Title Page

Subclonal somatic copy number alterations emerge and dominate in recurrent osteosarcoma

Authors

Michael D. Kinnaman[1], Simone Zaccaria[2,3], Alvin Makohon-Moore[4,5,6,7,8], Brian Arnold[9,10], Max Levine[1,11,12], Gunes Gundem[1,11], Juan E. Arango Ossa[1,11], Dominik Glodzik[1,11,13], M. Irene Rodríguez-Sánchez[1,14], Nancy Bouvier[1,15], Shanita Li[1], Emily Stockfisch[1], Marisa Dunigan[16], Cassidy Cobbs[16], Umesh Bhanot[17,18], Daoqi You[1], Michael V. Ortiz[1], Tara O'Donohue[1], Emily Slotkin[1], Leonard H. Wexler[1], Filemon S. Dela Cruz[1], Meera Hameed[17], Julia L. Glade Bender[1], William D. Tap[19], Paul A. Meyers[1], Elli Papaemmanuil[1,11], Andrew L. Kung[1], Christine A Iacobuzio-Donahue[4,5,6]

Affiliations

1. Department of Pediatrics, Memorial Sloan Kettering Cancer Center, New York, NY, USA
2. Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK
3. Computational Cancer Genomics Research Group, University College London Cancer Institute, London, UK
4. Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA
5. David M. Rubenstein Center for Pancreatic Cancer Research, Memorial Sloan Kettering Cancer Center, New York, NY, USA
6. Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York, USA
7. Hackensack Meridian Health Center for Discovery and Innovation, Nutley, NJ, USA (current affiliation)
8. Georgetown University Lombardi Comprehensive Cancer Center, Washington, DC, USA (current affiliation)
9. Department of Computer Science, Princeton University, Princeton, NJ, USA
10. Center for Statistics and Machine Learning, Princeton University, Princeton, NJ, USA
11. Department of Epidemiology & Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA
12. Isabl, New York, NY, USA (current affiliation)
13. Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA (current affiliation)
14. Wunderman Thompson Health, New York, NY, USA (current affiliation)
15. IT and Digital Initiatives, Memorial Sloan Kettering Cancer Center, New York, NY, USA (current affiliation)
16. Integrated Genomics Operation Core, Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA
17. Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

18. Precision Pathology Biobanking Center, Memorial Sloan Kettering Cancer Center, New York, NY, USA
19. Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Running Title: Clonal heterogeneity in osteosarcoma

Keywords: osteosarcoma, intratumor heterogeneity, clonal evolution, somatic copy number alterations, whole-genome sequencing

Funding Sources

- **Award Group**
  - Funder: Rally Foundation
  - Principal Award Recipient(s): Michael Kinnaman
- **Award Group**
  - Funder: Hyundai Hope on Wheels
  - Principal Award Recipient(s): Michael Kinnaman
- **Award Group**
  - Funder: Conquer Cancer Foundation of the American Society of Clinical Oncology/Quad W Foundation
  - Principal Award Recipient(s): Michael Kinnaman
- **Award Group**
  - Funder: National Cancer Institute (NCI)
  - Award Id(s): K12
  - Principal Award Recipient(s): Michael Kinnaman
- **Award Group**
  - Funder: National Institutes of Health
  - Award Id(s): R00 (R00CA229979)
  - Principal Award Recipient(s): Alvin Makohon-Moore

Corresponding Authors: Michael Kinnaman, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065. Phone: 212-639-8856; Fax: 929-321-7101; E-mail: kinnamam@mskcc.org; and Christine A. Iacobuzio-Donahue, Memorial Sloan Kettering Cancer Center, 417 East 68th Street, New York, NY 10065. Phone: 646-888-2239; Fax: 646-888-3235; E-mail: iacobuzc@mskcc.org

Author's Disclosures

M. D. Kinnaman reports grants from the Rally Foundation, Hyundai Hope on Wheels, Conquer Cancer Foundation of the American Society of Clinical Oncology/Quad W Foundation, and the NCI during the conduct of the study. S. Zaccaria is a Cancer Research UK Career Development Fellow (Award Reference RCCCDF-Nov21\100005) and is also supported by Rosetrees Trust (Grant Reference M917). A. Makohon-Moore reports research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R00CA229979 (A.M-M). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. M. Levine

reports equity in Isabl and has filed IP relating to whole genome analytics. W. D. Tap reports consulting fees from Eli Lilly, Daiichi Sankyo, Deciphera, Foghorn Therapeutics, AmMAx Bio, Novo Holdings, Servier, Medpacto, Ayala Pharmaceuticals, Kowa Research Institute, Epizyme (Nexus Global Group), Bayer, Cogent Biosciences, Amgen, Physicians' Education Resource, and BioAlta on the development of novel agents for the treatment of patients with sarcomas; has a patent Companion Diagnostic for CDK4 inhibitors (14/854,329) for the treatment of liposarcoma; holds stock options in and is a member of the scientific advisory boards for Certis Oncology Solutions, a life science technology company working on orthotopic patient-derived xenografts in oncology; and holds stock options in and is a co-founder and director of Atropos Therapeutics, a company that focuses on a drug discovery platform to identify chemical matter that can modulate the formation of senescent cells in relation to ageing and oncology. None of WDT's competing interests are related to the topic discussed here. WDT is the 2022 President of the Connective Tissue Oncology Society. E.P. is a founder, equity holder and has a fiduciary role in Isabl, a cancer whole genome sequencing analytics company. A.L. Kung is on the Scientific Advisory Board of Emendo Biotherapeutics, Karyopharm Therapeutics, Imago BioSciences, and DarwinHealth, is co-founder and on the board of directors of Isabl Technologies, and has equity interest in Imago BioSciences, Emendo Biotherapeutics, and Isabl Technologies. This research was also supported by the Cancer Center Support grant P30 CA008748. No disclosures were reported by the other authors.

Total Word Count: 6028
Figures: 6 main figures, 10 supplemental figures

Abstract

Multiple large-scale tumor genomic profiling efforts have been undertaken in osteosarcoma; however, little is known about the spatial and temporal intratumor heterogeneity and how it may drive treatment resistance. We performed whole-genome sequencing of 37 tumor samples from eight patients with relapsed or refractory osteosarcoma. Each patient had at least one sample from a primary site and a metastatic or relapse site. We identified subclonal copy number alterations in all but one patient. We observed that in 5 patients, a subclonal copy number clone from the primary tumor emerged and dominated at subsequent relapses. *MYC* gain/amplification was enriched in the treatment-resistant clone in 6/7 patients with more than one clone. Amplifications in other potential driver genes, such as *CCNE1*, *RAD21*, *VEGFA*, and *IGF1R*, were also observed in the resistant copy number clones. Our study sheds light on intratumor heterogeneity and the potential drivers of treatment resistance in osteosarcoma.

Significance

Subclonal copy number clones emerged and dominated in relapsed osteosarcoma, with *MYC* gain/amplification being the defining characteristic in our cohort. Selective pressure from neoadjuvant chemotherapy revealed this clone at the time of primary resection, highlighting that genomic profiling at this time may identify clones that are selected for, or determine innate resistance to primary chemotherapy.

Introduction

Osteosarcoma is an aggressive bone tumor which primarily affects children and young adults. Patients who present with metastatic disease at diagnosis have a poor overall prognosis and those with an inferior response to neoadjuvant chemotherapy have a high risk for recurrence[1–3]. Multiple large-scale tumor genomic profiling efforts have been undertaken to describe the genomic underpinnings and identify new potential therapeutic targets for osteosarcoma[4–9]. These studies revealed that osteosarcoma, typically characterized by a high degree of chromosomal instability, has a large number of chromosomal deletions, translocations, and amplifications. The common alterations present in osteosarcoma primarily involve tumor suppressor genes (e.g., *TP53*, *RB1*, *PTEN*), whereas targetable activating mutations are rare, making it challenging to link the mutational genotype to a broadly applicable treatment strategy[10,11]. However, recent studies have suggested that targeting focal gene amplifications in consensus driver genes may be an effective strategy for identifying precision-based therapies[4,12].

Recent genomic studies in osteosarcoma have suggested that metastatic clones do not correspond to the dominant clones present in the primary tumor,[13] and osteosarcoma may evolve via parallel evolution,[14] with evidence for both monoclonal[15] and polyclonal synchronous seeding of metastases[13,14]. Copy number alterations in consensus driver genes, such as *MYC* and *CDK4* were found to be likely early events[15]. Cisplatin-induced mutagenesis has also been highlighted as a potential driver of treatment resistance in recurrent osteosarcoma[15]. Despite these initial insights into the clonal heterogeneity of osteosarcoma, the extent to which neoadjuvant chemotherapy affects clonal selection in patients with a poor response to chemotherapy and the degree to which copy number alterations evolve from diagnosis to relapse remains unclear.

To address these open questions about clonal selection and heterogeneity in osteosarcoma, we performed whole-genome sequencing of 37 spatially and temporally separated tumor samples from eight patients with osteosarcoma who had a poor response to neoadjuvant chemotherapy (<90% necrosis). We describe spatial intermetastatic heterogeneity and temporal clonal evolutionary processes, with a focus on identifying and tracking unique copy number clones from diagnosis through relapse.

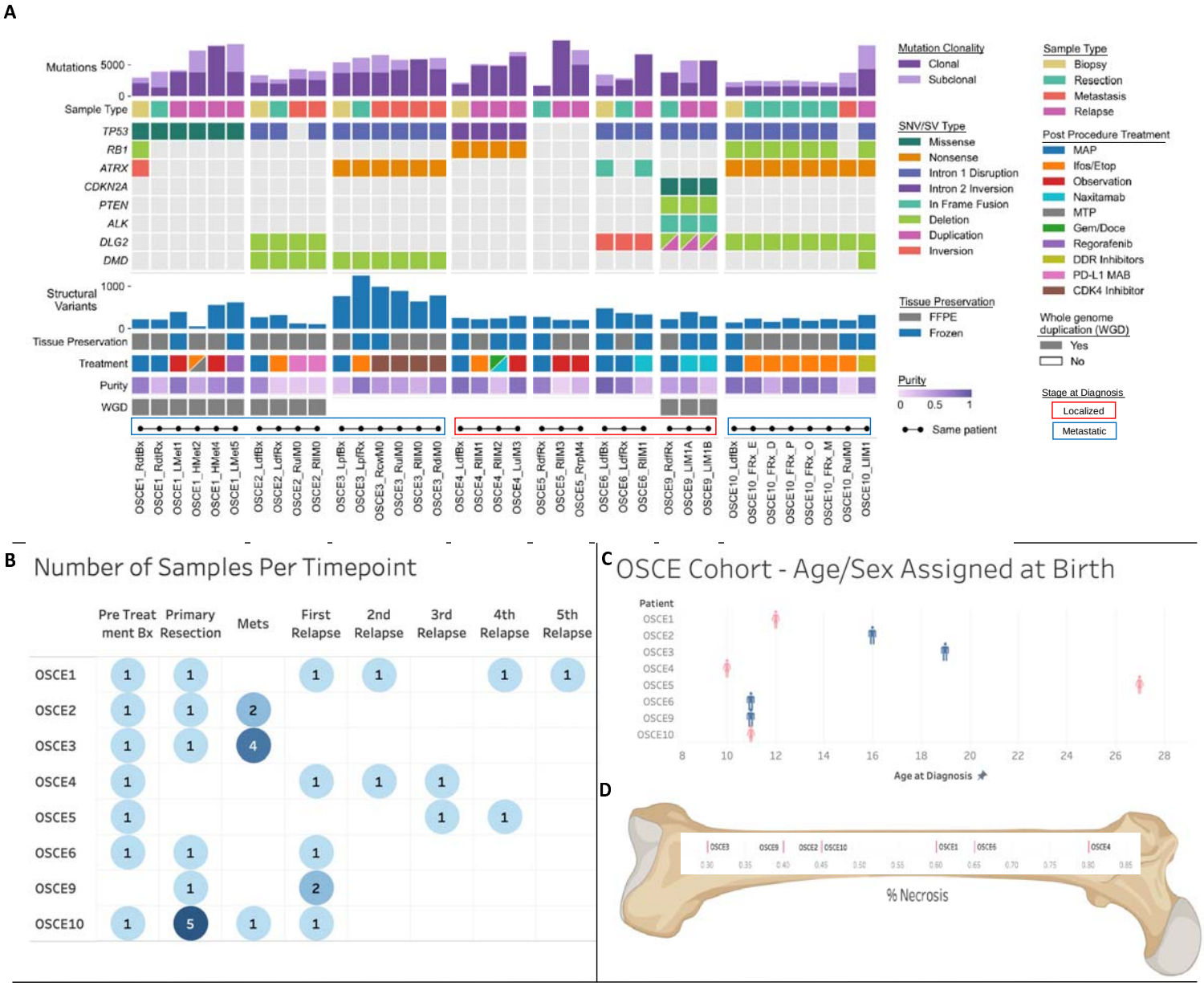
Results

Analysis of Single Nucleotide Variants reveal limited driver gene heterogeneity in temporally and spatially distinct osteosarcoma sample

To analyze clonal evolution and intratumoral heterogeneity in osteosarcoma, we performed whole-genome sequencing (WGS) of tumor tissues from multiple spatially and temporally distinct samples from eight individuals with relapsed/refractory osteosarcoma. DNA was extracted from 84 samples collected from 10 patients. After initial quality control, we sequenced 62 unique tumor samples from eight patients with WGS to a target depth of 80x (Supplementary Fig. 1). Of these eight patients, four had localized disease at diagnosis (OSCE4, OSCE5, OSCE6, OSCE9) and four had metastatic disease at diagnosis (OSCE1, OSCE2, OSCE3, OSCE10) (Fig. 1A), and the age at diagnosis was 11-27 years (four girls and four boys, Fig. 1C). All patients were treated at the Memorial Sloan Kettering Cancer Center and received methotrexate, cisplatin, and doxorubicin (MAP) chemotherapy (subsequent post-procedure treatment, Fig. 1A). Seven of the eight patients had a poor response to neoadjuvant chemotherapy (<90% necrosis at the time of primary resection, Fig. 1D), while OSCE5 had an upfront resection; therefore, the response to therapy could not be evaluated.

After sequencing was completed, we reviewed the quality of the sequencing data (purity, sequencing coverage) and determined that 37 of the 62 (59%) samples, similar to other studies[16], met our quality control requirements to proceed with further downstream analysis (Supplementary Fig. 1). Of these 37 samples, 17 came from primary sites, with seven of eight patients having a pretreatment sample, six of which were biopsies and one (OSCE5), which was a pretreatment primary resection (Fig. 1A/B). Of the seven patients who did not undergo an

Figure 1


A
Mutations
Sample Type
TP53
RB1
ATRX
CDKN2A
PTEN
ALK
DLG2
DMD
Structural Variants
Tissue Preservation
Treatment
Purity
WGD
Mutation Clonality
Clonal
Subclonal
SNV/SV Type
Missense
Nonsense
Intron 1 Disruption
Intron 2 Inversion
In Frame Fusion
Deletion
Duplication
Inversion
Tissue Preservation
FFPE
Frozen
Purity
Same patient
Sample Type
Biopsy
Resection
Metastasis
Relapse
Post Procedure Treatment
MAP
Ifos/Etop
Observation
Naxitamab
MTP
Gem/Doce
Regorafenib
DDR Inhibitors
PD-L1 MAB
CDK4 Inhibitor
Whole genome duplication (WGD)
Yes
No
Stage at Diagnosis
Localized
Metastatic
B
Number of Samples Per Timepoint
Pre Treatment Bx
Primary Resection
Mets
First Relapse
2nd Relapse
3rd Relapse
4th Relapse
5th Relapse
C
OSCE Cohort - Age/Sex Assigned at Birth
Patient
Age at Diagnosis
D
% Necrosis
E
Disease Course Over Time
Patient
Diagnosis
Time Until Event (Years)
Event
Primary Resection
Metastatectomy
Relapse 1
Relapse 2
Relapse 3
Relapse 4
Relapse 5
Disease Free Interval
Death

180

Figure 1. Characteristics of the patients and samples included in the analysis cohort. **A,** Oncoprint of sample and patient level details for each patient. Samples from the same patient are connected by dots and lines on bottom of figure and are in chronological order of time obtained (earliest on left). Sample Name Key: Bx = biopsy, Rx = Resection, Sample Ending in 0 = metastatic site present at diagnosis, sample ending in number >0 = number relapse. L/R = laterality, d = distal, p=proximal, f=femur, t=tibia, cw=chest wall, H=heart, ul=upper lobe, ll=lower lobe, di=diaphragm, rp=retroperitoneal, l=lobe. **B,** Summary of the number of samples per timepoint for each patient, darker shades of blue represent higher number of samples at a respective time point. **C,** Age and sex assigned at birth for each patient, patient is on the x-axis, age is on the y-axis, and sex assigned at birth is plotted on the chart as blue for male and pink for female. **D,** Percent necrosis at time of primary resection for each patient (Note OSCE2/OSCE10 both have 45% necrosis). **E,** Patients are listed on the y-axis, disease course are plotted with time to an event (in years). Events are marked as the time from the previous event, with the timing of the highlighted event marked by the right side of the bard. When looking at a colored bar, the event happens at the timepoint marked by the right side of the bar.

upfront resection, all had at least one on-therapy resection sample, and one patient (OSCE10) had multi-region sampling from the primary tumor (Fig. 1A/B). The other 20 samples came from metastatic sites, 15 of which were from lung metastases, and 7 of the 20 were metastatic sites that were present at diagnosis (Fig. 1A/B). Fresh frozen samples accounted for 15 of the 37 samples selected for downstream analysis, while the remaining 22 were formalin fixed paraffin embedded (FFPE) samples (Fig. 1A). The median purity of fresh frozen samples was 0.75 compared to 0.46 of FFPE samples (Fig. 1A). All eight patients in this analysis had matched normal blood samples sequenced at a target depth of 40x.

The clinical courses of OSCE2, OSCE3, and OSCE10 were defined as refractory disease with progression on MAP (methotrexate, doxorubicin, cisplatin) chemotherapy and extremely virulent disease courses, with time from diagnosis to death of 1.08, 1.3, and 1.75 years respectively (Fig. 1E). OSCE1 and OSCE5 had long protracted relapsing and remitting disease courses, with a time from diagnosis to death of seven and six years, respectively (Fig. 1E). OSCE4 also had a relapsing and remitting disease course but is currently in remission 8 years after diagnosis (Fig. 1E). OSCE6 is alive 4.5 years after diagnosis but has had a recent recurrence (Fig. 1E). OSCE9 is 4 years out from the initial metastatic recurrence (Fig. 1E).

After filtering and germline subtraction, whole-genome sequencing data identified between 1684 and 16,215 single nucleotide variants (SNVs) per sample (Fig. 1A). Of these, there were between 12 and 181 coding nonsynonymous SNVs per sample (median = 54, Supplementary Fig. 2A). The average number of nonsynonymous SNVs in primary site samples was 42, compared to 73 in metastatic or relapsed samples. SNVs were clustered across all samples for each patient using the DeCiFer[17] algorithm, which determines the descendant cell fraction (DCF) of all SNVs for a given cluster in each patient (analogous to the cancer cell

fraction but accounting for potential mutation losses[17]). Following previous approaches[17], SNV's were categorized as clonal if they belonged to a cluster in a sample with a DCF ≥ 90%, and subclonal if they belonged to a cluster with a DCF < 90%. The proportion of clonal SNV's ranged from 35.7-100% across all samples (median = 68.2%), with relapse samples having the highest proportion (median = 92.7%) of clonal SNVs compared with biopsy, resection, and metastatic samples (median = 66.1%, 64.7%, and 63.6%, respectively) (Fig. 1A and Supplementary Fig. 1B). Likely functional driver gene SNVs[18] were identified in five of the eight patients, including genes known to be frequently mutated in osteosarcoma, such as *TP53*, *ATRX*, *RB1*, and *CDKN2A* (Fig. 1A)[5–9]. These driver genes were clonal (shared) across all samples for each patient, and no new SNVs were likely functional drivers that were unique to any metastatic or recurrent samples (Fig. 1A, Supplementary Fig. 3A).

SNVs were clustered across all samples for each patient using the DeCiFer[17] algorithm, which determines the descendant cell fraction (DCF) of all SNVs for a given cluster in each patient (analogous to the cancer cell fraction but accounting for potential mutation losses[17]). Following previous approaches[17], SNV's were categorized as clonal if they belonged to a cluster in a sample with a DCF ≥ 90% and subclonal if they belonged to a cluster with a DCF < 90%. The proportion of clonal SNV's ranged from 35.7-100% across all samples (median = 68.2%), with relapse samples having the highest proportion (median = 92.7%) of clonal SNVs compared with biopsy, resection, and metastatic samples (median = 66.1%, 64.7%, and 63.6%, respectively) (Fig. 1A and Supplementary Fig. 1B).

Since there were no new SNVs in consensus driver genes unique to relapse or metastatic samples, we next examined structural variations and copy number alterations. Structural variants (SVs) in consensus driver genes were shared across all samples for each patient. *TP53* structural

variants involving intron 1 were observed in 5/8 patients (OSCE2, OSCE3, OSCE6, OSCE9, and OSCE10; Fig. 1A). A *TP53* intron 2 inversion was observed in OSCE4 (Fig. 1A). There was a *TP53* SNV in OSCE1 and while there was no *TP53* SNV or SV found in OSCE5 (Fig. 1B), there was amplification of *MDM2*, an important negative regulator of *TP53*. In OSCE1, there was a deletion event in *RB1* and an inversion in *ATRX* in the pretreatment sample (RdtBx) that was not seen in the primary resection or relapse samples (Fig. 1A). Disruptions in *DLG2* were observed in 4 patients (OSCE2, OSCE6, OSCE9, and OSCE10; Fig. 1A). Deletion events in *DMD* were observed in the three refractory cases (OSCE2, OSCE3, and OSCE9) and were present in all samples for OSCE2 and OSCE3; however, they were only detected in the relapse sample in OSCE10 (Fig. 1A). OSCE9 was found to have a deletion event in *PTEN* and an in-frame fusion event in *ALK* (Fig. 1A).

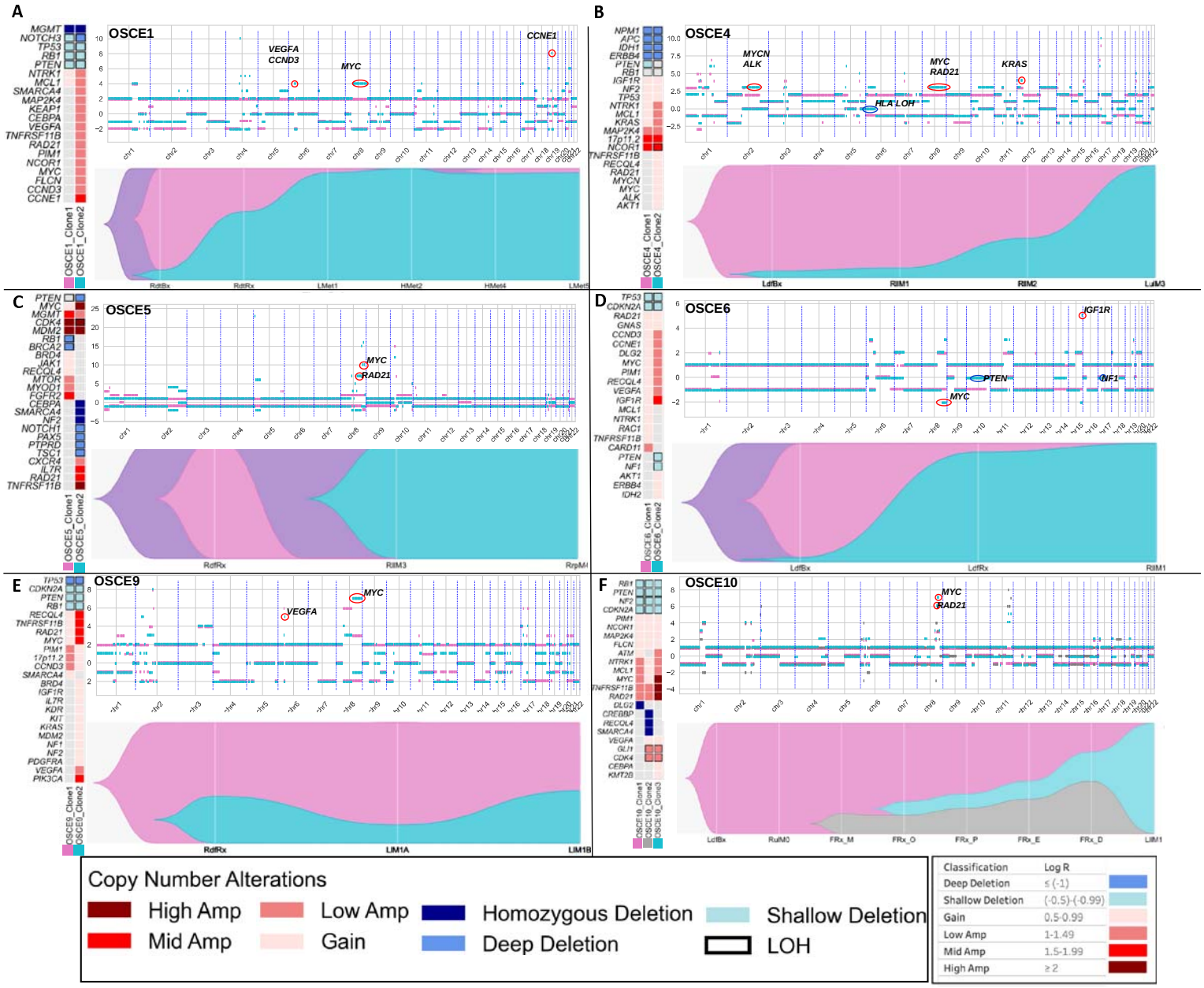
Subclonal Somatic Copy Number Alterations Emerge and Dominate in Recurrent Osteosarcoma

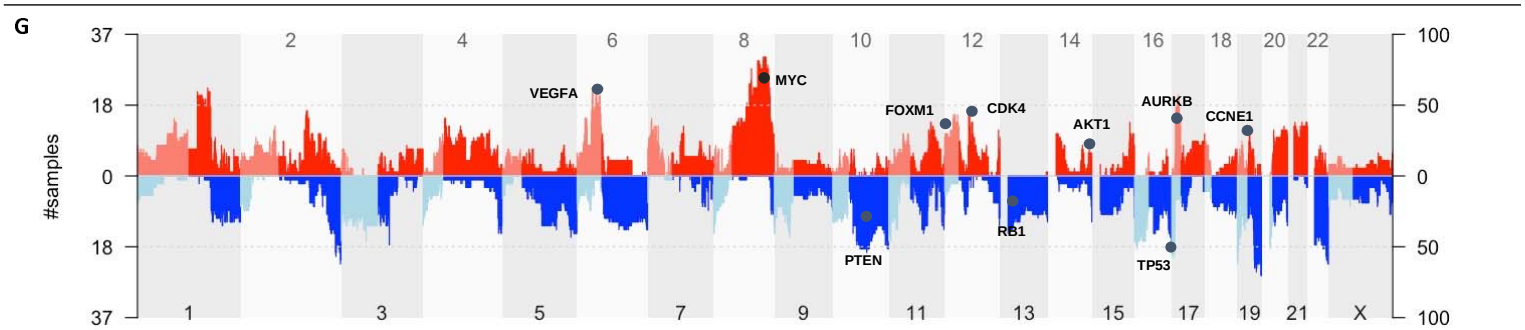
A high prevalence of somatic copy number alterations (CNAs) in osteosarcoma has been reported previously[5–9]. Therefore, we used the HATCHet[19] algorithm to infer both allele and clone-specific CNAs as well as their relative proportions across multiple samples from the same patient. The average tumor ploidy for each sample ranged from 1.7 in OSCE10 to 3.15 in OSCE3 (Supplementary Fig. 2B). HATCHet[19] identified subclonal CNAs in all but one patient (OSCE2) and whole genome duplications present at diagnosis in three of the eight patients (OSCE1, OSCE2, and OSCE9; Fig. 1A). Among the seven patients with subclonal CNAs, six were identified as having two major copy number clones (Fig. 2A-E, Supplementary Fig. 4A), and one patient (OSCE10) had three distinct copy number clones (Fig. 2F). In four of the seven

patients (OSCE1, OSCE4, OSCE6, and OSCE10), multiple distinct copy-number clones were identified to be simultaneously present in the primary tumor (Fig. 2A, 2B, 2D, 2F), but only one of these subclones emerged and dominated at subsequent relapses. In two of these patients (OSCE1, OSCE6), the emergence of this clone was observed in the primary resection sample when compared to the pretreatment biopsy for each patient (Fig. 2A, 2D). In OSCE5, a new copy number clone emerged in the late relapse samples, which was not identified as being present in the pretreatment sample (Fig. 2C). However, when combining the analysis of mutations and CNAs (see Online Methods), the dominant SNV-based clone in the relapse samples was found to be sub-clonal in the pretreatment sample, providing evidence that the dominant copy number clone in the relapse sample for OSCE5 was present at diagnosis (Supplementary Fig. 5A). In OSCE10, no subclonal copy number aberrations were identified in the pretreatment sample; however, a subclonal copy number clone was detected in the primary resection sample that emerged and dominated at relapse in this patient (Fig. 2F). In summary, we found that in most cases, a minor subclone present in the primary tumor emerged and dominated in patients with relapsed disease.

In 2/3 refractory cases (OSCE2, OSCE3), there was no subclonal copy number clone that was present in the primary that emerged and dominated in metastatic or relapse samples. In OSCE2, only a single major copy number clone was identified; in OSCE3, two copy number clones were identified in the pretreatment sample, with the copy number subclone emerging and dominating in one of the metastatic sites but in mixed proportion in the three other metastatic sites. OSCE9 did not have a pretreatment sample for comparison but did show two copy number clones in mixed proportions in the primary resection sample and the two relapse sites.

Figure 2


A
OSCE1
VEGFA
CCND3
MYC
CCNE1
B
OSCE4
MYCN
ALK
MYC
RAD21
KRAS
HLA LOH
C
OSCE5
MYC
RAD21
D
OSCE6
IGF1R
PTEN
NF1
MYC
E
OSCE9
VEGFA
MYC
F
OSCE10
MYC
RAD21
Copy Number Alterations
High Amp
Mid Amp
Low Amp
Gain
Homozygous Deletion
Deep Deletion
Shallow Deletion
LOH
Classification
Log R
Deep Deletion
≤ (-1)
Shallow Deletion
(-0.5)-(-0.99)
Gain
0.5-0.99
Low Amp
1-1.49
Mid Amp
1.5-1.99
High Amp
≥ 2


G
#samples
VEGFA
MYC
FOXM1
CDK4
AKT1
AURKB
CCNE1
RB1
PTEN
TP53

287

Figure 2. Subclonal copy number clones emerge at relapse. **A, B, C, D, E, and F,** For each patient there is a panel of three figures. The figure on the left is an oncoprint featuring clone specific copy number alterations in recurrently altered genes of interest in osteosarcoma. The top figure is a plot of allele specific copy number alterations for each clone with significant events for each clone circled and highlighted. Clone 1 is the magenta clone, clone 2 is the teal clone, and clone 3 in OSCE10 is the gray clone. The major allele is plotted above 0 and the minor allele is plotted below 0. The bottom figure in each panel is a TimeScape plot of the prevalence of each clone at different timepoints throughout a patient's disease course. **G,** Combined genome-wide copy number alterations across all patients in the cohort with recurrently altered genes highlighted.

Determinations regarding branched vs. parallel evolution (depicted in the TimeScape plots in Fig. 2A-F, Supplementary Fig. 4A) for patients with ≥2 copy number clones were based on a review of loss of heterozygosity (LOH) events in the dominant metastatic or recurrent copy number clone, using the rationale and methods outlined by Watkins et al[20]. Branched evolution with the emergence of the treatment resistant copy number clone from the dominant pretreatment copy number clone (clone 2 emerging from clone 1) was observed in three patients (OSCE4, OSCE9, and OSCE10). Parallel evolution, where the pretreatment and treatment resistant copy number clones share the same parent clone (clone 1 and clone 2 share the same parent clone) but evolve in parallel, was seen in four patients (OSCE1, OSCE3, OSCE5, OSCE6). LOH events were common in tumor suppressor genes such as *TP53, RB1, and PTEN* and were mostly shared between clones for each patient, with a median of 77.95% of LOH events shared between all clones (range 18.9%-86.82%, Supplementary Fig. 4C).

Copy number amplifications in recurrently altered oncogenes in osteosarcoma characterize chemoresistant copy number clones

Cohort-wide copy number alterations reflected what has been previously described in osteosarcoma[4], with gains and amplifications seen in *VEGFA*, *MYC*, *FOXM1*, *CDK4*, *AKT1*, *AURKB*, and *CCNE1*, and loss/deletion events in *PTEN*, *RB1*, and *TP53* (Fig. 2G). In contrast to previously identified SNV drivers, the relative proportions of alterations across tumor cells changed with time. Clones were classified as chemoresistant if they were present in the primary site and became dominant at relapse, and chemosensitive if it was the dominant clone in the primary site and became subclonal or eliminated in metastatic or relapse sites. When comparing the genomic alterations between copy number clones for each patient, deletion or LOH events in

tumor suppressor driver genes were often clonal in patients found to have ≥ 2 copy number clones (Fig. 2A-F, Supplementary Fig. 4A). In the 5 patients with clear emergence of a chemoresistant copy number clone (clone 2 in OSCE1, OSCE4, OSCE5, OSCE6 and clone 3 in OSCE10), the resistant clone had a higher degree of *MYC* gain or amplification then the dominant chemosensitive clone at diagnosis (Fig. 2A-F, Supplementary Fig. 5B). In OSCE10, high-level MYC-amplified clone 3 emerged at the time of primary resection and dominated at relapse (Fig. 2F). Notably, this treatment-resistant clone was present in 4/5 multi-region samples from the primary resection, suggesting intratumoral heterogeneity regarding copy number alterations depending on the site sampled. In addition to *MYC*, amplification of *CCNE1*/*CCND3* (OSCE1, OSCE6; Fig. 2A/D), *KRAS* (OSCE2, OSCE4; Supplementary Fig. 4B, Fig. 2), *IGF1R* (OSCE6), *CDK4* (OSCE10), and *VEGFA* (OSCE1, OSCE6, OSCE9) uniquely characterized treatment-resistant or metastatic copy number clones in these examples.

Subclonal selection/emergence evident at the time of primary resection

Three patients within our cohort had a pretreatment sample on therapy resection, and at least one relapse sample (OSCE1, OSCE6, OSCE10), which provided an opportunity to analyze the effect of neoadjuvant chemotherapy on subclonal copy number selection at the time of primary resection and whether this selection is reflective of the dominant clone at the time of relapse. OSCE1 and OSCE6 share a similar pattern, where a copy number subclone present at diagnosis emerges as the dominant clone at the time of primary resection and continues to dominate at the first and subsequent relapses. Chemoresistant clone 2 was characterized by mid-level *CCNE1* amplification in OSCE1 (Fig. 2A) and mid-level *IGF1R* amplification in OSCE6 (Fig. 2D). A slightly different pattern emerged in the refractory case of OSCE10, in which only a

single copy number clone was present at diagnosis. The primary resection sample underwent multi-region sequencing of five spatially separated sites, which revealed the emergence of two new copy number clones, with clone 2 present in all five samples and clone 3 in 4/5 samples. Clone 3 then became the dominant clone in the first relapse specimen, characterized by high-level *MYC* amplification.

Dormant copy number clone characterized by *KRAS* amplification and LOH at the HLA locus emerges after antibody therapy

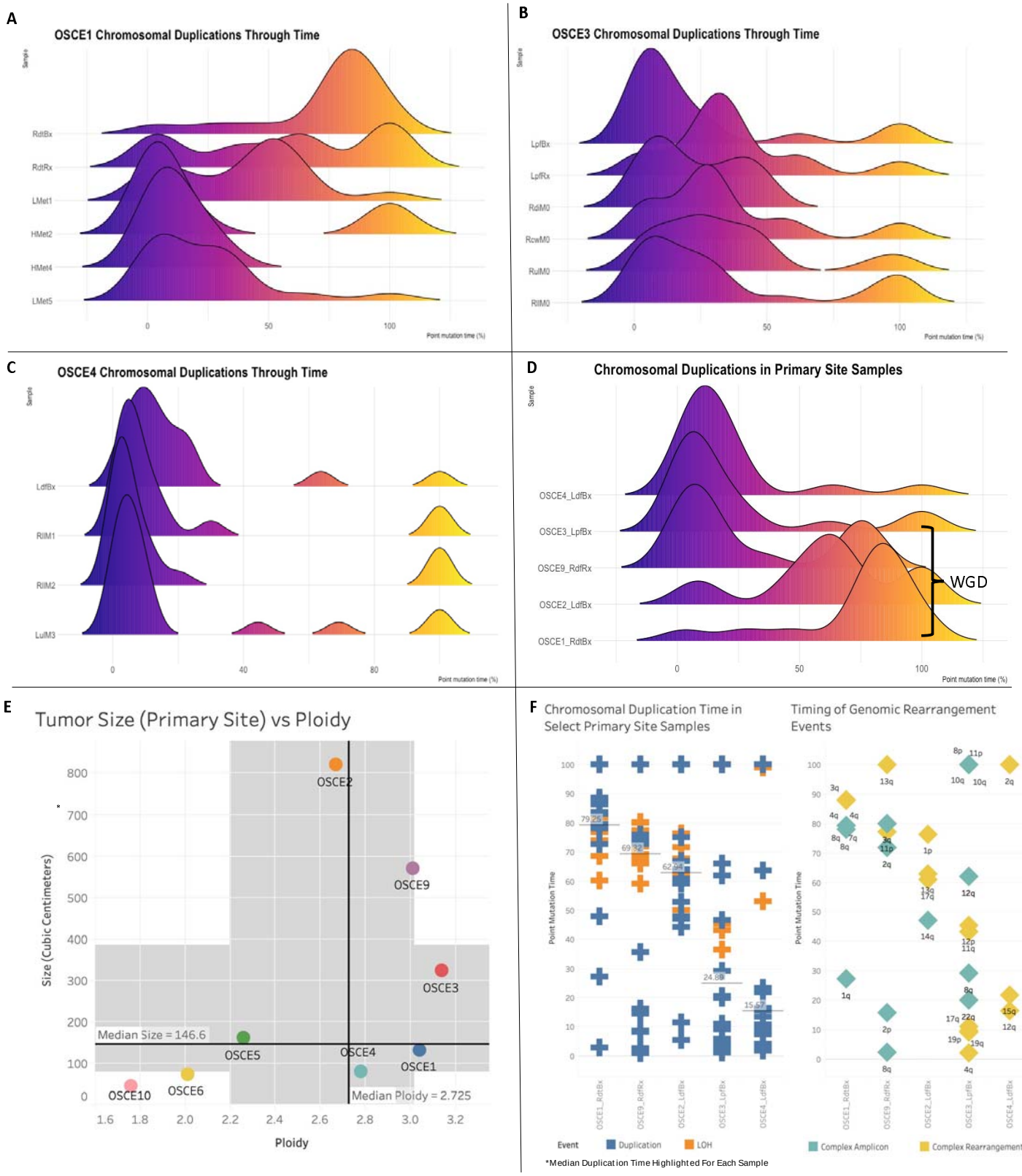
Within our cohort, three patients had ≥2 relapse samples from different time points (OSCE1, OSCE4, and OSCE5). While a single dominant copy number clone characterizes the relapse samples in OSCE1 and OSCE5 (Fig. 2A and 2C), the emergence of clone 2 in OSCE4 followed a uniquely different pattern, as it slowly emerged, and finally dominated in the third relapse sample. A potential explanation for this pattern is that clone 2 had an LOH event at the HLA locus (Fig. 2B, Supplementary Fig. 4D), which may have been selected because of monoclonal antibody therapy the patient received in the adjuvant setting between relapse 2 and relapse 3. LOH at the HLA locus has previously been described as an immune escape mechanism in patients undergoing immune-based therapies[21]. There were no other examples where the HLA LOH status was different between clones with different copy numbers, with three patients having LOH at HLA shared across all clones (OSCE1, OSCE9, OSCE10; Supplementary Fig. 4D) and four patients with no evidence of LOH at HLA (OSCE2, OSCE3, OSCE5, OSCE6; Supplementary Fig. 4D). The only patient with a post-immune-based treatment sample was OSCE1, who received mifamurtide after resection of the second relapse sample, HMet2. Also notable is *KRAS* amplification, which characterizes this late-emerging clone 2.

KRAS is not known to be recurrently altered in osteosarcoma; however, in our small cohort, it was also found to be amplified in the refractory case of OSCE2.

Timing of Copy Number Gains Reveals Large Bursts of Copy Number Gains Before Diagnosis

The HATCHet algorithm also infers whole-genome duplication (WGD) events jointly across all samples for each patient. Of the five patients with average ploidy ≥ 2.5 (OSCE1, OSCE2, OSCE3, OSCE4, OSCE9), WGD was identified in OSCE1, OSCE2, and OSCE9 (Fig. 1A and Supplementary Fig. 2B). The timing of these WGD events as well as other chromosomal duplications can be determined in "molecular time" using previously described methods which compares the number of duplicated vs non-duplicated mutations to estimate the timing of each duplication (Supplementary Fig. 6)[22,23]. For the patients who were determined to have WGD events, these appeared to be late events for each respective patient (median point mutation time (pmt) range = 62-79%, Fig. 3F), compared to the two patients with ploidy ≥2.5 without WGD, which appeared to have early synchronous duplications (median pmt range = 16-25%, Fig. 3F). The two patients with the lowest ploidy (OSCE10 and OSCE6) also had the smallest primary tumors (Fig. 3E), whereas all large primary tumors (≥300 cm^3) had ploidy ≥ 2.5 (OSCE2, OSCE9, OSCE3). For the five patients with ploidy ≥ 2.5, we analyzed their earliest primary site samples to assess the natural history of these duplication events in the context of tumorigenesis. Most of these duplication/LOH events clustered in a single burst of events (Fig. 3D) and were associated with complex rearrangement (chromoplexy, chromothripsis) or complex amplicon[24] (Tyfonas, breakage fusion bridge, double minute) events (Fig. 3F). When comparing longitudinal samples from the same patient (Fig. 3A-C), there was no subsequent burst that was

Figure 3


A
OSCE1 Chromosomal Duplications Through Time
Sample
RdtBx
RdtRx
LMet1
HMet2
HMet4
LMet5
0
50
100
Point mutation time (%)
B
OSCE3 Chromosomal Duplications Through Time
Sample
LpfBx
LpfRx
RdlM0
RcwM0
RulM0
RllM0
0
50
100
Point mutation time (%)
C
OSCE4 Chromosomal Duplications Through Time
Sample
LdfBx
RllM1
RllM2
LulM3
0
40
80
Point mutation time (%)
D
Chromosomal Duplications in Primary Site Samples
Sample
OSCE4_LdfBx
OSCE3_LpfBx
OSCE9_RdfRx
OSCE2_LdfBx
OSCE1_RdtBx
WGD
0
50
100
Point mutation time (%)
E
Tumor Size (Primary Site) vs Ploidy
Size (Cubic Centimeters)
800
700
600
500
400
300
200
100
0
OSCE2
OSCE9
OSCE3
Median Size = 146.6
OSCE5
OSCE4
OSCE1
OSCE10
OSCE6
Median Ploidy = 2.725
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
Ploidy
F
Chromosomal Duplication Time in Select Primary Site Samples
Timing of Genomic Rearrangement Events
Point Mutation Time
79.25
69.32
62.94
24.88
15.57
OSCE1_RdtBx
OSCE9_RdfRx
OSCE2_LdfBx
OSCE3_LpfBx
OSCE4_LdfBx
Event
Duplication
LOH
Complex Amplicon
Complex Rearrangement
*Median Duplication Time Highlighted For Each Sample

390

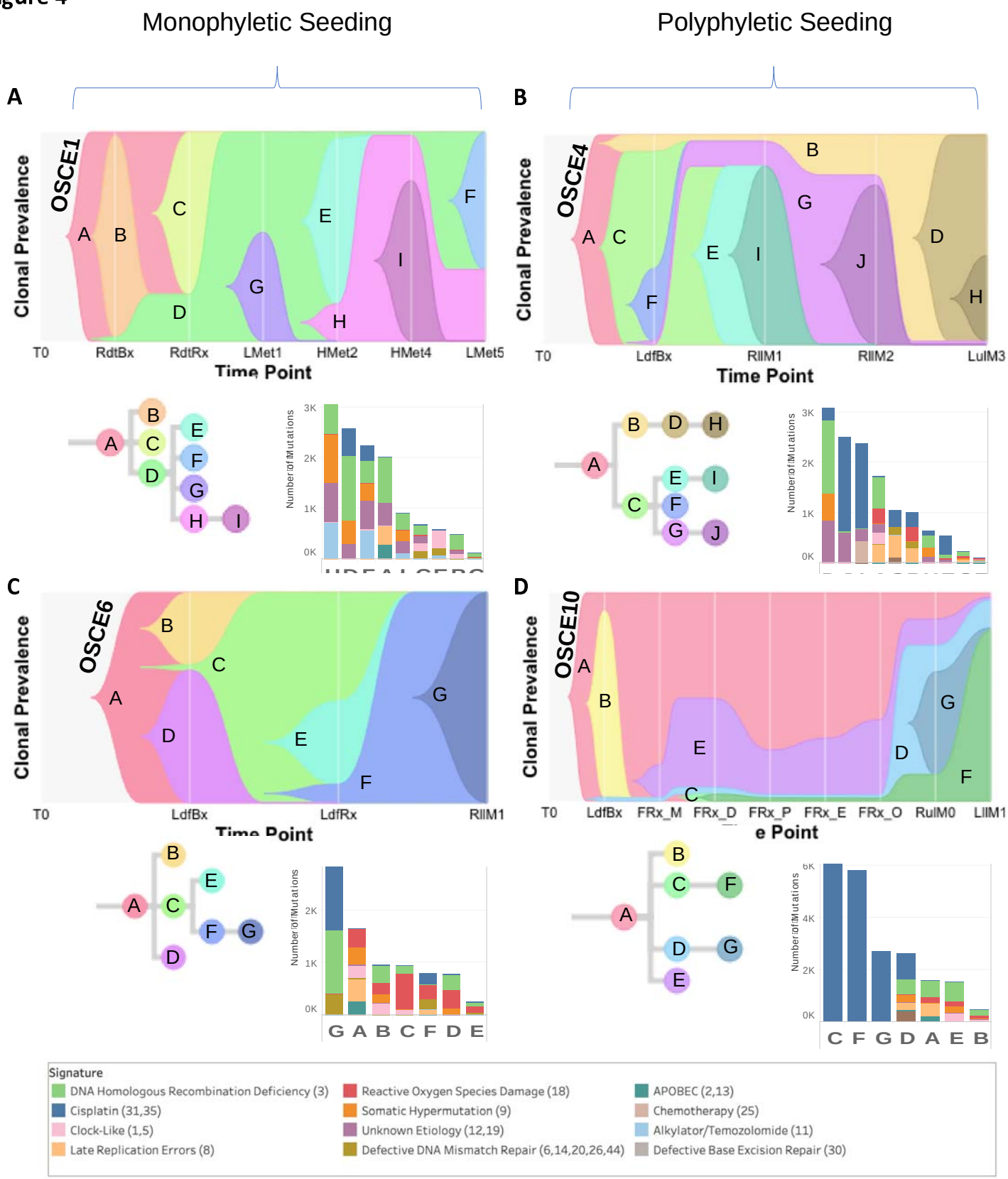
Figure 3. Chromosomal duplication timing analysis reconstructs evolutionary past of genomic instability events. **A, B, C,** Ridgeline plots of the density of duplication events over molecular time for each sample for the selected patients. Notably the highest peak in duplications occurs before diagnosis. **D,** Chromosomal duplication timing of 5 primary site samples. Whole genome duplication events as called by HATCHet appear to be a late event in our cohort. **E,** Plot of tumor size (by volume) verse ploidy called by HATCHet. Median size (y-axis) and ploidy (x-axis) values are plotted with dark black lines, with the shaded gray areas representing the range between the lower and upper quartile for each metric. **F,** Plot of duplication and rearrangement events in molecular time. Left side of figure is plot of duplication events in blue and LOH events in orange for select primary site samples. The median duplication time is highlighted for each sample. The plot on the right side of the figure are complex amplicon events in teal and complex rearrangement events in yellow. Y-axis for both plots is molecular time. The samples are in the same order for each plot. Each plotted event represents an affected chromosomal arm.

equal to or greater than the primary site burst. Across the cohort, duplication events appeared to be fixed during tumorigenesis and had decreasing average molecular times when comparing biopsy/resection samples with metastatic/relapse samples (Supplementary Fig. 7A/C). LOH events were consistently "late" events, with a median molecular time of 82.16 across the cohort (compared to 21.72 for duplication events, Supplementary Fig. 7B/D).

Heterogenous seeding patterns observed in metastatic and recurrent osteosarcoma

Clone-based phylogenies were created to explore the clonal architecture and track the spatial and temporal patterns of evolution. The SNVs across all samples for each patient were clustered using the DeCiFer[17] algorithm. After clustering, each SNV was assigned to a clone with an estimated descendant cell fraction (DCF) per sample, and clone-based phylogenetic trees were then constructed using the CALDER[25] algorithm, allowing for the assessment of modes of metastatic seeding and dissemination. The median number of clones per patient was eight (range = 5-12, Supplementary Fig. 8I). At the patient level, there was a heterogeneous mix of dissemination patterns (Fig. 4A-D, Fig. 5A-D, and Supplementary Fig. 8A-H), with complete agreement between clone and sample-based phylogenies (Supplementary Fig. 3A) regarding whether metastatic dissemination was monophyletic (all metastatic clones shared a common ancestral clone: Fig. 4A, 4C, 5E; Supplementary Fig. 6A, 6C, 6D, 6G, 6H)) vs. polyphyletic in origin (where ≥2 clones are present from distinct branches of phylogeny whose common ancestor represents the trunk of the primary tumor tree: Fig. 4B, 4D, 5B; Supplementary Fig. 6B, 6D, 6F). Of note, OSCE6 was the only patient among those with a monophyletic origin that had monoclonal seeding (one clone present in the sample, Fig. 4C and Supplementary Fig. 6H). The seeding and dissemination patterns were also assessed for each metastatic sample (Fig. 5E). In

Figure 4


Monophyletic Seeding
Polyphyletic Seeding
A
B
C
D
OSCE1
OSCE4
OSCE6
OSCE10
Clonal Prevalence
Time Point
T0
RdtBx
RdtRx
LMet1
HMet2
HMet4
LMet5
LdfBx
RIIM1
RIIM2
LuIM3
LdfRx
FRx_M
FRx_D
FRx_P
FRx_E
FRx_O
RuIM0
LIIM1
Number of Mutations
3K
2K
1K
0K
6K
4K
G A B C F D E
C F G D A E B
Signature
DNA Homologous Recombination Deficiency (3)
Cisplatin (31,35)
Clock-Like (1,5)
Late Replication Errors (8)
Reactive Oxygen Species Damage (18)
Somatic Hypermutation (9)
Unknown Etiology (12,19)
Defective DNA Mismatch Repair (6,14,20,26,44)
APOBEC (2,13)
Chemotherapy (25)
Alkylator/Temozolomide (11)
Defective Base Excision Repair (30)

427

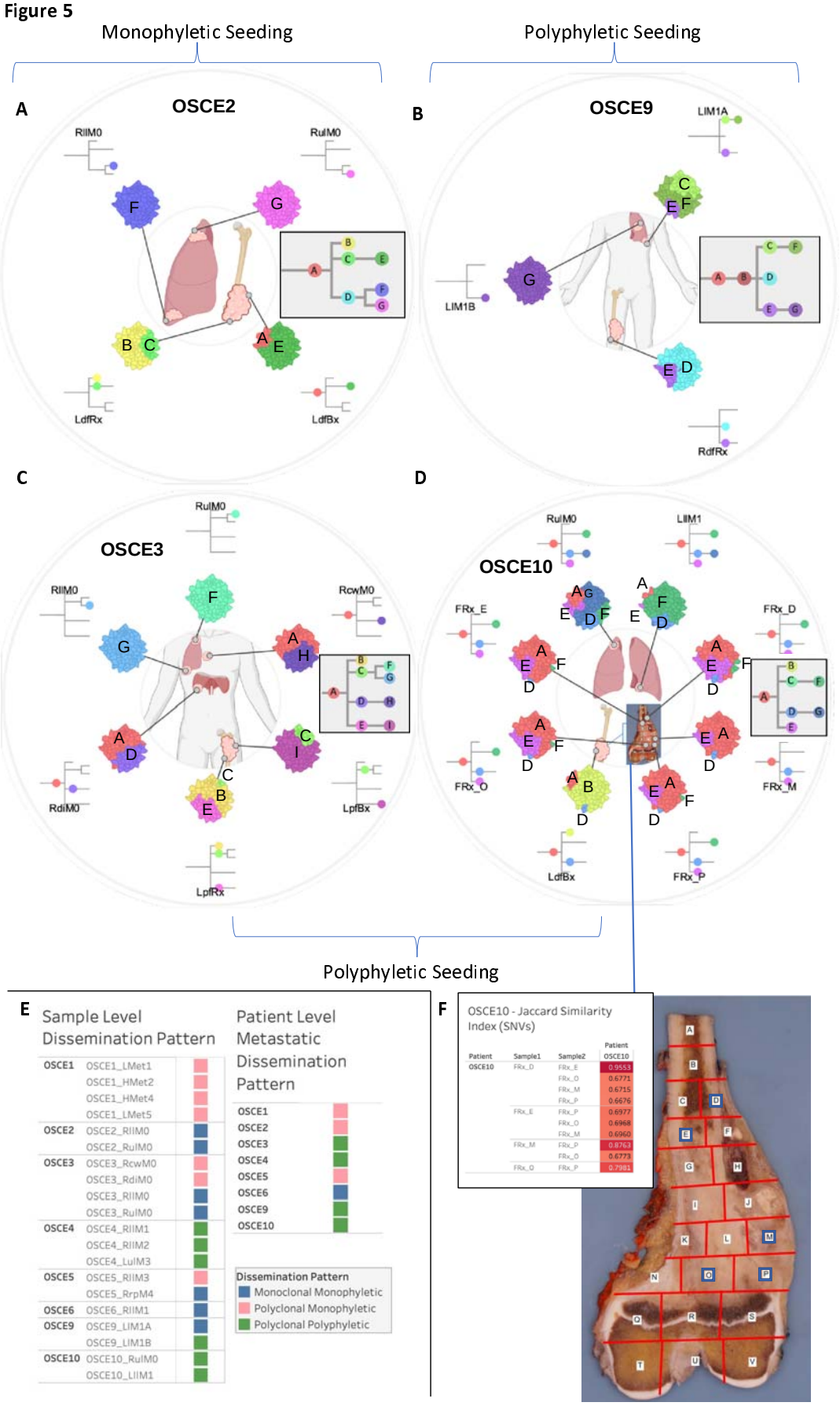
Figure 4. SNV based phylogenies highlighting temporal evolution with clonal mutational signature composition. **A, B, C, and D,** Upper figure in each panel is a TimeScape plot of the inferred evolutionary phylogeny, highlighting clonal proportions over time. The prevalence of different clones is shown over time on the vertical axis, with the different clones represented by different colors. The horizontal axis represents the timepoints, which are represented by gray lines. The evolutionary relationships between the clones are shown on the phylogenetic tree and in the TimeScape layout. The bottom right of each panel is a stacked bar plot of the total number of mutations assigned to each clone. Colors represent total number of mutations attributed to each mutational signature with color legend at bottom of figure. Patient level metastatic seeding patterns are denoted by the brackets at the top of the page.

two patients (OSCE3 and OSCE5), there were examples of different modes of metastatic dissemination among spatially separated metastases from the same resection. In OSCE3 for example, extra-pulmonary metastases (RcwM0, RdiM0) demonstrate polyclonal (≥2 clones present in the sample) monophyletic patterns of dissemination, while the pulmonary metastases (RllM0, RulM0) demonstrate monoclonal monophyletic dissemination, with both samples sharing the same ancestral clone (Fig. 5C, Supplementary Fig. 6B, 6I).

Limited Heterogeneity after Induction Chemotherapy

For OSCE10, we obtained a section of the primary tumor sample for multi-region sequencing after the patient had received 10 weeks of induction MAP chemotherapy. The specimen was mapped (Fig. 5F), DNA was extracted from the sections with viable tumor, and then five sections (E, D, M, O, P) with the highest DNA quanity/quality metrics underwent WGS. When examining the clonal architecture of each section (Fig. 4D/5F), each section had the same four clones (A, D, E, F), with truncal clone A dominating, except for FRx_M, which did not have clone F, which later became the clone that dominated at relapse. This pattern was also seen with the copy number clones, where clone 3 was present in all sections except for FRx_M and then dominated at relapse (Fig. 2F). When comparing Jaccard similarity coefficients based on SNV composition across the different sections, adjoining sections shared the highest coefficients (D/E=0.96, O/P=0.79, M/P=0.88, Fig. 5F), whereas non-adjoining sections had similar coefficients, regardless of distance from each other (range = 0.67-0.69, Fig. 5F). As highlighted previously, a single-sample sequencing strategy that sampled from FRx_M would have missed detecting the metastatic subclone that was present in the other four sections.

Figure 5


Monophyletic Seeding
Polyphyletic Seeding
A
OSCE2
RlIM0
RulM0
LdfRx
LdfBx
B
OSCE9
LIM1A
LIM1B
RdfRx
C
OSCE3
RulM0
RlIM0
RcwM0
RdiM0
LpfBx
LpfRx
D
OSCE10
RulM0
LIIM1
FRx_E
FRx_D
FRx_O
FRx_M
LdfBx
FRx_P
Polyphyletic Seeding
E
Sample Level Dissemination Pattern
Patient Level Metastatic Dissemination Pattern
OSCE1 OSCE1_LMet1 OSCE1_HMet2 OSCE1_HMet4 OSCE1_LMet5
OSCE2 OSCE2_RlIM0 OSCE2_RulM0
OSCE3 OSCE3_RcwM0 OSCE3_RdiM0 OSCE3_RlIM0 OSCE3_RulM0
OSCE4 OSCE4_RlIM1 OSCE4_RlIM2 OSCE4_LulM3
OSCE5 OSCE5_RlIM3 OSCE5_RrpM4
OSCE6 OSCE6_RlIM1
OSCE9 OSCE9_LIM1A OSCE9_LIM1B
OSCE10 OSCE10_RulM0 OSCE10_LIIM1
OSCE1 OSCE2 OSCE3 OSCE4 OSCE5 OSCE6 OSCE9 OSCE10
Dissemination Pattern
Monoclonal Monophyletic
Polyclonal Monophyletic
Polyclonal Polyphyletic
F
OSCE10 - Jaccard Similarity Index (SNVs)
Patient Sample1 Sample2 Patient OSCE10
OSCE10 FRx_D FRx_E 0.9553
FRx_O 0.6771
FRx_M 0.6715
FRx_P 0.6676
FRx_E FRx_P 0.6977
FRx_O 0.6968
FRx_M 0.6960
FRx_M FRx_P 0.8763
FRx_O 0.6773
FRx_O FRx_P 0.7981

460

Figure 5. SNV based phylogenies highlighting spatial evolution and descriptions of metastatic seeding patterns at the sample and patient level. **A, B, C, and D,** Spatially and in some cases temporally distinct samples are indicated on the anatomic sites from where the sample originated. The colors represent different clones, and the phylogenetic trees show the evolutionary relationships between these clones. The prevalence of each clone at a particular site is proportional to the colored area of the cellular aggregate representation. **E,** Sample and patient level dissemination patterns are characterized in these charts. *Monoclonal dissemination*: single subclone within the primary tumor seeds one or more metastatic lesions, *polyclonal dissemination*: multiple distinct subclones from the primary tumor seed one or more metastatic lesions, *monophyletic origin*: all metastatic clones are derived from a recent common ancestor, *polyphyletic origin*: metastasizing clones are more similar to other subclones within the primary tumor than they are to each other. These descriptions can be considered at the sample level, focused on the clonal make up of a single metastatic site compared to the primary tumor, or taken as a whole, evaluating all spatially or temporally separated samples and how they relate back to the primary tumor. **F,** Multi-region sequencing was performed on a primary resection sample from OSCE10. Regions D, E, M, O, P were sequenced from the specimen grid depicted. A table of Jaccard similarity indexes based on shared SNVs for these samples is shown in the upper left inset.

Most new SNVs in relapsed disease are attributed to HRD and cisplatin mutational signatures

To further characterize the potential drivers of clonal evolution, mutational signature analysis was performed for each clone (Fig. 4A-D, Supplementary Fig. 6A-D). In three patients, the largest clone by number of SNVs (OSCE5 – clone C, OSCE6 – clone G, and OSCE10 – clone C) had over half of the SNVs attributed to cisplatin (Supplementary Fig. 9A). In four patients, the largest clone by number of SNVs (OSCE1 – clone H, OSCE3 – clone A, OSCE4 – clone D, and OSCE9 – clone G) had a plurality of SNVs in each clone attributed to homologous repair deficiency (HRD, Supplementary Fig. 9A). In OSCE2 and OSCE3, both refractory cases, the largest number of SNVs was assigned to the truncal clone, clone A, which had a high number of SNVs (OSCE2 – 1731/2234, OSCE3 – 2464/4022) associated with HRD and late replication errors (Sig. 8). In patients with patient-level monophyletic seeding of metastases (OSCE1, OSCE2, OSCE5, OSCE6), where a single ancestral metastatic clone could be identified (OSCE1–clone D, OSCE2–clone D, OSCE5–clone C, OSCE6–clone C), there were no clear patterns regarding the signature composition identified (Supplementary Fig. 9B). OSCE1–clone D and OSCE2-clone D had a plurality of SNVs attributed to HRD, whereas OSCE5-clone C had the majority of SNVs attributed to cisplatin, and OSCE6 had the majority of its SNVs attributed to reactive oxygen species damage (Sig. 18). When looking at doublet base substitution (DBS) signatures at the clonal level, in clones with ≥10 DBS SNVs, cisplatin-associated DBS 5 was the largest contributor of DBS SNVs, accounting for 50% or more of the total SNVs in 17/26 clones (Supplementary Fig. 9C).

Emergence of cisplatin and alkylator signatures helps time the formation of metastases

A simple method for ascertaining whether a given metastasis arose before or after treatment with cisplatin or an alkylator is to find clonal SNVs attributed to the respective signature in a metastatic tumor sample[26]. When looking at the dominant clone in the first relapse sample for patients with recurrent disease (OSCE1-Clone D, OSCE4-Clone E/I, OSCE5-Clone C, OSCE6-Clone C/F/G, OSCE9-Clone C/G, OSCE10-Clone C/F), there was a cisplatin signature present in each of these respective clones/clades, indicating that the metastases arose after therapy (Fig. 4A-D, Supplementary Fig. 6C, 6D). Additionally, in OSCE1, after the patient received chemotherapy with ifosfamide and etoposide, OSCE1-Clone H and OSCE1-Clone F had mutations attributed to the alkylator signature, SBS11 (Fig. 4A). In the three patients with refractory disease (OSCE2, OSCE3, and OSCE10), we can demonstrate which metastatic samples were present at diagnosis and which developed while on therapy. Within OSCE2, Clone F and Clone G were the dominant subclones in RllM0 and RulM0 respectively in OSCE2 (Fig. 5A, Supplementary Fig. 6A). Clone F had 181/614 mutations attributed to cisplatin and Clone G had no mutations attributed to cisplatin, evidence Clone F was seeded on therapy as opposed to prior to therapy. In OSCE3, there was no evidence of a cisplatin signature in any of the subclones (Supplementary Fig. 6B), indicating that all metastatic sites had developed prior to initiating therapy. In OSCE10, the metastatic sample (RulM0), which has polyphyletic and polyclonal seeding (Fig. 4D, 5D), showed that the dominant clone/subclone pair of D/G likely seeded on therapy, given that there is a cisplatin signature attributed to approximately half of the mutations (1027/2644) and in all the mutations in subclone G (2727/2727).

HRD and Reactive Oxygen Species Damage Linked to Driver Gene Mutagenesis

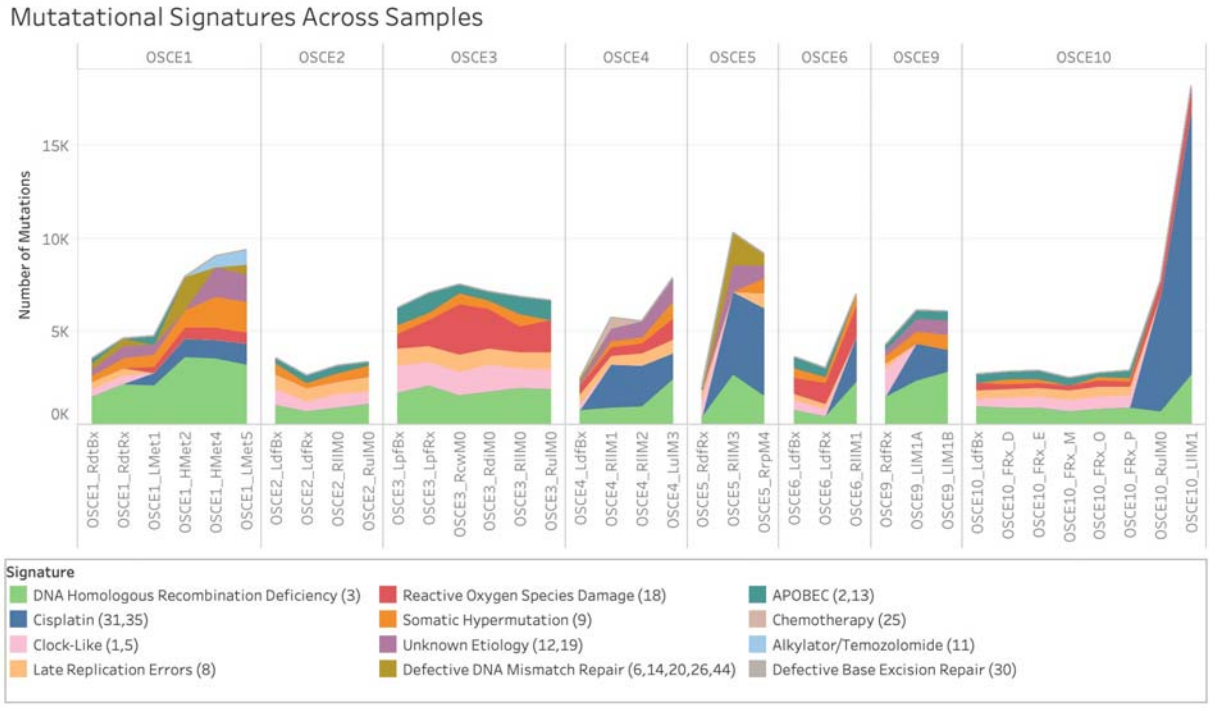
To identify the mutational processes most likely to be the origin of truncal driver gene SNVs, we calculated the likelihood that each individual SNV was caused by each signature, considering the mutation category and proportion of each mutational signature in the tumor genome[23]. To minimize the effect of treatment-related mutagenesis, we limited this analysis to the earliest primary site sample available for patients with truncal driver SNVs. Similar to the observations across all samples, HRD had the highest probability of attribution in the *TP53* driven OSCE1 sample and the *RB1* driven OSCE4 samples (OSCE1 probability=0.49, OSCE4 probability=0.44, Fig. 6D), and a slightly lower attribution probability than the clock-like signature in OSCE9 (HRD probability=0.32, clock-like probability=0.34, Fig. 6D). Both OSCE3 and OSCE10 had truncal ATRX SNVs, with reactive oxygen series damage accounting for the highest probability of attribution in both samples (OSCE3 probability=0.54, OSCE10 probability=0.42, Fig. 6D).

Cisplatin associated hypermutation in a case of refractory osteosarcoma

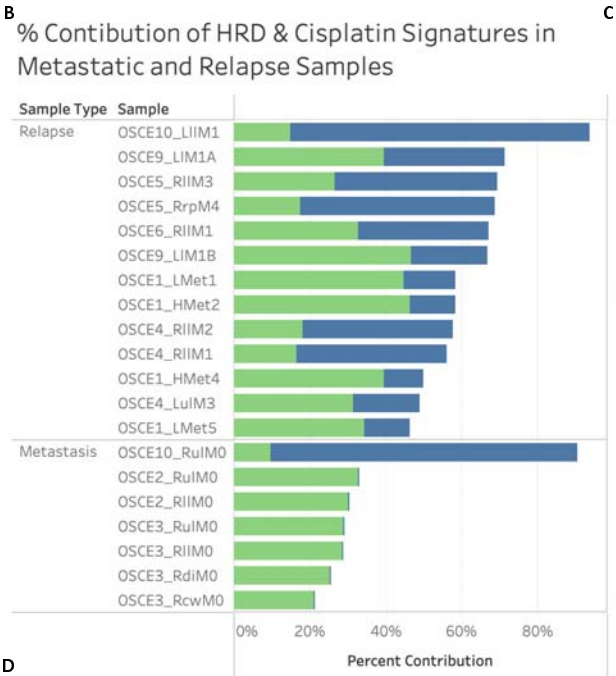
When evaluating mutational signatures at the sample level, the two most prevalent SBS signatures across all samples were HRD and cisplatin (Fig. 6A). In the 13 relapse samples, HRD and cisplatin accounted for more than half of all the SNVs (Fig. 6B). In the 17 primary site samples, the HRD, clock-like (Sig. 5), late replications errors (Sig. 8), ROS damage (Sig. 18), and APOBEC (Sig. 2,13) accounted for 75%–100% of all mutations (Fig. 6C). In 6/7 metastatic samples, HRD, clock-like, late replication errors, somatic hypermutation, ROS damage, and APOBEC accounted for all SNVs (Supplementary Fig. 10A). In the metastatic sample thought to have emerged on therapy for OSCE10 (RulM0) and the subsequent relapse sample (LllM1), SNVs attributed to cisplatin accounted for 6226/7724 (86%) and 14276/18216 (78%) SNVs,

Figure 6

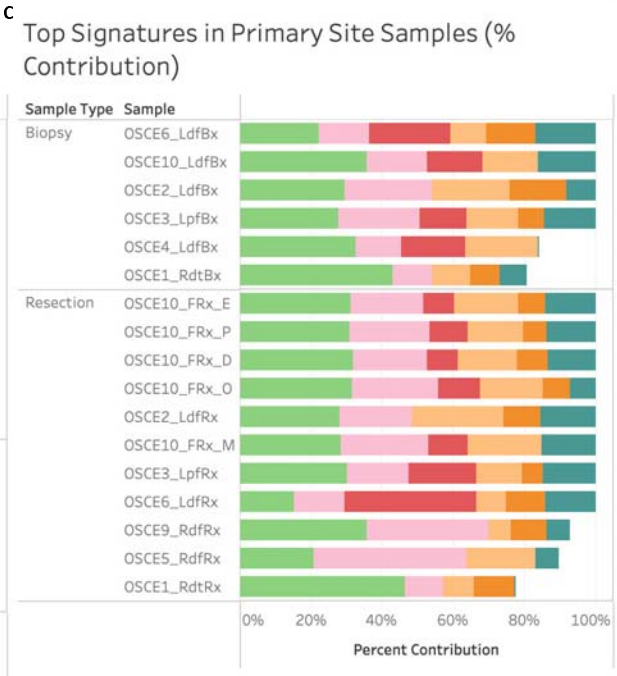
A



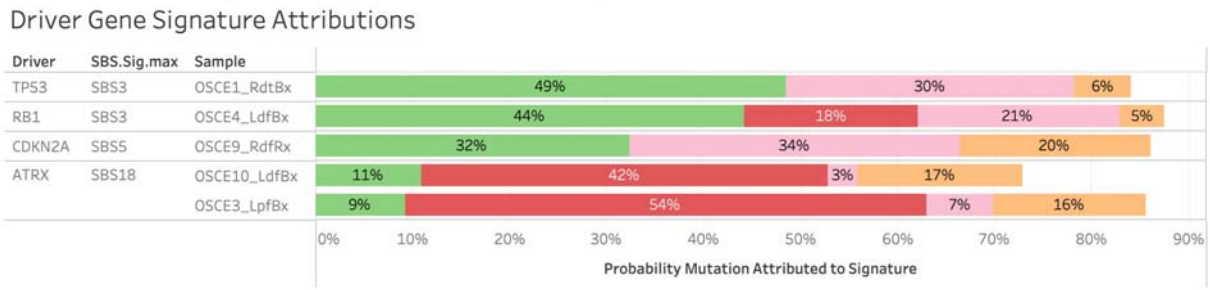
B



C



D



547

Figure 6. Mutational signature patterns across the cohort. **A,** Stacked line chart of mutational signature contribution by total number of mutations attributed to each signature. Colors represent the different signatures. **B,** Stacked bar chart of the relative contribution of HRD (green) & cisplatin (blue) signature in metastatic and relapse samples. **C,** Stacked bar chart of the relative contribution of HRD (green), clock-like (pink), reactive oxygen species damage (red), late replication errors (beige), somatic hypermutation (orange), and APOBEC (teal), in primary site samples. **D,** Stacked bar chart of the probability that driver gene SNVs were attributed to a mutational signature. Pretreatment samples from patients with driver SNVs were included and the primary resection sample from OSCE9 since no pretreatment sample was available. Colors represent the different signatures.

respectively (Fig. 10A). In the six patients with relapsed disease, the number of mutations attributed to HRD consistently increased from diagnosis to subsequent relapses, with at least 1000 new SNVs attributed to HRD when comparing diagnostic and relapse samples (Supplementary Fig. 10C).

Doublet base substitution (DBS) signatures were also evaluated across the cohort (Supplementary Fig. 10B). DBS signature 5, which is associated with cisplatin, was detected in 21 samples; however, to filter out false positives, a threshold of ≥ 5 DBS signature 5 SNVs was used to confirm the absence of the signature. Using this filter, DBS signature 5 was present in 15 samples, all metastatic or relapse samples, with complete overlap with the 14 samples that had SBS cisplatin signatures 31 or 35. Only OSCE3_RdiM0 had a DBS cisplatin signature but not an SBS cisplatin signature (Supplementary Fig. 10D).

Discussion

Tumor evolution and clonal heterogeneity have been increasingly recognized as major causes of therapeutic resistance to current anti-neoplastic therapies[27]. These findings extend our understanding of therapeutic resistance in spatially and temporally separated tumor samples from 8 patients with recurrent or refractory osteosarcoma. We found that while clonal driver gene SNVs and structural variants remain largely unchanged over the course of tumor progression, subclonal tumor populations with unique driver gene amplifications are present at diagnosis, emerge after treatment, and persist as the major clone at subsequent relapses.

Somatic copy number alterations are now increasingly recognized for their prognostic value over SNVs in multiple cancer types[28]. Oncogenic copy number alterations, while heterogeneous across osteosarcoma, represent attractive therapeutic targets, given the lack of recurrent targetable SNVs or structural variants[4–8]. Our study revealed that in our four patients who had relapsed osteosarcoma with a matched pretreatment sample (OSCE1, OSCE4, OSCE5, OSCE6), there was a subclonal treatment resistant copy number clone that emerged as the dominant clone in the relapsed setting. Furthermore, in the two patients with both a pre-treatment sample and on-therapy primary resection (OSCE1, OSCE6), this treatment resistant clone clonally expanded after 10 weeks of neoadjuvant chemotherapy. We believe this finding has important implications for molecular profiling strategies in osteosarcoma, as it suggests that the primary resection sample, and not the pretreatment biopsy, is more reflective of the metastatic potential for a tumor than the pretreatment biopsy, due to the selection pressure of neoadjuvant chemotherapy. Achieving a cure in osteosarcoma requires the extinction of all cancer cells with a successful “first-strike’ strategy with maximum tolerated doses of cisplatin, doxorubicin, and methotrexate[29]. For patients in whom this first strike fails (poor necrosis at the

time of primary resection), characterizing and targeting the treatment-resistant population of cancer cells using a second-strike strategy may prove to be an effective treatment strategy[29]. Our work highlights that molecular profiling of primary resection samples could allow for a more precise "genome-informed" approach[4], aimed at targeting resistant copy number alterations, to augment MAP chemotherapy.

The emergence of subclonal copy number alterations in primary tumors to fully clonal alterations in metastatic or recurrent samples, as demonstrated in our study in 6/7 patients with recurrent/refractory disease and pretreatment samples, has been previously described in a subset of adult cancers where analysis of matched primary and metastasis was performed[20]. We found *MYC* gain/amplification to be enriched in the treatment-resistant clone in 6/7 patients with more than one clone. Previous studies have shown that *MYC* amplification is often enriched in metastatic sites[20] and has been previously associated with poor outcomes and increased cell proliferation in osteosarcoma;[30–34] however, a recent study questioned its prognostic significance[35]. Our study may rectify this discrepancy, as we demonstrated that in patients with localized and metastatic disease at diagnosis, *MYC* amplification is subclonal in pretreatment samples and emerges after 10 weeks of neoadjuvant chemotherapy. As we begin to define molecular risk categories within osteosarcoma, our work demonstrates that profiling of post-treatment primary resection samples may reveal previous subclonal amplifications in driver oncogenes; thus, this time point would be more informative when assessing metastatic potential. Furthermore, while multi-region profiling of post-treatment resection from OSCE10 revealed limited heterogeneity among the different sites, there was one site where the metastatic clone was not present, highlighting the potential risk of failing to profile the metastatic clone with single-sample strategies. When considering future sequencing approaches, pooling DNA/RNA

extracts from multiple anatomically distinct tumor regions of the primary tumor could be a cost-effective way to improve DNA yield and variant detection, while providing a more complete picture of intratumoral heterogeneity[36].

Our chromosomal duplication timing analysis revealed that gains for the same patient often clustered around the same time point, regardless of whether whole-genome duplication was present. These bursts of duplications occurred prior to diagnosis, and there were no comparable bursts of duplications in resection, metastatic, or recurrent samples that would reflect ongoing instability. In a pan-cancer cohort, synchronous bursts of copy number gain were found to occur in 57% of diploid samples and 78% of WGD samples[37]. We found that these clustered duplication events were associated with catastrophic complex genomic rearrangement and amplicon events that occurred before diagnosis, such as chromothripsis and tyfonas. In contrast to a previous multi-region osteosarcoma study[15], we demonstrated that tumor ploidy remained consistent across all samples for each patient, which is likely because our copy number calls were inferred jointly across all samples for each patient, which can improve ploidy estimation in sample sets with wide ranges of purity[19]. These findings of pre-diagnostic duplication events followed by relative genomic stability support recent work that demonstrated that early catastrophic events are responsible for the structural complexity in the osteosarcoma genome, as opposed to sustained evolution and instability over time[38]. Whole-genome duplication was confirmed in a subset of patients and was found to be a late event in all three cases. Previous pan-cancer studies have found that WGD events are typically early events in a tumor's molecular time history, but are often preceded by TP53 inactivation[20,37,39]. Late WGD events have been described in a cohort of patients with hepatocellular carcinoma (HCC), and they are typically associated with larger tumors, leading to the conclusion that they may be the last step prior to

rapid growth and expansion[23]. Our data support a macroevolutionary model of evolution in osteosarcoma[40], with a large number of genomic aberrations acquired over a short period of time secondary to chromosomal instability events, followed by clonal selection, as opposed to ongoing evolution.

Large-scale genomic sequencing studies in osteosarcoma have revealed that there is significant inter-tumoral heterogeneity in osteosarcoma, with shared mutations typically in tumor suppressor genes rather than in targetable oncogenes[5–8]. We observed a heterogeneous mix of metastatic and recurrent seeding patterns in our cohort. We observed only one example of monoclonal, monophyletic dissemination in OSCE6, which is typically a result of a treatment-induced bottle-necking effect. There were three cases of polyclonal/monophyletic dissemination, in which multiple clones were present in the metastatic/recurrent samples, but they all shared a common ancestor, and there were four cases of polyclonal polyphyletic dissemination where multiple distinct clones from the primary tumor seeded a metastatic site. In previous studies, a de novo induced murine model of osteosarcoma demonstrated polyclonal seeding of metastases with ongoing parallel evolution[14], while studies using longitudinal and spatially separated samples have yielded mixed results, demonstrating both polyclonal seeding with parallel evolution[13] and monoclonal monophyletic seeding[15] in a majority of the respective cases from each study. These studies were limited by the lack of pretreatment primary tumors; therefore, the analysis relied on comparing metastatic and recurrent samples to post-treated primaries in many cases.

We demonstrate that while cisplatin and HRD are active mutagenic processes in osteosarcoma, accounted for most new mutations in relapsed disease, and we found no new driver SNVs in relapsed disease that could account for treatment resistance. The HRD signature

was conserved across all samples in our cohort, consistent with a recent pan-pediatric cancer study that found that 18/19 (95%) patients with osteosarcoma had mutations attributed to HRD[41]. The prevalence of the HRD signature in osteosarcoma is comparable to BRCA1 deficient cancers, suggesting that drugs that target homologous recombination deficient cells, such as PARP inhibitors, may have therapeutic value in osteosarcoma, a concept currently being evaluated in a phase II clinical trial (NCT04417062)[42]. A commonly cited limitation of using signature-based assays to assess HRD is that they reflect the HRD state prior to sample acquisition and may not reflect the current state, where HRD may have been restored[43]. Our study demonstrated that in osteosarcoma, the number of mutations attributed to HRD increases at each time point in patients with recurrent disease, suggesting that HRD continues to be an active mutagenic process after diagnosis.

The cisplatin signature was present in all relapse samples and metastatic sites that were thought to have developed during upfront therapy. The extent of cisplatin-induced mutagenesis has been previously described in osteosarcoma, where it was found that cisplatin therapy could potentially increase the mutational burden by two-fold[15]. Although we also found that cisplatin therapy led to large increases in mutational burden in recurrent samples, none of these mutations were likely drivers of treatment resistance, which is consistent with previous studies in patients with platinum-resistant ovarian cancer[44] and osteosarcoma[15]. Although we cannot account for copy number alterations or structural variants induced by cisplatin, recent cell line work in cisplatin exposed esophageal and liver tumors, found few copy number alterations or structural variants, suggesting that cisplatin does not contribute significantly to genomic instability[45].

Our findings highlight that the chemoresistant population of tumor cells in osteosarcoma is subclonal at diagnosis and is characterized by unique oncogenic amplifications. As our ability

to target these oncogenic amplifications improves, future studies aimed at identifying and targeting these oncogenic drivers during upfront therapy may be an effective strategy to eliminate chemoresistant tumor cells and improve survival.

Methods

Patient consent and tissue processing

This study was approved by the Institutional Review Board of the Memorial Sloan Kettering Cancer Center (New York, NY, USA) and conducted in accordance with the Declaration of Helsinki. Informed written consent was obtained from each subject or guardian. Tumor samples and matched normal samples were collected from 10 patients with a pathologically confirmed diagnosis of osteosarcoma, who were identified both retrospectively and prospectively for those who had their tumor banked at diagnosis and at least one other time point. Only patients who consented to an IRB-approved blood and tumor collection protocol were eligible for tumor sequencing. Fresh tumor samples were procured from the operating room in a sterile container. The tissue was processed using scalpels and divided into pea-sized pieces before being stored at -80°C. Frozen tissue samples from several patients were also available through our Precision Pathology Biobanking Center and were acquired using the same protocol. Additionally, in several patients, archival tumor specimens in the form of formalin-fixed paraffin-embedded (FFPE) specimens were obtained from both the internal and external pathology departments using the same protocol. Only FFPE samples that were not subjected to harsh decalcification techniques were selected. FFPE samples that had been decalcified using EDTA were deemed appropriate for further downstream analysis.

Each frozen tissue sample was submitted to our pathology core, where it was embedded in Tissue-Tek optimum cutting temperature compound and sectioned at 5–10 mm on a Leica Cryostat to create a hematoxylin and eosin–stained (H&E) slide for review. Each FFPE sample was sectioned using a Leica Microtome. H&E slides were evaluated by a trained pathologist to determine tumor content. After pathologic review, tumor samples were isolated via a 21-gauge

punch, curl biopsy, or macro-dissected from sectioned slides to remove non-neoplastic components. The neoplastic component of each tumor underwent genomic DNA extraction using a Qiagen DNAeasy Blood and Tissue Kit and protocol, whereas the FFPE samples were extracted using a QIAamp DNA FFPE Tissue Kit and protocol.

PBMCs utilized for matched normal sequencing were brought up to 15mL volume in cold PBS and isolated with the DNeasy Blood & Tissue Kit (QIAGEN catalog # 69504) according to the manufacturer's protocol and incubated at 55°C for digestion. DNA was eluted in 0.5X Buffer AE.

Whole-genome sequencing and alignment

DNA quantification, library preparation, and whole-genome sequencing were performed using the Integrated Genomics Operation at the Memorial Sloan Kettering Cancer Center (New York, NY). After PicoGreen quantification and quality control using an Agilent BioAnalyzer, 131-500ng of genomic DNA was sheared using an LE220-plus Focused-ultrasonicator (Covaris catalog # 500569), and sequencing libraries were prepared using the KAPA Hyper Prep Kit (Kapa Biosystems KK8504) with modifications. Briefly, libraries were subjected to a 0.5 × size selection using aMPure XP beads (Beckman Coulter catalog # A63882) after post-ligation cleanup. Libraries with < 500 ng of genomic DNA were amplified using 5-6 cycles of PCR and pooled equimolar amounts. Libraries containing at least 500 ng of genomic DNA were not amplified. Samples were run on a NovaSeq 6000 in a 150bp/150bp paired end run, using the NovaSeq 6000 SBS v1 Kit and an S1, S2, or S4 flow cell (Illumina). The average number of read pairs per normal was 614 million and the average number of read pairs per tumor was 1.3 billion.

Whole Genome Sequencing Pipeline

Whole genome sequencing data were analyzed using the ISABL[46] platform, with methods previously described in detail[47]. The additional downstream analyses are described below.

Single Nucleotide Variant Filtering

For all eight patients, single nucleotide variants (SNV) were called in triplicate by MuTect[48], Strelka[49], and Caveman[50]. Only mutations that had a "PASS" flag, were called by at least two mutation callers and observed in less than 2% of the reads of the matched normal sample with 10x coverage were considered for further analysis. It is notoriously difficult to extract high quality DNA from the osteoid matrix that surrounds viable tumor cells, often leading to sequenced samples with low purity. Across the eight patients, a mix of samples was prepared as either fresh-frozen or FFPE. FFPE samples are thought to be inferior to fresh frozen samples, as the formalin fixation process results in nucleic acid fragmentation, DNA crosslinks, and deamination, leading to C>T mutation artifacts. As a result, downstream analysis of FFPE can be challenging when filtering out artifacts from a true positive. Aggressive filtering of low-allele frequency variants has been shown to increase SNV overlap when comparing matched FFPE and FF samples[51]. Given these assumptions, a custom filtering approach was utilized to maximize the ability to utilize low-purity FFPE samples within a mix of higher-quality FFPE samples and fresh frozen samples for each patient. For fresh frozen samples with an estimated purity of 20%, mutations with a variant allele frequency (VAF) less than 5 were filtered out. If a frozen sample had a purity of less than 20%, no additional filtering was applied, given the potential for filtering out subclonal mutations in an otherwise high-quality frozen tissue sample. For FFPE samples,

we filtered out mutations with a VAF less than 20%. These thresholds for FFPE samples were then further purity adjusted, for example a sample with an estimated purity of 80%, would have mutations below a VAF of 0.16 (0.8 purity * 0.2 VAF filter) initially filtered out. Although this initial filtering for FFPE is strict, the advantage of having multiple samples per patient allows us to utilize the mutation calls in other high-quality samples for a patient to rescue mutations that may have been initially filtered in a low-quality sample. This is accomplished through a pile-up rescue, where all filtered mutations for a patient are combined and then specifically searched in the BAM file that was generated for each sample.

Driver Gene Analysis

All somatic variants that led to a frameshift insertion, frameshift deletion, in-frame insertion, in-frame deletion, missense, nonsense, nonstop, or splice site/region mutation, or a translation start site were considered. For variants identified as missense or nonsense, we required the variant to be considered a likely functional driver using the LiFD tool[18], which is a two-phase algorithm that pulls from various databases and bioinformatic methods to determine whether a given mutation is likely to be functional. We also considered genes that were significantly mutated in large pediatric cancer and osteosarcoma sequencing studies[4–9,52]. The final list consisted of 639 genes.

Evolutionary Analysis

To determine the spatial and temporal dynamics of subclonal diversity within a patient, we first used Treeomics v1.9.2[53] to derive phylogenies. Treeomics reconstructs the phylogeny of metastatic lesions and maps subclones to their anatomical locations. Treeomics utilizes a

Bayesian inference model to account for error-prone sequencing and differing neoplastic cell contents to calculate the likelihood that a specific variant is present or absent. Treeomics then infers a global optimal tree based on mixed-integer linear programming[53].

The HATCHet[19] v1.0.1 algorithm was used to infer allele and clone-specific copy-number aberrations, clone proportions, and whole-genome duplications (WGD) for each patient. HATCHet was run using the GATK4-CNV custom pipeline, with Battenberg copy number calls fitted to meet the input requirements for running the tool. Solutions were manually reviewed with the creator of the tool, Simone Zaccaria PhD, to allow for advanced fine-tuning and ensure that the most accurate solutions were selected.

The DeCiFer[17] v1.1.5 algorithm was used to cluster mutations across all samples for each patient, providing descendant cell fractions (DCF) and cancer cell fractions for each cluster. The copy number input for this algorithm is the output of the HATCHet algorithm. Custom state trees were generated utilizing a maximum copy number between 6-8 for each patient (lower maximum copy number states were selected if the runtime exceeded 48 h). After clustering of mutations, the CALDER[25] v0.11 algorithm was used to infer evolutionary phylogenies. To run the DeCiFer output through CALDER, the inferred cluster DCF was converted to a read count by multiplying the DCF by 1000 and then dividing by 2 (because CALDER assumes that all mutations are in heterozygous diploid regions). Therefore, if a mutation has 1000 reads and an inferred DCF of 40%, the corresponding input will have 200 variant reads and 800 reference reads. Longitudinal constraints were lifted when analyzing OSCE2 and OSCE3, because all analyzed samples were present at diagnosis.

The Palimpest[22] algorithm (version = github commit 4795da2) was used to characterize and visualize mutational signatures (using Cosmic SBS/DBS v3.2) at both clone and sample

levels. This algorithm also provided information regarding the timing of duplication and loss of heterozygosity events using previously described methods[23].

Structural Variant Analysis

Structural variants were annotated using iAnnotateSV software[54]. We used svpluscnv[55] and ShatterSeek[56] to identify regions of chromothripsis. The JaBba tool[24] was used to identify regions with complex rearrangements or amplicon events. The calls from all tools were combined for further downstream analysis.

Data Visualizations

Oncoprint was created using the CoMut[57] tool. Timescape[58] (https://github.com/shahcompbio/timescape) and Mapscape[58] (https://github.com/shahcompbio/mapscape) were used to visualize temporal and spatial clonal evolution. Tableau Desktop (v2021.4) was used to analyze and visualize the data with charts, bars, and line graphs. Ridgeline graphs were created using R utilizing the ggridges package (https://cran.r-project.org/web/packages/ggridges/vignettes/introduction.html). Sankey plots were created using SankeyMATIC (https://sankeymatic.com/build/). Anatomic cartoons were created using BioRender (https://biorender.com/).

Data Availability

The raw data for whole genome sequencing generated in this study have been deposited in the dbGAP database under the accession code _____. These data are available under restricted access because of privacy concerns. Permanent employees of an institution at a level equivalent

to a tenure-track professor or senior scientist with laboratory administration and oversight responsibilities may request access through the dbGAP. The requests, which are managed by the National Cancer Institute's (NCI) Data Access Committee, take less than two days for approval, and access is permitted for 12 months.

References

1. Kager L, Zoubek A, Pötschger U, et al. Primary metastatic osteosarcoma: Presentation and outcome of patients treated on neoadjuvant Cooperative Osteosarcoma Study Group protocols. *Journal of Clinical Oncology*. 2003;21(10):2011-2018. doi:10.1200/JCO.2003.08.132
2. Kim MS, Cho WH, Song WS, Lee SY, Jeon DG. Time dependency of prognostic factors in patients with stage II osteosarcomas. *Clin Orthop Relat Res*. 2007;463:157-165. doi:10.1097/BLO.0B013E318142B27D
3. Bielack SS, Kempf-Bielack B, Delling G, et al. Prognostic Factors in High-Grade Osteosarcoma of the Extremities or Trunk: An Analysis of 1,702 Patients Treated on Neoadjuvant Cooperative Osteosarcoma Study Group Protocols. *Journal of Clinical Oncology*. 2002;20(3):776-790. doi:10.1200/jco.2002.20.3.776
4. Sayles LC, Breese MR, Koehne AL, et al. Genome-Informed Targeted Therapy for Osteosarcoma. *Cancer Discov*. Published online September 28, 2018:CD-17-1152. doi:10.1158/2159-8290.CD-17-1152
5. Perry JA, Kiezun A, Tonzi P, et al. Complementary genomic approaches highlight the PI3K/mTOR pathway as a common vulnerability in osteosarcoma. *Proc Natl Acad Sci U S A*. 2014;111(51):E5564-73. doi:10.1073/pnas.1419260111
6. Suehara Y, Alex D, Bowman A, et al. Clinical Genomic Sequencing of Pediatric and Adult Osteosarcoma Reveals Distinct Molecular Subsets with Potentially Targetable Alterations. *Clinical Cancer Research*. 2019;25(21):6346-6356. doi:10.1158/1078-0432.CCR-18-4032
7. Behjati S, Tarpey PS, Haase K, et al. Recurrent mutation of IGF signalling genes and distinct patterns of genomic rearrangement in osteosarcoma. *Nat Commun*. 2017;8(May):1-8. doi:10.1038/ncomms15936
8. Chen X, Bahrami A, Pappo A, et al. Cell Reports Report Recurrent Somatic Structural Variations Contribute to Tumorigenesis in Pediatric Osteosarcoma. Published online 2014. doi:10.1016/j.celrep.2014.03.003
9. Kovac M, Blattmann C, Ribi S, et al. Exome sequencing of osteosarcoma reveals mutation signatures reminiscent of BRCA deficiency. *Nat Commun*. 2015;6(1):8940. doi:10.1038/ncomms9940
10. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz Jr. LA, Kinzler KW. Cancer Genome Landscapes. *Science (1979)*. 2013;339(6127):1546-1558. doi:10.1126/science.1235122
11. Meltzer PS, Helman LJ. New Horizons in the Treatment of Osteosarcoma. *New England Journal of Medicine*. 2021;385(22):2066-2076. doi:10.1056/NEJMRA2103423/SUPPL_FILE/NEJMRA2103423_DISCLOSURES.PDF
12. Persha HE, Kato S, De P, et al. Osteosarcoma with cell-cycle and fibroblast growth factor genomic alterations: case report of Molecular Tumor Board combination strategy resulting in long-term exceptional response. *J Hematol Oncol*. 2022;15(1):1-5. doi:10.1186/S13045-022-01344-X/FIGURES/2
13. Xu H, Zhu X, Bao H, et al. Genetic and clonal dissection of osteosarcoma progression and lung metastasis. *Int J Cancer*. 2018;143(5):1134-1142. doi:10.1002/ijc.31389

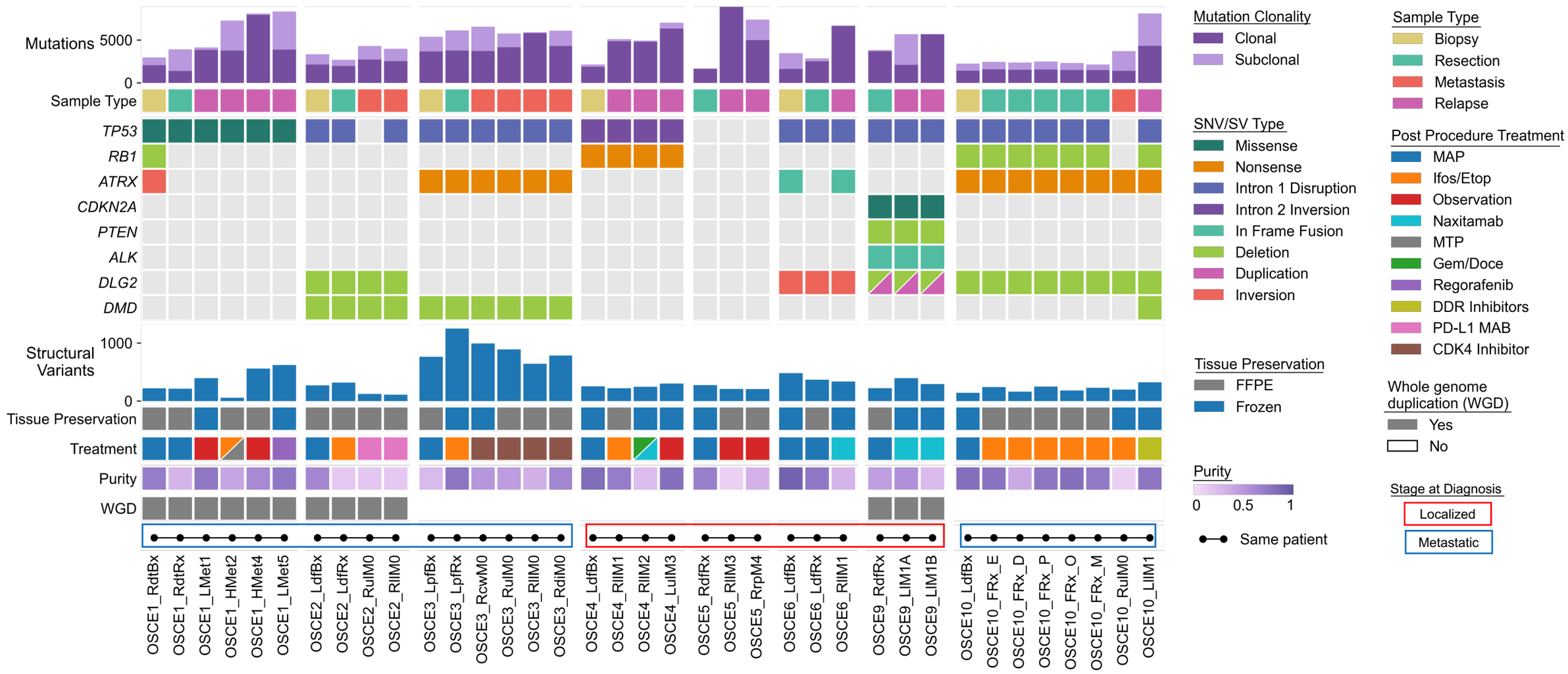
872 14. Gambera S, Abarrategi A, González-Camacho F, et al. Clonal dynamics in osteosarcoma
873 defined by RGB marking. *Nat Commun*. 2018;9(1). doi:10.1038/s41467-018-06401-z
874 15. Brady SW, Ma X, Bahrami A, et al. The clonal evolution of metastatic osteosarcoma as
875 shaped by cisplatin treatment. *Molecular Cancer Research*. 2019;17(4):895-906.
876 doi:10.1158/1541-7786.MCR-18-0620
877 16. Goswami RS, Luthra R, Singh RR, et al. Identification of Factors Affecting the Success of
878 Next-Generation Sequencing Testing in Solid Tumors. *Am J Clin Pathol*. 2016;145(2):222-
879 237. doi:10.1093/ajcp/aqv023
880 17. Satas G, Zaccaria S, El-Kebir M, Raphael BJ. DeCiFering the elusive cancer cell fraction in
881 tumor heterogeneity and evolution. *Cell Syst*. Published online August 2021.
882 doi:10.1016/j.cels.2021.07.006
883 18. Reiter JG, Baretti M, Gerold JM, et al. An analysis of genetic heterogeneity in untreated
884 cancers. *Nat Rev Cancer*. 2019;19(11):639-650. doi:10.1038/s41568-019-0185-x
885 19. Zaccaria S, Raphael BJ. Accurate quantification of copy-number aberrations and whole-
886 genome duplications in multi-sample tumor sequencing data. *bioRxiv*. Published online
887 2018:496174. doi:10.1101/496174
888 20. Watkins TBK, Lim EL, Petkovic M, et al. Pervasive chromosomal instability and karyotype
889 order in tumour evolution. *Nature*. 2020;587(7832):126-132. doi:10.1038/s41586-020-
890 2698-6
891 21. McGranahan N, Rosenthal R, Hiley CT, et al. Allele-Specific HLA Loss and Immune Escape
892 in Lung Cancer Evolution. *Cell*. 2017;171(6):1259-1271.e11.
893 doi:10.1016/j.cell.2017.10.001
894 22. Shinde J, Bayard Q, Imbeaud S, et al. Palimpsest: An R package for studying mutational
895 and structural variant signatures along clonal evolution in cancer. *Bioinformatics*.
896 2018;34(19):3380-3381. doi:10.1093/bioinformatics/bty388
897 23. Letouzé E, Shinde J, Renault V, et al. Mutational signatures reveal the dynamic interplay
898 of risk factors and cellular processes during liver tumorigenesis. *Nature Communications*
899 *2017 8:1*. 2017;8(1):1-13. doi:10.1038/s41467-017-01358-x
900 24. Hadi K, Yao X, Behr JM, et al. Distinct Classes of Complex Structural Variation Uncovered
901 across Thousands of Cancer Genome Graphs. *Cell*. 2020;183(1):197-210.e32.
902 doi:10.1016/j.cell.2020.08.006
903 25. Myers MA, Satas G, Raphael BJ. CALDER: Inferring Phylogenetic Trees from Longitudinal
904 Tumor Samples. *Cell Syst*. 2019;8(6):514-522.e5. doi:10.1016/j.cels.2019.05.010
905 26. Németh E, Krzystanek M, Reiniger L, et al. The genomic imprint of cancer therapies helps
906 timing the formation of metastases. *Int J Cancer*. 2019;145(3):694-704.
907 doi:10.1002/ijc.32159
908 27. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present,
909 and the Future. *Cell*. 2017;168(4):613-628. doi:10.1016/j.cell.2017.01.018
910 28. Smith JC, Sheltzer JM. Systematic identification of mutations and copy number
911 alterations associated with cancer patient prognosis. *Elife*. 2018;7.
912 doi:10.7554/ELIFE.39217
913 29. Reed DR, Metts J, Pressley M, et al. An evolutionary framework for treating pediatric
914 sarcomas. *Cancer*. 2020;126(11):2577-2587. doi:10.1002/cncr.32777

915 30. Ueda T, Healey JH, Huvos AG, Ladanyi M. Amplification of the MYC Gene in
916 Osteosarcoma Secondary to Paget's Disease of Bone. *Sarcoma*. 1997;1(3-4):131-134.
917 doi:10.1080/13577149778209
918 31. Shimizu T, Ishikawa T, Sugihara E, et al. c-MYC overexpression with loss of Ink4a/Arf
919 transforms bone marrow stromal cells into osteosarcoma accompanied by loss of
920 adipogenesis. *Oncogene*. 2010;29(42):5687-5699. doi:10.1038/onc.2010.312
921 32. Scionti I, Michelacci F, Pasello M, et al. Clinical impact of the methotrexate resistance-
922 associated genes C-MYC and dihydrofolate reductase (DHFR) in high-grade
923 osteosarcoma. *Annals of Oncology*. 2008;19(8):1500-1508. doi:10.1093/annonc/mdn148
924 33. Gamberi G, Benassi MS, Bohling T, et al. C-myc and c-fos in Human Osteosarcoma:
925 Prognostic Value of mRNA and Protein Expression. *Oncology*. 1998;55(6):556-563.
926 doi:10.1159/000011912
927 34. Ladanyi M, Park CK, Lewis R, Jhanwar SC, Healey JH, Huvos AG. Sporadic amplification of
928 the MYC gene in human osteosarcomas. *Diagn Mol Pathol*. 1993;2(3):163-167.
929 35. Taylor AM, Sun JM, Yu A, et al. Integrated DNA Copy Number and Expression Profiling
930 Identifies IGF1R as a Prognostic Biomarker in Pediatric Osteosarcoma. *Int J Mol Sci*.
931 2022;23(14):8036. doi:10.3390/ijms23148036
932 36. Litchfield K, Stanislaw S, Spain L, et al. Representative Sequencing: Unbiased Sampling of
933 Solid Tumor Tissue. *Cell Rep*. 2020;31(5):107550. doi:10.1016/J.CELREP.2020.107550
934 37. Gerstung M, Jolly C, Leshchiner I, et al. The evolutionary history of 2,658 cancers. *Nature*.
935 2020;578(7793):122-128. doi:10.1038/s41586-019-1907-7
936 38. Rajan S, Zaccaria S, Cannon M v, et al. Structurally complex osteosarcoma genomes
937 exhibit limited heterogeneity within individual tumors and across evolutionary time.
938 *bioRxiv*. Published online January 1, 2022:2021.08.30.458268.
939 doi:10.1101/2021.08.30.458268
940 39. Bielski CM, Zehir A, Penson A v., et al. Genome doubling shapes the evolution and
941 prognosis of advanced cancers. *Nat Genet*. 2018;50(8):1189-1195. doi:10.1038/S41588-
942 018-0165-1
943 40. Vendramin R, Litchfield K, Swanton C. Cancer evolution: Darwin and beyond. *EMBO J*.
944 2021;40(18):e108389. doi:10.15252/EMBJ.2021108389
945 41. Ma X, Liu Y, Liu Y, et al. Pan-cancer genome and transcriptome analyses of 1,699
946 paediatric leukaemias and solid tumours. *Nature 2018 555:7696*. 2018;555(7696):371-
947 376. doi:10.1038/nature25795
948 42. Forrest SJ, Kinnaman MD, Livingston JA, et al. Phase II trial of olaparib in combination
949 with ceralasertib in patients with recurrent osteosarcoma.
950 *https://doi.org/101200/JCO20213915_supplTPS11575*. 2021;39(15_suppl):TPS11575-
951 TPS11575. doi:10.1200/JCO.2021.39.15_SUPPL.TPS11575
952 43. Miller RE, Elyashiv O, El-Shakankery KH, Ledermann JA. Ovarian Cancer Therapy:
953 Homologous Recombination Deficiency as a Predictive Biomarker of Response to PARP
954 Inhibitors. *Onco Targets Ther*. 2022;15:1105-1117. doi:10.2147/OTT.S272199
955 44. Patch AM, Christie EL, Etemadmoghadam D, et al. Whole–genome characterization of
956 chemoresistant ovarian cancer. *Nature 2015 521:7553*. 2015;521(7553):489-494.
957 doi:10.1038/nature14410

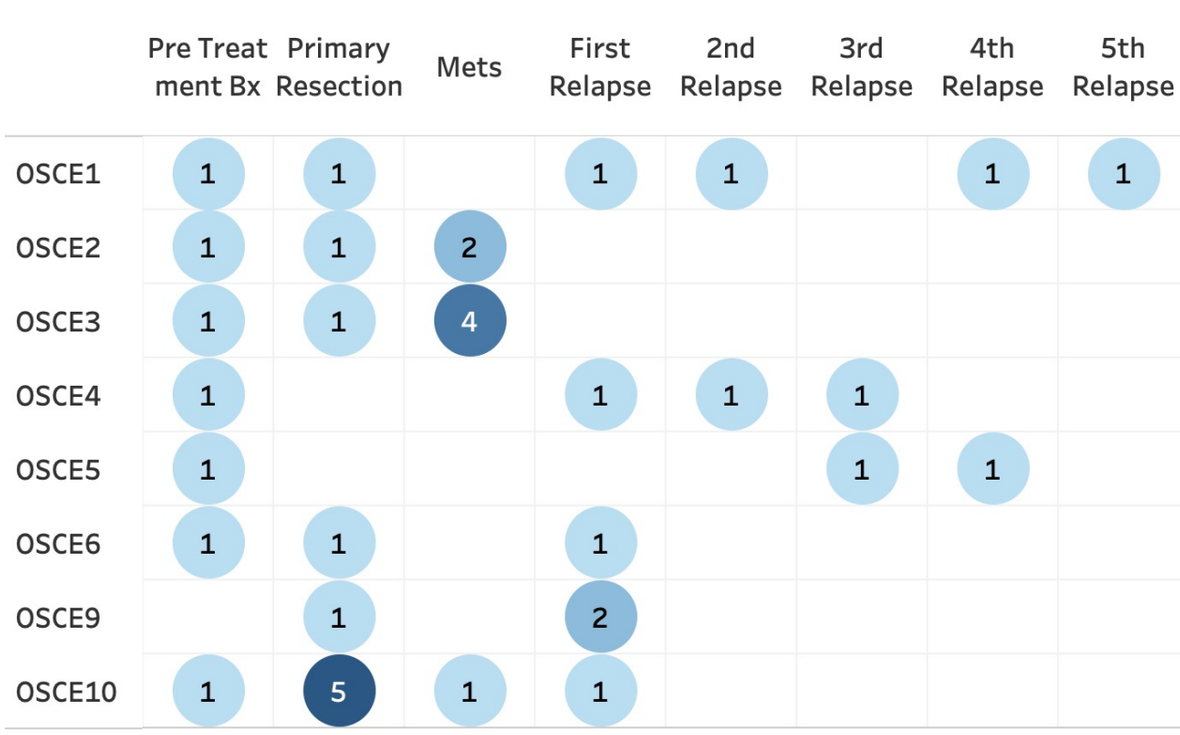
958 45. Boot A, Huang MN, Ng AWT, et al. In-depth characterization of the cisplatin mutational
959 signature in human cell lines and in esophageal and liver tumors. *Genome Res*.
960 2018;28(5):654-665. doi:10.1101/GR.230219.117
961 46. Medina-Martínez JS, Arango-Ossa JE, Levine MF, et al. Isabl Platform, a digital biobank
962 for processing multimodal patient data. *BMC Bioinformatics*. 2020;21(1):1-18.
963 doi:10.1186/S12859-020-03879-7/TABLES/1
964 47. Shukla N, Levine MF, Gundem G, et al. Feasibility of whole genome and transcriptome
965 profiling in pediatric and young adult cancers. *Nature Communications 2022 13:1*.
966 2022;13(1):1-15. doi:10.1038/s41467-022-30233-7
967 48. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations
968 in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013;31(3):213-219.
969 doi:10.1038/nbt.2514
970 49. Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and
971 somatic variants. *Nat Methods*. 2018;15(8):591-594. doi:10.1038/s41592-018-0051-x
972 50. Jones D, Raine KM, Davies H, et al. cgpCaVEManWrapper: Simple Execution of CaVEMan
973 in Order to Detect Somatic Single Nucleotide Variants in NGS Data. *Curr Protoc*
974 *Bioinformatics*. 2016;56:15.10.1-15.10.18. doi:10.1002/cpbi.20
975 51. Robbe P, Popitsch N, Knight SJL, et al. Clinical whole-genome sequencing from routine
976 formalin-fixed, paraffin-embedded specimens: pilot study for the 100,000 Genomes
977 Project. *Genetics in Medicine 2018 20:10*. 2018;20(10):1196-1205.
978 doi:10.1038/gim.2017.241
979 52. Gröbner SN, Worst BC, Weischenfeldt J, et al. The landscape of genomic alterations
980 across childhood cancers. *Nature*. 2018;555(7696):321-327. doi:10.1038/nature25480
981 53. Reiter JG, Makohon-Moore AP, Gerold JM, et al. Reconstructing metastatic seeding
982 patterns of human cancers. *Nat Commun*. 2017;8:1-10. doi:10.1038/ncomms14114
983 54. Shah RH. iAnnotateSV: Annotation of structural variants detected from NGS. Published
984 2021. Accessed November 2, 2021. https://github.com/rhshah/iAnnotateSV
985 55. Lopez G, Egolf LE, Giorgi FM, Diskin SJ, Margolin AA. Svpluscnv: Analysis and
986 Visualization of Complex Structural Variation Data . *Bioinformatics*. 2020;(October):1-3.
987 doi:10.1093/bioinformatics/btaa878
988 56. Cortés-Ciriano I, Lee JJK, Xi R, et al. Comprehensive analysis of chromothripsis in 2,658
989 human cancers using whole-genome sequencing. *Nature Genetics 2020 52:3*.
990 2020;52(3):331-341. doi:10.1038/s41588-019-0576-7
991 57. Crowdis J, He MX, Reardon B, van Allen EM. CoMut: visualizing integrated molecular
992 information with comutation plots. *Bioinformatics*. 2020;36(15):4348-4349.
993 doi:10.1093/BIOINFORMATICS/BTAA554
994 58. Smith MA, Nielsen CB, Chan FC, et al. E-scape: interactive visualization of single-cell
995 phylogenetics and cancer evolution. *Nature Methods 2017 14:6*. 2017;14(6):549-550.
996 doi:10.1038/nmeth.4303
997

Figure 1

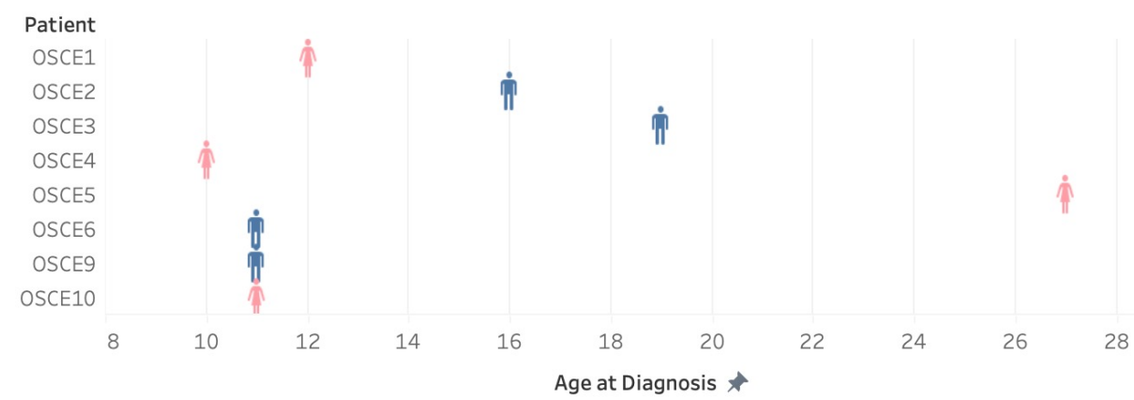
A



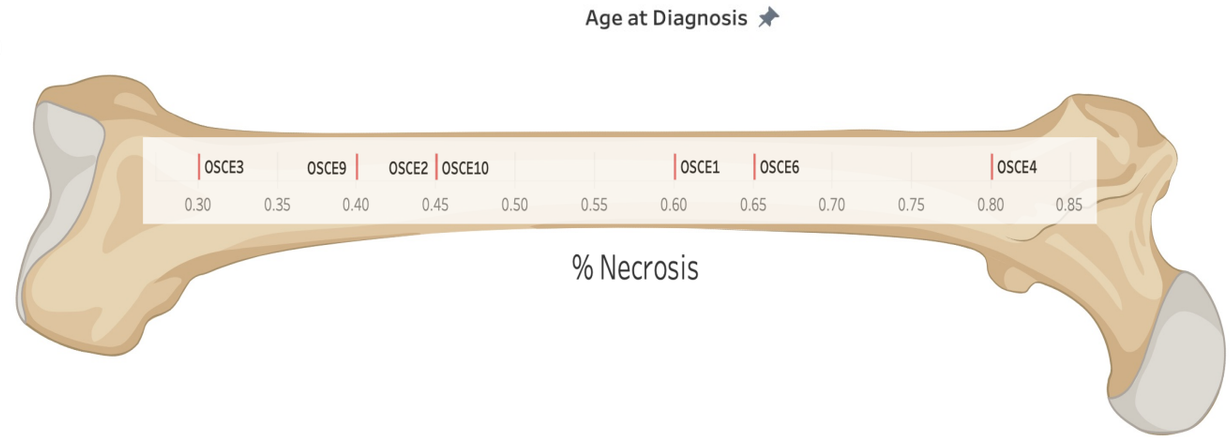
B



C



D



E

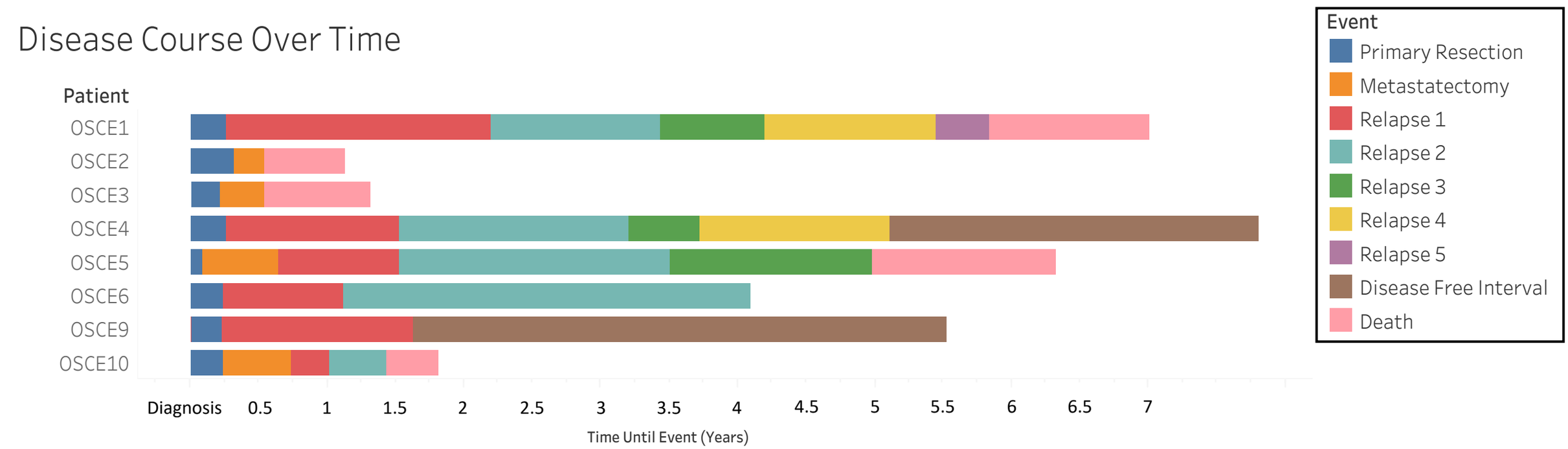


Figure 2

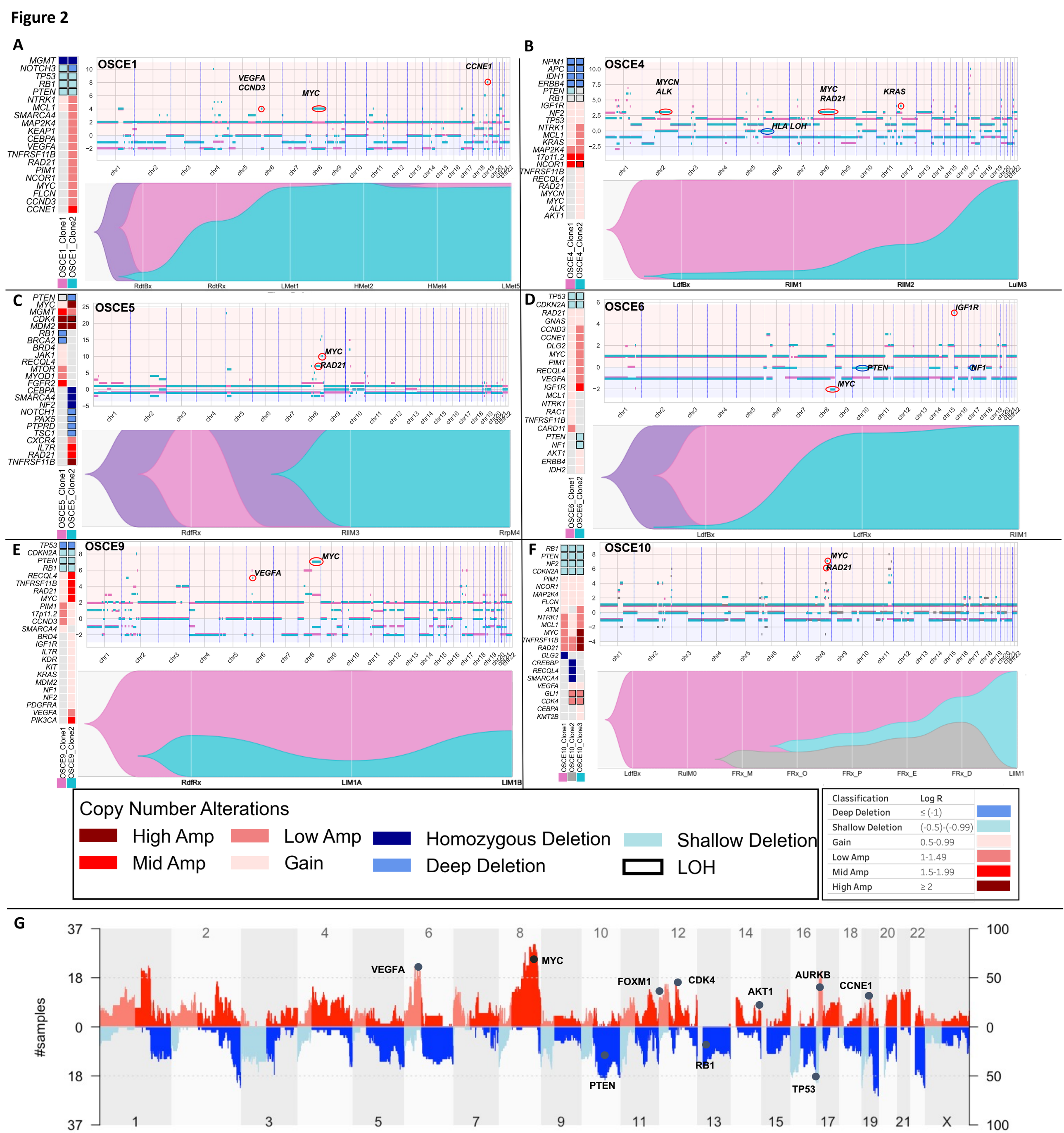

A
OSCE1
VEGFA
CCND3
MYC
CCNE1
OSCE1_Clone1
OSCE1_Clone2
RdtBx
RdtRx
LMet1
HMet2
HMet4
LMet5
B
OSCE4
MYCN
ALK
MYC
RAD21
KRAS
HLA LOH
OSCE4_Clone1
OSCE4_Clone2
LdfBx
RIIM1
RIIM2
LuIM3
C
OSCE5
MYC
RAD21
OSCE5_Clone1
OSCE5_Clone2
RdfRx
RIIM3
RrpM4
D
OSCE6
IGF1R
PTEN
NF1
MYC
OSCE6_Clone1
OSCE6_Clone2
LdfBx
LdfRx
RIIM1
E
OSCE9
MYC
VEGFA
OSCE9_Clone1
OSCE9_Clone2
RdfRx
LIM1A
LIM1B
F
OSCE10
MYC
RAD21
OSCE10_Clone1
OSCE10_Clone2
OSCE10_Clone3
LdfBx
RuIM0
FRx_M
FRx_O
FRx_P
FRx_E
FRx_D
LIIM1
Copy Number Alterations
High Amp
Low Amp
Homozygous Deletion
Shallow Deletion
Mid Amp
Gain
Deep Deletion
LOH
Classification
Log R
Deep Deletion
≤ (-1)
Shallow Deletion
(-0.5)-(-0.99)
Gain
0.5-0.99
Low Amp
1-1.49
Mid Amp
1.5-1.99
High Amp
≥ 2
G
#samples
VEGFA
MYC
FOXM1
CDK4
AKT1
AURKB
CCNE1
PTEN
RB1
TP53

Figure 3

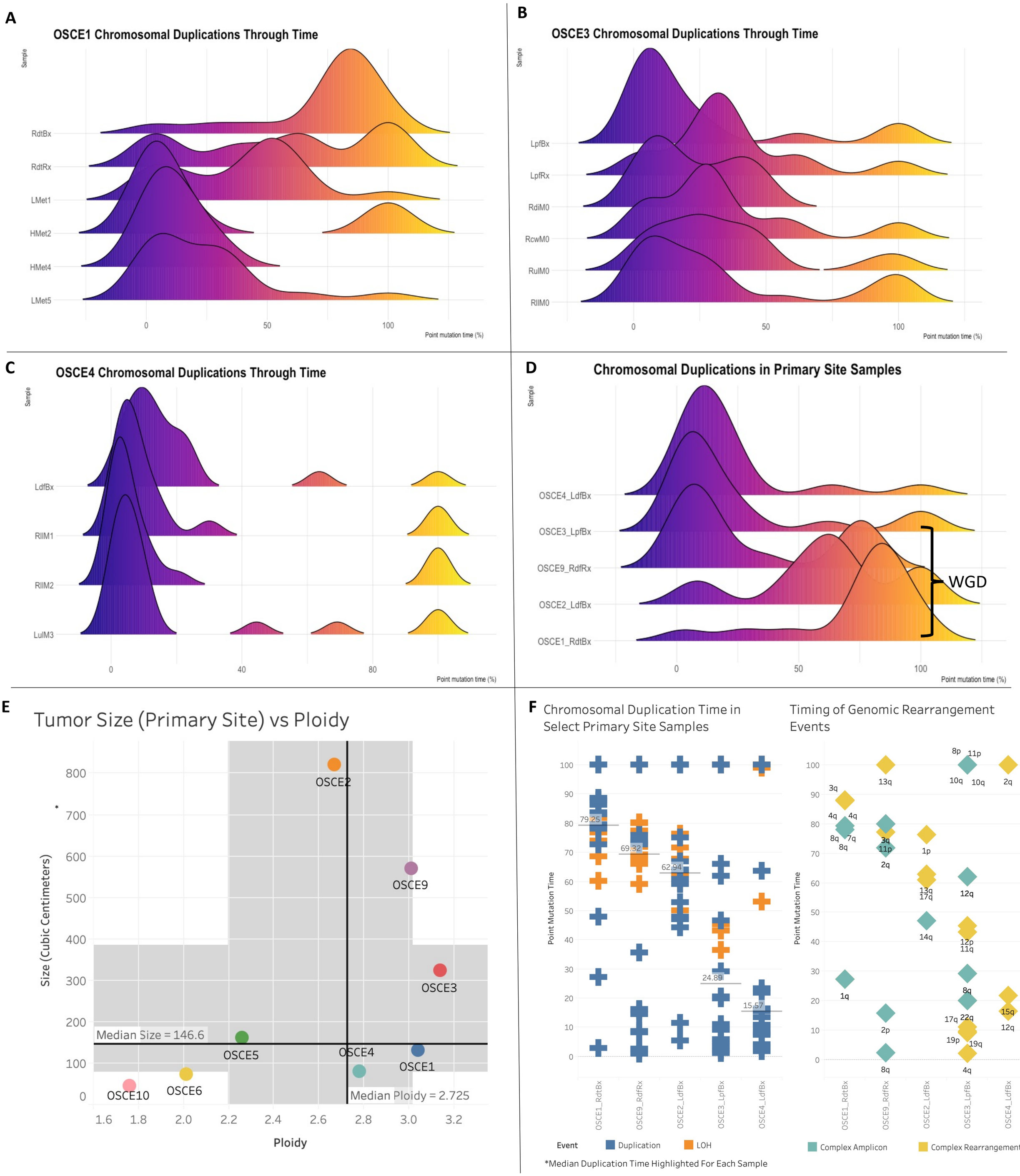

A
OSCE1 Chromosomal Duplications Through Time
Sample
RdtBx
RdtRx
LMet1
HMet2
HMet4
LMet5
0
50
100
Point mutation time (%)
B
OSCE3 Chromosomal Duplications Through Time
Sample
LpfBx
LpfRx
RdiM0
RcwM0
RulM0
RllM0
0
50
100
Point mutation time (%)
C
OSCE4 Chromosomal Duplications Through Time
Sample
LdfBx
RllM1
RllM2
LulM3
0
40
80
Point mutation time (%)
D
Chromosomal Duplications in Primary Site Samples
Sample
OSCE4_LdfBx
OSCE3_LpfBx
OSCE9_RdfRx
OSCE2_LdfBx
OSCE1_RdtBx
WGD
0
50
100
Point mutation time (%)
E
Tumor Size (Primary Site) vs Ploidy
Size (Cubic Centimeters)
800
700
600
500
400
300
200
100
0
OSCE2
OSCE9
OSCE3
Median Size = 146.6
OSCE5
OSCE4
OSCE1
OSCE10
OSCE6
Median Ploidy = 2.725
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
Ploidy
F
Chromosomal Duplication Time in Select Primary Site Samples
Point Mutation Time
79.25
69.32
62.94
24.89
15.57
OSCE1_RdtBx
OSCE9_RdfRx
OSCE2_LdfBx
OSCE3_LpfBx
OSCE4_LdfBx
Event
Duplication
LOH
*Median Duplication Time Highlighted For Each Sample
Timing of Genomic Rearrangement Events
Point Mutation Time
3q
4q
4q
8q
7q
8q
13q
3q
11p
2q
1p
13q
17q
14q
8p
11p
10q
10q
2q
12q
12p
11q
1q
8q
2p
8q
17q
22q
19p
19q
4q
15q
12q
OSCE1_RdtBx
OSCE9_RdfRx
OSCE2_LdfBx
OSCE3_LpfBx
OSCE4_LdfBx
Complex Amplicon
Complex Rearrangement

Figure 4

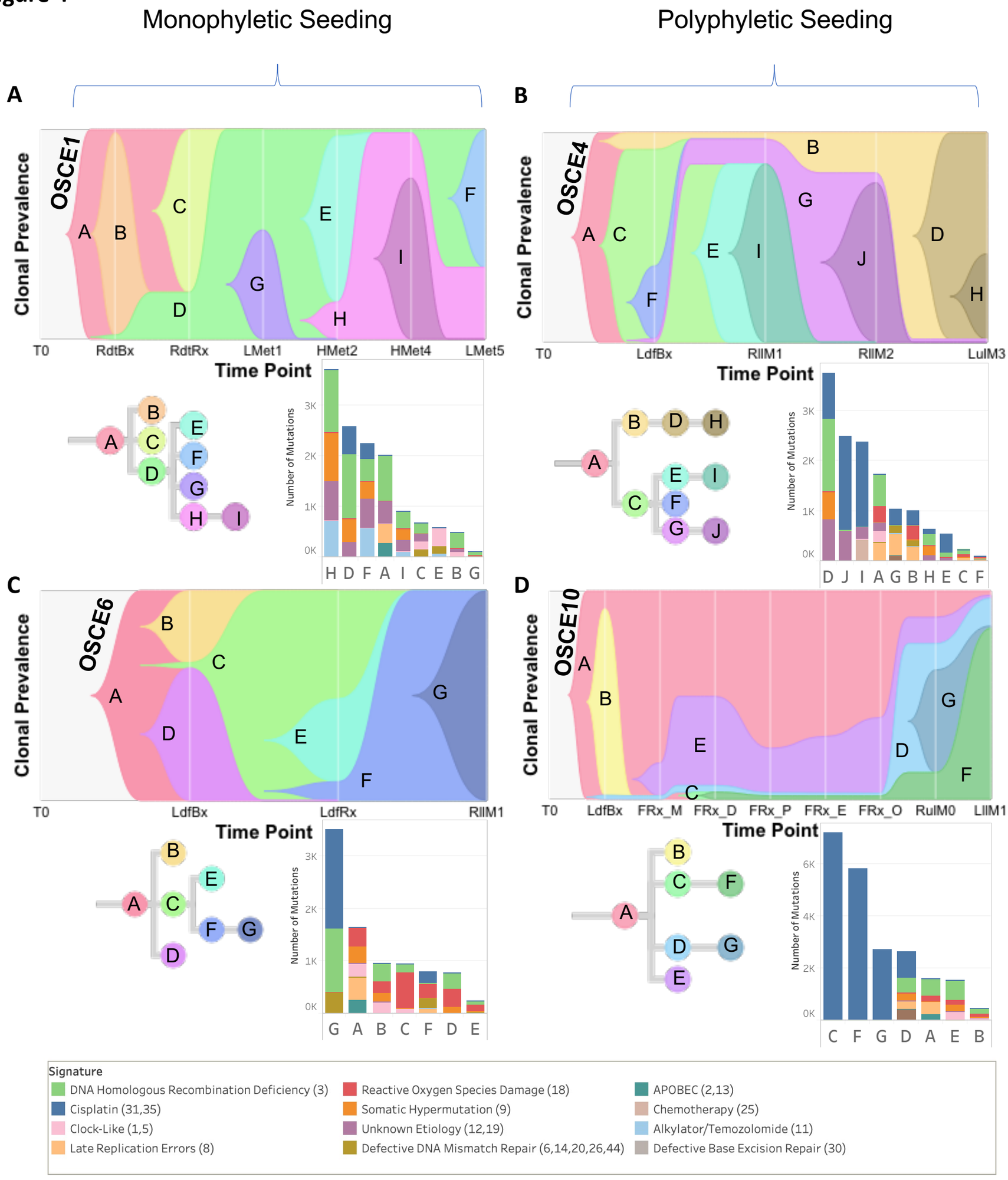
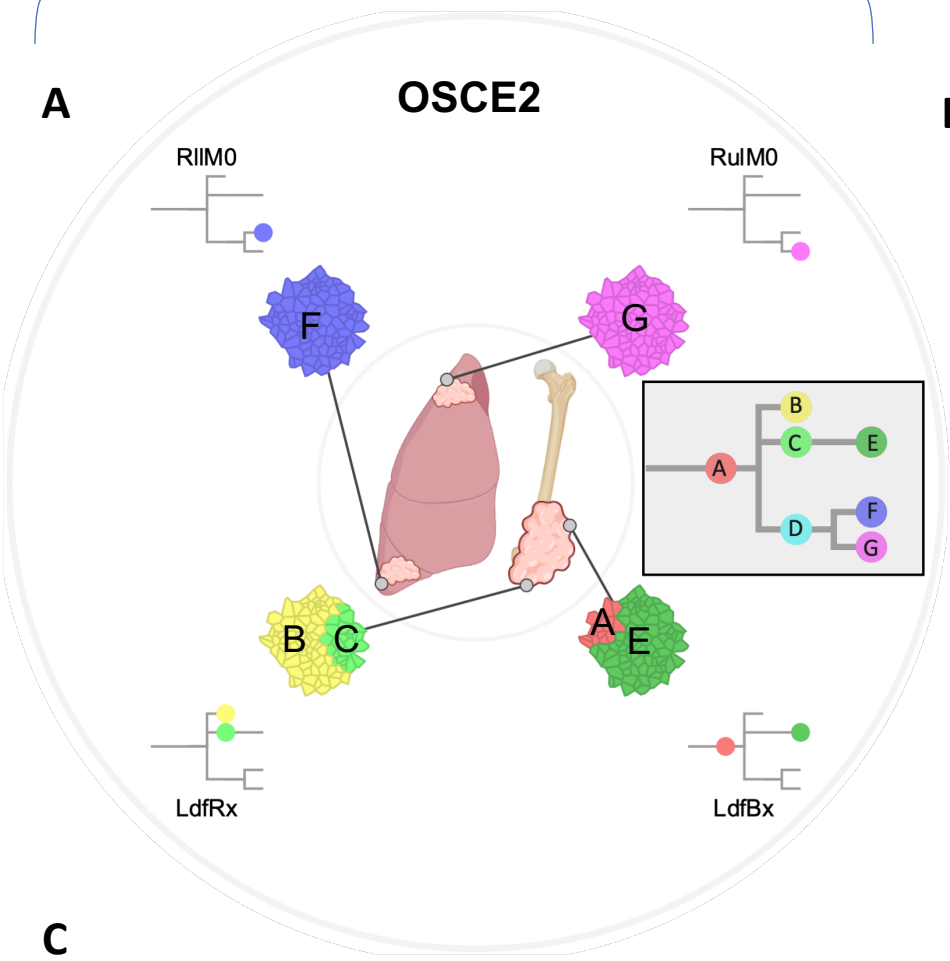
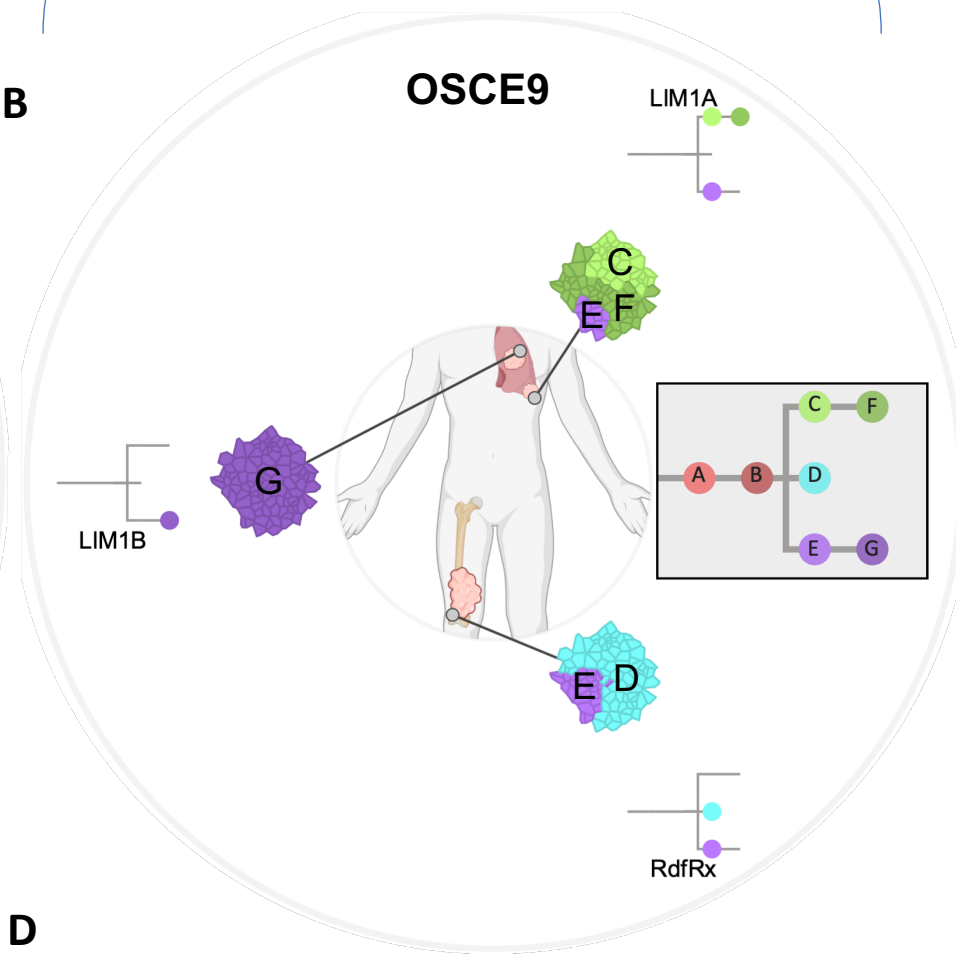

Monophyletic Seeding
Polyphyletic Seeding
A
B
C
D
OSCE1
OSCE4
OSCE6
OSCE10
Clonal Prevalence
Time Point
T0
RdtBx
RdtRx
LMet1
HMet2
HMet4
LMet5
LdfBx
RIIM1
RIIM2
LuIM3
LdfRx
RIIM1
FRx_M
FRx_D
FRx_P
FRx_E
FRx_O
RuIM0
LIIM1
Number of Mutations
0K
1K
2K
3K
4K
6K
H D F A I C E B G
D J I A G B H E C F
G A B C F D E
C F G D A E B
Signature
DNA Homologous Recombination Deficiency (3)
Cisplatin (31,35)
Clock-Like (1,5)
Late Replication Errors (8)
Reactive Oxygen Species Damage (18)
Somatic Hypermutation (9)
Unknown Etiology (12,19)
Defective DNA Mismatch Repair (6,14,20,26,44)
APOBEC (2,13)
Chemotherapy (25)
Alkylator/Temozolomide (11)
Defective Base Excision Repair (30)

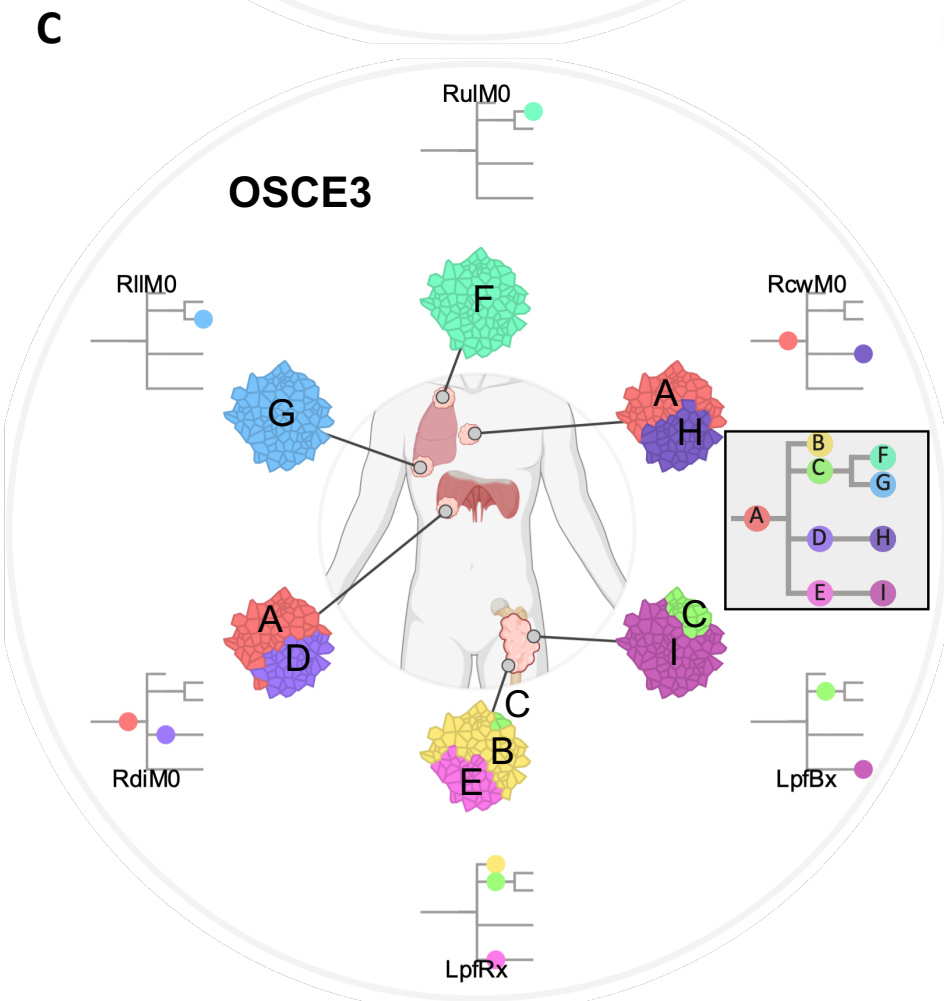
Figure 5

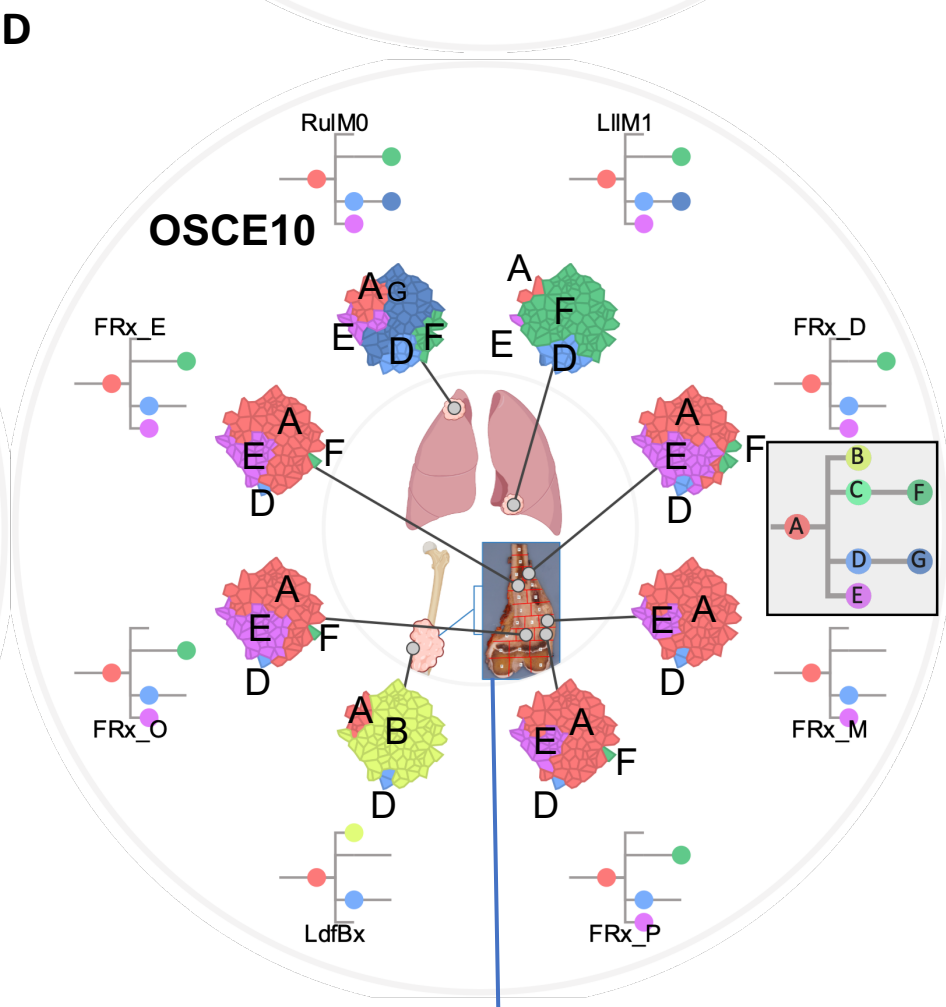
Monophyletic Seeding

Polyphyletic Seeding









Polyphyletic Seeding

E

Sample Level Dissemination Pattern

OSCE1	OSCE1_LMet1
	OSCE1_HMet2
	OSCE1_HMet4
	OSCE1_LMet5
OSCE2	OSCE2_RlIM0
	OSCE2_RulM0
OSCE3	OSCE3_RcwM0
	OSCE3_RdiM0
	OSCE3_RlIM0
	OSCE3_RulM0
OSCE4	OSCE4_RlIM1
	OSCE4_RlIM2
	OSCE4_LulM3
OSCE5	OSCE5_RlIM3
	OSCE5_RrpM4
OSCE6	OSCE6_RlIM1
OSCE9	OSCE9_LIM1A
	OSCE9_LIM1B
OSCE10	OSCE10_RulM0
	OSCE10_LIIM1

Patient Level Metastatic Dissemination Pattern

OSCE1
OSCE2
OSCE3
OSCE4
OSCE5
OSCE6
OSCE9
OSCE10

Dissemination Pattern
- Monoclonal Monophyletic
- Polyclonal Monophyletic
- Polyclonal Polyphyletic

F

OSCE10 - Jaccard Similarity Index (SNVs)

Patient	Sample1	Sample2	Patient OSCE10
OSCE10	FRx_D	FRx_E	0.9553
		FRx_O	0.6771
		FRx_M	0.6715
		FRx_P	0.6676
	FRx_E	FRx_P	0.6977
		FRx_O	0.6968
		FRx_M	0.6960
	FRx_M	FRx_P	0.8763
		FRx_O	0.6773
	FRx_O	FRx_P	0.7981

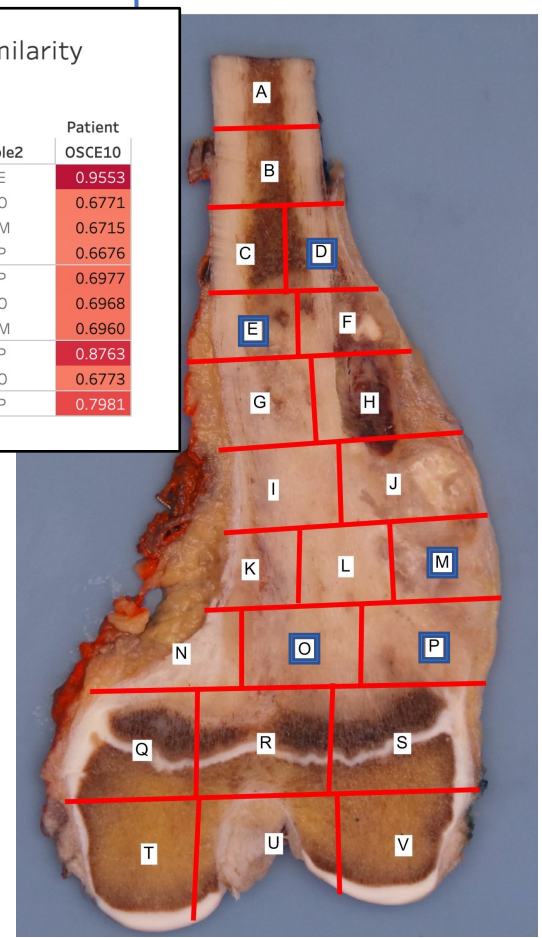
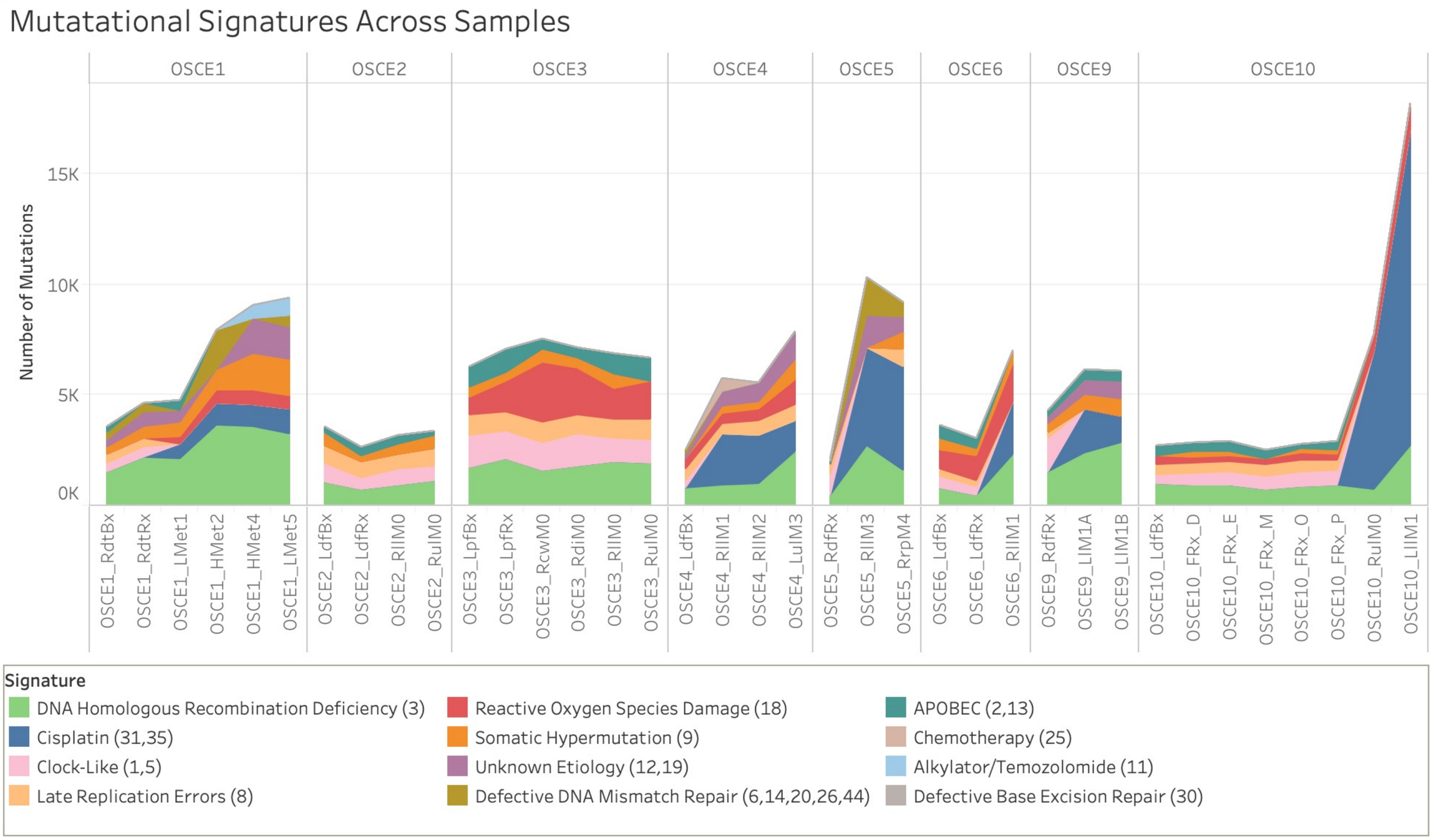
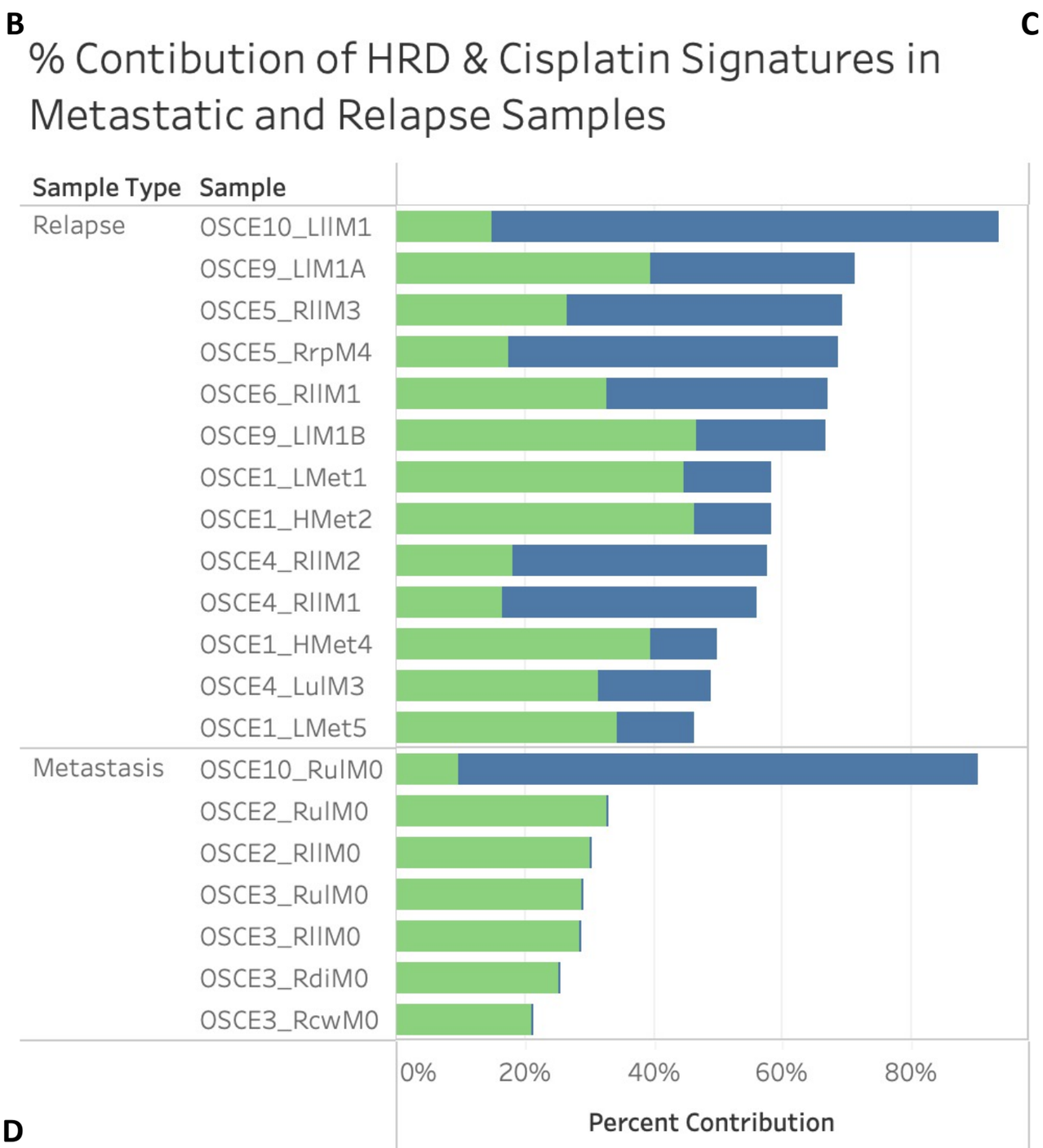


Figure 6

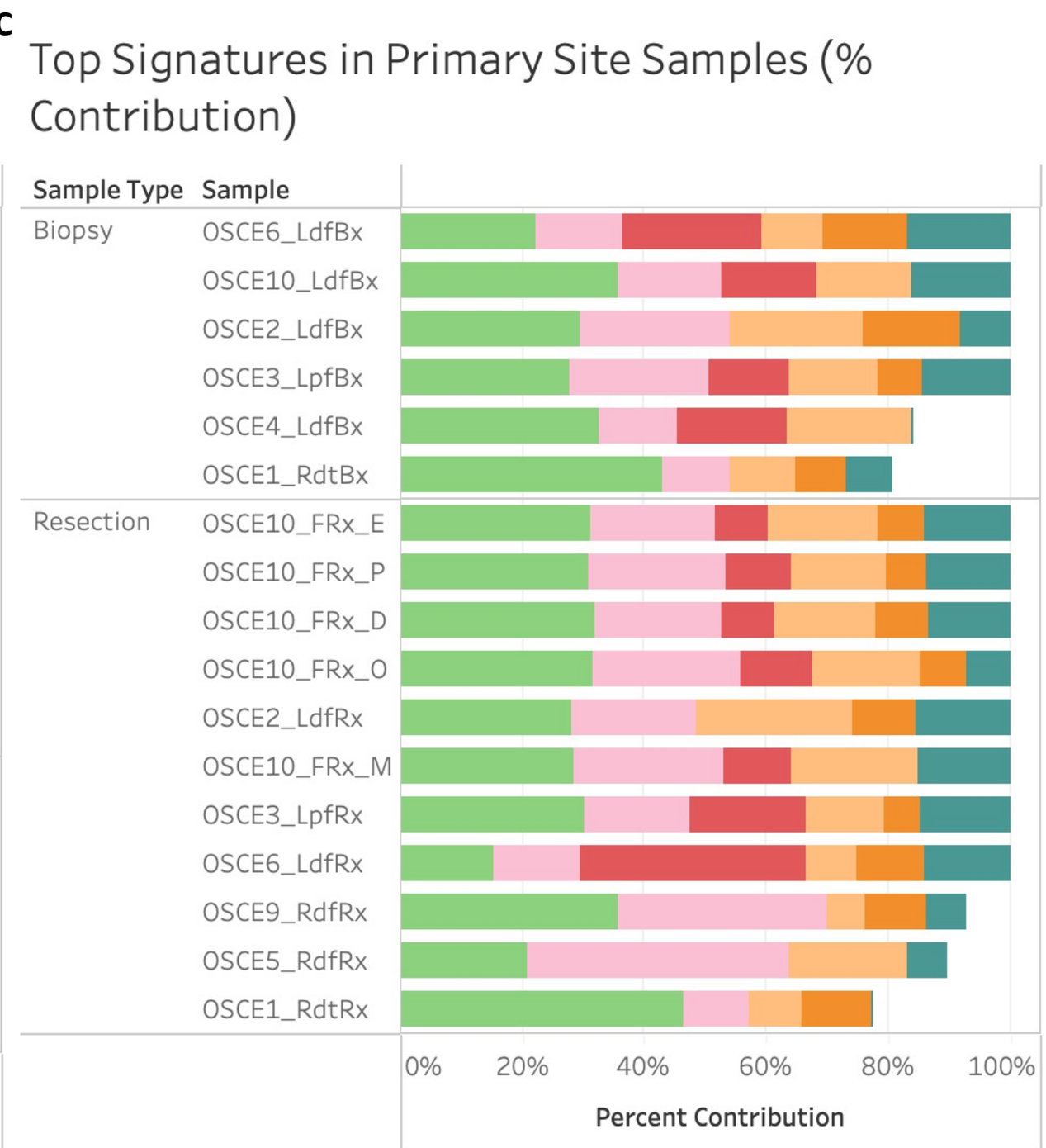
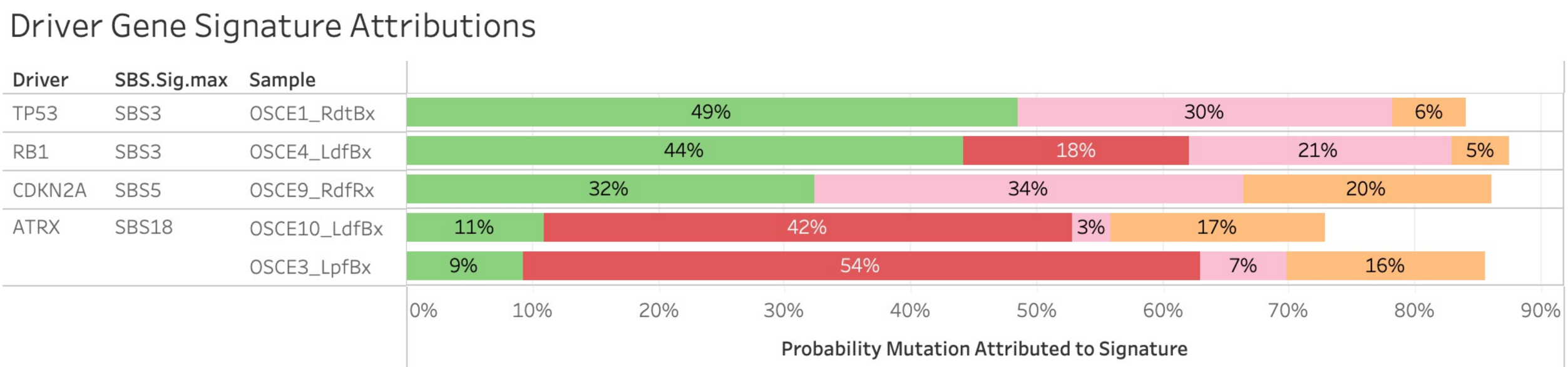
A



B



C



D