Osteosarcoma tumors maintain intratumoral heterogeneity, even while adapting to environmental pressures that drive clonal selection

Sanjana Rajan[1,2], Maren Cam[2], Amy C. Gross[2], Cenny Taslim[2], Meng Wang[2], Emily Franz[1,2], Ryan D. Roberts[2,3,4*]

INSTITUTIONS:

1. Molecular, Cellular, and Developmental Biology Program, The Ohio State University, Columbus, Ohio, USA.
2. Center for Childhood Cancers and Blood Diseases, Abigail Wexner Research Institute at Nationwide Children's Hospital, Ohio, USA.
3. The Ohio State University James Comprehensive Cancer Center, Columbus, Ohio, USA.
4. Division of Pediatric Hematology, Oncology, and BMT, Nationwide Children's Hospital, Columbus, Ohio, USA

Running title: Osteosarcoma retains phenotypic heterogeneity

Key words: Osteosarcoma, heterogeneity, phenotypic plasticity, clonal selection, microenvironment, metastasis

Financial support: This work was supported by funding provided by: NIH/NCI (K08CA201638, RDR), St Baldrick's Foundation Scholar Award (RDR), Hyundai Hope on Wheels Young Investigator Award (RDR), CancerFree Kids Foundation (RDR), Steps for Sarcoma Foundation (RDR), a Pelotonia Fellowship (SR), and an NIH CTSA Grant UL1TR002733.

Correspondence:

Ryan D. Roberts

Division of Pediatric Hematology, Oncology, and BMT

Nationwide, Children's Hospital

700 Children's Drive, Columbus, OH 43205

Email: Ryan.Roberts@nationwidechildrens.org

Phone: 614-722-29966; Fax: 614-355-2927

Disclosure of potential conflicts of interest: The authors declare no potential conflicts of interest.

Text words: 5850

Figures: 6

Supplementary figures and files: 2 figures, 3 files

Abstract

Osteosarcoma exhibits marked patient-to-patient heterogeneity, but little is known about heterogeneity within individual tumors. This study focuses on the roles that phenotypic plasticity and clonal selection play as tumors adapt to primary and metastatic microenvironments. We show that osteosarcomas have a high degree of transcriptional heterogeneity, like osteoblasts, that is retained even after prolonged cell culture or adaptation to changing microenvironments. We find that both cell lines and PDXs grown in cell culture or as flank tumors adopt markedly different transcriptional profiles when grown as primary bone tumors or metastatic lung lesions. By combining lineage tracing with single-cell transcriptomics, we find that very little clonal selection occurs when tumors grow in the tibia, but significant expansion of select clones occurs when grown as experimental metastases in the lung. Interestingly, the selective pressures that drive clonal expansion do not cause narrowing of transcriptional phenotypes. By comparing the phenotypes from transcriptional clusters in orthotopic/metastatic tumor pairs, we identify a transcriptional signature that is shared among clusters that become enriched during lung colonization. This includes responses to IFNG, TNF, PDGF, previously unidentified IL1B, and a shift away from genes associated with glycolysis and toward those associated with oxidative metabolism. The metastatic microenvironment enriches for phenotypically diverse clones that each display metabolic properties that engender fitness within the metastatic microenvironment. Together, these data suggest that an underlying program, possibly a developmental program retained from the tissue of origin, maintains phenotypic heterogeneity, even during adaptation to changing microenvironmental conditions.

Statement of significance: The work identifies a novel tumor intrinsic mechanism that allows for maintenance of phenotypic diversity while responding plastically to changing microenvironment conditions.

Introduction

Osteosarcoma is the most common malignant tumor of bone in pediatric patients[1]. This aggressive disease presents in the metaphyseal region of the long bones, coinciding with the anatomical location for bone growth[2]. Treatment consists of surgical resection of the primary lesion with adjuvant chemotherapy. This approach has not changed since the advent of chemotherapy in the 1980s, and as a consequence patient survival has not improved in forty years[3]. Indeed, over eighty percent of patients not treated with chemotherapy still develop—and die from—lung metastasis[4]. With a chaotic genomic landscape, high degree of genetic and phenotypic heterogeneity and no clear etiology, osteosarcoma is extremely challenging to treat and remains at a scientific and clinical impasse[5–7]. Whether the phenotypes that facilitate survival and growth within the metastatic lung environment arise from dynamic adaptive responses of individual tumor cells, or from the selection of clones with specific characteristics making them especially fit for the metastatic niche remains unclear. The development of innovative therapies for inhibiting metastatic progression will require a deeper and more detailed understanding of the role of intra-tumor heterogeneity as osteosarcoma cells adapt and evolve in the metastatic microenvironment. Although mathematical models are being developed to eliminate inconsequential mutational noise from genomic data, genetic heterogeneity is still a poor proxy for functional heterogeneity[8]. Phenotypic heterogeneity at the gene expression level is of greater clinical relevance, and may one day open the door to more effective treatments[9]. This work therefore focuses on phenotypes defined by transcription.

Sources of non-genetic heterogeneity in cancers are commonly attributed to the stochastic nature of biochemical processes[10], epigenetic variation[11], and differentiation hierarchies that arise from progenitor cells[12,13]. With the advent of single cell transcriptomics, functional heterogeneity can now be investigated at the non-genetic level by exploring variation in gene expression profiles and cellular states of tumor cells on a cell-by-cell basis[14]. These technologies have been applied to various cancers, including medulloblastomas and glioblastomas, to identify cellular states and their associated transcriptional signatures[15,16]. Similar studies in Ewing sarcoma, an otherwise genomically uncomplicated disease, have

revealed transcriptionally distinct cell states driven by oncoprotein expression levels[17]. A recent study revealed that osteosarcomagenesis can follow a neutral evolution model, with metastatic disease showing signs of polyclonality[18], but phenotypic heterogeneity associated with these clonal populations in osteosarcoma is yet to be investigated.

There are various models proposed to explain the evolutionary paradigms of clonal dynamics during cancer progression. The models are not mutually exclusive and include both the clonal selection and stem cell models. Both selection and neutral growth may operate simultaneously within the same tumor, and this may change dynamically over time[19]. While Nowell et al. formalized the concept of tumor evolution using the competitive linear model, Fidler et al. emphasized the importance of clonal diversity in metastasis[20]. Over the following decades, others have shown that selective pressures exerted by the microenvironment shape the evolutionary trajectory of both the tumor as a whole and the individual clonal lineages within it. These data show that malignant tumors are highly dynamic: they constantly evolve and adapt to changing microenvironment conditions, suggesting that the concept of a “magic bullet” for cancer is deeply flawed[21]. A deeper understanding of how tumors respond to and evolve within different microenvironments is necessary, especially in a heterogeneous cancer like osteosarcoma, if we are to develop treatments that anticipate the inevitable development of drug resistance.

To study how phenotypic heterogeneity evolves under various microenvironment conditions, we focus here on osteosarcoma cells grown in plastic, as subcutaneous patient-derived xenografts, as orthotopic primary, or metastatic tumors. Through longstanding protocols at our Children’s Hospital, we have amassed a large collection of osteosarcoma clinical specimens to develop patient-derived xenografts (PDXs). In comparing the phenotypic profiles of traditional osteosarcoma cell lines to that of PDXs, grown *in vitro* or *in vivo*, we show here that cancer cell phenotypes are highly influenced by the microenvironment. This adaptive response to changing microenvironment conditions was not associated with a narrowing but rather a maintenance of transcriptional heterogeneity. We combined a lentivirus-based lineage tracing system with single cell transcriptomics to study phenotypic and clonal dynamics

simultaneously during tumor formation and progression. Lentivirus infection introduced a unique, heritable and polyA-tail capture compatible lineage tag in each osteosarcoma cell. Single-cell libraries were processed bioinformatically to overlay phenotype and lineage information for each cell. We identified that most cells were capable of plastically switching between phenotypes, that while orthotopic tumor in the bone retained clonal diversity, those in the lung displayed clonal expansion. Although the metastatic lung microenvironment selects for clonal populations with diverse phenotypes, the most enriched clonal families display phenotypic heritability. By overlaying phenotypic profiles of matched primary and metastatic tumor pairs, we identified tibia-affinity and lung-affinity phenotypes defined by gene sets differentially expressed in osteosarcoma cells in each of the two microenvironment conditions. Longitudinal comparison of these gene sets across models identified shared activation of canonical pathways and response to upstream regulators in the metastatic tumor relative to matched primary tumors. This signature includes response to IFNG, TNF, PDGF, a previously unidentified IL1β and a shift from glycolytic to oxidative metabolism, possibly driven by a decrease in MYC activity. These studies establish the conservation of phenotypic heterogeneity within osteosarcoma tumors, reveal shared transcriptional modules that are recurrently activated in response to orthotopic and metastatic microenvironments, and identify specific upstream regulators that facilitate these transcriptional adaptations in tumor cells. Clinically, these findings underscore the importance of designing combination therapies to halt metastatic progression and highlight the transcriptional plasticity underlying tumor cell responses to microenvironmental selective pressures.

Methods

Experimental model - PDXs, Cell culture and murine studies.

Patient derived xenografts (PDXs). OS17 PDX tissue was obtained from a primary femur biopsy performed at St. Jude's Children's Research Hospital in Memphis and was a gift from Peter Houghton[44]. Patient-derived samples, NCH-OS-2 and NCH-OS-7 were obtained from patients consented under an

Institutional Review Board (IRB)-approved protocol IRB11-00478 at Nationwide Children's Hospital (Human Subject Assurance Number 00002860).

Cell culture. OS17 cells were derived from an early passage of the OS17 PDX (see above). These cells were grown in RPMI (Corning, 10-040-CV) supplemented with 10% FBS. 143B cells were obtained from the American Type Culture Collection (ATCC, CRL-8303) and grown in DMEM (Corning, 10-013-CV) supplemented with 10% FBS.

Murine Studies A. Flank. Cryopreserved viable tissue pieces from OS17, 143B, NCH-OS-2 and NCH-OS-7 PDX tumors were placed in the right flank of recipient C.B-17/IcrHsd-Prkdc^{scid} mice following IACUC guidelines. These subcutaneous tumors were allowed to grow to 300 mm3 before excision. These were then prepped for single-cell RNA-seq. *B. Orthotopic.* Single cell suspensions of OS17, 143B, NCH-OS-2 and NCH-OS-7 were injected intra-tibially in C.B-17/IcrHsd-Prkdc^{scid} mice as per IACUC guidelines. Primary tumors were excised once they grew to 800 mm3, and prepped for single-cell RNA-seq. *C. Experimental metastasis.* Our lab has previously established this model. Briefly, C.B-17/IcrHsd-Prkdc^{scid} mice received intra-venous injections of single cell suspensions of OS17, 143B, NCH-OS-2 and NCH-OS-7. Lungs were harvested once these mice reached endpoint. Endpoint criteria for euthanasia was defined as weight loss of >10% or a body condition score (BCS) of <10.

Single-cell RNA-seq

Tumor harvested from mice were processed using the human tumor dissociation kit (Miltenyi Biotec, 130-095-929) with a GentleMacs Octo Dissociator with Heaters (Miltenyi Biotec, 130-096-427). Single cell suspensions in 0.04% BSA-PBS of cell lines, and dissociated tumor tissues were generated and run on Chromium Single Cell 3′RNA-sequencing system (10x Genomics) with the Reagent Kit v3 according to the manufacturer's instructions. Briefly, cells were loaded into Chromium Single Cell Chip v3 (10x Genomics) with a targeted cell recovery of 5,000 cells per sample. After performing cDNA purification,

amplification, and library construction (sample index PCR 14 cycles) as instructed, we sequenced sample libraries on a half lane of HS4000 (Illumina) to yield (after QA) about 65,000 paired-end reads per cell.

Cellular Barcoding

We synthesized barcoded lentivirus library using CloneTracker XP™ 10M Barcode-3' Library with Venus-Puro (plasmid) (Cellecta, BCXP10M3VP-P) as described in Wang and McManus 2009 [45] with PEI as the transfection reagent. We infected OS17 cells with in-house prepared virus library in the presence of Polybrene (8 µg/ml). The barcoded plasmids contain ~10M 38-bp semi-random oligonucleotide sequence that is captured on the Chromium Single Cell 3′RNA-sequencing system. 48 h post-infection, OS17 cells were used to generate orthotopic and experimental metastasis models as described above.

Cellular Barcoding Computational Analysis

Raw sequencing data was pre-processed to extract the lineage tracing (LT) barcode and the matching cell ID (CID) using known flanking sequences. Lines without high-confidence cell ID assignment were removed and redundant reads were de-duplicated. The extracted LT barcode reads were matched against the know Cellecta barcode library. Reads with barcodes that didn't match the Cellecta library were eliminated. The LT barcode for each matching CID was integrated into the Seurat object metadata to allow for further analysis (see below).

Single-cell RNA-seq analysis

Cell Ranger version 3.0.2 (10x Genomics) was used to convert Illumina BCL files to FASTQ files. These FASTQ files were then de-multiplexed and aligned to the hg19 human reference genome provided by 10X Genomics, to generate gene-cell matrices. We used the Seurat R package[46–48] for quality control, filtering, and analysis of the data. Cells were filtered to remove doublets, low quality cells and mitochondrial genes. Cells with fewer than 800 expressed genes and genes expressed in fewer than 5 cells were filtered out. The total numbers of cells in each model after filtering were as follows: Osteoblasts (cell culture): 5735; OS17 (cell culture): 3,178; OS17 (tibia): 4,574; OS17(lung): 2,849; 143B (cell culture):

3,354; 143B (tibia): 6,187; 143B (lung): 5,134; NCH-OS-2 (flank): 6,749; NCH-OS-2 (tibia): 4,285; NCH-OS-2 (lung): 1,645; NCH-OS-7 (flank): 1,998; NCH-OS-7 (tibia): 2,891; NCH-OS-7 (lung): 5,133. Each model was downsampled to 1,500 for inter-model comparison. OS17 cells were downsampled to 2,800 cells per condition for inter-condition comparison. After log normalization of the dataset, we used a standard linear transformation ('scaling') step before performing dimensionality reduction techniques. We mitigated the effects of cell cycle heterogeneity by regressing out canonical G2/M- and S-Phase genes using the ScaleData function in Seurat. After pre-processing the data, we applied downstream analysis (both cell-level and gene-level approaches)[49] to extract underlying biological insight from our system.

The first 20 principal components were selected for each model based on the Elbow plot to capture majority of true signal. To cluster cells in an unsupervised manner, we applied the Louvain modularity optimization algorithm on the PCA reduced single cell dataset with resolution set between 0.1-2.5. Optimum resolution was set by maximizing silhouette scores calculated using the silhouette function in the R package cluster. Spatial distribution of the Louvain Clusters was visualized using UMAP, a non-linear dimensional reduction technique.

Marker genes were identified for each cluster relative to other clusters using the Wilcoxon rank-sum test, with p values adjusted for multiple testing. The cutoffs were set to adjusted P values < 0.01 and a fold change of >0.25. The enriched pathways and hallmarks (msigdb signatures) were identified using the enricher function in the R package clusterProfiler[50], with p values adjusted for multiple testing. The ITH score was calculated as described in Stewart et al.[23] Briefly, it was defined as the average Euclidean distance between the individual cells and all other cells, in terms of the first 20 principal components derived from the normalized expression levels of highly variable genes. Dispersion metric was calculated using the estimateDisp() function in EdgeR[51] package to measure biological coefficients of variations in RNA-seq expressions.

Statistical analysis

For gene expression differential analysis, the two-sided Wilcoxon rank-sum test was performed. For multiple testing, P values were adjusted using Bonferroni correction. P values were calculated using two-sided tests or the Wilcoxon rank-sum test. All statistical analyses were performed using the R software program[52].

Code availability

The bioinformatics analyses were performed using open-source software, including Seurat version 3.1.5[53], clusterProfiler version 3.14.3[50], and Cell Ranger version 3.0.2, as well as in-house R script that is available upon request.

Results

Transcriptional heterogeneity in cell line and PDX models of osteosarcoma

To investigate phenotypic heterogeneity within *in vitro* and *in vivo* models of osteosarcoma, we generated single-cell transcriptomic libraries from cell line (OS17) and PDX (NCH-OS-2) models grown in cell culture or as subcutaneous flank tumors respectively. Since osteosarcoma arises from transformed progenitor cells with osteoblastic differentiation and osteoid production[22], we compared the phenotypic heterogeneity in osteosarcoma cells relative to primary osteoblast cultures using our single-cell transcriptomics workflow (>2,500 cells per model). We aligned reads to the human genome, and this served as our in-built control to remove murine cells from downstream analysis. Using an unsupervised clustering approach, we identified clusters of cells with majority of the variance dominated by cell cycle associated genes, a common finding in highly proliferative malignant tissues (Figure S1A, B). To control for cell cycle effects, we normalized our data by generating a ‘corrected expression matrix’ so that effects directly related to cell cycle were reduced in downstream analysis (Figure S1C, D). After regressing out the effects of cell-cycle-related transcriptional programs, osteosarcoma cells and human osteoblasts revealed underlying transcriptional heterogeneity as identified by independent clusters within each

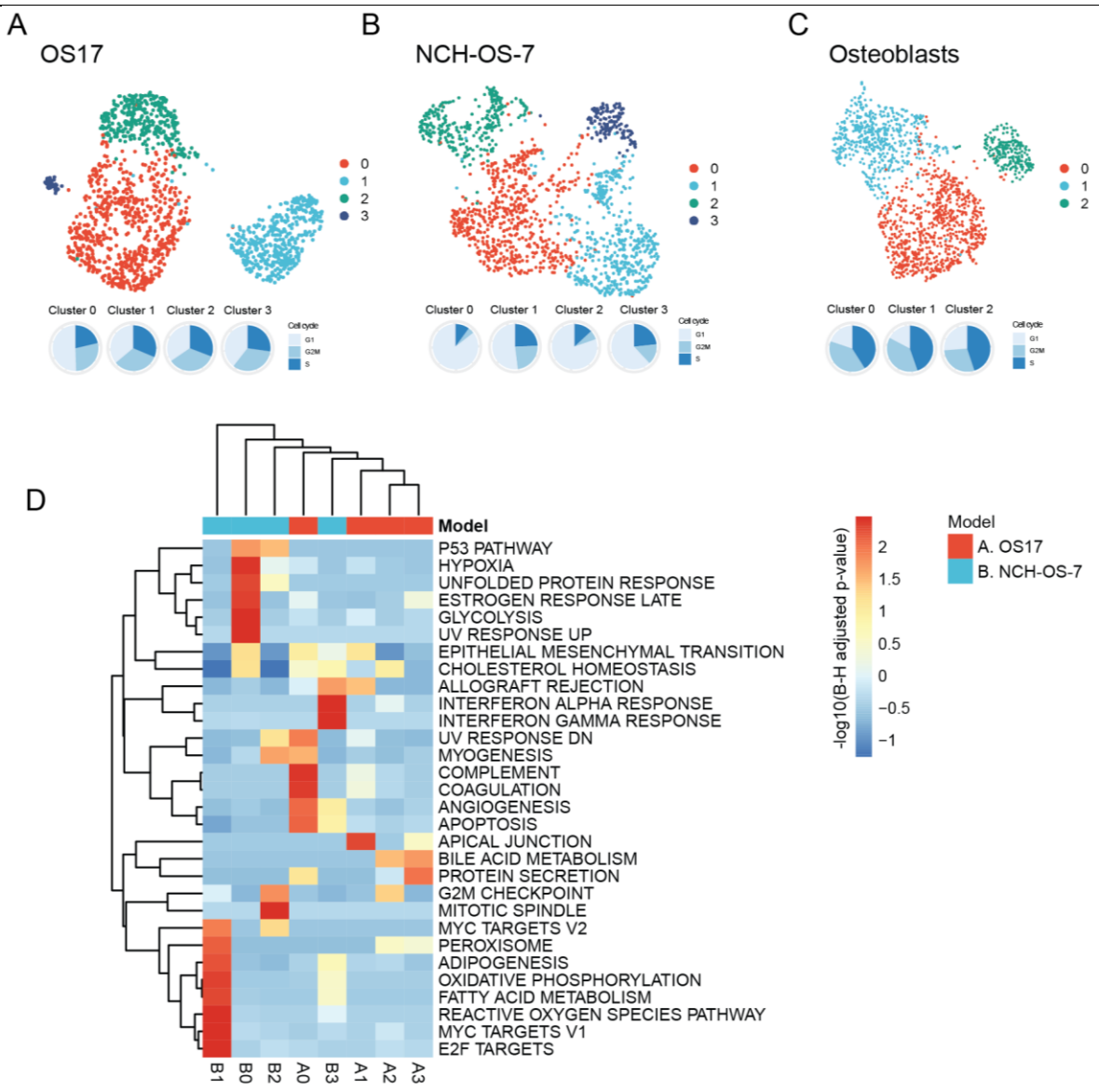

FIGURE 1: GSEA identifies functionally distinct groupings of cells in *in vitro* and *in vivo* models of osteosarcoma. A) UMAP analysis of OS17 cells grown in cell culture. Cell cycle distribution of cells is visualized as pie charts. B) UMAP analysis of NCH-OS-2 cells grown as subcutaneous flank tumor. Cell cycle distribution of cells is visualized as pie charts. C) UMAP analysis of human osteoblast cells grown in cell culture. Cell cycle distribution of cells is visualized as pie charts. D) GSEA for hallmark gene sets associated with clusters in the two osteosarcoma models grown in cell culture. Statistical significance was determined by enricher hypergeometric test and p values were adjusted for multiple comparisons. In A-D, n = 1,500 cells.

sample (Figure 1A, B, C). This observation provides evidence for the existence of transcriptional heterogeneity in osteosarcoma while also revealing underlying heterogeneity in normal osteoblast cells.

To identify the high-level features that defined each cluster within both of the osteosarcoma models, OS17 and NCH-OS-2, we performed Gene set enrichment analysis (GSEA). This allowed us to assess gene expression pathways associated with each cluster (Figure 1D). While certain clusters in OS17 cells were associated with functions such as protein secretion, angiogenesis and apoptosis, clusters in NCH-OS-7 were associated with estrogen response, unfolded protein response and fatty acid metabolism. Unsupervised hierarchical clustering revealed that the gene sets defining each cluster were largely mutually exclusive between the two models. In general, identification of gene sets exclusive to select clusters validated the distinct phenotypic identities of the sub-populations of cells in each of the models.

Model	Species	Type	Host	Endpoint	Characteristics
OS17	Human	Cell Line	Icr-SCID	70 days	low passage, very well-characterized, reliable mets
143B	Human	Cell Line	Icr-SCID	65 days	widely utilized, very well-characterized, reliable mets
NCH-OS-2	Human	PDX	Icr-SCID	200 days	Very low-passage, highly penetrant, from an untreated synchronous metastasis
NCH-OS-7	Human	PDX	Icr-SCID	184 days	Very low-passage, highly penetrant, from a heavily-pretreated, multiply-relapsed metastasis

Table 1. Characteristics of osteosarcoma models used in this study

Inter- and Intra- tumor heterogeneity in osteosarcoma models grown in different microenvironments.

To determine how osteosarcoma adapts to different microenvironments, we compared single cell gene expression libraries of cell lines and PDXs (**Error! Reference source not found.**; refer to Supplementary File S1 for full characteristics as per PDX-MI standards) grown as primary orthotopic tumors in the bone or as metastases in the lung (Figure 2A). Single cell RNA-seq analyses revealed that clustering of orthotopic and metastatic osteosarcoma cells were predominantly driven by the tumor model of origin (Figure 2B). Existence of such inter-tumor heterogeneity is a known characteristic of osteosarcoma tumors. Notably, osteosarcoma cells grown in cell culture or as flank tumors clustered distinctly separate

from their tibia and lung counterparts, except for NCH-OS-7 where clusters from the subcutaneous PDX overlapped with cells from orthotopic and metastatic tumors. This suggests that osteosarcoma cells, whether passaged *in vitro* or *in vivo,* adapted their transcriptional profiles in response to pressures imposed by each environment.

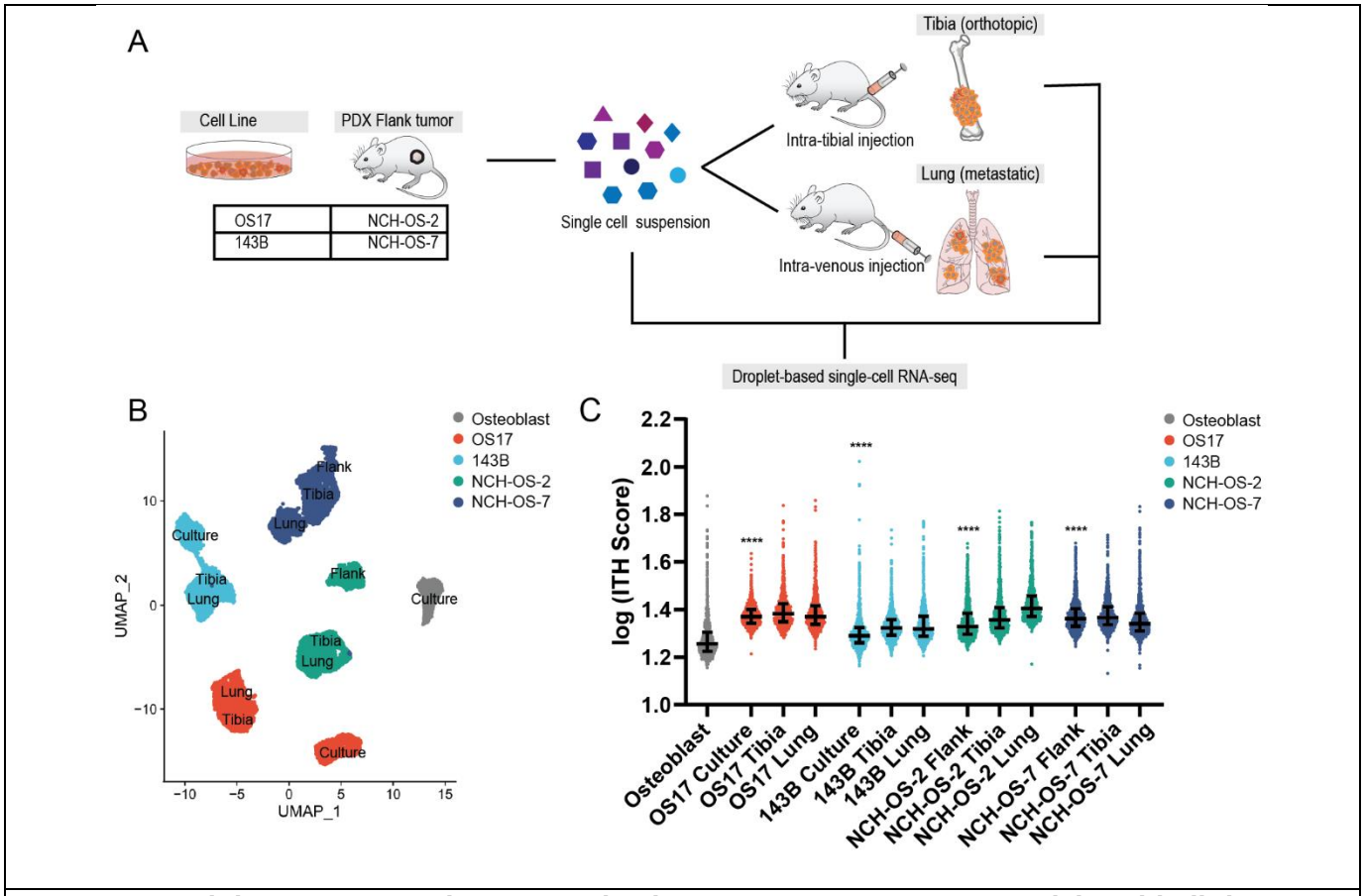

FIGURE 2: High Inter tumor heterogeneity between osteosarcoma models with little to no narrowing of transcriptional heterogeneity across environmental selective pressures. A) Schematic of study design depicting when samples were harvested for single-cell RNA-seq for the different models and conditions. B) UMAP analysis of four osteosarcoma models grown in three separate environments. Osteosarcoma cells from each model were more similar to themselves than to cells from other models. C) Osteosarcoma cells exhibited higher ITH scores than that of osteoblasts grown in cell culture, as determined by two-sided Wilcoxon rank-sum test (****p <0.0001, n = 1500 per model).

To estimate the degree of Intra-tumor heterogeneity (ITH) in our models, we computed ITH scores using genes with highly variable expression within individual samples after correcting for cell cycle heterogeneity[23]. When we compared the distribution of heterogeneity scores between human osteoblast cells and cancer cells, we found that irrespective of growth environment conditions, osteosarcoma cells

had significantly higher ITH scores (Figure 2C). Interestingly, osteoblasts themselves demonstrated some degree of heterogeneity. We hypothesized that changes in microenvironment conditions exert a selective pressure on osteosarcoma cells and cause a narrowing of ITH from cell culture to orthotopic tumors and from cell culture to metastatic lesions. Surprisingly, we found that osteosarcoma cells grown in bone and lung microenvironments had significantly higher ITH scores compared to those grown *in vitro* or as flank tumors, with NCH-OS-7 being the only exception--NCH-OS-7 showing slightly lower ITH in the lung tumors relative to the subcutaneous tumor (though still more heterogeneous than the osteoblast sample). Overall, these results show that osteosarcoma cells are more transcriptionally heterogeneous than normal osteoblast cells, that clustering of cells was primarily driven by tumor of origin, and that irrespective of model or microenvironment, osteosarcoma cells retained a high degree of intra-tumor transcriptional heterogeneity with little to no narrowing of heterogeneity in the orthotopic or lung microenvironments.

Shared transcriptional modules activated upon lung colonization in osteosarcoma

Over eighty percent of patients not treated with chemotherapy will develop metastasis in the lung[4]; this strict tissue tropism is a defining characteristic of osteosarcoma[24]. The metastatic bottleneck exerts a selective pressure that enriches for small groups of cells with specific traits that confer survival or growth advantage within the lung. To test if the lung microenvironment causes a convergent shift in phenotypic profiles of osteosarcoma cells, we compared transcriptional datasets from pairs of primary and metastatic tumors generated in each cell-line and PDX model. Certain clusters demonstrated higher fitness for growth within the lung environment while others for growth in the bone environment (Figure 3A). When we compared differentially upregulated genes between these lung-affinity phenotypes relative to tibia-affinity phenotypes, we found only two genes, EGR1 and FOS, that were shared as upregulated across the four models (Supplementary File S2). This observation corroborated the inter-tumor transcriptional heterogeneity identified in the previous figures.

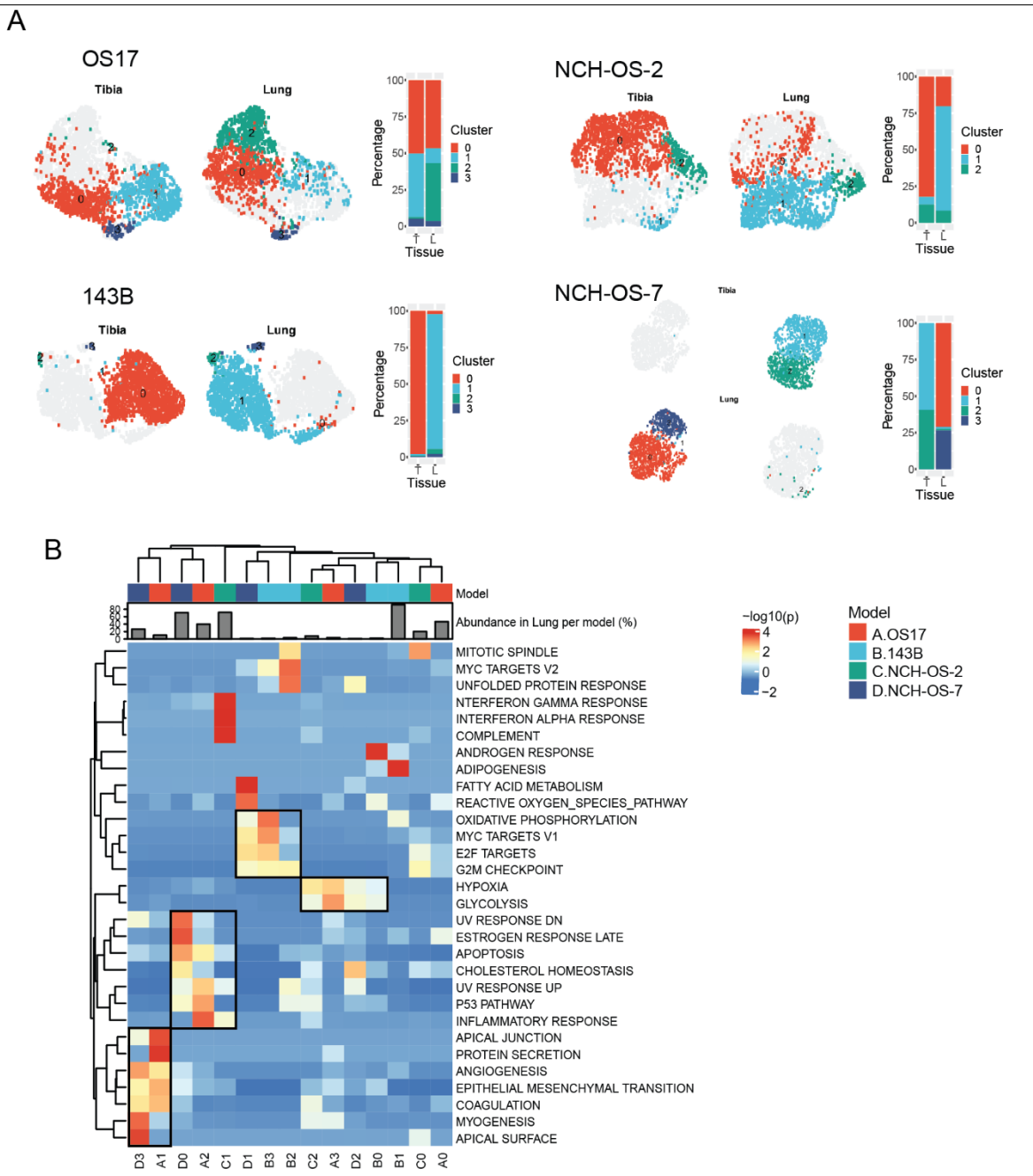

FIGURE 3: Metastatic lung enriches for features that are shared across different cell line-derived and PDX-derived models of metastatic OS. 1) UMAP analysis for merged orthotopic and metastatic samples in each of the four cell-line derived and PDX-derived models. Cluster enrichment analysis shows distribution of cells in each cluster in the two microenvironment conditions. In tumors that colonized lung tissues, certain clusters (namely cluster 2 in OS17 cells, cluster 1, 2 and 3 in 143B cells, cluster 1 in NCHOS2 cells, cluster 0 and 3 in NCHOS7 cells) demonstrate higher fitness for growth within the lung environment relative to other, less-enriched clusters. B) GSEA for hallmark gene sets associated with clusters in the four osteosarcoma models. Hierarchical clustering of the two axes identified blocks of transcriptional signatures that are shared in at least three of the models grown in either orthotopic bone or metastatic lung conditions. Statistical significance was determined by enricher

hypergeometric test and p values were adjusted for multiple comparisons. In A-B, n = 1,500 per condition.

Identification of individual shared genes can fail to detect biological processes, such as metabolic pathways and stress responses, that are regulated by network of genes due to redundancy in individual gene functionality[25]. This can be especially true when using techniques that generate sparse data, such as single-cell RNA-seq. To overcome this, we identified differentially expressed gene sets between the lung-affinity and tibia-affinity clusters and compared them across the four models. Using GSEA to evaluate the differentially expressed genes, we identified shared hallmark gene sets that are predicted to be activated in lung-affinity clusters compared to tibia-affinity clusters (Figure 3B). Hierarchical clustering of gene sets identified functional modules that are shared within either lung-affinity or tibia-affinity clusters. Tibia-affinity clusters were associated with functions such as hypoxia and glycolysis, while lung-affinity clusters were associated with UV response down, apoptosis and inflammatory response. These results suggest that core transcriptional modules make certain osteosarcoma cells more fit for survival/growth within the lung environment, irrespective of their overall transcriptional phenotype. These results suggest that core transcriptional modules make certain osteosarcoma cells more fit for survival/growth within the lung environment, irrespective of their overall transcriptional phenotype.

Lung-specific changes in osteosarcoma cells identified across multiple models

Adaptive changes in response to the microenvironment molds cancer cell behavior by altering intra-cellular signaling mechanisms. In order to identify changes in canonical signaling cascades, we performed IPA analysis using differentially up- and downregulated genes in lung-affinity clusters relative to tibia-affinity clusters in each model. In orthotopic tumors, the gene expression signature of osteosarcoma cells is associated with activated glycolysis and gluconeogenesis pathways with downregulation of oxidative phosphorylation (Figure 4A). In metastatic lesions, osteosarcoma cells downregulate genes associated with glycolysis and gluconeogenesis, with some models showing frank upregulation of genes associated with oxidative phosphorylation (Figure 4B). The lung microenvironment

is known to exert unique set of metabolic stressors, such as redox stress, that is distinct from the primary tumor[22]. The finding of glycolysis and gluconeogenesis as differentially regulated in osteosarcoma is consistent with previous work[26], however, identification of a metabolic switch away from glycolysis and gluconeogenesis in metastatic osteosarcoma is a novel finding.

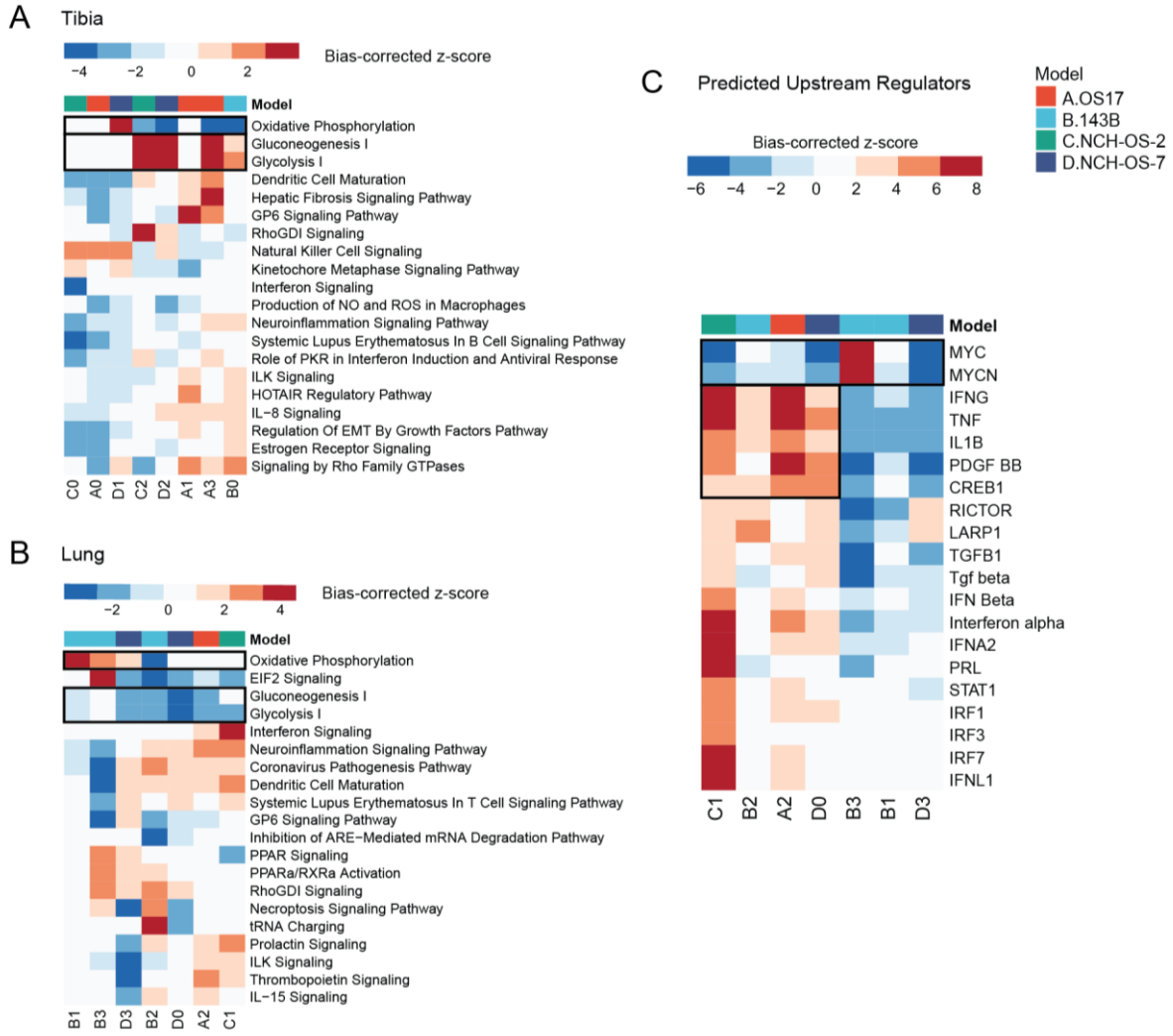

FIGURE 4: Lung-affinity signature identified across multiple models of osteosarcoma A) Canonical pathway analysis using genes differentially regulated in tibia-enriched clusters. Hierarchical clustering of both rows and columns and reporting top 20. B-H p value cutoff: 1.3. B) Canonical pathway analysis using genes differentially regulated in lung-enriched clusters. Hierarchical clustering of both rows and columns and reporting top 20. B-H p value cutoff: 1.3. D) Upstream regulator analysis identified block of common regulators responsible for activation of lung-affinity transcriptional signature. Hierarchical clustering of both rows and columns. In A-C, n = 1,500 per condition.

Given that transcriptional profiles of osteosarcoma cells undergo a plastic shift in the lung microenvironment, we hypothesized that lung-affinity adaptations in osteosarcoma cells occurred in response to environment cues. Using The Ingenuity Pathway Analysis (IPA)[27] analysis tool, we identified common upstream regulators that induce this lung-affinity transcriptional signature (Figure 4C). While we and others have identified many of these pathways and/or upstream regulators to be associated with malignant progression (IFNG, TNF, PDGF)[28–30], coexistence within individual tumor cells is a key finding. These results strongly suggest combination therapies should be rationally designed to target these specific subsets of cells with increased fitness for metastatic growth in lung.

Lineage tracing reveals that metastatic lung enriches for phenotypically distinct clonal families that demonstrate a proliferative advantage

The lung-affinity phenotypes can be targeted only if we identify if these transcriptional programs are inherent genetically encoded programs or if these are acquired in response to microenvironment cues. To monitor clonal dynamics in the orthotopic bone and metastatic lung microenvironment, we used the CloneTracker system. This is a high-diversity lentivirus-based lineage tracing strategy that is 10X-compatible and heritable. We used a parent population of uniquely tagged cells in cell culture to sample from and obtain cells for intra-tibial and intra-venous injections (Figure 5A). Tables showing patterns of lentivirus integration can be found in Supplementary File S3. We reasoned that if the bone or lung microenvironment promoted the survival of pre-existing clones, selective enrichment of the same set of lineage tags would be observed. If survival was mostly driven by adaptive responses, one could expect to see a maintenance of diversity with no specific enrichments.

We identified that phenotypic profiles of these lineage-tagged osteosarcoma cells *in vitro* are distinct from osteosarcoma cells in the orthotopic and metastatic lesions, as observed previously. While there are shared features between orthotopic and metastatic osteosarcoma cells, the latter develop transcriptional signatures specific to the lung microenvironment (Figure 5B). Surprisingly, tumors grown orthotopically in the bone exhibited very little evidence of clonal selection. They retained most of the clonal diversity

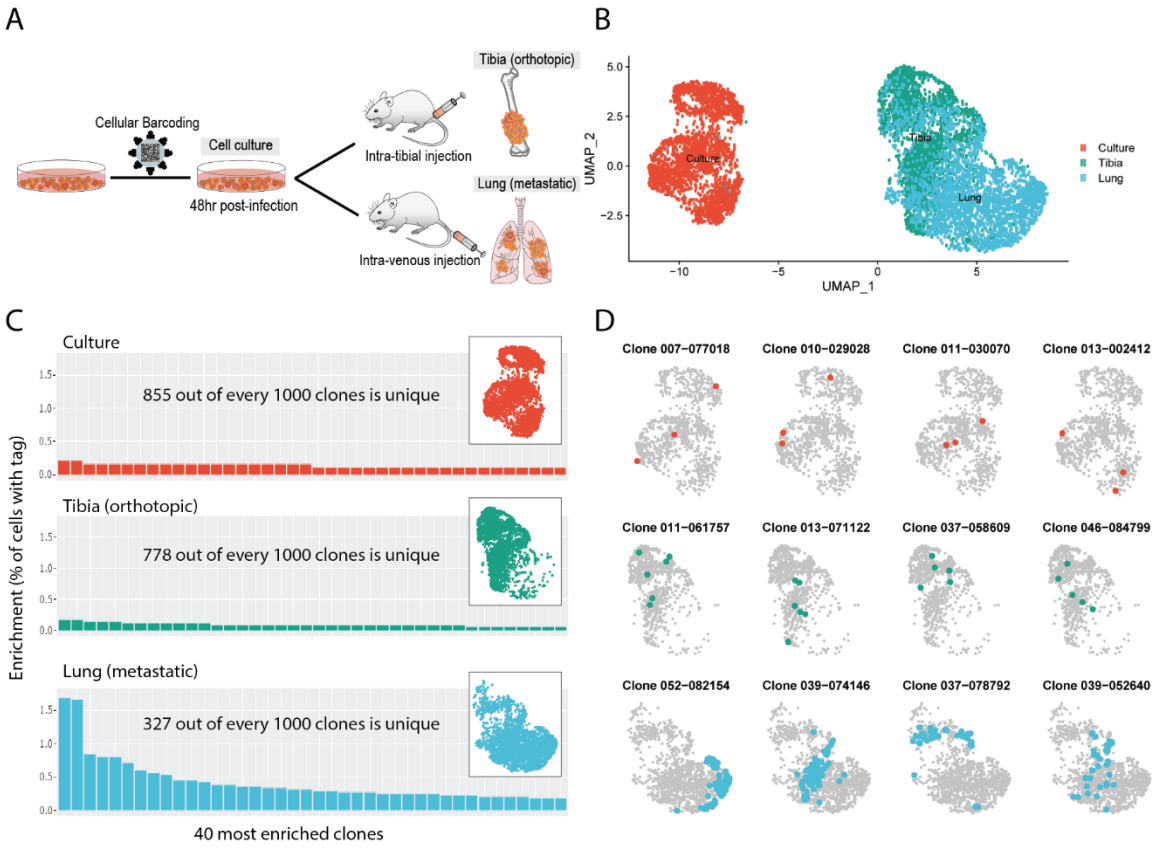

FIGURE 5: Lineage tracing reveals that metastatic lung enriches for phenotypically distinct clonal families that demonstrate a proliferative advantage. A) Schematic of experimental set-up. B) UMAP dimensional reduction plot of OS17 cells tagged with a high-diversity lineage tracing barcode library, then grown in culture, implanted into tibia, or inoculated to generate lung metastases. Metastatic and orthotopic samples share many cells with similar phenotypes, though cultured cells are dramatically different (n = 2800 per condition) C) Frequency distribution of clones identified in each of the conditions. A high level of clonal diversity is maintained in the orthotopic tumors (78% unique clones compared to 81% unique clones in cell culture); though dominant clones emerge in the lung (only 38% with unique lineage tag). D) Overlay of cells sharing a common ancestor onto the dimensional reduction plots (with tagged cells) shows that highly enriched clones in the metastatic lung maintain heritable phenotypes. In contrast, clones in the tibia seem to display a much higher degree of plasticity.

observed in cell culture and we did not observe significant enrichments for any lineage tags (Figure 5C). In contrast, the process of lung colonization selected for distinct clonal populations of tumor cells (Figure 5C). Surprisingly, while these clonal populations clustered separately, cells within individual enriched clonal populations clustered together (Figure 5D). We confirmed this pattern of clonal selection was reproducible in two independent replicates (Figure S2A). Admixing of OS17 cells from distinct, biological

replicates grown *in vitro* or as metastatic lesions suggests that intra-tumor transcriptional heterogeneity of osteosarcoma cells is not driven by individual tumor identity (Figure S2B).

It is important to recognize that our approach does not determine whether the enriched clones composed individual metastatic lesions, or whether different clones colonized individual lesions cooperatively--it simply shows an overall snapshot of emergent clones within all of the metastases of an individual lung. This is important, given that other studies have suggested that OS lung metastases are often polyclonal[18,31]. These results do, however, suggest that, clones emerging from several distinct phenotypic clusters can exhibit increased fitness for growth within the lung and that these diverse transcriptional phenotypes are maintained as distinct across multiple generations without convergence.

Discussion

Single cell transcriptomics has enabled the study of phenotypic evolution in the context of tumor progression. These studies provide insight into the roles that clonal fitness, intra-tumor heterogeneity, and dynamic transcriptional adaptation play as osteosarcoma cells respond to conditions within the primary and metastatic tumor microenvironments. We show that osteosarcomas maintain transcriptional heterogeneity as they respond to changing microenvironment conditions while also plastically adapting their transcriptional profiles. Our data show that the metastatic lung microenvironment enriches for shared features that are conserved across different cell line and PDX models of metastatic osteosarcoma. Features of both clonal selection and dynamic adaptive behavior contribute to metastatic colonization of the lung. Put together, our study suggests a novel tumor intrinsic mechanism that allows for maintenance of phenotypic diversity while responding plastically to changing microenvironment conditions.

While developmental process in normal cells maintain an overall degree of heterogeneity[10,32], we found it interesting that tumor cells displayed a higher degree of overall heterogeneity than their presumptive normal precursors – osteoblasts. Even more interesting, while expected to find that selective pressures associated with adaptation to the bone and lung microenvironments would cause a narrowing of

phenotype, we found no loss of heterogeneity (even after correcting for cell-cycle effects and even in samples that demonstrate high levels of clonal expansion). Each tumor model displayed a transcriptional signature that is distinct, a finding that corroborates previous findings that report high degrees of inter-tumor heterogeneity in osteosarcoma[33]. Overall, these data suggest that maintenance of cell populations with distinct transcriptional phenotypes is a fundamental property of osteosarcoma (similar to findings suggested in other tumor types). One might speculate that this results from the continued influence of normal developmental processes designed to maintain populations of cells that cooperate to support normal bone growth. Indeed, these cooperative mechanisms may provide a survival advantage.

Osteosarcoma cells maintain an overall degree of heterogeneity with a range of phenotypic manifestations that are specific to each microenvironment (even in plastic – a presumably homogeneous environment). Osteosarcoma cells undergo overall phenotypic change in the bone and lung microenvironments.

One of the most striking findings in these within-model comparisons was a consistent enrichment for clusters that exhibit a decrease in glycolysis-related genes and an increase in genes associated with oxidative phosphorylation. This, in some ways, goes against current dogma. However, this finding is not without precedent[34], including similar findings recently reported in breast cancer lung metastasis identified using single-cell sequencing[35]. This shift in bioenergetics coordinates (and is inter-related) with our finding that these same clusters that show increased lung fitness also display decreased MYC-related signaling. One may see this finding as initially counter-intuitive given recent reports that MYC amplification identifies a subset of osteosarcoma patients with particularly poor outcomes[36,37]. However, it is rational that disseminated tumor cells that can alter their bioenergetics to take advantage of the relative abundance of oxygen within the lung would acquire a significant growth and survival advantage.

We show here that at least one subset of cells in each of the osteosarcoma models adopt a shared transcriptional signature as a response to similar upstream signals in the metastatic lung. While others have reported the roles of IFNG, TNF, PDGF and to some extent IL1B in malignant progression, the

novel finding in the study is the co-existence of multiple activated pathways by upstream regulators within the same subset of cells. Whether the observed heterogeneity in lung-affinity transcriptional profiles occurs as a result of variability in response to upstream regulators, or the availability of upstream regulators due to heterogeneity in microenvironment conditions, is unclear. While we acknowledge that our results pertain to transcriptional data that is a snapshot of a single time-point, adoption of shared lung-affinity malignant phenotypes opens the possibility of improving patient survival with targeted therapy against this shared biology.

Combining the power of single cell RNA-sequencing with lineage tracing (CloneTracker system), we show that the majority of osteosarcoma cells can be viewed as plastic (or stem-like), where the plasticity is influenced by microenvironment cues and perhaps some stochastic cell-autonomous mechanisms[38,39]. The dual phenotype-associated lineage tracking method represents an unbiased approach to study tumor evolution, clonal development, response to selective pressures such as microenvironment cues as it does not require any preselected marker and allows the direct study of tumor clones and progeny and their transcriptional profiles. The CloneTracker system has been successfully used to study complexities of clonal behavior in metastatic triple negative breast cancer[40]. Marrying lineage tracing with single cell transcriptomics has allowed us to make inferences about the contributions of clonal selection and transcriptional plasticity to the emergence of traits that facilitate survival and proliferation under different environmental stressors. Given their high tumor-metastatic potential, our model, OS17 cells, represents a powerful tool to test the cohabitation of different cancer clones and the possible dynamics of competition among them. We show that, while orthotopic tumors maintain clonal diversity after a population-shift in phenotype relative to cell culture, metastatic lesions demonstrate clonal expansion and an overall adoption of lung-affinity phenotypic states (Figure 6). These findings corroborate previous studies showing that osteosarcomagenesis follows initial polyclonal dynamics, where different clones can co-exist, followed by clonal dominance with adaptation to the microenvironment as a result of extra-compartmental growth[18]. In interpreting these results, we must acknowledge that the study design

employed remains agnostic of the multi-step process required to effect metastasis in a patient. The design very simply asks whether the capacity to colonize these different tissues is contained within a few rare clones that are pre-destined to become metastases or whether diverse groups of cells have the capacity to adapt to the stressors imposed by the lung environment. Our results suggest that both answers have some truth. When seeded into the lung, some clones exhibit significantly higher fitness relative to other clones. However, these clones come from highly diverse overall transcriptional phenotypes. Overall, the drastic shift in phenotype from *in vitro* to *in vivo* growth conditions endorses the role of microenvironment cues in shaping tumor cell behavior. Additionally, maintenance of range/degree of intra-tumor heterogeneity endorses the role of stochastic cell-autonomous mechanisms in holding up a very basic and evolutionarily ancient feature - existence of meta-stable configuration of intracellular networks[38].

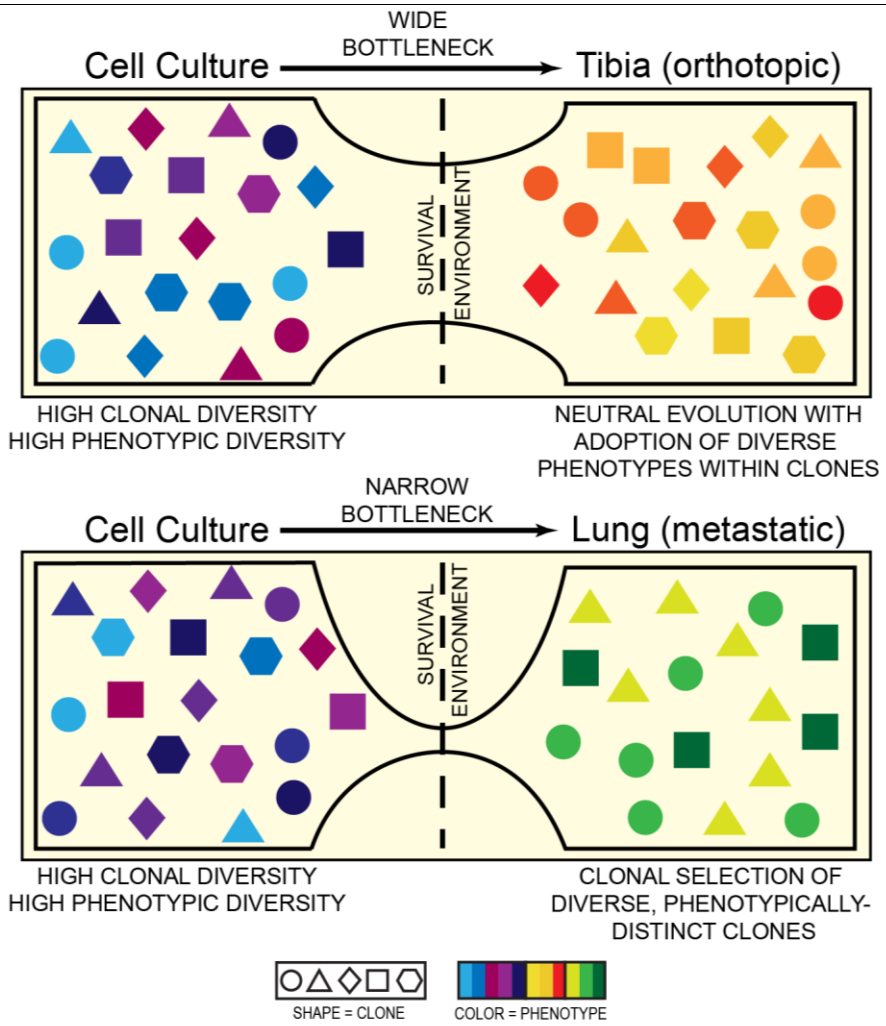

FIGURE 6: Schematic representation of the contributions of clonal diversity and phenotypic plasticity as osteosarcoma cells adapt to changing microenvironments. Osteosarcoma cells demonstrate high transcriptional heterogeneity in plastic, orthotopic bone and lung microenvironments. Tumors that grow in the orthotopic bone follow dynamics of neutral evolution where different cancer clones co-exist and propagate simultaneously. Metastatic tumors that grow in the orthotopic lung demonstrate selection of phenotypically distinct clones, where clonal families with enhanced fitness for the lung microenvironment displayed phenotypic heritability.

Findings from our study have significant implications for modeling metastatic disease (*in vitro* vs *in vivo)*. Lung metastatic progression involves complex 3 dimensional (3D) interactions between tumor cells and the lung microenvironment. Indeed, several studies have demonstrated that tumor cell behavior differs when comparing 2D vs. 3D growth conditions[41–43]. We show here that these differences arise from dynamic (and extensive) transcriptional state changes induced in the transition from a cell-culture to an in vivo model and not from selection for certain subclones that exhibit fitness on plastic or within an

animal. The methodology outlined in this study can serve as a process of ascertaining the validity and applications of various metastatic models of disease.

In conclusion, our study shows that the degree of intra-tumor transcriptional heterogeneity in osteosarcomas is retained through PDX passage, decades of cell culture, and through application of selective pressures of changing microenvironmental conditions. Despite this maintenance of ITH, drastic phenotypic changes follow as tumors transition from a laboratory maintenance model such as cell culture or flank/PDX to orthotopic tumors in the bone or in the lung. We show that these phenotypic changes do not arise primarily through clonal selection, but they arise primarily through change in the transcriptional program within individual cells. While the process of lung colonization enriches for certain clonal populations, the breadth of phenotypic diversity generated recapitulates that found in the primary tumors. Interestingly, the most enriched clonal families displayed robust heritability in phenotype while low-frequency populations demonstrated fewer instances of this phenomenon. Whether this phenotypic heritability has an underlying genetic or epigenetic basis, remains a point of active study.

Authors' contributions

Conception and design: S. Rajan, RD. Roberts

Development of methodology: S. Rajan, RD. Roberts

Acquisition of data: S. Rajan, M. Cam, AC. Gross

Analysis and interpretation of data: S. Rajan, C. Taslim, E. Franz, RD. Roberts

Writing and revision of the manuscript: S. Rajan, RD. Roberts

Study supervision: RD. Roberts

Acknowledgements

The authors would like to thank John Hinckley and Camille McAloney for their assistance in technical editing of the manuscript.

References

1. Casali, P. G. *et al.* Bone sarcomas: ESMO-PaedCan-EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **29**, iv79–iv95 (2018).

2. Savage, S. A. & Mirabello, L. Using epidemiology and genomics to understand osteosarcoma etiology. *Sarcoma* **2011**, 14–16 (2011).

3. Allison, D. C. *et al.* A meta-analysis of osteosarcoma outcomes in the modern medical era. *Sarcoma* **2012**, (2012).

4. Aljubran, A. H., Griffin, A., Pintilie, M. & Blackstein, M. Osteosarcoma in adolescents and adults: survival analysis with and without lung metastases. *Ann. Oncol.* **20**, 1136–1141 (2009).

5. Martin, J. W., Squire, J. A. & Zielenska, M. The genetics of osteosarcoma. *Sarcoma* **2012**, (2012).

6. Lorenz, S. *et al.* Unscrambling the genomic chaos of osteosarcoma reveals extensive transcript fusion, recurrent rearrangements and frequent novel TP53 aberrations. *Oncotarget* **7**, 5273–5288 (2016).

7. Perry, J. A. *et al.* Complementary genomic approaches highlight the PI3K/mTOR pathway as a common vulnerability in osteosarcoma. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E5564–E5573 (2014).

8. Williams, M. J., Werner, B., Barnes, C. P., Graham, T. A. & Sottoriva, A. Identification of neutral tumor evolution across cancer types. *Nat. Genet.* **48**, 238–244 (2016).

9. Gay, L., Baker, A.-M. & Graham, T. A. Tumour Cell Heterogeneity. *F1000Research* **5**, (2016).

10. Altschuler, S. J. & Wu, L. F. Cellular Heterogeneity: Do Differences Make a Difference? *Cell* vol. 141 559–563 (2010).

11. Brock, A., Chang, H. & Huang, S. Non-genetic heterogeneity a mutation-independent driving force for the somatic evolution of tumours. *Nature Reviews Genetics* vol. 10 336–342 (2009).

12. Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* **3**, 730–737 (1997).

13. Almendro, V., Marusyk, A. & Polyak, K. Cellular heterogeneity and molecular evolution in cancer. *Annual Review of Pathology: Mechanisms of Disease* vol. 8 277–302 (2013).

14. Baslan, T. & Hicks, J. Unravelling biology and shifting paradigms in cancer with single-cell sequencing. *Nature Reviews Cancer* vol. 17 557–569 (2017).

15. Hovestadt, V. *et al.* Resolving medulloblastoma cellular architecture by single-cell genomics. *Nature* **572**, 74–79 (2019).

16. Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835-849.e21 (2019).

17. Aynaud, M. M. *et al.* Transcriptional Programs Define Intratumoral Heterogeneity of Ewing Sarcoma at Single-Cell Resolution. *Cell Rep.* **30**, 1767-1779.e6 (2020).

18. Gambera, S. *et al.* Clonal dynamics in osteosarcoma defined by RGB marking. *Nat. Commun.* **9**, (2018).

19. Mcgranahan, N. & Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. (2017) doi:10.1016/j.cell.2017.01.018.

20. Fidler, I. J. Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer Res.* **38**, 2651–2660 (1978).

21. Charifson, M. A. & Trumble, B. C. Cancer treatment innovators discover Charles Darwin. *Evol. Med. Public Heal.* **2019**, 50–63 (2019).

22. Fan, T. M., Roberts, R. D. & Lizardo, M. M. Understanding and Modeling Metastasis Biology to Improve Therapeutic Strategies for Combating Osteosarcoma Progression. *Frontiers in Oncology* vol. 10 (2020).

23. Chalishazar, M. D. *et al.* Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. *Nat. cancer* (2020).

24. Geller, D. S. & Gorlick, R. Osteosarcoma: a review of diagnosis, management, and treatment strategies. *Clin. Adv. Hematol. Oncol.* **8**, 705–718 (2010).

25. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).

26. Chaiyawat, P. *et al.* Protein profiling of osteosarcoma tissue and soft callus unveils activation of the unfolded protein response pathway. *Int. J. Oncol.* **54**, 1704–1718 (2019).

27. Krämer, A., Green, J., Pollard, J. & Tugendreich, S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* **30**, 523–530 (2014).

28. Mojic, M., Takeda, K. & Hayakawa, Y. The dark side of IFN-γ: Its role in promoting cancer immunoevasion. *International Journal of Molecular Sciences* vol. 19 89 (2018).

29. Botter, S. M., Neri, D. & Fuchs, B. Recent advances in osteosarcoma. *Current Opinion in Pharmacology* vol. 16 15–23 (2014).

30. Xu, J., Xie, L. & Guo, W. PDGF/PDGFR effects in osteosarcoma and the “add-on” strategy. *Clin. Sarcoma Res.* **8**, 15 (2018).

31. Brady, S. W. *et al.* The clonal evolution of metastatic osteosarcoma as shaped by cisplatin treatment. *Mol. Cancer Res.* **17**, 895–906 (2019).

32. Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E. & Huang, S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**, 544–547 (2008).

33. Wang, D. *et al.* Multiregion sequencing reveals the genetic heterogeneity and evolutionary history of osteosarcoma and matched pulmonary metastases. *Cancer Res.* **79**, 7–20 (2019).

34. LeBleu, V. S. *et al.* PGC-1α mediates mitochondrial biogenesis and oxidative phosphorylation in cancer cells to promote metastasis. *Nat. Cell Biol.* **16**, 1-15,992-1003 (2014).

35. Davis, R. T. *et al.* Transcriptional diversity and bioenergetic shift in human breast cancer metastasis revealed by single-cell RNA sequencing. *Nat. Cell Biol.* **22**, 310–320 (2020).

36. Feng, W. *et al.* Myc is a prognostic biomarker and potential therapeutic target in osteosarcoma. *Ther. Adv. Med. Oncol.* **12**, 1758835920922055 (2020).

37. Sayles, L. C. *et al.* Genome-Informed Targeted Therapy for Osteosarcoma. *Cancer Discov.* **9**, 46–63 (2019).

38. Marusyk, A. & Polyak, K. Tumor heterogeneity: causes and consequences. *Biochim. Biophys. Acta* **1805**, 105–117 (2010).

39. Hill, R. P. Identifying cancer stem cells in solid tumors: case not proven. *Cancer Res.* **66**, 1891–5; discussion 1890 (2006).

40. Merino, D. *et al.* Barcoding reveals complex clonal behavior in patient-derived xenografts of metastatic triple negative breast cancer. *Nat. Commun.* **10**, 766 (2019).

41. Ekert, J. E. *et al.* Three-Dimensional Lung Tumor Microenvironment Modulates Therapeutic Compound Responsiveness In Vitro – Implication for Drug Development. *PLoS One* **9**, e92248

(2014).

42. Luca, A. C. *et al.* Impact of the 3D Microenvironment on Phenotype, Gene Expression, and EGFR Inhibition of Colorectal Cancer Cell Lines. *PLoS One* **8**, e59689 (2013).

43. Pickl, M. & Ries, C. H. Comparison of 3D and 2D tumor models reveals enhanced HER2 activation in 3D associated with an increased response to trastuzumab. *Oncogene* **28**, 461–468 (2009).

44. Houghton, P. J. *et al.* The pediatric preclinical testing program: Description of models and early testing results. *Pediatr. Blood Cancer* (2007) doi:10.1002/pbc.21078.

45. Wang, X. & McManus, M. Lentivirus production. *J. Vis. Exp.* (2009) doi:10.3791/1499.

46. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).

47. Haber, A. L. *et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).

48. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).

49. Luecken, M. D. & Theis, F. J. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.* **15**, e8746 (2019).

50. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* **16**, 284–287 (2012).

51. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).

52. Foundation, T. R. The R Project for Statistical Computing. https://www.r-project.org/.

53. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).