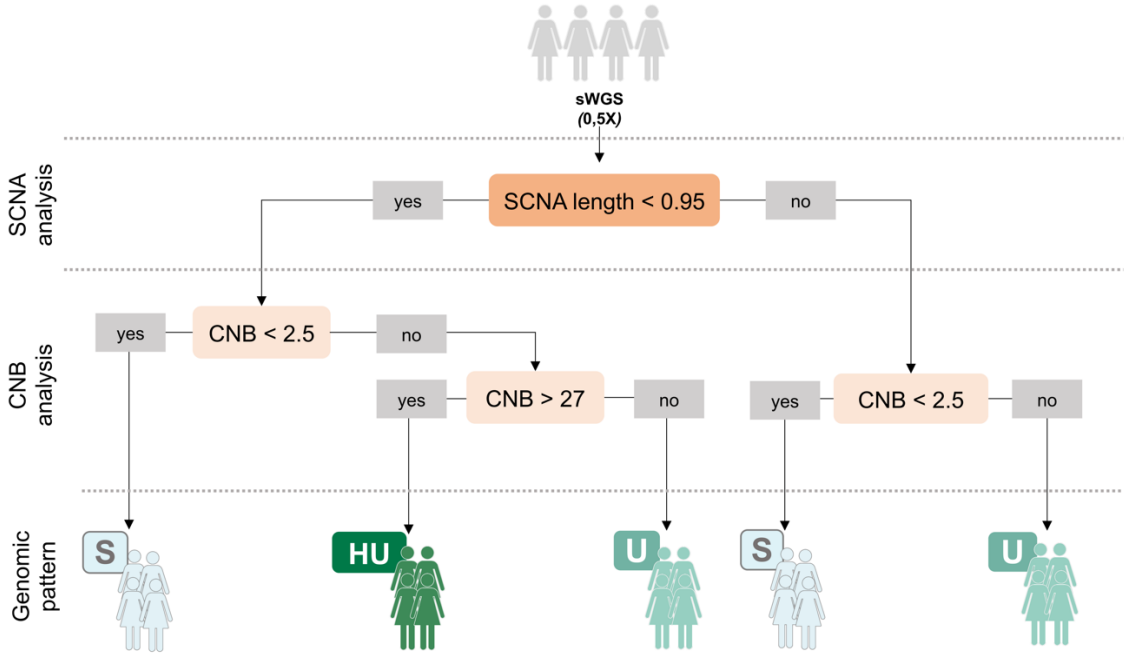


Figure 9: Decision binary tree to classify Stage I EOC on the bases of SCNAs length and copy number burden (CNB) derived from sWGS experiment.

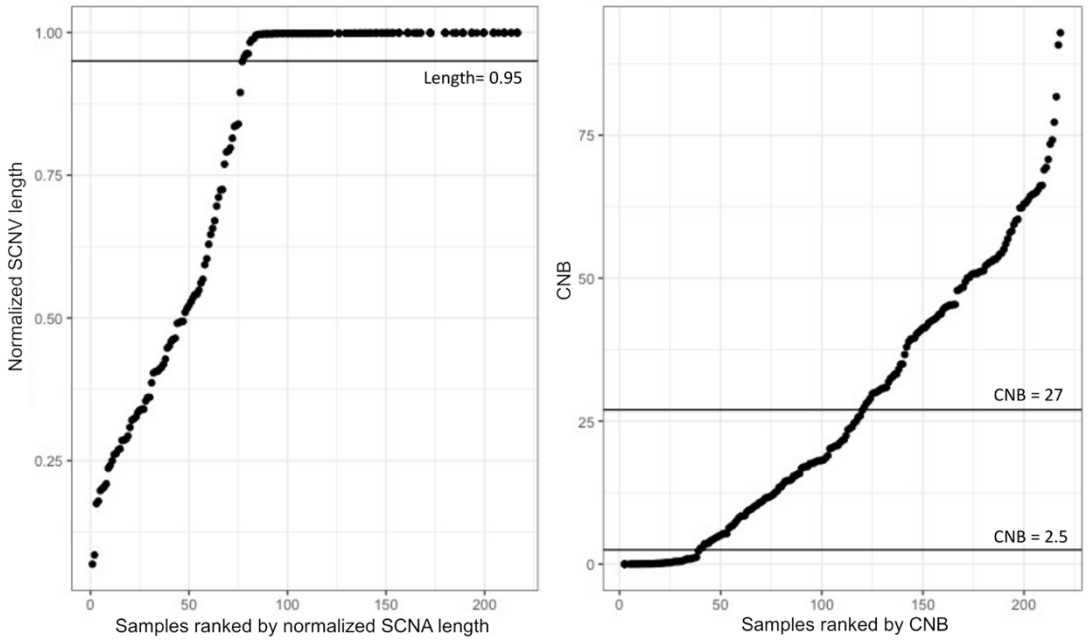


Figure 10. Distributions of the two decision tree variables, segment length (left hand side) and CNB (right hand side), in the sample cohort. Samples are ranked from the lowest values to the highest. The horizontal lines indicate the cut-offs used by the selection algorithm. These cut-offs

were directly estimated by the algorithm on the training set of 130 manually curated samples to construct a decision tree able to predict the SCNA pattern with the lowest error rate. The class of each sample is assigned based on the values of the two variables.

Finally, all cases were revised after the automatic classification and six of them with borderline values between HU and U patterns were manually assigned.

The S pattern had a median CNB of 0.13% (IQR 0.03-0.4%) and was characterized by a stable genome at the predefined sWGS resolution with the absence of any SCNA (Figure 11) at both focal and arm-level resolution (Figure 8). This pattern was identified in 40 cases (19% of tumours).

The U pattern had median CNB of 17.65% (IQR 10.72-33.2%) and was constituted by cases that exhibited large arm-level copy number rearrangements (Figure 8) with a median number of SCNA of 14 per sample (Figure 11). Approximately half of the cohort presented this pattern (104 cases, 51%).

The third SCNA pattern displayed a very high level of genomic instability and was called Highly Unstable (HU). The HU cases had a median CNB of 50.38% (IQR 37.9-59.61%) and were characterized by a huge number of copy number alterations (median 88 events per sample, IQR 68-120, Figure 11), at both focal- and arm-level resolution, involving substantially all chromosomes (Figure 8). This pattern belonged to 61 cases, most of which were HGSOC.

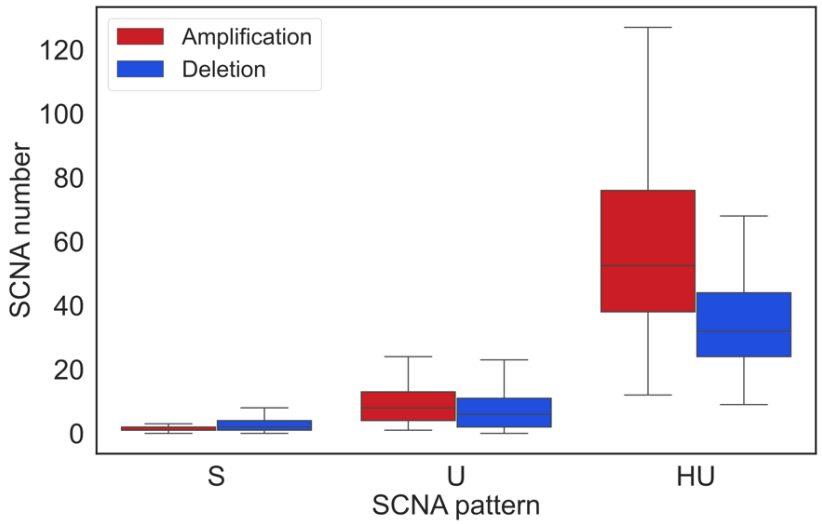


Figure 11: Number of SCNAs is different across the three patterns of genomic instability. Box plots of detected SCNA in each SCNA pattern. The red box plots show the distribution of amplifications, and blue box plots of deletions.

Since no SCNAs were identified in genome of tumours exhibiting the S pattern, HE-staining was exploited to assess the biopsy tumour cell content, and thus exclude that the S pattern was due to low tumour purity or sampling artifacts. Moreover, the sWGS was repeated on FFPE biopsies when possible, to confirm the results in samples with known tumour purity (at least 60%, as assessed by pathologists).

1) Evaluation of the S pattern through HE-staining on snap-frozen tumours:

The HE-staining was performed on 29 out of 40 S cases. Eleven cases were excluded due to the paucity of snap-frozen material. In 22 samples the tumour content was greater than 50%. Figure 12 shows two representative cases (namely, 132 and 135). Only in seven cases the tumour content was lower than 50%, but in five of these cases (namely, 111, 161_right, 185, 191 and 204) the S SCNA pattern was confirmed in the matched FFPE samples with assured tumour content of at least 60% (see next paragraph). Moreover, the S pattern was confirmed on matched FFPE tumour also in one of the cases excluded from HE-staining, (case 75). Therefore, the presence of sufficient tumour cells in S

samples was verified in 28 out of 40 S cases, certifying the existence of this pattern in stage I EOCs.

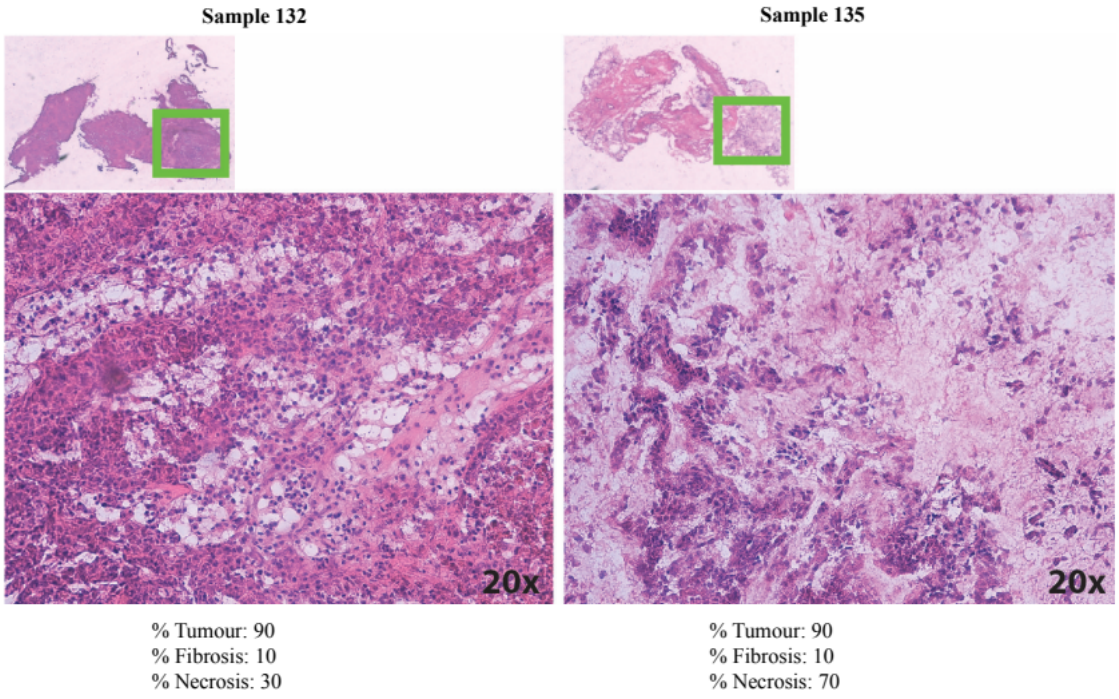


Figure 12. Haematoxylin-Eosin (HE) staining of two representative S samples (132 and 135). The upper panel provides an overview of the sample. The green square indicates the area of the sample then magnified (bottom panel) for each sample. The number in the lower right corner shows the magnification factor.

2) Comparison of fresh frozen and FFPE samples

Matched FFPE samples from 20 cases were collected and analysed by sWGS. A correlation analysis between the FFPE and the snap-frozen samples using the 30kbp bin level data was performed to assure the suitability of the technique on both biological materials. The median Pearson correlation was 74% (IQR 55-59; Table 15), indicating a good reproducibility of the results. Interestingly, 14 of these cases had an S pattern and showed high degree of correlation between snap-frozen and FFPE samples, including the five cases with snap-frozen tumour content less than 50% and the sample without HE-staining described above.

All the subsequent analyses were executed on FFPE samples for those cases where both a snap-frozen and a FFPE sample were available.

Table 15: Percentage of Pearson correlation coefficient between matched snap-frozen and FFPE tumour samples. Snap-frozen tumour content assessed by HE-staining is reported; S Stable; U Unstable; HU Highly Unstable; NA Not Analysable; NP Never Performed.

Sample ID	SCNA pattern	Snap-frozen tumor cell content (%) by HE-staining	Correlation between snap-frozen and FFPE samples (%)
75	S	NA	77
97	U	90	74
103	S	50	71
111	S	25	76
112	S	50	42
113_left	HU	NP	82
115	U	90	68
116	S	50	80
117	S	70	79
129	S	90	52
132	S	90	53
150	S	90	80
158	S	90	84
161_left	S	100	40
161_right	S	5	45
180	U	NP	76
181	U	NP	87
183_left	HU	NP	69
185	S	35	75
191	S	40	55
204	S	35	73

Since the absence of SCNAs in S cases, an accurate prediction of ploidy by ACE was prevented, thus ploidy was measured by flow cytometry. The analysis was performed on 17 S samples with sufficient sample quality. All of them were diploid (median DNA-index 1.06, IQR 0.96-1.08), with the exception of a moderate hyperdiploid case, with a DNA-index of 1.44 (data not shown).

Figure 12 reports the distribution of SCNAs patterns across the histotypes, approximately half of stage I EOC cases belonged to the U SCNA pattern, which characterized the majority of LGSOC (63%), EC (53%), OCCC (59%) and MOC (74%) cases. Data revealed that 34 out of 39 HGSOC stage I EOCs (87%) had a HU pattern, with the remaining five HGSOC cases with a U SCNA pattern. The absence of S cases in the HGSOC histotype, as well as the prevalence of the HU SCNA pattern, strongly suggests that the genomic instability, well-established in late stages, is peculiar to this tumour subtype since its early stages. The high-level of genomic instability in the HGSOC cases, due to great prevalence of HU SCNA pattern, is also translated in a significantly higher CNB than the other histotypes (median CNB approximately 60%, Mann-Whitney test $p < 0.001$) (Figure 12), which showed a lower median CNB, less than 20% in all but one histotype (OCCC), and largely similar across distributions (Mann-Whitney test $p > 0.05$). The HU SCNAs pattern encompassed also about 19% and 35% of EC and OCCC cases, respectively, while only one MOC (2.6%) and one LGSOC (5%) belonged to this category. The S pattern represented approximately 30% of LGSOC and EC tumours and 24% of MOCs. Only two OCCC samples (7%) had an S pattern. Therefore, OCCC group was the second histotype almost characterized by genomic instability, particularly, the U SCNA pattern was prevalent in this tumour type. It is noteworthy that the same SCNA pattern is always shared by both tumour samples of bilateral cases as well as by primary and relapse tumours couples.

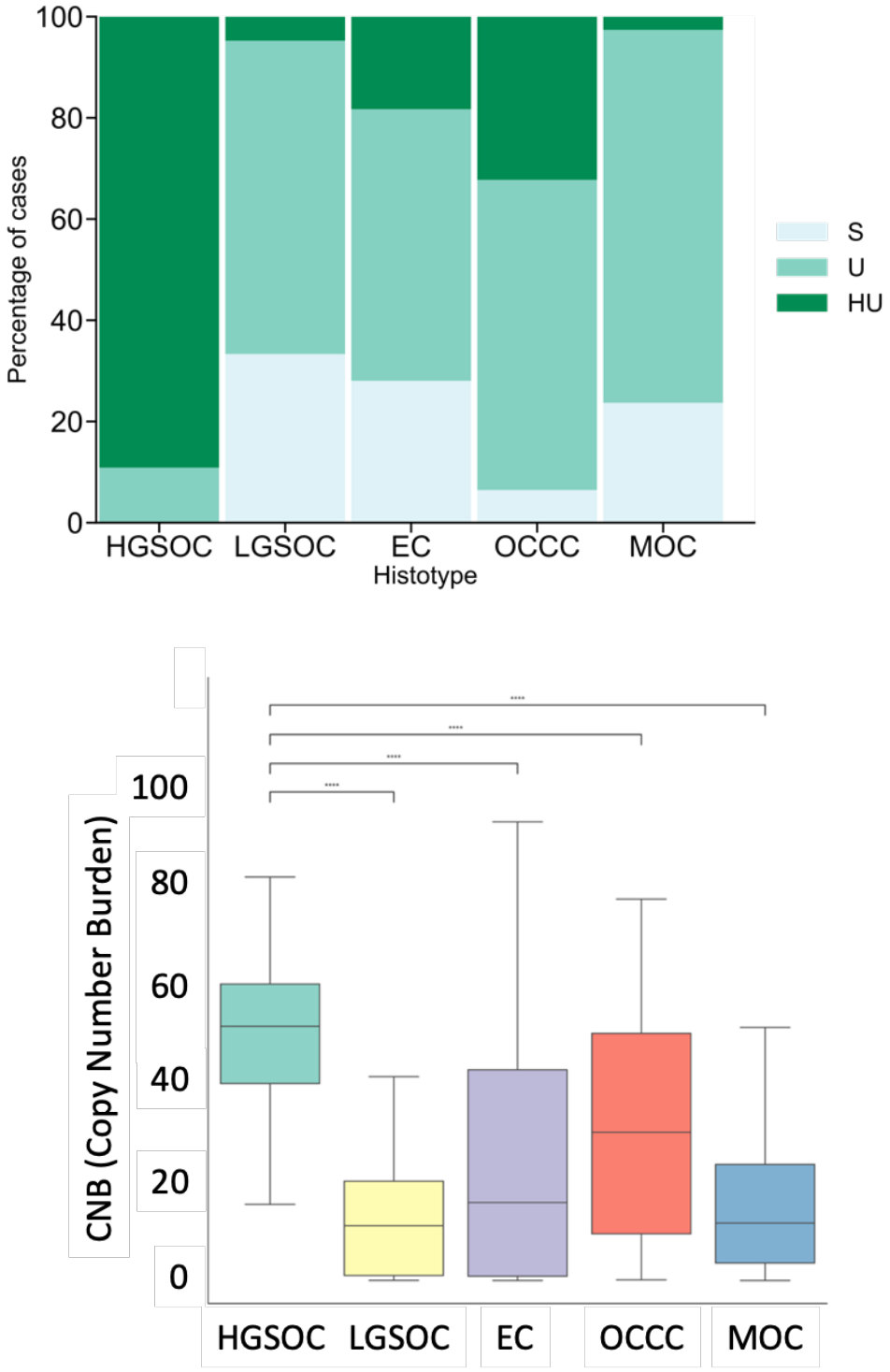


*Figure 12: Each histotype shows a different prevalence of genomic instability patterns and distinct distribution of CNB. Top plot: Distribution of SCNA patterns per Stage I EOC histotype. Bottom plot: Copy number burden (CNB) distribution across Stage 1 EOC histotypes. Each histotype is represented by a boxplot. ****, Mann-Whitney test p-value < 0.001.*

As HGSOC samples included most of the HU samples, the SCNAs distribution of HGSOC HU samples were compared with non-HGSOC HU samples testing for differences that could have been masked by the higher number of HGSOC cases in this genomic instability pattern. However, no differences in SCNA distribution were detected in HGSOC versus non-HGSOC HU cases (Figure 13), suggesting that this pattern is histotype-independent. Moreover, since HGSOC tumour is driven by the etiopathogenetic mutation of *TP53*, related itself to genomic instability, the presence of mutations in this gene that could explain why the same SCNA pattern was assessed in non-HGSOC HU cases. However, 14 out of 27 non-HGSOC HU samples had a *TP53* mutation, highlighting that the HU pattern is probably driven also by other etiopathogenetic mechanisms.

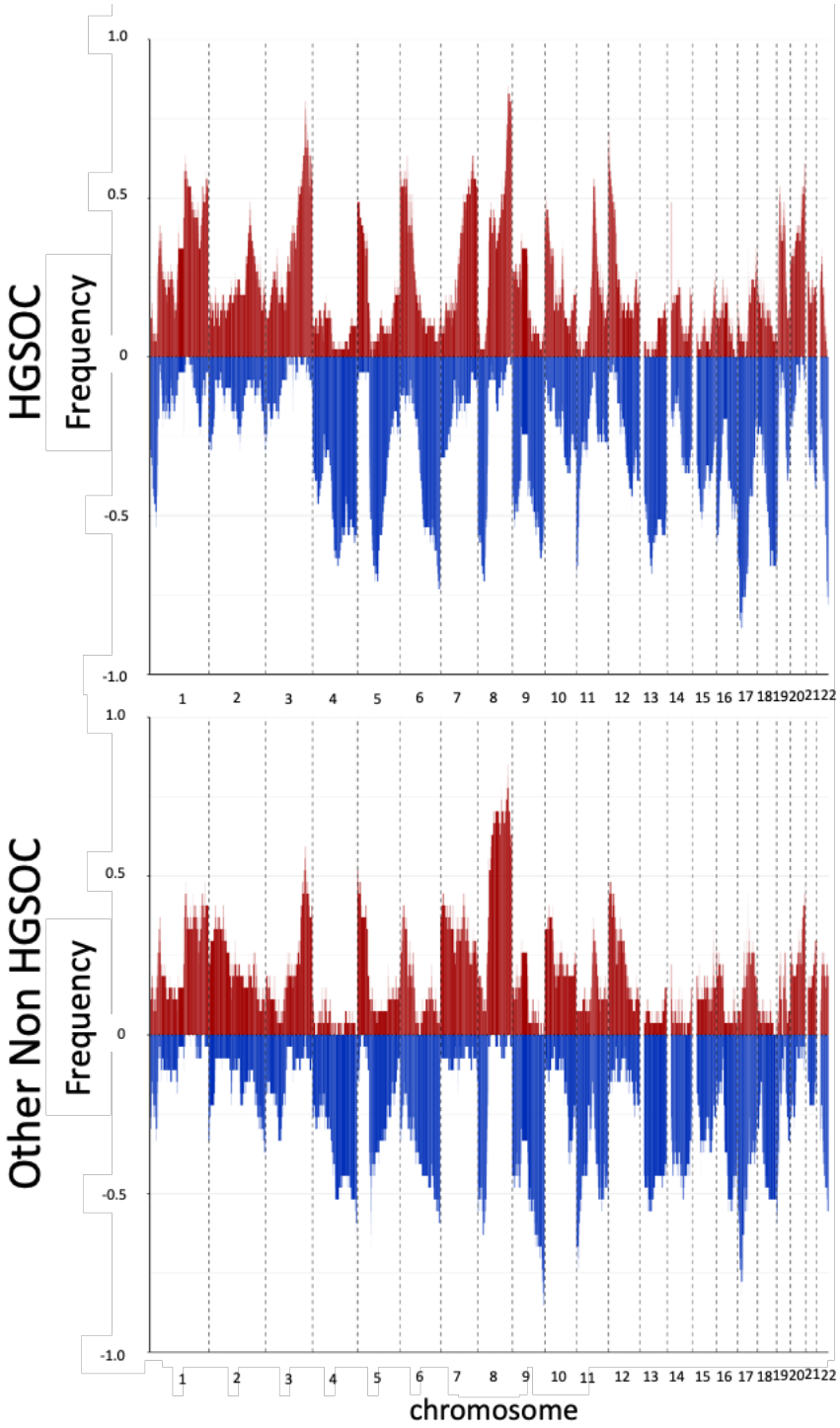


Figure 13. Frequency plot of SCNAs in highly unstable (HU) samples, either belonging to the HGSOC histotype, or to any other histotype. red, amplification; blue; deletion.

Five SCNAs signatures define the genomic instability features in U and HU patterns

Following previously published methods[70], sWGS data were used to infer the potential mutational processes underlying the three different SCNA patterns previously identified. Since the S cases exhibited little to no SCNAs, this analysis was restricted to the U and HU patterns. The genome-wide distribution of six CN features (the breakpoint count per 10MB, the copy-number of segments, the difference in CN between adjacent segments, the breakpoint count per chromosome arm, the lengths of oscillating CN segment chains and the size of segments) considered the distinctive marks of different processes causing the copy number alterations along the genome, were computed in each sample applying mixture modelling. A total of 39 mixture components was calculated. For each sample, the posterior probability of copy-number events arising from these components was computed and summed. Therefore, to identify SCNA signatures, non-negative matrix factorization (NMF) was used and, finally, to determine the optimal number of signatures to use, 1000 random permutations of the data were generated and used as a null measure. Therefore, the stability of cophenetic, dispersion and silhouette coefficients and the maximum sparsity achievable above this null model were measured and compared to those of observed data (Figure 14). The optimal number of signatures was set to five, as this is the optimal factorization rank able to explain the observed data but not the randomized ones.

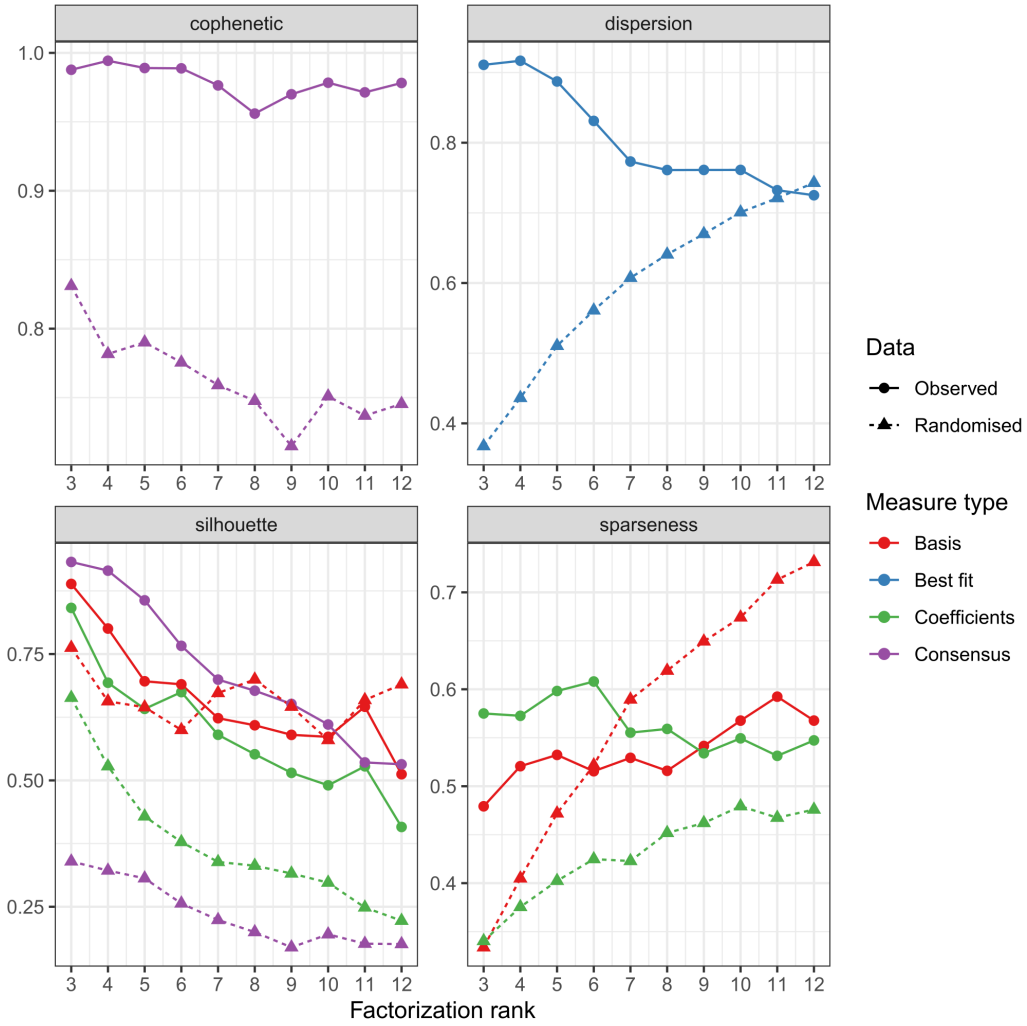
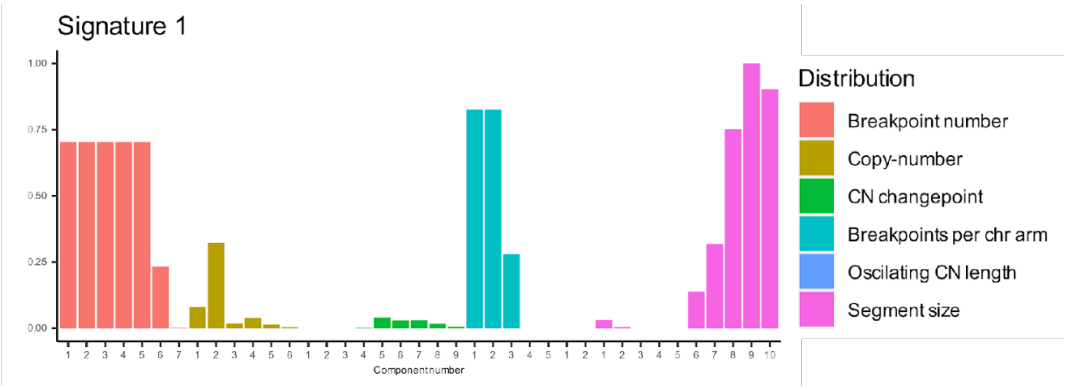
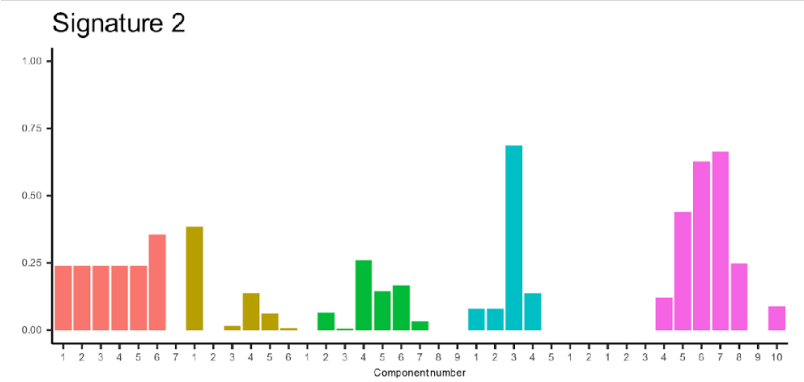
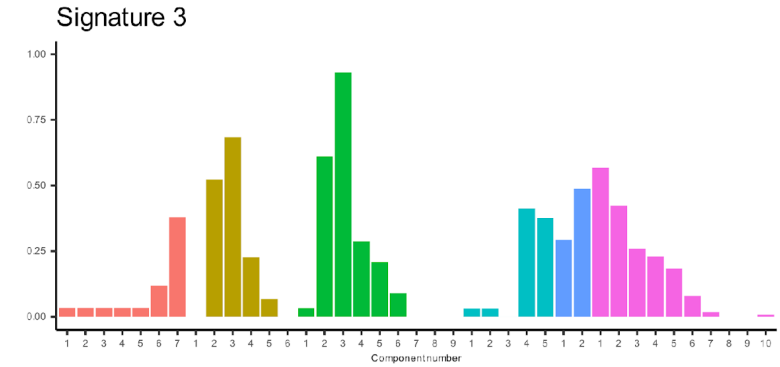


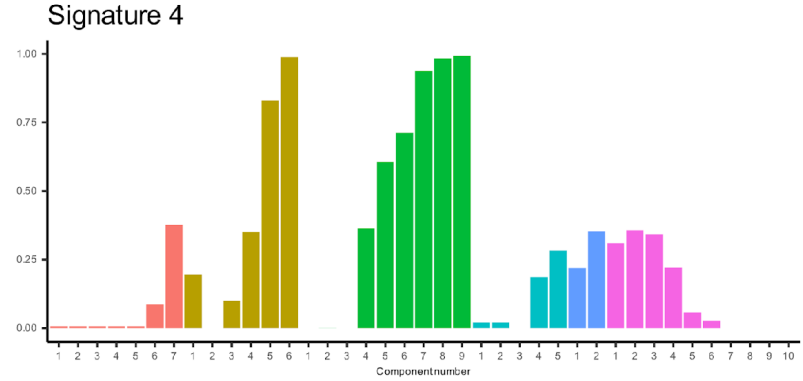
Figure 14. Selection of the optimal number of SCNA signatures. The circles indicate the data from the actual samples; the triangles, randomized data (1000 random permutations from the input data set) to test the features of the signatures against random noise. Basis refers to the matrix associating signatures to specific copy number features, coefficients to the matrix associating signatures to patients, and best fit to the run that showed the lowest objective score across 1000 runs.

The contribution or the weight of each component of the six SCNA features (breakpoint number, Copy-number, copy number changepoints, breakpoints per chromosome arm, oscillating copy number length and segment size) to each signature was evaluated and summarized in bars plot in Figure 15.


Signature 1
Distribution
Breakpoint number
Copy-number
CN changepoint
Breakpoints per chr arm
Oscilating CN length
Segment size
Componentnumber


Signature 2
Componentnumber


Signature 3
Componentnumber


Signature 4
Componentnumber

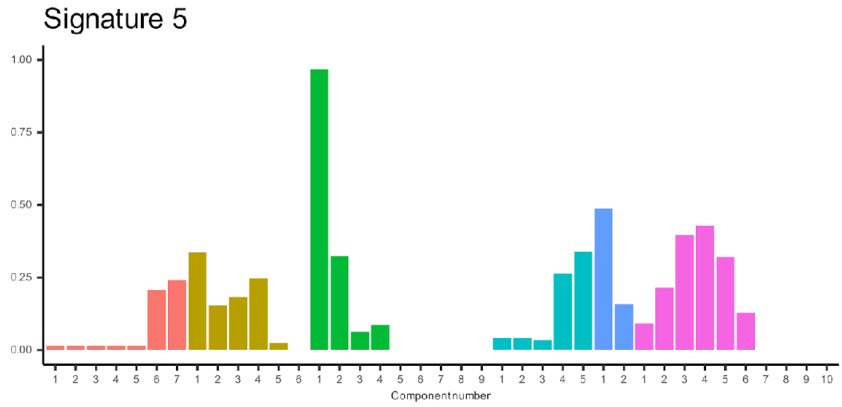


Figure 15. Features of mutational signatures. Coloured bars are grouped by colour, with each colour representing a category of alterations. Each signature is characterized by different weights of the SCNA components derived from the six copy number features listed in the legend (Breakpoint number, Copy-number, copy number change-points, breakpoints per chromosome arm, oscillating copy number length and segment size). The x-axis of each graph reports the SCNAs components used to compute the signature, grouped by category (i.e. breakpoint number, copy-number). Each category is indicated by one colour reported in the legend. The y-axis shows the weights, that is the contributions of a specific feature to the signature.

Signature 1 was mainly characterized by relatively low breakpoints per chromosome arm and large segment sizes. Signature 2 exhibited a higher (but still modest) incidence of breakpoints per chromosome arm and a higher incidence of segments with intermediate sizes. Signature 3 had a higher number of breakpoints, low copy number changepoints, the length of copy number segments was shorter than Signatures 1 and 2, and exhibited longer oscillating copy number segments. Signature 4 had a high number of breakpoints per 10Mbp, high copy number level and copy number changepoints, and generally shorter segment sizes. Lastly, signature 5 had a high number of breakpoints per 10Mb and per chromosome arm, but low copy number levels, copy number changepoints, and oscillating copy number length and on average short segment size.

Figure 16 shows the unsupervised distribution of U and HU cases among the identified SCNA signatures and the corresponding exposure, the measure indicating the prevalence of a signature, of each sample.

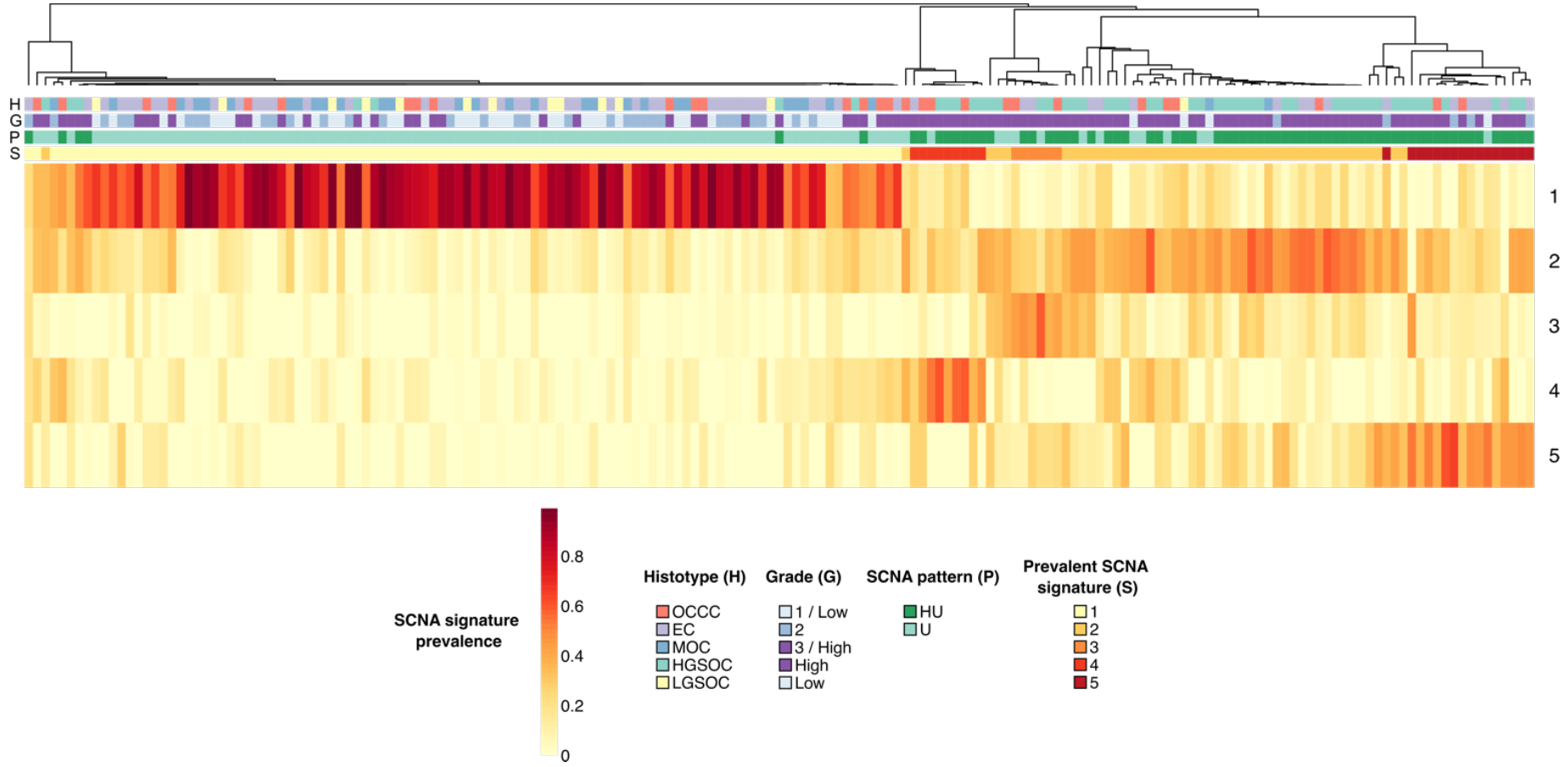

H
G
P
S
1
2
3
4
5
SCNA signature prevalence
0.8
0.6
0.4
0.2
0
Histotype (H)
OCCC
EC
MOC
HGSOC
LGSOC
Grade (G)
1 / Low
2
3 / High
High
Low
SCNA pattern (P)
HU
U
Prevalent SCNA signature (S)
1
2
3
4
5

Figure 16: Clustering of U and HU samples according to their SCNA signature. The rows indicate the five SCNA signatures identified from the data (Methods and Supplementary Methods), while columns show each individual sample. Colours show the prevalence (expressed on a yellow-red scale: the redder the colour, the higher the prevalence) of each signature in each sample. i) histotype (H): yellow, LGSOC; green, HGSOC; grey, EC; red, OCCC; blue, MOC; ii) grade (G): light-blue, one/Low; blue, two; purple; three/High; iii) SCNA pattern (P): light-green, U; green, HU. The prevalent signature (S) is indicated: yellow,1; light-orange, 2; orange, 3; red, 4; dark-red, 5. The values in the heatmap correspond to the consensus matrix obtained after obtaining copy number components, selecting the number of signatures, and applying non-negative matrix factorization (NMF). The clustering is computed using the consensus matrix itself as a similarity matrix and average linkage.

By this approach, it is possible to notice that each sample was exposed to multiple SCNAs signatures and was then associated with the most prevalent one (Materials and Methods). Cases belonging to the U pattern were mainly associated with SCNA signature 1, whilst cases belonging to HU were mainly exposed to signatures 2, 3, 4 and 5. Overall, the exposure is stronger in SCNA signature 1 than in other SCNA signatures (Mann-Whitney p-value versus all other signatures < 0.001). This suggests the prevalence of a common molecular mechanism driving genome instability for cases with prevalent SCNA signature 1, while the other cases are likely characterized by multiple and heterogeneous mechanisms.

To determine whether specific gene-level alterations were related to specific copy number signatures, it was investigated if the exposure (a measure of how strongly each sample was associated to a specific signature) for each signature was associated to the presence of mutations or SCNAs in specific genes. Figure 17 shows the result of the comparison of exposure levels in samples with or without altered genes for each signature (see Materials and Methods for a description of the approach used). With regards to signature 1, samples with alterations in *PTEN*, *PIK3CA*, *KRAS* and *CTNNB1* had significantly higher (q-value < 0.05) exposure than samples without alterations. Signature 2 was characterized by higher exposures in samples with *BRCA1*, *KDM5A*, *MYC* and *TP53* alterations. No altered genes with significantly different signature exposure were found

in Signature 3. Signature 4 had more exposure in *MAPK1*, *CCNE1* and *MYC*, and Signature 5 in *TP53*. The strongest differences were observed in Signatures 1 and 2.

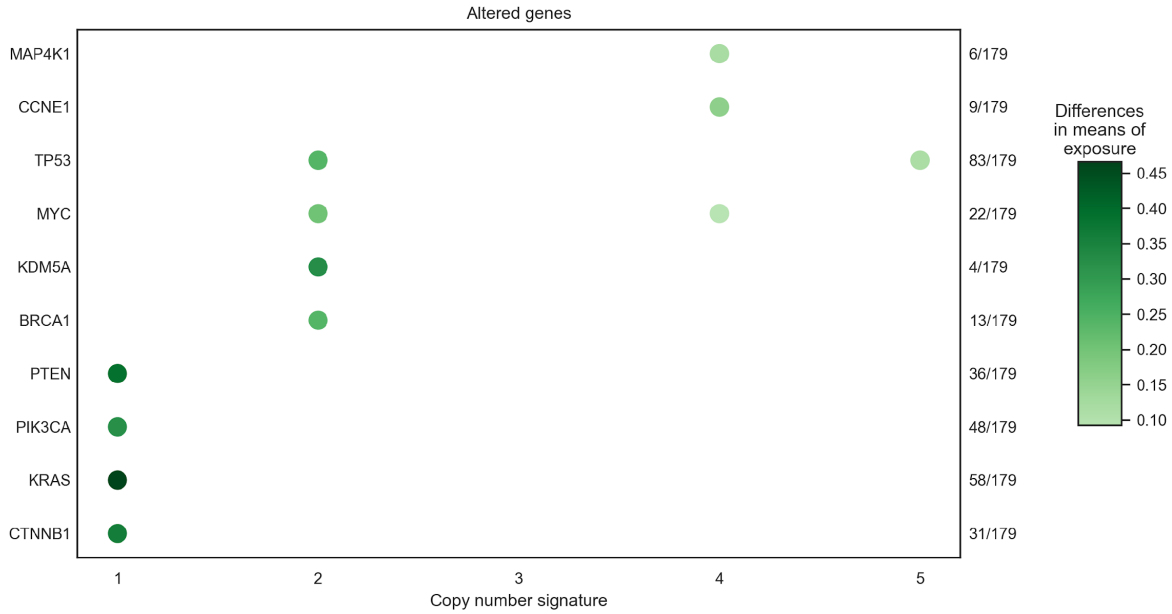


Figure 17. Differential exposure per signature on a set of altered genes. The colour indicates the median difference between the altered and non-altered groups for each gene (see Materials and Methods).

In brief, the major characteristics of each signature are:

Signature 1 showed few breakpoints per chromosome arm and large segment sizes (Figure 15). This result confirmed that in these cases the detected SCNAs were mainly arm-level events. The signature was prevalent in the majority of EC, MOC and LGSOC cases and about half of the OCCC cases. Interestingly, higher exposure to signature 1 was strictly associated with variants in *KRAS*, *PTEN*, *CTNNB1*, and *PIK3CA* (Figure 17). Mutations in these genes are known to cause chromosomal instability due to impaired mitotic chromosome segregation, genome doubling, telomere stability[128–131]. Moreover, signature 1 was weakly inversely correlated with age at diagnosis (Pearson's r=-0.17), suggesting that younger patients may exhibit higher prevalence of this signature. However, this signature is enriched in histotypes usually related to younger age at diagnosis[132], i.e. LGSOC (T-test, p-value = 0.02), EC (T-test, p-value <0.001) and MOC ((T-

test, p-value < 0.0001), thus it is not possible to establish if the inverse correlation with younger age at diagnosis is in fact a consequence of the histotypes prevalence of the cases affected by signature 1.

Signature 2 was characterized by a modest incidence of breakpoints per chromosome arm and intermediate segment size (Figure 15). This signature was prevalent in the majority of HGSOC cases (n = 23) and was enriched with variants in *BRCA1* along with *TP53* and amplifications in *MYC* and *KDM5A* (Figure 17). Moreover, amplification of *MYC* is associated to structural and numerical chromosomal aberrations, particularly in the formation of extrachromosomal elements, as well as errors in DNA replication, increasing replication origins firing, replication forks collapse and formation of Double Strand Breaks (DSBs) [133,134]. Amplifications in *KDM5A* have been involved in abnormal mitosis, disruption of the centrosome homeostasis and aneuploidies[135].

Signature 3 showed a high number of breakpoints per 10Mb and per chromosome arm, low copy number level and copy number changepoints, short segment sizes, and long oscillating copy number segments (Figure 15). It was prevalent in only six cases (2 OCCCs, 2 ECs and 2 HGSOCs) and the signature exposure was not related to any specific mutational or SCNA feature.

Signature 4 had a high number of breakpoints per 10Mb, high copy number level and copy number changepoints, and generally short segment sizes (Figure 15). Its exposure was positively correlated with age at diagnosis (Person's r= 0.25). Like signature 3, its prevalence was restricted to few cases (3 OCCCs, 1 EC and 4 HGSOCs), and enriched in amplifications in *CCNE1, MYC* and *MAP4K1* (Figure 17). *CCNE1* has been involved in precocious S-phase entry, centrosome amplifications and chromosome missegregation[134,136]. *MAP4K1* is involved in stress response, proliferation and apoptosis, its upregulation has been already related to MYC overexpression[137].

Signature 5 had a high number of breakpoints per 10Mb and per chromosome arm, but low copy number levels, copy number changepoints, and oscillating copy number length

and on average short segment size (Figure 14). It was associated with variants in the *TP53* gene (Figure 16), and it was prevalent in 15 cases (7 HGSOCs, 6 ECs and 2 OCCCs).

To better define the association between genomic instability characteristic of each signature and presence of variants in the identified genes, the prevalence of SNVs or SCNAs in signature associated genes was compared to those in the S cases. Each sample was associated with its most prevalent signature (Materials and Methods) and then prevalence of each altered gene was compared between the two groups (Table 16). Signature 1 had a larger proportion of *KRAS* mutated samples than S cases. Regarding the other altered genes associated with signature 1, they were almost equally distributed between signature 1 prevalent cases and S cases. Nevertheless, it was not possible to rule out that a combination between these altered genes and other processes not analysed in this study could be responsible for the SCNA pattern of signature 1. On the other hand, Signatures 2 and 5 showed larger proportions of *TP53* mutations compared to the S cases.

Table 16. Results of the analysis on comparison between altered genes in signatures 1, 2, 4 and 5 and the S samples. n.s., not significant.

Signature	Gene	Samples altered in signature	Samples altered in S group	Fisher test p-value
Signature 1	*CTNNB1*	19	11	n.s.
	KRAS	43	12	0.1
	PIK3CA	26	13	n.s.
	PTEN	21	14	n.s.
Signature 2	*BRCA1*	5	0	n.s.
	TP53	38	3	< 0.001
	KDM5A	4	0	0.06
	MYC	10	2	0.045
Signature 4	*CCNE1*	4	0	< 0.001
	MYC	4	2	0.01
	MAP4K1	2	0	0.02
Signature 5	*TP53*	11	3	< 0.001

Genomic instability correlates with different levels of global genome hypomethylation

Genomic instability is often correlated to epigenomic events, and a genome-wide hypomethylation is a frequent and crucial event in carcinogenesis and recently, it was demonstrated that in HGSOC, LINE1 retrotransposons were activated during progression from the p53 signature to STIC stages[92]. Thus, alteration of global methylation status and particularly deregulation of LINE1 is one of the major features that lead to genomic

instability[93]. The global methylation status, by means of the LINE1 promoter methylation, was evaluated to look for differences between the identified genomic patterns. Indeed, the LINE1 promoter methylation resulted significantly decreased in cases with HU pattern (Mann-Whitney test $p < 0.001$), compared both those with U and S patterns. This significant hypomethylation seems not to be related to *TP53* mutation status. Indeed, the TP53 mutated cases were similarly distributed between the three genomic patterns (Figure 18). Moreover, since the majority of cases with HU pattern belonged to HGSOC histotype, the analysis was repeated excluding this histotype. Even in this case the correlation between LINE1 hypomethylation and HU was significant ($p < 0.001$; Figure 19). The genome-wide hypomethylation of many HU cases could be one of the events that favours the genomic instability, nevertheless it could also be a consequence of some chromosomal rearrangements that cause change in the regulation of DNA methylation. The deregulation of LINE1 surely indicates that epigenetic alterations, such as DNA methylation, is strictly associated with the HU pattern, the most complex degree of genomic instability found in stage I EOC; further studies will be necessary to deepen this relation.

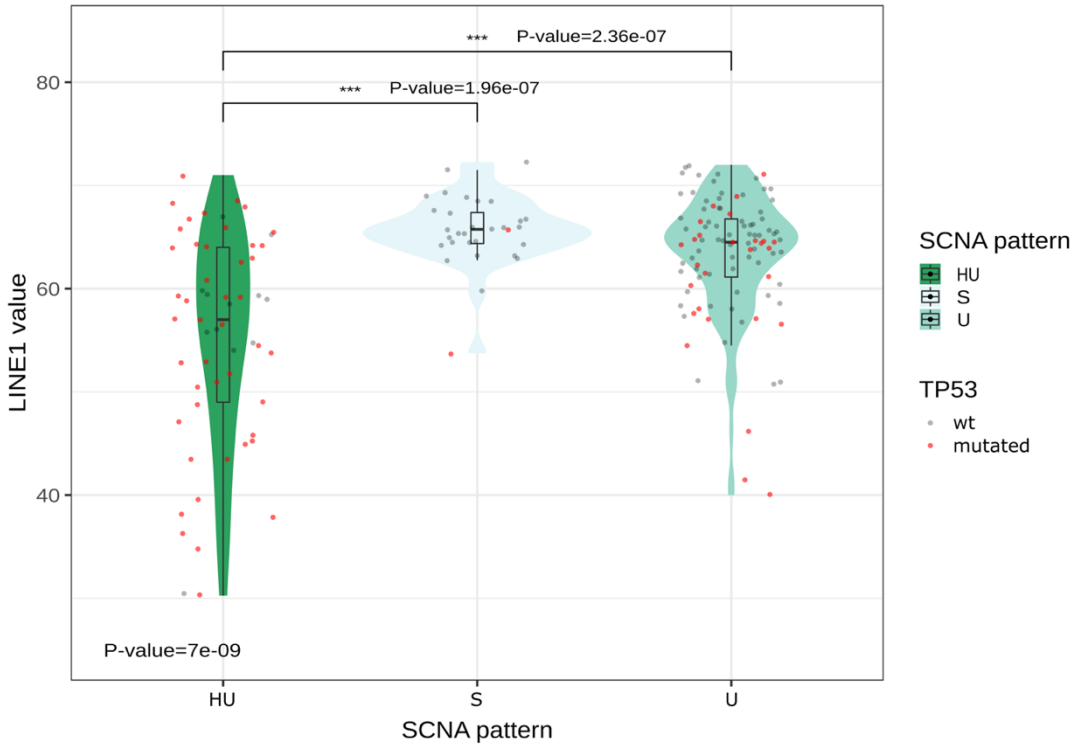


Figure 18. Violin plots of LINE1 methylation status in stable (S), unstable (U) and highly unstable (HU) sample groups. The Y axis refers to the percentage of LINE1 promoter methylation; the dots indicate individual samples; a red dot indicates a sample with a driver mutation in TP53. P-values refer to the Mann-Whitney test.

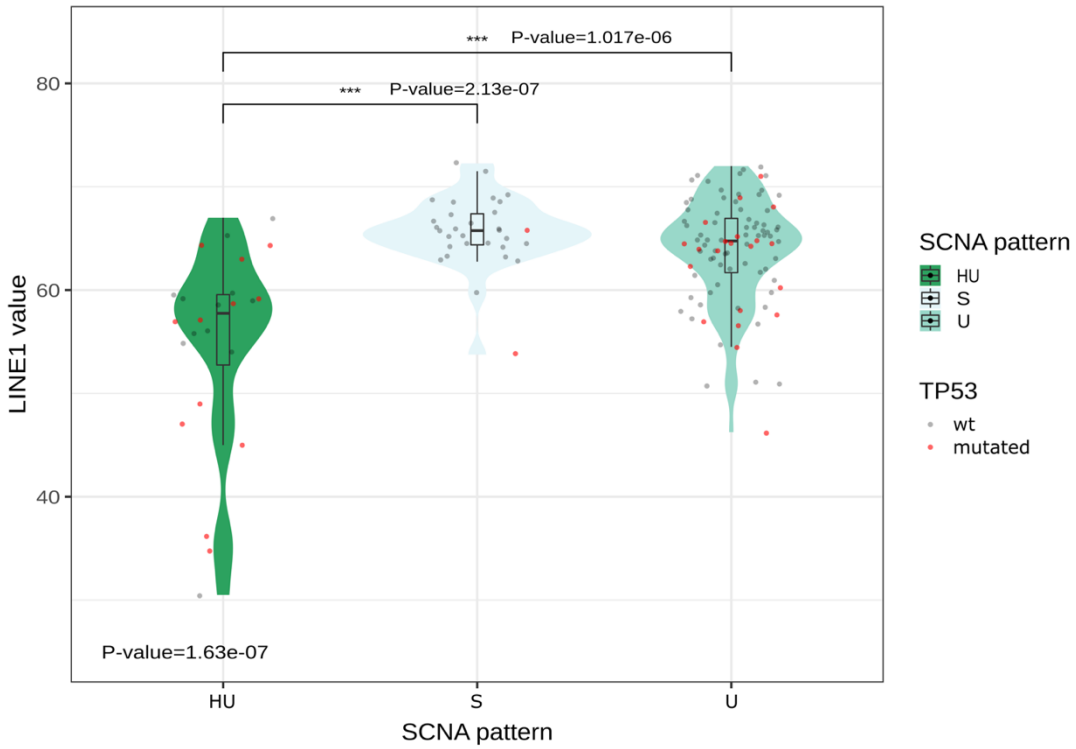


Figure 19. Violin plots of LINE1 methylation status in stable (S), unstable (U) and highly unstable (HU) sample groups excluding HGSOC samples. The Y axis refers to the percentage of LINE1 promoter methylation; the dots indicate individual samples; a red dot indicates a sample with a driver mutation in TP53. P-values refer to the Mann-Whitney test.

The SCNA genomic patterns (HU, U, S) are exploitable as prognostic factors for progression free survival and overall survival

Since one of the unmet clinical needs for stage I EOC is the implementation of prognostic markers able to better define the relapse risk of each patient; I tried to investigate if some of the genomic features that characterize tumour biopsy at diagnosis analysed in this study could be associated with clinical variables used for prognostic purposes.

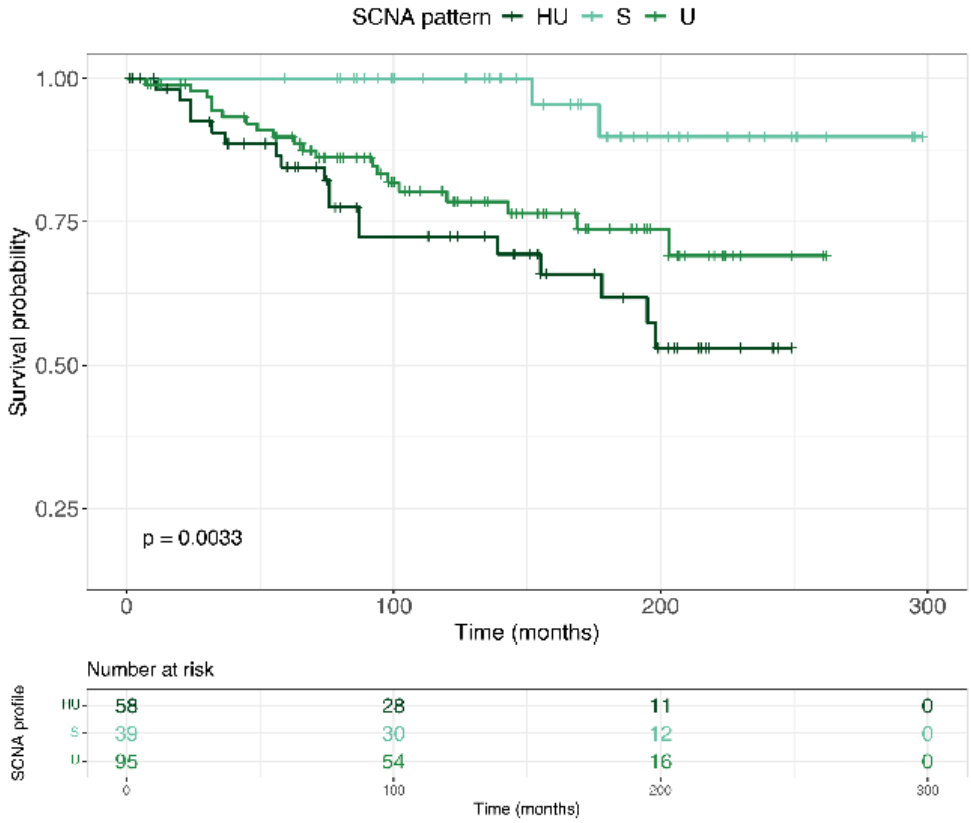
First, I verified if some of the recurrent SNVs could be related to OS or PFS. Among the most recurrent SNVs for each histotype (i.e. *TP53*, *ARID1A*, *KRAS*), only *TP53* exhibited a moderately significant association with patients' survival in PFS (HR=1.7, CI95%=0.91-3.13, p=0.09) and OS (HR=1.8, CI95%=0.97-3.31, p=0.06; Table 17), but this association was lost excluding the HGSOC cases.

Table 17: Univariate survival models (OS and PFS) for the most recurrent mutations. Hazard Ratio (HR) with Confidence Interval at 95% of confidence (CI95%) and p-value of the log-rank test are reported. MUT, mutated; WT, wild type.

	PFS			OS		
	Hazard Ratio	CI (95%)	P-value	Hazard Ratio	CI (95%)	P-value
TP53						
WT (ref)	-	-	-	-	-	-
MUT	1.69	0.91-3.13	0.095	1.79	0.97-3.31	0.063
KRAS						
WT (ref)	-	-	-	-	-	-
MUT	0.83	0.41-1.69	0.607	0.93	0.45-1.90	0.833
ARID1A						
WT (ref)	-	-	-	-	-	-
MUT	0.95	0.50-1.82	0.885	0.77	0.39-1.50	0.440
PIK3CA						
WT (ref)	-	-	-	-	-	-
MUT	0.99	0.49-2.02	0.977	0.77	0.36-1.67	0.505

The three SCNA patterns S, U, HU were significantly associated with survival in terms of PFS or OS, using univariate models (Figure 20).

OS


SCNA pattern HU S U
1.00
0.75
0.50
0.25
Survival probability
p = 0.0033
0
100
200
300
Time (months)
Number at risk
SCNA profile
HU 58 28 11 0
S 39 30 12 0
U 95 54 16 0
0
100
200
300
Time (months)

PFS

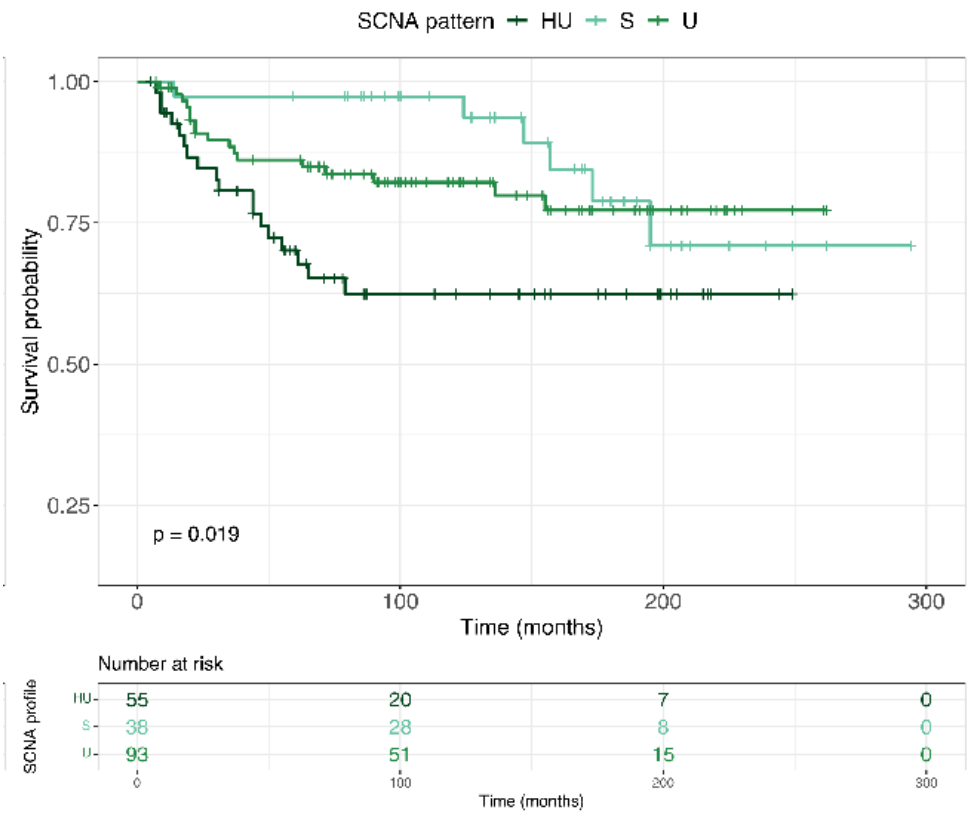

SCNA pattern HU S U
1.00
0.75
0.50
0.25
Survival probability
p = 0.019
0
100
200
300
Time (months)
Number at risk
SCNA profile
HU 55 20 7 0
S 38 28 8 0
U 93 51 15 0
0
100
200
300
Time (months)

Figure 20. Kaplan-Meier curves showing survival in patients belonging to the HU, U and S groups. OS (panel on the top) and PFS (panel at the bottom) survival analyses among the three groups HU (dark green), U (green) and S (light green) on the entire samples cohort. Univariate log-rank test p-values are reported within each plot along with the risk table.

Association with OS was significant even in multivariate models when FIGO substage, grading, age, chemotherapy treatment and *TP53* status were included (Table 18).

Table 18: Multivariate Cox proportional hazard models (OS and PFS) comparing SCNA patterns adjusted for Grade, FIGO2, chemotherapy (CT), TP53 mutation status and age. Hazard Ratio with Confidence Interval at 95% of confidence (CI95%) and p-value are reported. Highly (p < 0.05) and moderately significant (p < 0.1) variables are marked in bold.

	PFS			OS		
	Hazard Ratio	CI (95%)	P-value	Hazard Ratio	CI (95%)	P-value
SCNA pattern						
Stable (reference)	-	-	-	-	-	-
Unstable	1.48	0.57-3.82	0.417	5.10	1.16-22.36	**0.031**
Highly Unstable	1.54	0.48-4.96	0.470	4.39	0.78-24.70	**0.093**
Grade						
Low (reference)	-	-	-	-	-	-
High	1.78	0.83-3.80	0.138	1.67	0.76-3.68	0.204
FIGO2						
a (reference)	-	-	-	-	-	-
b	0.47	0.10-2.26	0.349	2.97	0.80-10.94	0.103
c	1.34	0.62-2.93	0.460	3.06	1.26-7.42	**0.013**
TP53						
WT (reference)	-	-	-	-	-	-
YES	1.36	0.59-3.13	0.471	1.07	0.45-2.55	0.883
CT						
NO (reference)	-	-	-	-	-	-
YES	5.52	1.20-25.43	**0.028**	0.61	0.25-1.48	0.274
Age	1.00	0.98-1.03	0.863	1.05	1.02-1.08	**0.001**

Patients with both U and HU patterns had significantly worse OS compared to those in the S group (HU vs S HR= 4.39, CI95%=0.78-24.70, p= 0.09; U vs S HR= 5.1, CI95%=1.16-

22.36, p = 0.03), while such differences were not observed between patients in HU and U (HR=0.90, CI95%=0.43 -1.88, p=0.78).

Since the HGSOCs were prevalent in the HU pattern (56%), the effect of histological subtype on OS was assessed for each SCNA pattern. OS rates were significantly different between SCNA groups even when the HGSOC samples were excluded and *TP53* status was included as an additional covariate (Table 19). Moreover, also in multivariate analysis integrating the SCNA patterns with histotypes, age and FIGO 2 grade as covariates the association of U and HU SCNA patterns to a worse OS was confirmed (Table 20). These results strengthen the prognostic role of the SCNA pattern, independently of tumour histotype.

*Table 19: Cox proportional multivariate hazard models (with OS and PFS) comparing SCNAs patterns (excluding Serous High patients) adjusted for Grade, FIGO2, CT, AGE, and TP53 mutational status. Hazard Ratio with Confidence Interval at 95% of confidence (CI95%) and p-value are reported. * Given the small number of cases of FIGO2b patients (n=8) and the absence of events in PFS the parameter cannot be estimated properly. In the PFS model these patients have been removed.*

	PFS			OS		
	Hazard Ratio	CI (95%)	P-value	Hazard Ratio	CI (95%)	P-value
SCNA pattern						
Stable (ref)	-	-	-	-	-	-
Unstable	1.42	0.54-3.76	0.481	4.18	0.95-18.34	**0.058**
Highly Unstable	1.26	0.29-5,52	0.755	2.65	0.41-16.92	0.304
Grade						
Low (ref)	-	-	-	-	-	-
High	1.45	0.62-3.42	0.392	1.64	0.71-3.75	0.244
FIGO2*						
a (ref)	-	-	-	-	-	-
b	NA	NA	NA	1.44	0.15-13.43	0.749
c	2.45	0.79-7.55	0.120	3.06	1.03-9.07	**0.043**
CT						
NO (ref)	-	-	-	-	-	-
YES	5.86	0.70-48.99	0.102	0.97	0.30-3.11	0.957
TP53						
WT (ref)	-	-	-	-	-	-
MUT	0.46	0.11-1.88	0.281	1.53	0.54-4.33	0.421
Age	0.98	0.95-1.02	0.330	1.05	1.01-1.09	**0.006**

Table 20: Cox proportional multivariate hazard models (with OS) comparing SCNAs patterns adjusted for Grade, FIGO2, Histotype and AGE. Hazard Ratio with Confidence Interval at 95% of confidence (CI95%) and p-value are reported. Highly (p<0.05) and moderately (p<0.1) significant variables are highlighted in bold. Ref, reference.

	OS		
	Hazard Ratio	**CI (95%)**	**P-value**
SCNA pattern			
Stable (ref)	-	-	-
Unstable	4.20	0.97-18.21	**0.055**
Highly Unstable	4.24	0.85-21.04	**0.077**
Histotype			
Clear Cell (ref)	-	-	-
Endometrioid	0.25	0.06-0.95	**0.042**
Mucinous	0.43	0.08-2.36	0.329
Serous High	0.51	0.19-1.39	0.190
Serous Low	0.21	0.03-1.66	0.137
Grade			
Low (ref)	-	-	-
High	0.69	0.18-2.74	0.600
FIGO 2			
a (ref)	-	-	-
b	3.82	1.01-14.50	**0.049**
c	2.66	1.14-6.22	**0.024**
Age	1.06	1.03-1.09	**0.0004**

In order to understand if SNCAs patterns can be markers of tumour progression and patients' prognosis in the OCCC subtype, one of the EOC histotypes for which an establishment of prognostic parameters in stage I is still an unmet need, I performed the Kaplan-Meier curve OCCCs divided on the basis of the SCNA patterns (Figure 21). Unfortunately, the number of cases included in the analysis (n = 28), is too low to reach any statistically significant result. Particularly, there are only two cases with the S pattern, that survived more than the majority of U and HU cases, but they are not sufficient to say that S pattern is a positive prognostic marker. Regarding U and HU cases, both showed a worse prognosis, moreover, it seems from the curve that the U cases had the worst survival. However, the paucity of cases who survived more than 8 years, less than five patients both for HU and U patterns, preventing to assess which genomic instability pattern had the worst prognosis.

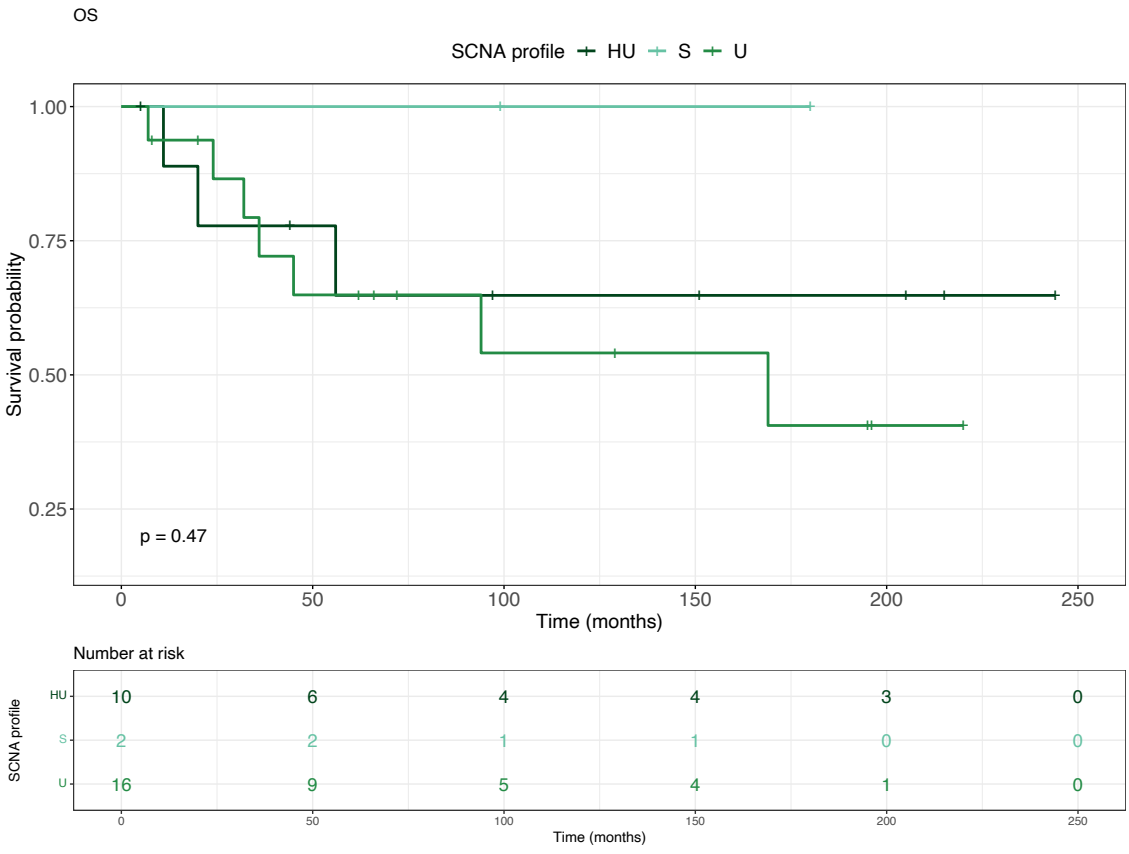


Figure 21. Kaplan-Meier curves showing OS in OCCC patients belonging to the HU, U and S groups. HU (dark green), U (green) and S (light green) on the entire samples cohort.

Lastly, since in this study the cases were collected in two different clinical centers, the analysis of the relation between the genomic features and survival parameters was performed also considering the two sample cohorts separately. The clinicopathological characteristics of each cohort were reported in Tables 21 and 22.

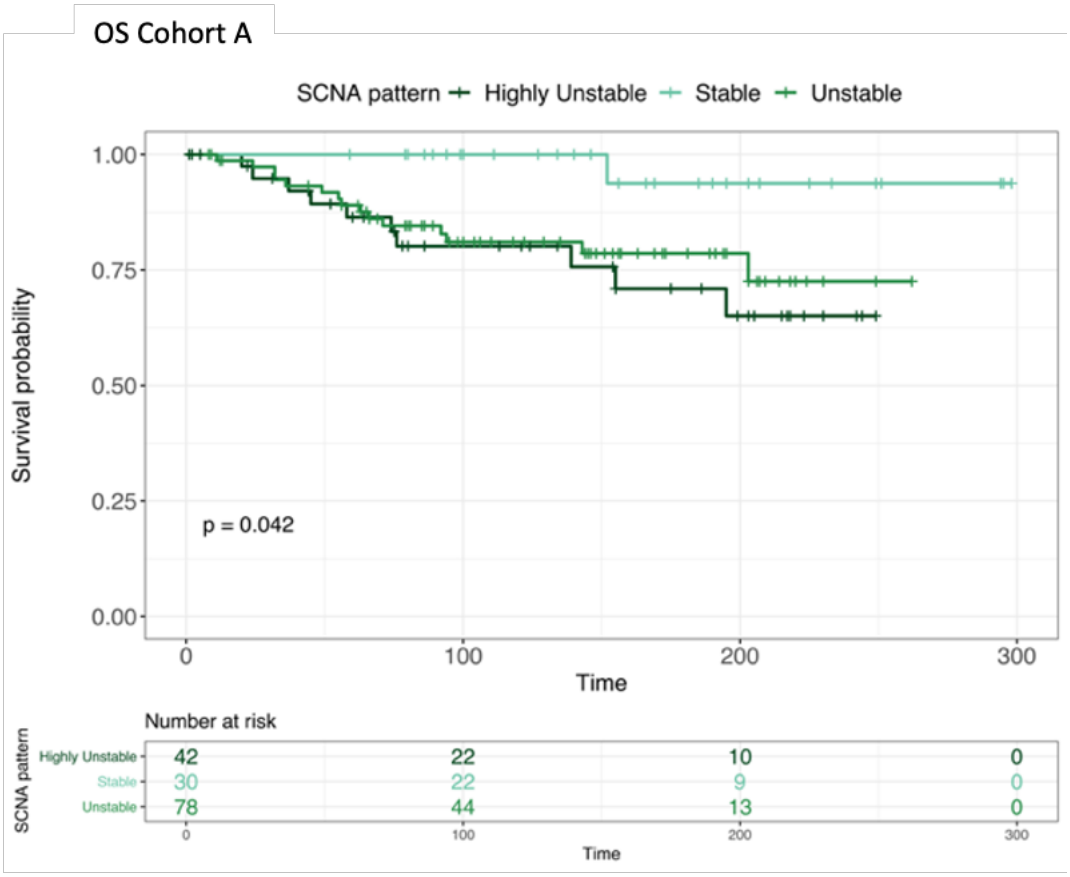
Table 21. Summary of the patients' clinicopathological features in cohort A.

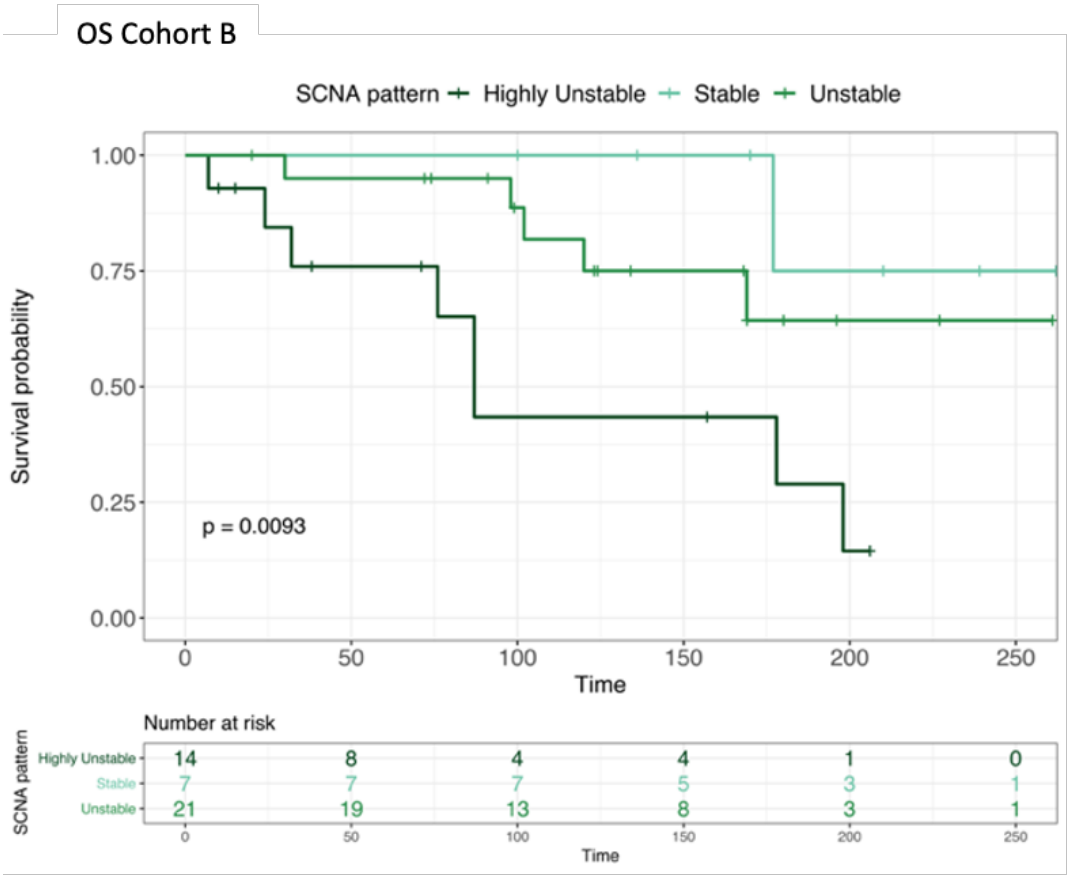
Clinical Annotations	Number of patients	% of patients	N. patients with FU	Median FU [IQR 1-3]
Histology and Grade				
Mucinous			30	12.49y [7.18y-15.18y]
G1	20	12.5		
G2	8	5		
G3	1	0.6		
n.d	1	0.6		
Clear Cells	22	13.8	21	5.13y [2.64y-12.62y]
Endometroid			61	9.44y [6.55y-14.08y]
G1	6	3.8		
G2	34	21.3		
G3	22	13.8		
Low Grade Serous	16	10	16	15.58y [10.13y-17.28y]
High Grade Serous	30	18.8	30	11.35y [5.30y-16.08y]
FIGO substages				
A	55	34.4		
B	11	6.9		
C	93	58.1		
n.d	1	0.6		
Median age at diagnosis [min-max];	52.38 y [16.49y – 89.33y]			
Total number of patients	160			

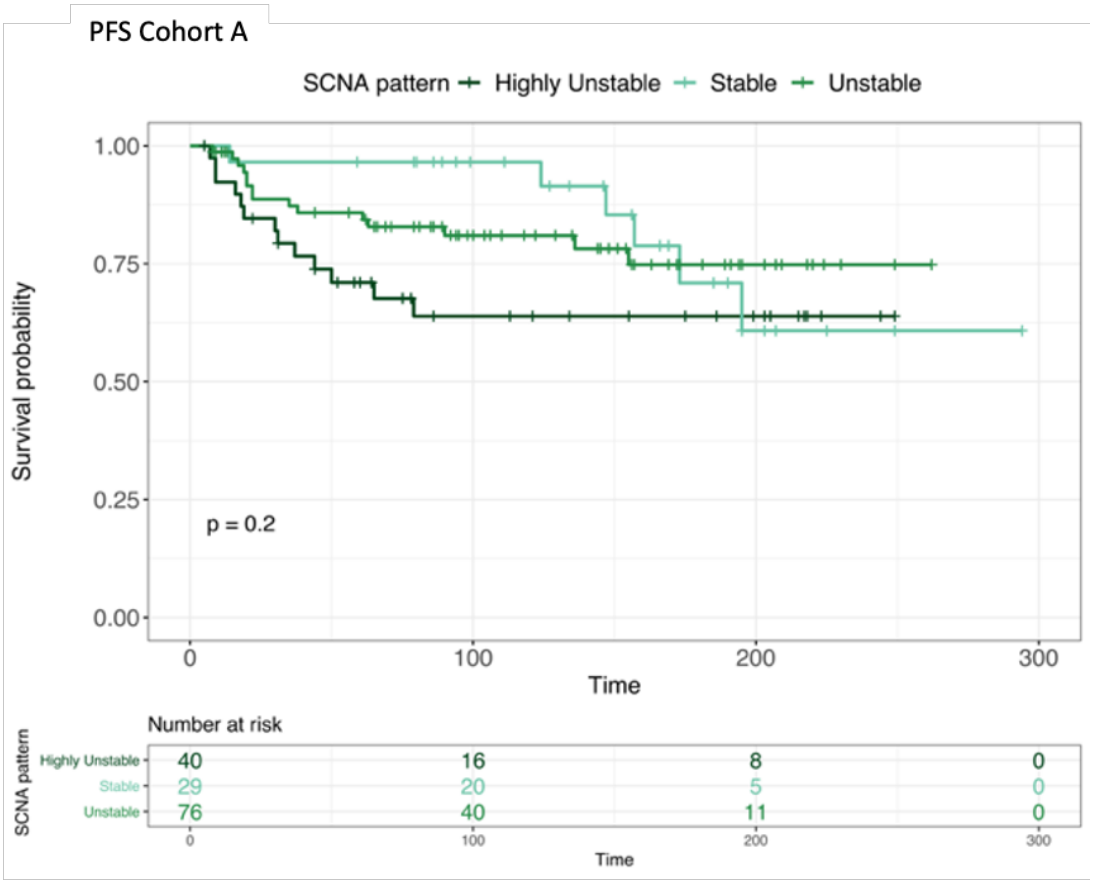
Table 22. Summary of the patients' clinicopathological features in cohort B.

Clinical Annotations	Number of patients	% of patients	N. patients with FU	Median FU [IQR 1-3]
Histology and Grade				
Mucinous			8	9.09y [7.25y-11.30y]
G1	6	13.3		
G2	2	4.4		
G3	0	0		
n.d	0	0		
Clear Cells	7	15.6	7	8.09y [3.83y-14.53y]
Endometroid			18	12.21y [8.24y-18.58y]
G1	8	17.8		
G2	4	8.9		
G3	6	13.3		
Low Grade Serous	3	4.4	2	12.22y [11.23y-13.20y]
High Grade Serous	9	22.2	9	6.79y [3.10y-10.20y]
FIGO substages				
A	13	28.9		
B	6	13.3		
C	23	51.1		
n.d	3	6.7		
Median age at diagnosis [min-max];	59.45 y [25.84y – 77.27y]			
Total number of patients	45			

The analysis confirmed the association of the genomic pattern and prognosis, particularly, except for PFS in cohort A, U and HU showed a significantly worse prognosis than S in both patient groups (Figure 22).


OS Cohort A
SCNA pattern Highly Unstable Stable Unstable
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.042
0
100
200
300
Time
Number at risk
SCNA pattern
Highly Unstable 42 22 10 0
Stable 30 22 9 0
Unstable 78 44 13 0
0
100
200
300
Time


OS Cohort B
SCNA pattern Highly Unstable Stable Unstable
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.0093
0
50
100
150
200
250
Time
Number at risk
SCNA pattern
Highly Unstable 14 8 4 4 1 0
Stable 7 7 7 5 3 1
Unstable 21 19 13 8 3 1
0
50
100
150
200
250
Time


PFS Cohort A
SCNA pattern Highly Unstable Stable Unstable
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.2
0
100
200
300
Time
Number at risk
SCNA pattern
Highly Unstable 40 16 8 0
Stable 29 20 5 0
Unstable 76 40 11 0
0
100
200
300
Time

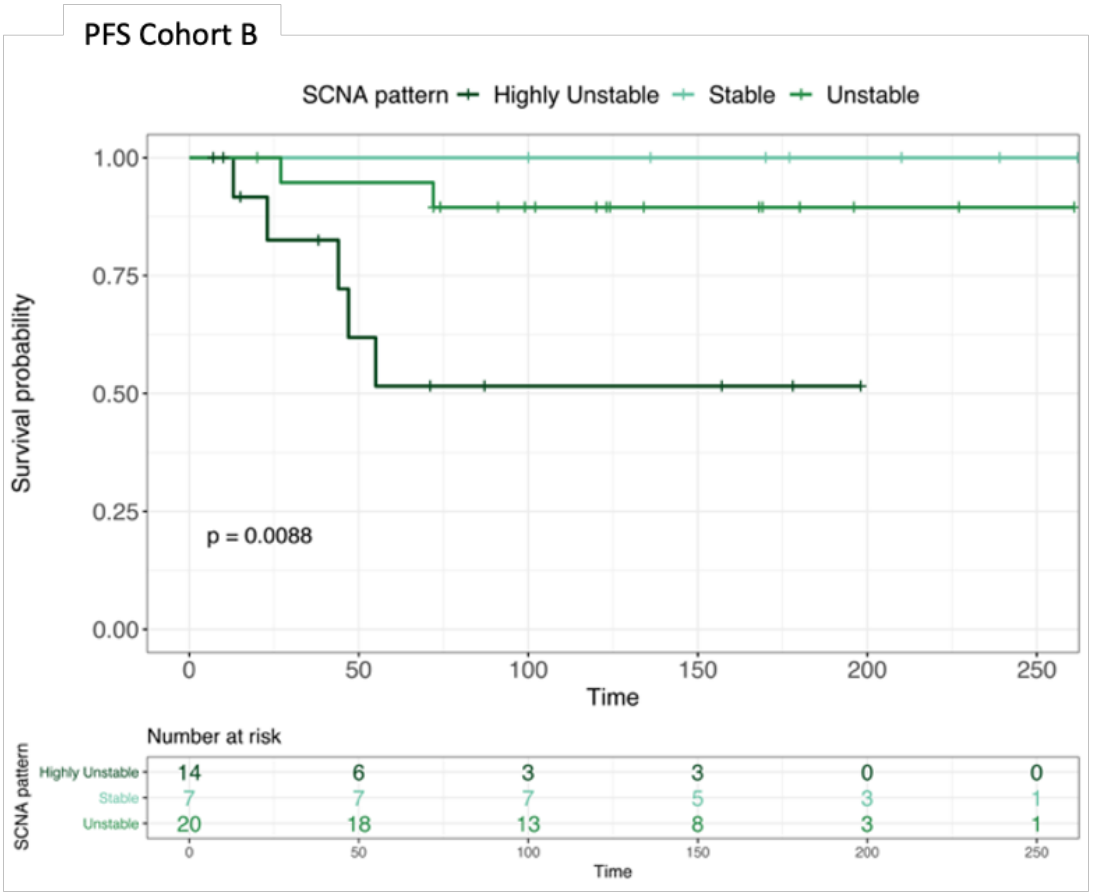

PFS Cohort B
SCNA pattern Highly Unstable Stable Unstable
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.0088
0
50
100
150
200
250
Time
Number at risk
SCNA pattern
Highly Unstable 14 6 3 3 0 0
Stable 7 7 7 5 3 1
Unstable 20 18 13 8 3 1
0
50
100
150
200
250
Time

Figure 22. Survival analyses among HU U and S patterns considering cohorts A and B separately. The p-value of the log-rank test is reported in the plot along with the risk table.

The last parameter that was analysed in association with survival was LINE1 methylation. It resulted significantly associated with PFS both in univariate (HR= 0.96, CI 95%=0.93-0.98, p= 0.003) and multivariate analysis (Table 23).

Table 23. Multivariate Cox proportional hazard models (PFS) comparing LINE1 values adjusted for SCNA pattern and age. Hazard Ratio with Confidence Interval at 95% of confidence (CI95%) and p-value are reported. Highly (p<0.05) and moderately (p<0.1) significant variables are highlighted in bold.

	PFS		
	Hazard Ratio	CI (95%)	P-value
LINE1	0.97	0.94-1.00	**0.085**
SCNA pattern			
S (reference)	-	-	-
U	1.02	0.40-2.63	0.967
HU	1.84	0.66-5.15	0.247
Age	1.00	1.00-1.03	0.818

In conclusion, the survival analyses showed that molecular features and particularly, the SCNAs patterns, could be improved markers to predict more precisely patient prognosis.

DISCUSSION

My PhD project aimed at deepening the molecular knowledge of stage I EOC, both from a genomic and an epigenomic perspective. The final purpose of this molecular characterization is to propose novel potential prognostic markers that could help clinicians to predict patients' outcome and better identify those eligible for chemotherapy. This study extends the current knowledge of the complex and heterogenous biology underlying stage I EOCs and defines novel molecular markers that, despite the peculiar features of each tumour subtype, could be indicative of a common tumour behaviour and thus patient's prognosis, i.e. the SCNAs patterns.

SNVs analysis

Through targeted analysis of SNVs among frequently mutated genes in cancer, it was possible to define recurrent altered genes responsible for the etiopathogenetic processes that drive and sustain tumour evolutionary routes. Moreover, many of these recurrent altered genes are also currently targetable by precise medicine, opening novel potential therapeutic windows for the patients.

The morphological differences existing between the subtypes of stage I EOCs are also recapitulated by the specific landscape of recurrent SNVs of each histotype. However, despite the intrinsic heterogeneity across the subtypes, many recurrent SNVs are shared by more than one histotype and are potentially exploited for tailored therapeutic approaches. HGSOC is the subtype that showed the tumour genome with the lowest number of recurrent SNVs, the only recurrent altered gene being clonal mutations in *TP53*, as occurs in stage III-IV EOC[138]. Differently, LGSOCs show a prevalence of hotspot mutations in the *BRAF* gene, which is amenable to therapeutic targeting[139], or in the *KRAS* gene, inhibitors of which are currently undergoing clinical trials[140,141]. Even MOC samples show a prevalence in *KRAS* hotspot mutations, along with mutations in *TP53*, although with a lower MAF than that seen in HGSOCs. These different MAFs corroborate the fact that in HGSOCs the loss of TP53 functionality represents the early event present also in precancerous lesions, while in MOCs *TP53* mutations originate later during the tumour

evolution and are thus subclonal[65]. Out of all histological subtypes, EC cases present the highest incidence in mutational rate, with SNVs in most of the analysed genes, particularly in *ARID1A*, *KRAS*, *PIK3CA* and *CTNNB1*. The OCCC subtype exhibits mutations in almost the same genes as EC cases, reflecting the closely related etiopathogenesis of the two subtypes, even if with a different prevalence, with *ARID1A* being the major recurrently mutated gene. Both subtypes are characterized by a small incidence in MSI phenotype or *TP53* mutation. Moreover, both histotypes showed a co-occurrence in *ARID1A* and *PIK3CA* mutations, even if only in ECs cases the statistical significance of this association was reached. The co-existence of *ARID1A* and *PIK3CA* mutations was already reported by several studies [114,115]. It is particularly relevant in tumours derived from endometriosis, because it is related to tumorigenic processes that favour tumour growth and progression. Chandler and colleagues[114] demonstrated the epistatic relationship between SWI/SNF chromatin remodelling and PI3K pathway mutations in OCCCs, alterations in these pathways convergingly caused high levels of tumour cell-derived IL-6, a pro-tumorigenic cytokine signalling which promote tumour growth. Few years later, Wilson et al. showed that coexistent ARID1A and PI3K mutations promote epithelial transdifferentiation and collective invasion in EC tumours [115]. The molecular mechanisms driven by the co-occurrence of *ARID1A* and *PIK3CA* mutations are particularly relevant also from a therapeutic point of view, not only for the treatment with PI3K inhibitors, but also because anti-IL-6 antibody (Siltuximab) therapy may be a potential and effective treatment strategy for these patients. Regarding molecular classification of endometrioid tumours, following the one proposed by ProMisE for endometrial carcinoma[76], all the four molecular classes (POLE ultramutated tumours, MMRd hypermutated tumours, TP53-mutated cases and NSMP cases) were identified also in EC stage I EOCs, further studies will be necessary to understand if these classes could be used to set different therapeutic regimens. Together, these observations indicate that most of stage I EOC tumours carry a targetable driver mutation in a known onco- or suppressor gene. This suggests that a more thorough characterization could reveal novel bespoke clinically exploitable therapeutic windows.

SCNAs analysis

As for the SNVs, the analysis of SCNAs recapitulated the heterogeneity among the stage I EOC subtypes. Particularly, HGSOCs showed the highest degree of genomic instability with a huge number of SCNAs covering almost all chromosomes. Many of these SCNAs were focal and recurrent in more than half of the cases. These recurrent SCNAs often involved loci encoding for onco- or suppressor genes and could thus further contribute to genomic instability and tumour evolution. Particularly, most of HGSOCs were affected by chromosome 17p deletion, that causes the complete loss of the *TP53* gene [142]. Another recurrent SCNA in HGSOC cases is the deletion of cytoband 13q14.11 involving *RB1*. Loss of this tumor suppressor gene has been reported to occur early in tumorigenesis and to correlate with different clinical outcomes and therapy response [117]. The most frequent focal amplification in HGSOCs included *MYC*, already reported to correlate with poor prognosis in FIGO stages I and II EOCs [143]. Other genes involved in frequent copy number changes in HGSOCs are *STK11* (or *LKB1*) and *ZNF516*, both included into two regions where broad deletions occurred in about 60% of cases, chromosomes 19p13.3 and 18q23, respectively. *STK11* is the causative gene of the Peutz-Jeghers syndrome, a germline syndrome associated with increased risk of several tumours [144]. The loss of *LKB1* is considered an early event in tumorigenesis, and has indeed been reported to cause the development of HGSOC from ovarian surface epithelial cells in murine models along with the deletion of *PTEN* [123] and the upregulation of the NF-κB pathway in HGSOC cells models [145]. ZNF516 is a Zinc Finger protein involved in transcription regulation, but its cellular functions are not completely elucidated yet, it has been involved in replication stresses and particularly its loss was associated to increase chromosome instability [125], but its role in genomic instability has to be clarified yet. The HGSOC was the only histotype characterized by high levels of genomic instability in all the cases, unlike what was observed in other histotypes where not all cases had many SCNAs and often they covered the entire chromosome arm. Only few cases showed the same SCNAs pattern of HGSOCs. This consequently explains why few strongly recurrent SCNAs emerged in the other histotypes. Among the recurrent SCNAs in the other stage I EOC subtypes, there were the amplification of the two oncogenes *MYC* and

ZNF217 in OCCCs, the deletion of the oncosuppressor gene *CDKN2A* and the amplification of the oncogene *ERBB2* in MOCs. The amplification of *MYC* and *ZNF217*, as well as the deletion of *CDKN2A*, and the amplification of *ERBB2* were already reported not only in EOCs, but also in many other cancers and these alterations seemed strictly related to prognosis and potentially to be exploitable as targets for precise treatments[124,126,127,146–152].

However, observing the distribution of the SCNAs along the genome, it was possible to notice that there were differences in SCNAs profiles between cases belonging to the same histotype and similarities among cases of different histotypes. Indeed, it was possible to identify three common patterns of genomic instability, namely S, U and HU, that reflect different and increasing degrees of genomic instability. These patterns were based on the different incidence of copy number events and their extent in the genome. Indeed, an algorithm based on the values of the length of individual SCNAs (normalized by chromosome arm) and CNB was developed to assign the SCNA pattern to each tumour. Exploiting this algorithm, it is possible to easily classify each tumour to one of the three identified SCNAs patterns. These patterns were heterogeneously distributed among the five stage I EOC subtypes. The genome of HGSOCs shows a marked prevalence of the HU pattern and a complete absence of the S one, remarking that this histotype is characterized by genomic instability even at early disease stages. EC cases mutated in *TP53* showed a HU genomic pattern, confirming the close relationship between these EC cases and HGSOC histotype. Nevertheless, HU pattern was not associated only to *TP53* alterations, indeed, in MOCs and OCCCs several cases with HU pattern did not have *TP53* alterations, parallelly, tumours with *TP53* pathogenic variants, showed a U pattern. Therefore, other mechanisms might be responsible for this pattern of genomic instability.

In an effort to unravel the mutational pathways that could generate specific type and distribution of SCNAs, peculiar features of SCNAs, such as length and copy number value, were exploited to reconstruct five distinct signatures which finely describe the pattern of genome instability. Having defined the five SCNAs signatures that could explain different model of generation of SCNAs, the presence of a signature in a tumour genome was

correlated to the SNVs found in the sample. *KRAS*, along with *PTEN*, *CTNNB1*, and *PIK3CA*, were associated to signature 1, which encompasses the vast majority of MOCs, LGSOCs, OCCCs and ECs, corresponding mainly to U cases. Mutations in these genes are known to be involved in chromosomal instability due to impaired mitotic chromosome segregation, genome doubling and telomere instability[125,128]. In contrast, HU SCNAs pattern can be divided into a combination of four signatures, highlighting the extreme heterogeneity of mechanisms responsible for this highest level of genome instability in stage I EOC. Among the altered genes related to these four signatures there were amplification of *MYC*, as well as amplification of *CCNE1*. These alterations have been associated to structural and numerical chromosomal aberrations, due to abnormal DNA replication and mitosis, such as precocious S-phase entry, centrosome amplifications and chromosome missegregation[133,134,136]. It is noteworthy that each tumour represents a mosaic of different signatures, thus indicating that multiple mutational pathways could contribute to the genomic instability.

Interestingly, even if in the previous study, Macintyre and colleagues [70] identified seven SCNA signatures only in advanced stages HGSOC samples, the components weights in each of their seven signatures showed several similarities with those of the SCNA signatures identified in my PhD project in all histotypes of EOCs at stage I. Indeed, Stage I SCNA signature 1 is similar in components distribution to signature 1 of Macintyre [70], such as in segment size distribution. This resemblance is further supported by the fact that in both signatures the prevalence of mutated pathways involved genes as *KRAS* and thus it could be speculated that also Stage I signature 1 is related to breakage-fusion-bridge (BFB) events. The signature 2 features instead partially resembled those of SCNA signature 3 described by Macyntire et al.[70], associated with mutations in the *BRCA1* gene and HR-related genes, as stage I EOC signature 2. Signature 3 components pattern was partially reminiscent of Macyntire et al.'s signature 2 [70], especially in high number of breakpoints, long chains of oscillating copy-number and small segment size. The Macyntire et al.'s signature 2 [70] was related to *CDK12* mutations and tandem duplicator phenotype[70]. Stage I signature 4 is similar to the same number signature of Macyntire [70], indeed both were associated to *CCNE1* and *MYC* amplification and thus probably related

to whole genome duplication due to failure of cell cycle control. Finally, signature 5 profile is similar to the previously reported signature 5, related to chromothriptic-like events[70].

Despite the necessity to confirm these SCNA signatures on a larger cohort of cases and to better characterize the molecular processes at the bases of the different types of SCNAs events along tumour genome, the SCNA signatures undoubtedly demonstrate a clear clinical utility. Indeed, they could be used at the time of diagnosis to infer, without SNVs analysis on a panel of genes or exome, the mutated pathways that are driving tumour progression and thus could be targeted by precise therapeutic treatments, such as *KRAS* in signature 1, or HR in signature 2. Moreover, further studies will be fundamental to understand if this SCNA signature could also have a prognostic role.

LINE1 methylation

To further deepen potential mechanisms related to genomic instability, LINE1 methylation was assessed in stage I EOC. HU cases were characterized by lower levels of LINE 1 promoter methylation compared to the samples of other SCNA patterns, suggesting that higher levels of genomic instability might be strictly related also to epigenetic phenomena, such as the activation of transposable elements[92,93].

Survival analysis

The tight association between specific mutations in driver genes, SCNAs patterns and tumour behaviour suggested that these molecular aspects could also be exploited as prognostic markers able to predict tumour evolution. To evaluate possible association between molecular features and prognosis, I considered two parameters, the OS and the PFS. The OS was defined as the total period of time from tumour diagnosis to death, thus the event used for the univariate and multivariate analyses was the patients 'death as registered in the Pandora biobank; patients were considered alive until the time of the last follow up visit as registered in the biobank. The PFS was defined as the time intercourse between the tumour diagnosis and the relapse insurgence determined by clinical parameters (i.e. CA125 levels, MRI, etc) and/or by the date of second surgery

and/or by the beginning of a novel therapeutic treatment, all these information were retrieved from the Pandora biobank. In case no information about relapse, second surgery or treatments were available, cases were excluded from the analyses with the PFS.

None of the recurrent altered genes resulted strongly associated to prognosis. Otherwise, the identified genomic instability patterns (S, U and HU) were associated to survival, particularly HU and U patterns were correlated to worse OS and PFS. This association with OS was maintained also in multivariate analyses with FIGO substage, grading, age, chemotherapy treatment, the presence of *TP53* mutation and histotype or excluding the HGSOCs from the analysis since they are the most instable tumours in the cohort.

LINE1 methylation resulted associated with PFS as well both in univariate and in multivariate analysis with SCNAs patterns and age, corroborating the intimate relationship between genome instability and LINE1 methylation.

Limitations

The work has some limitations:

i) The size of the cohort was not sufficient to verify and validate both the prognostic role of the SCNAs patterns and the potential of the SCNAs signatures in stage I EOCs and particularly in each of their subtypes. Indeed, I tried to verify the prognostic role of the SCNA patterns for cases belonging to the OCCC subtype, for which no precise prognostic parameters are currently available to avoid detrimental chemotherapeutic treatments. However, the paucity of cases impeded to demonstrate that HU cases had the worst prognosis. Nevertheless, although no statistically significant results had been obtained, it seems that could be worthwhile to expand the study of these SCNA patterns on larger cohort of patients affected by OCCCs. Further retrospective and prospective studies on larger cohorts will be necessary to better confirm the utility of the discrete genome instability categories (HU, U and S) as predictors. Moreover, regarding the SCNAs signatures, their utility

both in terms of tumours classification and in terms of genomic scars that reflect the mutational pathway that cause them and that could be potential targetable markers was not completely assessed due to the paucity of cases for this kind of analyses. Nevertheless, to the best of my knowledge, this is the large stage I EOCs cohort. Further multicenter studies collecting higher number of cases and will be required to deepen the genomic aspects and their link to patient's prognosis.

ii) The number of genes analysed by targeted sequencing was not sufficiently large to make a comprehensive mutational analysis, able to unveil all the mutated pathways and their association with SCNA patterns.

iii) The survival analysis was not run on an independent dataset to confirm the results, but, sufficiently detailed and large public sequencing datasets of stage I EOC are currently not available.

Conclusion

Despite the limitations described above, this project reveals for the first time potential genomic markers that could be exploited to better predict the prognosis in stage I EOCs. Interestingly, these genomic markers are a sort of genomic scars, the SCNAs patterns S, U and HU, produced by the mutational pathways that drive tumour growth. Even if stage I EOC, as advanced EOC, could be considered as a heterogeneous group of distinct tumour type, some alterations are present in more than one histotype and could similarly guide tumour evolution. Indeed, to date one of the strategies to improve cancer diagnosis and treatment is to define common molecular markers that could be shared also by tumours of different origin. Several other publications identified signatures based on SCNAs to classify tumours and predict tumour prognosis, independently from tumour location and type[71,86,87,153].

Moreover, since the determination of SCNA patterns through sWGS could become a suitable surrogate marker to help clinicians assess tumour aggressiveness, it is relevant that this type of analysis can be performed on tumour biopsies which are routinely stored

in hospitals and the relatively straightforward nature of data evaluation are clinically feasible.

Overall, the data presented in this work extend the knowledge on molecular landscape of stage I EOCs and define novel potential markers that could be useful for better stratification of patients, opening the possibility of an improved prognosis and differential therapy in the management of the disease.

REFERENCES

1. Colombo, N. *et al.* International collaborative ovarian neoplasm trial 1: A randomized trial of adjuvant chemotherapy in women with early-stage ovarian cancer. *J. Natl. Cancer Inst.* (2003). doi:10.1093/jnci/95.2.125

2. Cree, I. A., White, V. A., Indave, B. I. & Lokuhetty, D. Revising the WHO classification: female genital tract tumours. *Histopathology* (2020). doi:10.1111/his.13977

3. McMullen, M., Karakasis, K., Rottapel, R. & Oza, A. M. Advances in ovarian cancer, from biology to treatment. *Nature Cancer* (2021). doi:10.1038/s43018-020-00166-5

4. Kim, M. *et al.* Major clinical research advances in gynecologic cancer in 2018. *Journal of Gynecologic Oncology* (2019). doi:10.3802/jgo.2019.30.e18

5. Stefania, G. *et al.* I numeri del cancro in Italia. *Rep. AIOM-AIRTUM* 1–232 (2015).

6. Prat, J. Staging classification for cancer of the ovary, fallopian tube, and peritoneum. *Int. J. Gynecol. Obstet.* (2014). doi:10.1016/j.ijgo.2013.10.001

7. Ovaio, T. D., Operativa, U. & Ginecologica, M. Linee guida. (2019).

8. Norquist, B. M. *et al.* Inherited mutations in women with ovarian carcinoma. *JAMA Oncol.* (2016). doi:10.1001/jamaoncol.2015.5495

9. Ramus, S. J. *et al.* Germline mutations in the BRIP1, BARD1, PALB2, and NBN genes in women with ovarian cancer. *J. Natl. Cancer Inst.* (2015). doi:10.1093/jnci/djv214

10. Jervis, S. *et al.* Ovarian cancer familial relative risks by tumour subtypes and by known ovarian cancer genetic susceptibility variants. *J. Med. Genet.* (2014). doi:10.1136/jmedgenet-2013-102015

11. Bewtra, C. D., Watson, P., Conway, T., Read-Hippee, C. & Lynch, H. T. Hereditary

ovarian cancer: A clinicopathological study. *Int. J. Gynecol. Pathol.* (1992). doi:10.1097/00004347-199207000-00003

12. Vasen, H. F. A., Watson, P., Mecklin, J. P. & Lynch, H. T. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative Group on HNPCC. in *Gastroenterology* (1999). doi:10.1016/S0016-5085(99)70510-X

13. Folkins, A. K. & Longacre, T. A. Hereditary gynaecological malignancies: Advances in screening and treatment. *Histopathology* (2013). doi:10.1111/his.12028

14. Reid, B. M., Permuth, J. B. & Sellers, T. A. Epidemiology of ovarian cancer: a review. *Cancer Biology and Medicine* (2017). doi:10.20892/j.issn.2095-3941.2016.0084

15. Jacobs, I. J. *et al.* Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): A randomised controlled trial. *Lancet* (2016). doi:10.1016/S0140-6736(15)01224-6

16. Buys, S. S. *et al.* Effect of screening on ovarian cancer mortality: The Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening randomized controlled trial. *JAMA - J. Am. Med. Assoc.* (2011). doi:10.1001/jama.2011.766

17. Moore, R. G. *et al.* A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. *Gynecol. Oncol.* (2009). doi:10.1016/j.ygyno.2008.08.031

18. Vergote, I. *et al.* Neoadjuvant Chemotherapy or Primary Surgery in Stage IIIC or IV Ovarian Cancer. *N. Engl. J. Med.* (2010). doi:10.1056/nejmoa0908806

19. Trimbos, J. B. *et al.* Impact of adjuvant chemotherapy and surgical staging in early-stage ovarian carcinoma: European organisation for research and treatment of cancer-adjuvant chemotherapy in ovarian neoplasm-trial. *J. Natl. Cancer Inst.* (2003). doi:10.1093/jnci/95.2.113

20. Trimbos, B. *et al.* Surgical staging and treatment of early ovarian cancer: Long-term analysis from a randomized trial. *J. Natl. Cancer Inst.* (2010). doi:10.1093/jnci/djq149

21. Oseledchyk, A. *et al.* Adjuvant chemotherapy in patients with stage I endometrioid or clear cell ovarian cancer in the platinum era: A Surveillance, Epidemiology, and End Results Cohort Study, 2000-2013. *Ann. Oncol.* (2017). doi:10.1093/annonc/mdx525

22. Gore, M. *et al.* An international, phase III randomized trial in patients with mucinous epithelial ovarian cancer (mEOC/GOG 0241) with long-term follow-up: and experience of conducting a clinical trial in a rare gynecological tumor. *Gynecol. Oncol.* (2019). doi:10.1016/j.ygyno.2019.03.256

23. Kurnit, K. C. *et al.* Effects of Gastrointestinal-Type Chemotherapy in Women with Ovarian Mucinous Carcinoma. in *Obstetrics and Gynecology* (2019). doi:10.1097/AOG.0000000000003579

24. Burger, R. A. *et al.* Incorporation of Bevacizumab in the Primary Treatment of Ovarian Cancer. *N. Engl. J. Med.* (2011). doi:10.1056/nejmoa1104390

25. Perren, T. J. *et al.* A Phase 3 Trial of Bevacizumab in Ovarian Cancer. *N. Engl. J. Med.* (2011). doi:10.1056/nejmoa1103799

26. Armstrong, D. K. *et al.* Intraperitoneal Cisplatin and Paclitaxel in Ovarian Cancer. *N. Engl. J. Med.* (2006). doi:10.1056/nejmoa052985

27. Walker, J. L. *et al.* Randomized trial of intravenous versus intraperitoneal chemotherapy plus bevacizumab in advanced ovarian carcinoma: An NRG oncology/Gynecologic oncology group study. *J. Clin. Oncol.* (2019). doi:10.1200/JCO.18.01568

28. Van Driel, W. *et al.* A phase 3 trial of hyperthermic intraperitoneal chemotherapy (HIPEC) for ovarian cancer. *J. Clin. Oncol.* (2017).

doi:10.1200/jco.2017.35.15_suppl.5519

29. Lim, M. C. *et al.* Randomized trial of hyperthermic intraperitoneal chemotherapy (HIPEC) in women with primary advanced peritoneal, ovarian, and tubal cancer. *J. Clin. Oncol.* (2017). doi:10.1200/jco.2017.35.15_suppl.5520

30. Coleman, R. L. *et al.* Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* (2017). doi:10.1016/S0140-6736(17)32440-6

31. Pujade-Lauraine, E. *et al.* Olaparib tablets as maintenance therapy in patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-Ov21): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Oncol.* (2017). doi:10.1016/S1470-2045(17)30469-2

32. Mirza, M. R. *et al.* Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer. *N. Engl. J. Med.* (2016). doi:10.1056/nejmoa1611310

33. Moore, K. *et al.* Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. *N. Engl. J. Med.* (2018). doi:10.1056/nejmoa1810858

34. González-Martín, A. *et al.* Niraparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. *N. Engl. J. Med.* (2019). doi:10.1056/nejmoa1910962

35. Coleman, R. L. *et al.* Veliparib with First-Line Chemotherapy and as Maintenance Therapy in Ovarian Cancer. *N. Engl. J. Med.* (2019). doi:10.1056/nejmoa1909707

36. Ray-Coquard, I. *et al.* Olaparib plus Bevacizumab as First-Line Maintenance in Ovarian Cancer. *N. Engl. J. Med.* (2019). doi:10.1056/nejmoa1911361

37. Armstrong, D. *et al.* NCCN Clinical Practice Guidelines in Oncology. Ovarian Cancer. *NCCN* (2020).

38. Disis, M. L. *et al.* Avelumab (MSB0010718C; anti-PD-L1) in patients with

recurrent/refractory ovarian cancer from the JAVELIN Solid Tumor phase Ib trial: Safety and clinical activity. *J. Clin. Oncol.* (2016). doi:10.1200/jco.2016.34.15_suppl.5533

39. Hamanishi, J. *et al.* Safety and antitumor activity of Anti-PD-1 antibody, nivolumab, in patients with platinum-resistant ovarian cancer. *J. Clin. Oncol.* (2015). doi:10.1200/JCO.2015.62.3397

40. Varga, A. *et al.* Pembrolizumab in patients (pts) with PD-L1–positive (PD-L1 +) advanced ovarian cancer: Updated analysis of KEYNOTE-028. . *J. Clin. Oncol.* (2017). doi:10.1200/jco.2017.35.15_suppl.5513

41. Shih, I. M. & Kurman, R. J. Ovarian Tumorigenesis: A Proposed Model Based on Morphological and Molecular Genetic Analysis. *American Journal of Pathology* (2004). doi:10.1016/S0002-9440(10)63708-X

42. Koshiyama, M., Matsumura, N. & Konishi, I. Recent concepts of ovarian carcinogenesis: Type i and type II. *BioMed Research International* (2014). doi:10.1155/2014/934261

43. De Leo, A. *et al.* What is new on ovarian carcinoma: Integrated morphologic and molecular analysis following the new 2020 world health organization classification of female genital tumors. *Diagnostics* (2021). doi:10.3390/diagnostics11040697

44. Soslow, R. A. *et al.* Morphologic patterns associated with BRCA1 and BRCA2 genotype in ovarian carcinoma. *Mod. Pathol.* (2012). doi:10.1038/modpathol.2011.183

45. Laokulrath, N., Warnnissorn, M., Chuangsuwanich, T. & Hanamornroongruang, S. Sectioning and extensively examining the fimbriated end (SEE-FIM) of the fallopian tube in routine practices, is it worth the effort? *J. Obstet. Gynaecol. Res.* (2019). doi:10.1111/jog.13845

46. Shih, I. M., Wang, Y. & Wang, T. L. The Origin of Ovarian Cancer Species and

Precancerous Landscape. *American Journal of Pathology* (2021). doi:10.1016/j.ajpath.2020.09.006

47. Mehra, K. *et al.* STICS, SCOUTs and p53 signatures; A new language for pelvic serous carcinogenesis. *Front. Biosci. - Elit.* (2011). doi:10.2741/e275

48. Mittal, N. *et al.* Secretory cell outgrowths, p53 signatures, and serous tubal intraepithelial carcinoma in the fallopian tubes of patients with sporadic pelvic serous carcinoma. *Indian J. Pathol. Microbiol.* (2016). doi:10.4103/0377-4929.191789

49. Soong, T. R. *et al.* Evidence for lineage continuity between early serous proliferations (ESPs) in the Fallopian tube and disseminated high-grade serous carcinomas. *J. Pathol.* (2018). doi:10.1002/path.5145

50. Ahn, G., Folkins, A. K., McKenney, J. K. & Longacre, T. A. Low-grade Serous Carcinoma of the Ovary: Clinicopathologic Analysis of 52 Invasive Cases and Identification of a Possible Noninvasive Intermediate Lesion. *Am. J. Surg. Pathol.* (2016). doi:10.1097/PAS.0000000000000693

51. Santandrea, G. *et al.* Immunohistochemical Biomarkers as a Surrogate of Molecular Analysis in Ovarian Carcinomas: A Review of the Literature. *Diagnostics* (2021). doi:10.3390/diagnostics11020199

52. Kurman, R. J. & Shih, I. M. The dualistic model of ovarian carcinogenesis revisited, revised, and expanded. *American Journal of Pathology* (2016). doi:10.1016/j.ajpath.2015.11.011

53. Vang, R., Shih, I. M. & Kurman, R. J. Fallopian tube precursors of ovarian low- and high-grade serous neoplasms. *Histopathology* (2013). doi:10.1111/his.12046

54. Matias-Guiu, X. & Stewart, C. J. R. Endometriosis-associated ovarian neoplasia. *Pathology* (2018). doi:10.1016/j.pathol.2017.10.006

55. Sainz De La Cuesta, R. *et al.* Histologic transformation of benign endometriosis to early epithelial ovarian cancer. *Gynecol. Oncol.* (1996). doi:10.1006/gyno.1996.0032

56. Jiang, X., Morland, S. J., Hitchcock, A., Thomas, E. J. & Campbell, I. G. Allelotyping of endometriosis with adjacent ovarian carcinoma reveals evidence of a common lineage. *Cancer Res.* (1998).

57. Köbel, M. *et al.* Differences in tumor type in low-stage versus high-stage ovarian carcinomas. *Int. J. Gynecol. Pathol.* (2010). doi:10.1097/PGP.0b013e3181c042b6

58. MacKay, H. J. *et al.* Prognostic relevance of uncommon ovarian histology in women with stage III/IV epithelial ovarian cancer. *Int. J. Gynecol. Cancer* (2010). doi:10.1111/IGC.0b013e3181dd0110

59. Chan, J. K. *et al.* Do clear cell ovarian carcinomas have poorer prognosis compared to other epithelial cell types? A study of 1411 clear cell ovarian cancers. *Gynecol. Oncol.* (2008). doi:10.1016/j.ygyno.2008.02.006

60. Wentzensen, N. *et al.* Ovarian cancer risk factors by histologic subtype: An analysis from the Ovarian Cancer Cohort Consortium. *J. Clin. Oncol.* (2016). doi:10.1200/JCO.2016.66.8178

61. Chiang, Y. C. *et al.* Trends in incidence and survival outcome of epithelial ovarian cancer: 30-year national population-based registry in Taiwan. *J. Gynecol. Oncol.* (2013). doi:10.3802/jgo.2013.24.4.342

62. Kim, S. I. *et al.* Incidence of epithelial ovarian cancer according to histologic subtypes in Korea, 1999 to 2012. *J. Gynecol. Oncol.* (2016). doi:10.3802/jgo.2016.27.e5

63. Lee, K. R. & Young, R. H. The distinction between primary and metastatic mucinous carcinomas of the ovary: Gross and histologic findings in 50 cases. *Am. J. Surg. Pathol.* (2003). doi:10.1097/00000478-200303000-00001

64. Bassiouny, D. *et al.* Comprehensive Clinicopathologic and Updated Immunohistochemical Characterization of Primary Ovarian Mucinous Carcinoma. *Int. J. Surg. Pathol.* (2018). doi:10.1177/1066896917752861

65. Cheasley, D. *et al.* The molecular origin and taxonomy of mucinous ovarian carcinoma. *Nat. Commun.* (2019). doi:10.1038/s41467-019-11862-x

66. Jordan, S. J., Green, A. C., Whiteman, D. C. & Webb, P. M. Risk factors for benign, borderline and invasive mucinous ovarian tumors: Epidemiological evidence of a neoplastic continuum? *Gynecol. Oncol.* (2007). doi:10.1016/j.ygyno.2007.06.006

67. Mandai, M. *et al.* Heterogeneous distribution of K-ras-mutated epithelia in mucinous ovarian tumors with special reference to histopathology. *Hum. Pathol.* (1998). doi:10.1016/S0046-8177(98)90387-2

68. TCGA. Integrated Genomic Analyses of Ovarian Carcinoma The Cancer Genome Atlas Research Network. *Nature* (2011). doi:10.1038/nature10166.Integrated

69. Patch, A. M. *et al.* Whole-genome characterization of chemoresistant ovarian cancer. *Nature* (2015). doi:10.1038/nature14410

70. Macintyre, G. *et al.* Copy number signatures and mutational processes in ovarian carcinoma. *Nat. Genet.* (2018). doi:10.1038/s41588-018-0179-8

71. Graf, R. P., Eskander, R., Brueggeman, L. & Stupack, D. G. Association of Copy Number Variation Signature and Survival in Patients with Serous Ovarian Cancer. *JAMA Netw. Open* **4**, 1–12 (2021).

72. Miller, R. E. *et al.* ESMO recommendations on predictive biomarker testing for homologous recombination deficiency and PARP inhibitor benefit in ovarian cancer. *Ann. Oncol.* (2020). doi:10.1016/j.annonc.2020.08.2102

73. Jones, S. *et al.* Low-grade serous carcinomas of the ovary contain very few point mutations. *J. Pathol.* (2012). doi:10.1002/path.3967

74. Hunter, S. M. *et al.* Molecular profiling of low grade serous ovarian tumours identifies novel candidate driver genes. *Oncotarget* (2015). doi:10.18632/oncotarget.5438

75. Etemadmoghadam, D. *et al.* EIF1AX and NRAS mutations co-occur and cooperate in low-grade serous ovarian carcinomas. *Cancer Res.* (2017). doi:10.1158/0008-5472.CAN-16-2224

76. Cybulska, P. *et al.* Molecular profiling and molecular classification of endometrioid ovarian carcinomas. *Gynecol. Oncol.* (2019). doi:10.1016/j.ygyno.2019.07.012

77. Hollis, R. L. *et al.* Molecular stratification of endometrioid ovarian carcinoma predicts clinical outcome. *Nat. Commun.* (2020). doi:10.1038/s41467-020-18819-5

78. Howitt, B. E. *et al.* Clear cell ovarian cancers with microsatellite instability: A unique subset of ovarian cancers with increased tumor-infiltrating lymphocytes and PD-1/PD-L1 expression. *Oncoimmunology* (2017). doi:10.1080/2162402X.2016.1277308

79. Stewart, C. J. R., Bowtell, D. D. L., Doherty, D. A. & Leung, Y. C. Long-term survival of patients with mismatch repair protein-deficient, high-stage ovarian clear cell carcinoma. *Histopathology* (2017). doi:10.1111/his.13040

80. Wang, Y. K. *et al.* Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. *Nat. Genet.* (2017). doi:10.1038/ng.3849

81. Kim, I. *et al.* Genomic landscape of ovarian clear cell carcinoma via whole exome sequencing. *Gynecol. Oncol.* (2018). doi:10.1016/j.ygyno.2017.12.005

82. Samartzis, E. P., Noske, A., Dedes, K. J., Fink, D. & Imesch, P. ARID1A mutations and PI3K/AKT pathway alterations in endometriosis and endometriosis-associated ovarian carcinomas. *International Journal of Molecular Sciences* (2013). doi:10.3390/ijms140918824

83. Tropé, C. *et al.* Randomized study on adjuvant chemotherapy in stage I high-risk ovarian cancer with evaluation of DNA-ploidy as prognostic instrument. *Ann. Oncol.* (2000). doi:10.1023/A:1008399414923

84. Kristensen, G. B. *et al.* Large-scale genomic instability predicts long-term outcome for women with invasive stage I ovarian cancer. *Ann. Oncol.* (2003). doi:10.1093/annonc/mdg403

85. Zaal, A. *et al.* Genomic aberrations relate early and advanced stage ovarian cancer. *Cell. Oncol.* (2012). doi:10.1007/s13402-012-0077-5

86. Essers, P. B. M. *et al.* Ovarian cancer-derived copy number alterations signatures are prognostic in chemoradiotherapy-treated head and neck squamous cell carcinoma. *Int. J. Cancer* (2020). doi:10.1002/ijc.32962

87. Gao, B. & Baudis, M. Signatures of Discriminative Copy Number Aberrations in 31 Cancer Subtypes. *Front. Genet.* (2021). doi:10.3389/fgene.2021.654887

88. Pattamadilok, J. *et al.* LINE-1 hypomethylation level as a potential prognostic factor for epithelial ovarian cancer. *Int. J. Gynecol. Cancer* (2008). doi:10.1111/j.1525-1438.2007.01117.x

89. Dammann, R. H. *et al.* Frequent aberrant methylation of the imprinted IGF2/H19 locus and LINE1 hypomethylation in ovarian carcinoma. *Int. J. Oncol.* (2010). doi:10.3892/ijo-00000488

90. Zeimet, A. G. *et al.* DNA ploidy, nuclear size, proliferation index and DNA-hypomethylation in ovarian cancer. *Gynecol. Oncol.* (2011). doi:10.1016/j.ygyno.2010.12.332

91. Senthong, A. *et al.* Differences in LINE-1 methylation between endometriotic ovarian cyst and endometriosis-associated ovarian cancer. *Int. J. Gynecol. Cancer* (2014). doi:10.1097/IGC.0000000000000021

92. Ishak, C. A. & de Carvalho, D. D. DNA methylation profiling of premalignant lesions as a path to ovarian cancer early detection. *Clin. Cancer Res.* (2020). doi:10.1158/1078-0432.CCR-20-3331

93. Pisanic, T. R. *et al.* Long Interspersed Nuclear Element 1 Retrotransposons Become Deregulated during the Development of Ovarian Cancer Precursor Lesions. *Am. J. Pathol.* (2019). doi:10.1016/j.ajpath.2018.11.005

94. Marchini, S. *et al.* Analysis of gene expression in early-stage ovarian cancer. *Clin. Cancer Res.* (2008). doi:10.1158/1078-0432.CCR-08-0523

95. Marchini, S. *et al.* Association between miR-200c and the survival of patients with stage I epithelial ovarian cancer: A retrospective study of two independent tumour tissue collections. *Lancet Oncol.* (2011). doi:10.1016/S1470-2045(11)70012-2

96. Calura, E. *et al.* MiRNA landscape in stage i epithelial ovarian cancer defines the histotype specificities. *Clin. Cancer Res.* (2013). doi:10.1158/1078-0432.CCR-13-0360

97. Calura, E. *et al.* Wiring miRNAs to pathways: A topological approach to integrate miRNA and mRNA expression profiles. *Nucleic Acids Res.* (2014). doi:10.1093/nar/gku354

98. Calura, E. *et al.* A prognostic regulatory pathway in stage I epithelial ovarian cancer: New hints for the poor prognosis assessment. *Ann. Oncol.* (2016). doi:10.1093/annonc/mdw210

99. Martini, P. *et al.* LncRNAs as novel indicators of patients' prognosis in stage i epithelial ovarian cancer: A retrospective and multicentric study. *Clin. Cancer Res.* (2017). doi:10.1158/1078-0432.CCR-16-1402

100. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* (2009). doi:10.1093/bioinformatics/btp324

101. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* (2013). doi:10.1038/nbt.2514

102. Lai, Z. *et al.* VarDict: A novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkw227

103. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* (2016). doi:10.1186/s13059-016-0974-4

104. Paila, U., Chapman, B. A., Kirchner, R. & Quinlan, A. R. GEMINI: Integrative Exploration of Genetic Variation and Genome Annotations. *PLoS Comput. Biol.* (2013). doi:10.1371/journal.pcbi.1003153

105. Macintyre, G., Ylstra, B. & Brenton, J. D. Sequencing Structural Variants in Cancer for Precision Therapeutics. *Trends in Genetics* (2016). doi:10.1016/j.tig.2016.07.002

106. Scheinin, I. *et al.* DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res.* (2014). doi:10.1101/gr.175141.114

107. Olshen, A. B., Venkatraman, E. S., Lucito, R. & Wigler, M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* (2004). doi:10.1093/biostatistics/kxh008

108. Poell, J. B. *et al.* ACE: Absolute copy number estimation from low-coverage whole-genome sequencing data. *Bioinformatics* (2019). doi:10.1093/bioinformatics/bty1055

109. Mermel, C. H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* (2011). doi:10.1186/gb-2011-12-4-r41

110. Gaujoux, R. & Seoighe, C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* (2010). doi:10.1186/1471-2105-11-367

111. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* (1995). doi:10.1111/j.2517-6161.1995.tb02031.x

112. Ciriello, G. *et al.* Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* (2013). doi:10.1038/ng.2762

113. Ahmed, A. A. *et al.* Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J. Pathol.* (2010). doi:10.1002/path.2696

114. Chandler, R. L. *et al.* Coexistent ARID1A-PIK3CA mutations promote ovarian clear-cell tumorigenesis through pro-tumorigenic inflammatory cytokine signaling. *Nat. Commun.* **6**, (2015).

115. Wilson, M. R. *et al.* ARID1A and PI3-kinase pathway mutations in the endometrium drive epithelial transdifferentiation and collective invasion. *Nat. Commun.* **10**, (2019).

116. Ballabio, S. *et al.* Multisite analysis of high-grade serous epithelial ovarian cancers identifies genomic regions of focal and recurrent copy number alteration in 3q26.2 and 8q24.3. *J. Oncol.* **145**, 2670–2681 (2019).

117. Milea, A. *et al.* Retinoblastoma pathway deregulatory mechanisms determine clinical outcome in high-grade serous ovarian carcinoma. *Mod. Pathol.* **27**, 991–1001 (2014).

118. Shariska Petersena, Andrew J. Wilsonb, Jeff Hirsta, Katherine F. Robyc, g, Oluwole Fadared, Marta A. Crispense, Alicia Beeghly-Fadiele, f, D. K. CCNE1 and BRD4 co-amplification in high-grade serous ovarian cancer is associated with poor clinical outcomes. *Gynecol Oncol* **157**, 405–410 (2020).

119. Karst, A. M. *et al.* Cyclin E1 deregulation occurs early in secretory cell transformation to promote formation of fallopian tube-derived high-grade serous ovarian cancers. *Cancer Res.* **74**, 1141–1152 (2014).

120. Midorikawa, Y. *et al.* Accumulation of molecular aberrations distinctive to hepatocellular carcinoma progression. *Cancer Res.* (2020). doi:10.1158/0008-5472.CAN-20-0225

121. Möller, K. *et al.* Chromosome 5 harbors two independent deletion hotspots at 5q13 and 5q21 that characterize biologically different subsets of aggressive prostate cancer. *Int. J. Cancer* (2021). doi:10.1002/ijc.33344

122. Huang, J., Zhao, Y. L., Li, Y., Fletcher, J. A. & Xiao, S. Genomic and functional evidence for an ARID1A tumor suppressor role. *Genes Chromosom. Cancer* (2007). doi:10.1002/gcc.20459

123. Tanwar, P. S. *et al.* Loss of LKB1 and PTEN tumor suppressor genes in the ovarian surface epithelium induces papillary serous ovarian cancer. *Carcinogenesis* (2014). doi:10.1093/carcin/bgt357

124. Appay, R. *et al.* CDKN2A homozygous deletion is a strong adverse prognosis factor in diffuse malignant IDH-mutant gliomas. *Neuro. Oncol.* (2019). doi:10.1093/neuonc/noz124

125. Burrell, R. A. *et al.* Replication stress links structural and numerical cancer chromosomal instability. *Nature* (2013). doi:10.1038/nature11935

126. Mohammed, R. A. A. *et al.* Pattern of HER-2 Gene Amplification and Protein Expression in Benign, Borderline, and Malignant Ovarian Serous and Mucinous Neoplasms. *Int. J. Gynecol. Pathol.* **36**, 50–57 (2017).

127. Chao, W. R., Lee, M. Y., Lee, Y. J., Sheu, G. T. & Han, C. P. Comparing the 2017 ASCO/CAP guideline for gastroesophageal adenocarcinoma surgical specimen to the 2018 ASCO/CAP guideline for breast cancer in assessing the HER2 status in

primary mucinous ovarian carcinoma. *Virchows Arch.* (2022). doi:10.1007/s00428-022-03285-9

128. Berenjeno, I. M. *et al.* Oncogenic PIK3CA induces centrosome amplification and tolerance to genome doubling. *Nat. Commun.* (2017). doi:10.1038/s41467-017-02002-4

129. Vanhaesebroeck, B. *et al.* Perspective: Potential impact and therapeutic implications of oncogenic PI3K activation on chromosomal instability. *Biomolecules* (2019). doi:10.3390/biom9080331

130. Hingorani, S. R. *et al.* Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. *Cancer Cell* (2005). doi:10.1016/j.ccr.2005.04.023

131. Hou, S. Q., Ouyang, M., Brandmaier, A., Hao, H. & Shen, W. H. PTEN in the maintenance of genome integrity: From DNA replication to chromosome segregation. *BioEssays* (2017). doi:10.1002/bies.201700082

132. Huang, Y., Ming, X., Li, B. & Li, Z. Histological Characteristics and Early-Stage Diagnosis Are Associated With Better Survival in Young Patients With Epithelial Ovarian Cancer: A Retrospective Analysis Based on Surveillance Epidemiology and End Results Database. *Front. Oncol.* **10**, 1–10 (2020).

133. Macheret, M. & Halazonetis, T. D. Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. *Nature* (2018). doi:10.1038/nature25507

134. Bester, A. C. *et al.* Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell* (2011). doi:10.1016/j.cell.2011.03.044

135. Zhou, J., Zhou, X. A., Zhang, N. & Wang, J. Evolving insights: how DNA repair pathways impact cancer evolution. *Cancer Biology and Medicine* (2020). doi:10.20892/j.issn.2095-3941.2020.0177

136. Ferguson, R. L. & Maller, J. L. Cyclin E-dependent localization of MCM5 regulates centrosome duplication. *J. Cell Sci.* (2008). doi:10.1242/jcs.034702

137. Wang, Q., Zhang, Y. & Yang, H. S. Pdcd4 knockdown up-regulates MAP4K1 expression and activation of AP-1 dependent transcription through c-Myc. *Biochim. Biophys. Acta - Mol. Cell Res.* (2012). doi:10.1016/j.bbamcr.2012.07.004

138. Vang, R. *et al.* Molecular Alterations of TP53 are a Defining Feature of Ovarian High-Grade Serous Carcinoma. *Int. J. Gynecol. Pathol.* (2016). doi:10.1097/pgp.0000000000000207

139. Chapman, P. B. *et al.* Improved Survival with Vemurafenib in Melanoma with BRAF V600E Mutation. *N. Engl. J. Med.* (2011). doi:10.1056/nejmoa1103782

140. Moore, A. R., Rosenberg, S. C., McCormick, F. & Malek, S. RAS-targeted therapies: is the undruggable drugged? *Nature Reviews Drug Discovery* (2020). doi:10.1038/s41573-020-0068-6

141. Goebel, L., Müller, M. P., Goody, R. S. & Rauh, D. KRasG12C inhibitors in clinical trials: A short historical perspective. *RSC Medicinal Chemistry* (2020). doi:10.1039/d0md00096e

142. Chien, J. *et al.* TP53 mutations, tetraploidy and homologous recombination repair defects in early stage high-grade serous ovarian cancer. *Nucleic Acids Res.* **43**, 6945–6958 (2015).

143. Skírnisdóttir, I. A., Sorbe, B., Lindborg, K. & Seidal, T. Prognostic impact of p53, p27, and C-MYC on clinicopathological features and outcome in early-stage (FIGO IYII) epithelial ovarian cancer. *Int. J. Gynecol. Cancer* **21**, 236–244 (2011).

144. Piombino, C. *et al.* Secondary Prevention in Hereditary Breast and/or Ovarian Cancer Syndromes Other Than BRCA. *J. Oncol.* **2020**, (2020).

145. Buensuceso, A. *et al.* Loss of LKB1-NUAK1 signalling enhances NF-κB activity in a

spheroid model of high-grade serous ovarian cancer. *Sci. Rep.* **12**, 1–16 (2022).

146. Rahman, M. T. *et al.* Prognostic and therapeutic impact of the chromosome 20q13.2 ZNF217 locus amplification in ovarian clear cell carcinoma. *Cancer* (2012). doi:10.1002/cncr.26598

147. Jiménez-Sánchez, A. *et al.* Unraveling tumor–immune heterogeneity in advanced ovarian cancer uncovers immunogenic effect of chemotherapy. *Nat. Genet.* (2020). doi:10.1038/s41588-020-0630-5

148. Shen, Y. A. *et al.* Inhibition of the MYC-regulated glutaminase metabolic axis is an effective synthetic lethal approach for treating chemoresistant ovarian cancers. *Cancer Res.* (2021). doi:10.1158/0008-5472.CAN-19-3971

149. Konstantinopoulos, P. A. *et al.* Berzosertib plus gemcitabine versus gemcitabine alone in platinum-resistant high-grade serous ovarian cancer: a multicentre, open-label, randomised, phase 2 trial. *Lancet Oncol.* (2020). doi:10.1016/S1470-2045(20)30180-7

150. Aftab, A. *et al.* CDKN2A/P16INK4A variants association with breast cancer and their in-silico analysis. *Breast Cancer* (2019). doi:10.1007/s12282-018-0894-0

151. Massó-Vallés, D. & Soucek, L. Blocking Myc to Treat Cancer: Reflecting on Two Decades of Omomyc. *Cells* (2020). doi:10.3390/cells9040883

152. Głodzik, D. *et al.* Mutational mechanisms of amplifications revealed by analysis of clustered rearrangements in breast cancers. *Ann. Oncol.* (2018). doi:10.1093/annonc/mdy404

153. Asaka, S. *et al.* Analysis of Telomere Lengths in p53 Signatures and Incidental Serous Tubal Intraepithelial Carcinomas Without Concurrent Ovarian Cancer. *Am. J. Surg. Pathol.* (2019). doi:10.1097/PAS.0000000000001283