Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Lineage plasticity and tumor microenvironment transcriptional features define a novel aggressive state in pediatric rhabdomyosarcoma

Davide Pernici
Fondazione IRCCS Istitituto Nazionale dei Tumori di Milano
Stefano Percio
Fondazione IRCCS Istituto Nazionale dei Tumori https://orcid.org/0000-0002-9362-2639
Gabriele Tiné
Fondazione IRCCS Istitituto Nazionale dei Tumori di Milano
Elena Di Blasi
Fondazione IRCCS Istitituto Nazionale dei Tumori di Milano
Bahar Fayaztorshizi
Fondazione IRCCS Istitituto Nazionale dei Tumori di Milano
Renata Ferri
Fondazione IRCCS Istitituto Nazionale dei Tumori di Milano
Yiyi Zhu
Fondazione IRCCS Istitituto Nazionale dei Tumori di Milano
Loris De Cecco
Fondazione IRCCS Istituto Nazionale dei Tumori di Milano
Paola Collini
Fondazione IRCCS Istituto Nazionale dei Tumori
Stefano Chiaravalli
Fondazione IRCCS Istitituto Nazionale dei Tumori di Milano
Michela Casanova
Fondazione IRCCS Istituto Nazionale dei Tumori
Rosalba Miceli
Fondazione IRCCS Istituto Nazionale dei Tumori https://orcid.org/0000-0003-0954-375X
Sandro Pasquali
Fondazione IRCCS Istituto Nazionale dei Tumori https://orcid.org/0000-0003-4815-6293
Andrea Ferrari
Patrizia Gasparini
Fondazione IRCCS Istituto Nazionaledei Tumori
Roberta Sulsenti
Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy https://orcid.org/0000-0002-2247-056X

Elena Jachetti

elena.jachetti@istitutotumori.mi.it

Fondazione IRCCS Istitituto Nazionale dei Tumori di Milano https://orcid.org/0000-0001-6483-7350

Article

Keywords:

Posted Date: October 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11429245/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Abstract

Pediatric rhabdomyosarcoma (RMS) features aberrant myogenic differentiation, pointing to lineage plasticity involvement in its development. Tumor-intrinsic regulators of epithelial-to-neuroendocrine plasticity also play roles in RMS, suggesting shared plasticity programs. Plasticity is also shaped by the tumor microenvironment; yet, mast cells (MCs), which restrain epithelial-to-neuroendocrine differentiation, remain unexplored in RMS.

Despite multimodal therapy and current risk stratification, survival remains poor in RMS patients, highlighting the need for biologically informed biomarkers.

Applying gene signatures of neuroendocrine-lineage promoting factors ("NE-LIN") and MCs in a cohort of 61 pediatric and adolescent/young adults RMS patients, we identified two clusters, NE-LIN^{high}/MC^{low} and NE-LIN^{low}/MC^{high}, independent of histology and fusion status. The former displayed aggressive transcriptional programs. Findings were confirmed in an independent cohort and supported by two additional single-cell/single-nuclei transcriptomic datasets and by immunohistochemistry. Furthermore, the NE-LIN^{high}/MC^{low} state was significantly associated with adverse overall survival beyond conventional risk classification in both localized and metastatic RMS patients.

This work identifies a novel Plasticity-Associated Aggressive State (PAAS) in RMS, with distinct lineage plasticity and microenvironmental features, revealing parallels between plasticity programs in epithelial and mesenchymal cancers. PAAS-RMS defines a biologically and clinically relevant state that may inform patient stratification and guide the development of personalized therapeutic approaches in RMS.

INTRODUCTION

Lineage plasticity is the ability of cancer cells to alter their phenotype and function, and plays a crucial role in cancer progression and therapy resistance by allowing tumors to adapt to diverse environmental pressures and evade conventional and targeted treatments [1]. Besides a few driving genomic alterations, lineage plasticity is mainly governed by epigenetic and transcriptional regulators, as well as from signals coming from the tumor microenvironment [2]. Key pan-cancer lineage determining factors, have been specifically associated to different cell plasticity trajectories [3]. These include for example SNAIL, SLUG, and ZEB for epithelial-to-mesenchymal transition, a process ubiquitous in epithelial cancers, and EZH2, AURKA, N-MYC, and ASCL1 for epithelial-to-neuroendocrine differentiation, mainly observed in prostate and lung cancer [3].

Being now recognized as one of the hallmarks of cancer [4], lineage plasticity has been extensively described in epithelial cancers, with well-characterized mesenchymal, neuroendocrine, squamous, or endothelial trans-differentiation trajectories in different tumor types/settings [3]. In contrast, whether mesenchymal tumors undergo lineage plasticity remains less well explored, although emerging evidence suggest it may be a relevant biological feature in these tumors. Also, whether the same molecular

players governing lineage plasticity in epithelial cancers may operate in mesenchymal tumors remains undefined, raising the possibility of either shared or distinct regulatory mechanisms across tumor types.

Rhabdomyosarcoma (RMS), the most common soft-tissue sarcoma in children [5,6], provides a compelling context in which to investigate these hypotheses. RMS is a mesenchymal malignant neoplasm, characterized by the aberrant differentiation of skeletal muscle cells [7]. Despite their diverse molecular and histological features, RMS tumors generally retain aspects of the myogenic program while failing to complete terminal differentiation [7]. RMS cells instead remain locked in immature progenitor-like states maintained by dysregulated transcriptional and epigenetic programs [8]. Also, RMS may arise in anatomical sites lacking skeletal muscle tissue [9], suggesting that the disease is defined more by activation of a myogenic transcriptional program than by its tissue of origin. Together, these features raise the possibility that lineage plasticity may play an important role in RMS biology. Notably, several key regulators implicated in neuroendocrine differentiation in prostate cancer have also been linked to functional roles in RMS biology, including EZH2, N-MYC, and AURKA [7,8,10–12]. Taken together, these lines of evidence raise the possibility that molecular programs commonly associated with epithelial-to-neuroendocrine cell trajectories may capture broader mechanisms of tumor cell lineage plasticity that extend beyond epithelial cancers.

Pediatric type RMS are historically classified into two prevalent histological subtypes, embryonal (ERMS) and alveolar (ARMS) [13]. The latter is now further stratified by molecular features, with up to 80% of ARMS harboring the *PAX3/7::FOXO1* translocation, associated with poor outcome [13]. The presence of lymph node and distant metastases, the histology or rather the genomic and molecular profile, tumor size, the site of origin, and the surgical stage are the factors used to define risk group stratification. Among these variables, patient age is also a prognostic factor, with adolescents and young adults characterized by a worse prognosis than children [13]. Although multimodal therapy combining chemotherapy, surgery, and/or radiotherapy has improved the survival, substantial heterogeneity in clinical outcomes persists within established risk groups [5]. Also, 5-years survival rates remain below 30% in metastatic patients [5]. These observations underscore the need for biomarkers to improve prognostic stratification in RMS, and suggest that lineage plasticity may represent an additional relevant RMS feature not captured by current risk-classification systems.

Tumor cell plasticity is not solely governed by cell-intrinsic regulatory programs but can also be shaped by the tumor microenvironment [14,15]. Our recent work unveiled that mast cells (MCs), a population of innate immune cells with pleiotropic and context-dependent roles in cancer [16], restrain the development of neuroendocrine prostate cancer and are consequently markedly reduced when this tumor type outgrow [17]. The role of MCs in pediatric RMS remains uncharacterized, and whether MCs may similarly shape lineage plasticity in this setting is unknown.

In this study we investigated whether a lineage-determining regulatory program associated with neuroendocrine prostate cancer identifies a distinct transcriptomic state in pediatric RMS, its

relationship with MC infiltration, and its prognostic relevance.

RESULTS

Neuroendocrine lineage and mast cells related signatures define a plasticity-associated aggressive state (PAAS) in RMS associated with poor prognosis.

To understand whether lineage plasticity operates also in RMS, we took advantage of an institutional patient cohort (the "INT" cohort), including a transcriptomic dataset comprising 61 tumor samples from pediatric and adolescent/young adult patients with RMS [18]. Knowing that several key regulators implicated in neuroendocrine lineage plasticity in prostate cancer have biological relevance in RMS [7,8,10–12], we analyzed our data set with a gene signature that characterizes the transcriptomic profiling of neuroendocrine prostate cancer [19]. This signature, hereafter named "NE-LIN", comprises genes associated with the neuroendocrine lineage, most of which encode for transcription factors or epigenetic regulators (Supplementary Table S1). Based on our previous findings showing that MCs are negative regulators of neuroendocrine differentiation [17], we also investigated an in-house generated [20] signature of tumor-infiltrating MCs (Supplementary Table S1). Unsupervised analyses showed that the combination of these two signatures stratifies RMS patients into two clusters: NE-LIN^{high}/MC^{low} and NE-LIN^{low}/MC^{high} (Fig. 1A, heatmap). Notably, the NE-LIN^{high}/MC^{low} cluster is enriched in tumors from patients with reduced OS, as reflected by a higher proportion of dead patients in the NE-LIN^{high}/MC^{low} cluster (Fig. 1A, pie charts). Furthermore, the NE-LIN^{high}/MC^{low} cluster included patients across tumor histopathological subtypes (ERMS and ARMS), translocation status, age, and risk groups, with no apparent concentration of any of these features within either cluster (Fig. 1A, heatmap and pie charts).

When patients were further examined within individual EpSSG risk categories (A-H subgroups [21]), we observed a higher number of deaths among patients belonging to the NE-LIN^{high}/MC^{low} cluster across risk groups (Fig. 1B), suggesting that this molecular state may identify patients at higher risk of death, refining conventional risk classification. In particular, risk subgroups B, E, F, and G showed a consistently significantly higher percentage of deaths among patients in the NE-LIN^{high}/MC^{low} cluster (Fig. 1C). Patients in subgroup C showed good prognosis irrespective of cluster assignment, with no disease-related deaths observed, whereas patients classified in the A risk subgroup were not present in the NE-LIN^{high}/MC^{low} cluster (Fig. 1C). In contrast, in subgroup H deaths were only associated with the NE-LIN^{low}/MC^{high} cluster (Fig. 1C). Patients with metastatic RMS, who are considered separately from the localized EpSSG risk classification, also showed a slightly higher number of deaths in the NE-LIN^{high}/MC^{low} cluster (Fig. 1B). These results further support an association between the NE-LIN^{high}/MC^{low} cluster and adverse OS across different clinical risk categories based on clinic-pathological prognostic features.

To analyze the enrichment of the two signatures in individual samples, we then computed a single-sample (ss) score for each of them. The difference between signature scores was also calculated

(ssDelta = ssNE-LIN – ssMC). Consistent with the unsupervised analysis (Fig. 1A), the two signatures were inversely correlated within the clusters, with ssMC scores higher in the NE-LIN^{low}/MC^{high} cluster, whereas ssNE-LIN and ssDelta scores were higher in the NE-LIN^{high}/MC^{low} one (Fig. 1D).

These data indicate that the NE-LIN^{high}/MC^{low} transcriptional features identify a new RMS state, which we termed as Plasticity Associated Aggressive State (PAAS), associated with poor clinical outcome across common histological/molecular classification and clinical risk groups.

Differential expression analysis reveals enrichment of developmental and gene-regulatory programs coupled with reduced immune signaling in PAAS-RMS.

To infer about biological characteristics of PAAS-RMS, we then evaluated differentially expressed genes between the NE-LIN^{high}/MC^{low} and the NE-LIN^{low}/MC^{high} cluster in our dataset (Fig. 2). We found 650 upregulated and 91 downregulated genes in the former ($|\log_2 FC| > 1$, FDR < 0.05; Supplementary Table S2). Gene set enrichment pathway analyses (GSEA) showed that the NE-LIN^{high}/MC^{low} cluster was predominantly enriched for pathways related to ribosome biogenesis, mRNA processing, and protein translation (Fig. 2A-B). Molecular function analysis further highlighted enrichment of ribosomal components, RNA-binding activities, and protein folding machinery, indicating enhanced translational output and an increased burden on protein homeostasis. Pathways associated with cell cycle progression and DNA replication were also significantly enriched (Fig. 2A-B). These results are consistent with a highly proliferative tumor state in the NE-LIN^{high}/MC^{low} subgroup, primarily associated with enhanced protein synthesis. Moreover, filtered analysis looking for terms related to muscle- and developmental- related pathways revealed significant enrichment of myogenic programs, together with pathways involved in developmental and transcriptional regulation, including WNT, NOTCH, TGF-β, RUNX family-mediated transcriptional regulation, and epigenetic regulation of gene expression (Fig. 2C).

In contrast, gene sets downregulated in the NE-LIN^{high}/MC^{low} cluster were predominantly associated with immune- and inflammatory- related pathways, including cytokine and chemokine activity, type I interferon signaling, MHC class I, and complement-associated functions, as well as pathways related to arachidonic acid metabolism and GPCR signaling (Fig. 2A-B). This downmodulation of immune-related transcriptional profile is consistent with the reduced expression of the MC signature in PAAS-RMS.

Overall, these findings identify PAAS-RMS as a transcriptionally distinct state characterized by coordinated remodeling of developmental and regulatory programs, consistent with lineage plasticity, increased proliferation, and aberrant myogenic program, and by reduced immune-related transcriptional signatures.

The PAAS state is a prognostic factor for OS in patients with RMS.

To test the translational relevance of our results, we assessed the prognostic value of the PAAS-RMS classification in our RMS patient cohort for OS. Kaplan-Meyer analysis comparing patients in the NE-

LIN^{high}/MC^{low} and NE-LIN^{low}/MC^{high} groups revealed that patients in the NE-LIN^{high}/MC^{low} cluster had significantly shorter OS compared to those in the other group (p = 0.0033; Fig. 3A). Univariate Cox proportional hazard models further confirmed the significant association between the NE-LIN^{high}/MC^{low} cluster and poorer survival (Fig. 3B; HR = 3.3; 95% CI 1.4–7.5; p = 0.005).

To further assess the prognostic relevance of PAAS-RMS using a score-based approach with potential clinical applicability, we next assessed the association between the ssDelta score and survival. In the univariate analysis, the ssDelta score was significantly associated with OS (Fig. 3C, HR = 2.0; 95% CI 1.3-3.0; p = 0.001), indicating a higher hazard of death with increasing ssDelta score. Importantly, in multivariate Cox regression analysis, the association between the ssDelta signature score and OS remained significant after adjustment for known clinical-pathological prognostic factors, namely EpSSG risk classification for patients with localized disease, and metastatic status (Fig. 3D; HR = 2.2; 95% CI 1.4–3.5, p < 0.001), indicating that the PAAS-RMS associated signatures add prognostic information to the established clinical risk classification.

Finally, to further characterize the prognostic contribution of individual genes within our signatures, a separate multivariable Cox proportional hazard model was fitted for each gene, adjusting for clinical-pathological variables (EpSSG risk group). Four out of twelve genes in the NE-LIN signature (*EZH2*, *AURKA, MELK*, and *AURKB*) showed a statistically significant association with shorter OS independently of established prognostic features (Fig. 3E). An independent and statistically significant association with OS was also found for one out of ten genes in the MC signature (*HPDGS*) and this correlated with a longer OS (Fig. 3E).

Taken together these findings indicate that the PAAS-RMS state, as identified by the NE-LIN^{high}/MC^{low} signature, is associated with worse clinical outcome and provides additional prognostic information beyond the current clinical risk classification in our RMS cohort.

Prognostic relevance of the PAAS state in patients with localized RMS.

To further investigate the prognostic relevance of the PAAS-RMS state within clinically actionable risk categories, we repeated the molecular clustering, survival, and Cox regression analyses after excluding patients with metastatic disease and focusing on patients with localized RMS in the INT dataset (n = 49). Unsupervised analysis showed that the two molecular clusters were preserved also in this non-metastatic population (Fig. 4A, heatmap), with the NE-LIN^{high}/MC^{low} cluster comprising again a higher proportion of patients with unfavorable clinical outcome (Fig. 4A, pie charts). As observed in the overall cohort (Fig. 1A), cluster assignment was not associated with tumor histology, age, or EpSSG risk subgroup (Fig. 4A, heatmap and pie charts). The enrichment of the ssNE-LIN and ssDelta scores in the NE-LIN^{high}/MC^{low} cluster was maintained (Fig. 4B).

Kaplan-Meier analysis confirmed the association between cluster assignment and OS also when analyzing only localized RMS, with patients in the NE-LIN^{high}/MC^{low} cluster showing significantly poorer

outcome (Fig. 4C; p = 0.043). Consistently, Cox regression analysis showed that ssDelta scores were associated with poorer OS even in the localized RMS cohort (Fig. 4D; HR = 2.2; 95% CI 1.3–3.9; p = 0.004), with a slightly stronger effect than what observed in the overall population. This negative prognostic association was retained in multivariable Cox regression after adjustment for EpSSG risk classification (Fig. 4E; HR = 2.1; 95% CI 1.2–3.6; p = 0.008). These results support the prognostic relevance of the PAAS state in patients with localized RMS.

PAAS-RMS is broadly represented across different RMS cohorts.

To assess the reproducibility of the PAAS-RMS state, we applied the NE-LIN/MC signature analyses on an independent cohort of 141 RMS patients collected in the US, with tumor microarray and matched clinical and follow up data, described by Davicioni et al. [22,23]. Unsupervised analyses identified the NE-LIN^{high}/MC^{low} and NE-LIN^{low}/MC^{high} clusters also in this external cohort, alongside an intermediate NE-LIN^{low}/MC^{low} cluster, which included the majority of the patients (Fig. 5A, heatmap). Consistently to what observed in the INT cohort, the NE-LIN^{high}/MC^{low} cluster included a higher proportion of dead patients, albeit the difference did not reach statistical significance (Fig. 5A, heatmap and pie charts), and was not associated with histology, age, and COG/IRS Clinical Group (I–IV). The NE-LIN^{high}/MC^{low} cluster was likewise characterized by significantly elevated ssNE-LIN and ssDelta scores in the Davicioni cohort (Fig. 5B).

Biologically (Supplementary Fig. S1 and Supplementary Table S3), GSEA performed comparing the NE-LIN^{high}/MC^{low} cluster with remaining samples of the Davicioni dataset, recapitulated the finding obtained in the INT cohort, confirming enrichment for gene sets related to transcriptional, translation, and proliferative programs (Fig. S1A-B) alongside with the enrichment of developmental, myogenic, and epigenetic regulatory pathways (Fig. S1C), and the depletion of immune-related programs (Fig. S1A-B). These results confirm the biological features of the PAAS-RMS state across the two cohorts.

For survival analyses, the NE-LIN^{high}/MC^{low} cluster was compared with the remaining two clusters combined. Consistent with the findings in the INT cohort, also in this data set the NE-LIN^{high}/MC^{low} cluster showed an association with poorer OS, albeit it did not reach statistical significance, both in Kaplan-Meier (Fig. 5C; p = 0.13) and in univariate Cox regression analyses (Fig. 5D; HR1.9; 95% CI 0.83–4.2; p = 0.135). Similarly, the ssDelta score showed a trend toward increased risk of death, although not statistically significant (Fig. 5E; HR = 1.2; 95% CI 0.86–1.7; p = 0.276), supporting, at least in part, the results obtained in the INT cohort.

We then analyzed a s.c./s.n. RNA-Seq dataset of RMS described by Danielli et al. [24]. Among the different analyses performed in their paper, Danielli et al. identified, across an harmonized and integrated dataset comprising both human tumor samples and patient-derived cell lines and xenografts, a cluster of proliferative tumor cells characterized by the enrichment of cell-cycle programs, which the authors defined as “cycling”, compared to the rest of “non cycling” tumor cells [24]. We here restricted the analysis to the 21 RMS tumor samples of the Danielli data set, which were integrated from two independent

patients cohorts [25,26], for a total of n = 31676 cells. First, we visualized the cycling state of the cells, as defined in the paper by Danielli, on a new computed UMAP (Fig. 6A). We then assessed the distribution of the NE-LIN signature across “cycling” and “non cycling” tumor cell populations. To this aim, we computed a NE-LIN gene signature score (gssNE-LIN) analogous to those used in microarray analyses. UMAP visualization showed that the gssNE-LIN score was higher in the cycling cell compartment (Fig. 6B), while it was unrelated to histology (Fig. 6C) or fusion status (Fig. 6D) of the tumors. As the samples in this dataset were enriched for tumor cells, assessment of the MC signature was not possible.

Altogether, these findings support the presence of the PAAS-RMS state across independent RMS patient cohorts, confirming that it does not reflect a specific histological subtype or fusion status. Furthermore, survival analyses in the Davicioni cohort confirmed a trend towards association of PAAS-RMS with poor prognosis, consistent with the findings in the INT cohort.

EZH2 protein upregulation and reduced MC infiltration mirror the NE-LIN ^{high} /MC ^{low} tumor cluster in RMS.

Finally, to further corroborate our *in-silico* findings, we validated the expression of key selected markers of the PAAS-RMS state by immunohistochemistry on tumor RMS samples from our institutional cohort. Among the genes within the NE-LIN signature, we prioritized *EZH2*, based on its association with OS in multivariable Cox regression, its relatively strong effect size (Fig. 3E), and its known biological relevance as a key epigenetic regulator of linage plasticity and tumor aggressiveness in several solid tumors, comprising neuroendocrine prostate cancer [3,27], breast cancer [28], and some soft tissue sarcomas [29–31], including RMS [8,10]. Notably, analyzing the expression of this single gene in our dataset we found that *EZH2* transcript is significantly upregulated in samples belonging to the NE-LIN^{high}/MC^{low} cluster (Fig. 7A). Conversely, *CPA3*, a gene encoding for a specific MC marker [32] selected as representative of the MC signature, shows opposite expression patterns (Fig. 7A).

We therefore performed immunohistochemistry for EZH2 on RMS tumor samples from the INT cohort, analyzing n = 5–8 samples per condition (NE-LIN^{high}/MC^{low} and NE-LIN^{low}/MC^{high} group), and then quantified results by computing a composite score considering both staining intensity and percentage of positive cells (Fig. 7B). In parallel, we also evaluated the number of infiltrating MC with staining for the specific MC marker tryptase (Fig. 7C). We found that samples belonging to the NE-LIN^{high}/MC^{low} cluster showed increased EZH2 staining (Fig. 7B, histogram) and lower MC number (Fig. 7C, histogram) compared to the NE-LIN^{low}/MC^{high} group. However, differences did not reach statistical significance, likely due to the low number of tested tumor samples. Nevertheless, correlation analysis revealed a statistically significant negative association between EZH2 protein expression and MC number (Fig. 7D). These results suggest that the relationship between EZH2 tumor immunostaining and MC infiltration partially reflects the PAAS state. However, these immunohistochemical analyses are based on a limited sample size (n = 5–8 per group), and thus warrant cautious interpretation.

Conclusions

In conclusion, we unveiled the *in-silico* defined NE-LIN^{high}/MC^{low} cluster as an unrecognized state in RMS, associated with poorer patient outcomes and able to add prognostic relevance beyond established risk subgroups. This newly defined state, which we label as PAAS-RMS, is independent of current histological and molecular classification systems and instead characterized by distinct lineage plasticity and tumor microenvironment related traits. These results pave the way for further exploring biological mechanisms and clinical relevance of the PAAS state in RMS.

DISCUSSION

Despite its established role in promoting tumor aggressiveness and therapy resistance in epithelial malignancies [3,4], lineage plasticity in mesenchymal tumors has remained less explored. In this work, we uncover the involvement in RMS, of lineage-determining factors typically associated with neuroendocrine differentiation of adenocarcinoma, revealing unexpected biological commonalities between epithelial and mesenchymal cancers. Moreover, we define a correlation between this peculiar lineage plasticity associated traits and the tumor microenvironment, another still underdeveloped area of research in RMS.

Our work identifies across independent patient cohorts the existence of a previously unrecognized plasticity-associated aggressive state (PAAS) of RMS, defined by the NE-LIN^{high}/MC^{low} signatures, and transcriptionally characterized by increased proliferative and translational activity and by epigenetic regulation, along with retention of myogenic programs and remodeling of developmental signaling pathways. RMS cells are known to preserve myogenic lineage features while failing to undergo terminal differentiation [8]. Since normal myogenesis couples activation of myogenic programs with cell-cycle withdrawal, the coexistence of myogenic features and increased proliferation in PAAS-RMS may suggest an uncoupling between lineage identity and terminal differentiation [33]. Thus, PAAS-RMS may represent a hybrid state in which myogenic identity is preserved while cells acquire highly proliferative and plastic properties.

The PAAS state is transversal to any common histological or molecular classification for RMS and correlates with poor survival. Furthermore, in our cohort it adds prognostic value beyond current EpSSG risk stratification systems. These findings suggest that aggressiveness and prognosis in RMS may be governed by dynamic functional states, such as lineage plasticity and microenvironmental features, rather than fixed histological or genomic classifications, and are consistent with previous s.c RNA-Seq studies showing that transcriptional programs reflecting different myogenic differentiation states are associated with clinical outcome in RMS [34], further supporting the concept that the differentiation state of tumor cells may contribute to disease aggressiveness and prognosis.

Our work opens the possibility to investigate therapeutic vulnerabilities specifically associated to the PAAS state and the use of NE-LIN/MC signatures as an additional stratification tool for patients with RMS. Our results indicating a higher proportion of death events associated to PAAS-RMS even in

low/intermediate risk subgroups suggest that the PAAS classification may have particular clinical value in these patients, potentially identifying individuals who, despite favorable conventional risk classification, may benefit from closer monitoring and therapy escalation. On the other side, the development of new therapeutic strategies specific for PAAS-RMS have the potential to improve the survival of high-risk or metastatic RMS patients, for which specific and effective therapeutic options are lacking. This personalized medicine approach might also help reducing overtreatment in less aggressive states, potentially limiting the severe toxicity associated with conventional chemotherapy and radiotherapy regimens, which can significantly impair patient quality of life and have long-term consequences for survivors. This issue is particularly important given the young age of the pediatric and adolescent/young adult population affected by RMS.

Having identified that the PAAS-RMS state is characterized by a MC-excluded microenvironment, we also foresee further studies to investigate the protective role of MC against this tumor state, paralleling our work in neuroendocrine prostate cancer [17] and similar data obtained in nasopharyngeal cancer [35]. Furthermore, knowing that MCs are key cellular players able to shape the immune tumor microenvironment with context-dependent roles [16], our results also underlie the relevance of comparing the immune-contexture of NE-LIN^{high}/MC^{low} and NE-LIN^{low}/MC^{high} tumors. Indeed, our GSEA analyses showing downregulation of immune-related genes suggest a distinct immune microenvironment in PAAS-RMS. Defining the molecular networks through which MCs modulate tumor aggressiveness and characterizing the microenvironmental peculiarities of the PAAS state, could pave the way for exploring immunotherapy-based strategies in patients with RMS, where immunotherapy has so far been scarcely used and provided limited benefit [36,37].

Our findings do not indicate that PAAS-RMS should be classified as neuroendocrine. Such possible classification would require a comprehensive histologic characterization of canonical neuroendocrine markers (e.g. CgA, SYP, NSE), which is beyond the scope of this work. Yet, this possibility cannot be excluded. To date, only three studies, including two small case series [38,39] and one larger cohort of ARMS [40], have shown histologic expression of neuroendocrine markers in a subset of RMS cases. Furthermore, two different recent s.c./s.n. RNA-Seq analyses revealed the existence of tumor cell subpopulations with neuronal features in RMS patient samples, even if results were partially contrasting. Danielli et al. described a transcriptomic neuronal cell state, identifiable in a subset of fusion-positive RMS, while absent in fusion-negative samples [24]. On the contrary, Whitfield et al. revealed a convergent high-risk associated cell state across fusion-positive and lethal fusion-negative RMS, characterized by neuronal features [41]. This second paper is consistent with our observation that the PAAS state, characterized by high expression of neuroendocrine lineage associated molecular drivers, is present across histological and molecular RMS classification and negatively associated with prognosis. Nevertheless, rather than pointing to a “neuroendocrine” classification for RMS, our data suggest instead that both the PAAS state and the epithelial-neuroendocrine differentiation are governed by shared epigenetic regulators. It is conceivable that across different tumor contexts (e.g. epithelial versus mesenchymal), the same epigenetic mechanisms converge on different lineage plasticity endpoints

which determine different phenotypic outcomes, such as neuroendocrine differentiation in epithelial cancers and the PAAS state in sarcomas. Both the aforementioned hypothesis warrants further investigation.

Among epigenetic regulators, our data point to EZH2 as one of the main drivers of the PAAS state, having a significant negative prognostic role in our RMS cohort. These results are in line with a previous study showing that EZH2 overexpression in different pediatric soft tissues sarcomas, including RMS, correlates with increased aggressiveness, metastatic disease, and poor survival [42], and suggest of evaluating the antitumor activity of drugs against EZH2 or its complex in RMS patients stratified according to the PAAS state.

However, while adding the evaluation of NE-LIN/MC signatures to common criteria holds potential clinical relevance in RMS, the prognostic value of the PAAS-RMS state will require validation in independent cohorts. Furthermore, the implementation of transcriptomic signatures in clinical settings may be challenging. In this regard, our results suggest that immunohistochemistry evaluation of selected signature-associated markers (e.g. EZH2 and tryptase) could be a more feasible surrogate. Yet, in this work immunohistochemistry analyses were restricted to a very small sample size due to limited tissue availability; thus, staining on large-scale tumor specimens will be required to validate this approach.

The main limitation of our study is the lack of an independent validation of the prognostic value of the PAAS-RMS state in an external cohort. To our knowledge, besides the two analyzed in this project, only another European RMS datasets with annotated survival data was described by Williamson et al. [43], which however mainly included patients who had received treatment before sample collection, and was thus not considered suitable to be compared with our institutional one. This reflects the paucity of clinically annotated datasets for this rare disease, as well as the fact that existing cohorts are often assembled to enrich for specific patient characteristics or address defined research questions. For example, the INT cohort includes both pediatric and young adult patients, whereas the Davicioni cohort exclusively comprises pediatric patients. Furthermore, the latter includes patients recruited between 1990s-2006, whereas our cohort spans 1999–2020, reflecting differences in clinical practice and treatment protocols over time and countries (EU and US, for the INT and Davicioni cohort, respectively). These differences might have limited the power and the interpretation of prognostic comparisons. Nevertheless, albeit non statistically significant, the survival analyses in the Davicioni dataset are in line with our results, and thus corroborate the possibility that the prognostic value of PAAS-RMS can be extended beyond the INT cohort. Also, it is worth noting that, considering the rarity of this cancer type, our cohort of 61 RMS samples is sizeable, thereby adding strength and robustness to our findings.

In conclusion, our work identifies a previously undescribed aggressive state of RMS, characterized by specific lineage plasticity and tumor microenvironment related transcriptional features, which transcends histological and molecular classifications and is associated with adverse clinical outcomes and prognostic value beyond current risk stratification systems. These findings provide a framework for

future studies aimed at elucidating the biological underpinnings of the newly defined PAAS state and its potential clinical relevance in RMS.

MATERIAL AND METHODS

Datasets used

The in-house RMS microarray data set (which we here define as the “INT” dataset”) has been previously described [18] and deposited in the Gene Expression Omnibus (GEO) with the accession number GSE135517. Briefly, data for pediatric, adolescent and young adult (AYA) patients with primary, consecutive RMS treated between 1999 and 2020 at Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Milan, Italy, were extracted from a prospectively maintained research database. As for inclusion criteria, FFPE tumoral samples were eligible if they met the following criteria: (1) histologically confirmed diagnosis of primary RMS; (2) availability of FFPE tumor tissue obtained from surgical resection specimen and/or diagnostic biopsy; (3) sample obtained at initial diagnosis (pre-treatment whenever available); (4) adequate and sufficient tumor material for molecular and/or histopathological analyses; (5) patients treated consecutively at the institution between January 1999 and December 2020. Samples were excluded if they met any of the following criteria: (1) tissue obtained from relapsed or recurrent disease; (2) insufficient tumor content or inadequate FFPE material for analysis; (3) incomplete clinical data in institutional database. Tumors were classified according to the Intergroup Rhabdomyosarcoma Study (IRS) postsurgical clinical grouping system (Groups I–IV) [21]. Histological subtype was determined by expert pathological review and categorized according to WHO criteria. Risk stratification was performed following European pediatric Soft tissue sarcoma Study Group (EpSSG) recommendations, integrating histological subtype, IRS post-surgical clinical group, primary tumor site, nodal involvement, tumor size, and age [21]. In this paper, only patients with either ERMS (37 patients) or ARMS (24 patients) were analyzed.

The Davicioni et al. RMS microarray data set [22,23] including 141 patients with RMS (71 ERMS and 70 ARMS) was downloaded from GEO (GSE92689). As described by the authors, frozen tumor samples were obtained from patients enrolled in the Intergroup Rhabdomyosarcoma Study Group (IRSG IV-V) Children’s oncology Group (COG) clinical trials, conducted in the US predominantly during the 1990s and early 2000s. Tumor specimens were retrieved from the IRSG/Pediatric Cooperative Human Tissue Network tumor repository (Columbus, OH, USA) and the institutional tumor bank at Children’s Hospital Los Angeles (CHLA, Los Angeles, CA, USA). Clinical covariates and outcome information were provided by the COG Statistics and Data Center (Arcadia, CA, USA). Only tumor samples containing > 80% tumor cells were analyzed for gene expression [22,23]. Clinical characteristics included age, sex, stage, and COG/IRS Clinical Group (I–IV) [44], as annotated in the dataset (GSE92689).

Publicly available single-cell and/or single-nuclei RNA sequencing (s.c. RNA-Seq; s.n. RNA-Seq) datasets from two RMS cohorts were also analyzed (GSE195709 and GSE174376), the first from the

Massachusetts General Hospital, Charleston, MA, USA [25], and the second from the St. Jude Children's Hospital, Memphis, TN, USA [26]. For the present analyses, we used the pre-processed and uniformly processed Seurat objects generated in the analyses of the aforementioned datasets by Danielli et al. [24] and made publicly available through Figshare (https://figshare.com/projects/RMS_consensus_analysis/194417). The deposited objects included data from multiple RMS s.c./s.n. RNA-Seq datasets analyzing patient RMS samples alongside with patient-derived xenografts, and patient-derived and commercial cell lines. Among these, for this paper the pre-processed Seurat object generated by Danielli et al. was retrieved and subset to retain only the 21 RMS tumor samples from patients contained in GSE195709 [25] and GSE174376 [26], and particularly 4 patients (all with ERMS histology) for the GSE195709 dataset and 17 patients (6 with ARMS and 11 with ERMS) for the GSE174376. These datasets are enriched for tumor cells and do not allow a comprehensive assessment of the tumor microenvironment. No additional preprocessing of the raw sequencing data was performed.

Data preprocessing

Pre-processed gene expression data provided on GEO/Fighshare were used for each data set. Normalized matrix was calculated using $\log_2$(intensity) for microarray data.

Bioinformatic analyses on microarray data

Bioinformatic analyses were conducted using R (v. 4.6.0; https://www.R-project.org; [45]). To analyze the enrichment of the NE-LIN and MC signatures in individual samples, we calculated a single-sample (ss) score for each of the two signatures (ssNE-LIN and ssMC) using singscore (v. 1.32.0) [46], with default settings. A difference signature score was also evaluated (ssDelta = ssNE-LIN – ssMC; calculated by setting NE-LIN as upSet and MC as downSetp parameters in the singscore function).

Heatmaps were generated with ComplexHeatmap (v. 2.28.0) [47], scaling expression values by rows (Z-score) and using euclidean distance and Ward.D2 linkage to cluster samples. For each heatmap, the optimal number of clusters was determined as the most frequent across all indexes using NbClust (v. 3.0.1) [48].

Differential expression analysis was performed with the limma R package (v. 3.68.4) [49] and measured with fold change (FC) and t-statistics, while pre-ranked Gene Set Enrichment Analysis was executed with clusterProfiler (v. 4.20.0, Reactome and GO Molecular Function (MF) databases), ranking genes by the moderated t-statistic derived from limma, which accounts for both the magnitude and precision of differential expression.

Bioinformatic analyses on s.c./s.n. RNA-Seq data

The retained subset of the s.c./s.n. RNA-Seq dataset was analyzed through Seurat (v. 5.4.0) [50]. Gene signature scores (gss) were calculated applying the AddModuleScore function to the original log-normalized data layer of the RNA assay. The dataset was split by patient to preserve sample-level structure prior to normalization. Each sample was normalized and variance-stabilized using SCTransform [51], and subsequently integrated using Harmony [52] to correct for inter-patient technical variation. Dimensionality reduction was performed using Uniform Manifold Approximation and Projection (UMAP), and the resulting embedding was used to visualize the calculated gene set scores on the tumor cell population.

Ethics

Formalin-fixed paraffin-embedded (FFPE) RMS samples from patients were obtained from the archives of Fondazione IRCCS Istituto Nazionale dei Tumori (INT, Milan, Italy). The study and all experimental protocols were approved by Ethics Committee of INT (approval number INT 132/16 and INT160/24), and all experiments were performed following relevant national ethical guidelines and regulations, in accordance with the principles of the Helsinki Declaration. All patients, and/or their legal guardians provided written informed consent for research activities at the time of hospital admission.

Histopathological evaluation

Tissue collection, fixation, processing and hematoxylin and eosin (H&E) staining were performed according to standardized protocols as part of routine clinical practice. Cases were selected from NE-LIN^{high}-MC^{low}(n = 8) and NE-LIN^{low}-MC^{high} (n = 5) clusters, as defined by prior *in silico* bioinformatic analyses, based on tissue availability and residual material within formalin-fixed paraffin-embedded (FFPE) blocks. H&E-stained sections, were independently reviewed by a board-certified pathologist (Dr. Elena Di Blasi) who identified and annotated a representative viable tumor area for each sample. These regions were defined as regions of interest (ROIs) for all downstream analyses.

Immunohistochemistry

FFPE samples were cut at 5 µm thickness, de-paraffined and re-hydrated. Heat-induced antigen retrieval was performed using Target Retrieval Solution Citrate (Dako) pH6 for EZH2, and the EnVision FLEX Target Retrieval Solution, High pH (Dako) for tryptase, autoclaving at 96°C for 15 minutes. Sections were cooled to room temperature and washed in PBS. Endogenous peroxidase activity was quenched using 3% H_2O_2, followed by Fc blocking with a protein blocking solution. Slides were incubated 15 minutes at room temperature with primary antibodies: (i) rabbit anti-mouse/human EZH2 monoclonal antibody (clone D2C9; cell signaling cat. no. 5246; dilution 1:50) overnight at 4°C or (ii) mouse anti-human mast cell Tryptase monoclonal antibody (clone AA1; Abcam cat. no. ab2378; dilution 1:100.000). Detection was performed using polymer-based systems (EnVision FLEX/HRP, Dako for EZH2; Novolink Polymer Detection System, Leica for Typtase) with 3,3'-diaminobenzidine (DAB) as chromogen. Slides were counterstained with Gill III hematoxylin (Leica Biosystems) and mounted with Eukitt (Sigma-Aldrich).

Slides were visualized using a Leica DM4 B microscope equipped with a Leica DFC450 C digital camera and LAS software (v.4.8). Whole-slide scanning was performed using an Aperio GT 450 DX (Leica).

Digital image analysis

Whole-slide images were analyzed using QuPath software (v0.4.4). Following EZH2 staining, the pathologist identified the previously defined tumor ROI on H&E-stained sections and mapped these regions onto corresponding serial sections. For EZH2, both staining intensity (I) and abundance (A) of the positive cells were evaluated and scored on a 0–3 scale. EZH2 expression was then evaluated within these ROIs using a composite score obtained by multiplying the two parameters (IxA), resulting in scores ranging from 0 to 9. On serial sections, tryptase staining was performed, and tryptase-positive MCs were manually counted within the same tumor ROIs used for EZH2 evaluation.

Statistical analyses

Statistical analyses were performed with Fisher's exact test, Wilcoxon rank-sum test for pie charts and boxplots, respectively. Analyses were conducted using GraphPad Prism software (version 9.4; GraphPad Prism Inc., San Diego, CA, USA) or R (v. 4.6.0). Associations between two continuous variables were assessed using Spearman's rank correlation coefficient (ρ). Threshold for assess statistical significance was set to 0.05. In each figure, either the exact P value was indicated.

Overall survival (OS) was defined as the time to death from any cause, calculated from the date of diagnosis. For alive patients the time censored at the last date of follow-up. OS curves were estimated using the Kaplan-Meier method and statistically compared with the log-rank test.

Cox proportional hazard regression models were used to evaluate the association between the delta score (ssDelta) of the NE-LIN/MC transcriptomic signature (continuous variable) and OS. Univariable and multivariable models were fitted to data. Multivariable models were adjusted for known clinic-pathologic prognostic factors, namely EpSSG risk classification status [21] for patients with localized disease and metastatic status. For the analyses, we merged patients in the low/intermediate and the high/very high group. Patients in the low/intermediate EpSSG risk group were used as the reference category. The outcome measure was the hazard ratio (HR), reported with 95% confidence intervals (CI).

Further analyses were performed at the individual-gene level for genes included in the NE-LIN and MC signatures. For each gene, a separate Cox proportional hazards model was fitted and hazard ratios (HRs) with 95% confidence intervals (CIs) were reported. Cox models were stratified by EpSSG risk group, allowing the association between gene expression and OS to be estimated independently of established clinical prognostic strata while accounting for EpSSG risk group-specific baseline hazards.

DATA AVAILABILITY

All the dataset analyzed in this paper are publicly available. The INT and the Davicioni microarray datasets are deposited on GEO: INT dataset GSE135517, Davicioni dataset GSE92689. For analyses of the two s.c./s.n. RNA-Seq datasets (GSE195709 and GSE174376), in this paper we used the pre-

processed and uniformly processed Seurat objects generated by Danielli et al. [24] and made publicly available through Figshare (https://figshare.com/projects/RMS_consensus_analysis/194417).

Declarations

ACKNOWLEDGEMENTS

We thank from Fondazione IRCCS Istituto Nazionale dei Tumori: Claudia Chiodoni for critical reading of the manuscript, Ester Grande for administrative assistance, Marco Gentili and Francesca Broglia for technical assistance. AI tools (ChatGPT, OpenAI; Mocrosoft Copilot) were used solely to assist with language refinement, improve clarity of expression, and grammar checking. No confidential information was provided to AI tools.

FUNDING

This work was funded by 5xmille funds for healthcare research (Ministry of Health; INT “Bandi di Linea” to R.S), Grants from Associazione Italiana per la Ricerca sul Cancro (AIRC; Investigator Grant ID# 27469 to E.J. and Individual Grant–Next Gen Clinician Scientist “Fondazione 13 Marzo” ID# 28546 to S.P.) and Italian Ministry of Health (“Ricerca corrente” Funds).

DISCLOSURES

The authors declare no potential conflicts of interest.

AUTHORS’ CONTRIBUTIONS

Conceptualization: E. Jachetti, R. Sulsenti, S. Pasquali, P. Gasparini, A. Ferrari,

Investigation: D. Pernici, S Percio, G. Tinè, E. Di Blasi, B. Fayaztorshizi, R Ferri, Y Zhu, R. Sulsenti

Resources: A. Ferrari, P. Collini, R. Miceli, L. De Cecco, S. Chiaravalli, M. Casanova

Formal analysis: D. Pernici, S Percio, G. Tinè,

Visualization: D. Pernici, R. Sulsenti, S Percio G. Tinè,

Funding acquisition: E. Jachetti, R. Sulsenti,

Study Supervision: E. Jachetti, R. Sulsenti, P. Gasparini

Writing (original draft/ review&editing): E. Jachetti, R. Sulsenti, P. Gasparini, S. Pasquali, A. Ferrari.

References

1. Quintanal-Villalonga A et al (2020) Lineage plasticity in cancer: a shared pathway of therapeutic resistance. Nat Rev Clin Oncol 17:360–371. https://doi.org:10.1038/s41571-020-0340-z
2. Davies AH, Beltran H, Zoubeidi A (2018) Cellular plasticity and the neuroendocrine phenotype in prostate cancer. Nat Rev Urol 15:271–286. https://doi.org:10.1038/nrurol.2018.22
3. Davies A, Zoubeidi A, Beltran H, Selth LA (2023) The Transcriptional and Epigenetic Landscape of Cancer Cell Lineage Plasticity. Cancer Discov 13:1771–1788. https://doi.org:10.1158/2159-8290.CD-23-0225
4. Hanahan D (2026) Hallmarks of cancer-Then and now, and beyond. Cell. https://doi.org:10.1016/j.cell.2025.12.049
5. Amer KM et al (2019) Epidemiology, Incidence, and Survival of Rhabdomyosarcoma Subtypes: SEER and ICES Database Analysis. J Orthop Res 37:2226–2230. https://doi.org:10.1002/jor.24387
6. Yohe ME et al (2019) Insights into pediatric rhabdomyosarcoma research: Challenges and goals. Pediatr Blood Cancer 66:e27869. https://doi.org:10.1002/pbc.27869
7. Gustafson AL, Durbin AD, Artinger KB, Ford HL (2024) Myogenesis gone awry: the role of developmental pathways in rhabdomyosarcoma. Front Cell Dev Biol 12:1521523. https://doi.org:10.3389/fcell.2024.1521523
8. Pomella S et al (2023) Genomic and Epigenetic Changes Drive Aberrant Skeletal Muscle Differentiation in Rhabdomyosarcoma. Cancers (Basel) 15. https://doi.org:10.3390/cancers15102823
9. Chen S, Kelsey AM, Rudzinski ER (2025) Rhabdomyosarcoma in children and young adults. Virchows Arch 486:101–116. https://doi.org:10.1007/s00428-024-03961-y
10. Ciarapica R et al (2014) The Polycomb group (PcG) protein EZH2 supports the survival of PAX3-FOXO1 alveolar rhabdomyosarcoma by repressing FBXO32 (Atrogin1/MAFbx). Oncogene 33:4173–4184. https://doi.org:10.1038/onc.2013.471
11. Ommer J et al (2020) Aurora A Kinase Inhibition Destabilizes PAX3-FOXO1 and MYCN and Synergizes with Navitoclax to Induce Rhabdomyosarcoma Cell Death. Cancer Res 80:832–842. https://doi.org:10.1158/0008-5472.CAN-19-1479
12. Tonelli R et al (2012) Antitumor activity of sustained N-myc reduction in rhabdomyosarcomas and transcriptional block by antigene therapy. Clin Cancer Res 18:796–807. https://doi.org:10.1158/1078-0432.CCR-11-1981
13. Gartrell J, Pappo A (2020) Recent advances in understanding and managing pediatric rhabdomyosarcoma. *F1000Res* 9 https://doi.org:10.12688/f1000research.22451.1
14. Zhang J, Hu Z, Horta CA, Yang J (2023) Regulation of epithelial-mesenchymal transition by tumor microenvironmental signals and its implication in cancer therapeutics. Semin Cancer Biol 88:46–66. https://doi.org:10.1016/j.semcancer.2022.12.002
15. Zhou H et al (2022) Focus on the tumor microenvironment: A seedbed for neuroendocrine prostate cancer. Front Cell Dev Biol 10:955669. https://doi.org:10.3389/fcell.2022.955669

16. Sulsenti R, Jachetti E (2023) Frenemies in the Microenvironment: Harnessing Mast Cells for Cancer Immunotherapy. Pharmaceutics 15. https://doi.org:10.3390/pharmaceutics15061692
17. Sulsenti R et al (2024) Intracellular Osteopontin Promotes the Release of TNFalpha by Mast Cells to Restrain Neuroendocrine Prostate Cancer. Cancer Immunol Res 12:1147–1169. https://doi.org:10.1158/2326-6066.CIR-23-0792
18. Gasparini P et al (2019) Age-Related Alterations in Immune Contexture Are Associated with Aggressiveness in Rhabdomyosarcoma. Cancers (Basel) 11. https://doi.org:10.3390/cancers11091380
19. Chedgy EC et al (2018) Biallelic tumour suppressor loss and DNA repair defects in de novo small-cell prostate carcinoma. J Pathol 246:244–253. https://doi.org:10.1002/path.5137
20. Jachetti E et al (2018) Cross-Talk between Myeloid-Derived Suppressor Cells and Mast Cells Mediates Tumor-Specific Immunosuppression in Prostate Cancer. Cancer Immunol Res 6:552–565. https://doi.org:10.1158/2326-6066.CIR-17-0385
21. De Salvo GL et al (2024) Reappraisal of prognostic factors used in the European Pediatric Soft Tissue Sarcoma Study Group RMS 2005 study for localized rhabdomyosarcoma to optimize risk stratification and generate a prognostic nomogram. Cancer 130:2351–2360. https://doi.org:10.1002/cncr.35258
22. Davicioni E et al (2006) Identification of a PAX-FKHR gene expression signature that defines molecular classes and determines the prognosis of alveolar rhabdomyosarcomas. Cancer Res 66:6936–6946. https://doi.org:10.1158/0008-5472.CAN-05-4578
23. Davicioni E, Anderson JR, Buckley JD, Meyer WH, Triche TJ (2010) Gene expression profiling for survival prediction in pediatric rhabdomyosarcomas: a report from the children's oncology group. J Clin Oncol 28:1240–1246. https://doi.org:10.1200/JCO.2008.21.1268
24. Danielli SG et al (2024) Single cell transcriptomic profiling identifies tumor-acquired and therapy-resistant cell states in pediatric rhabdomyosarcoma. Nat Commun 15:6307. https://doi.org:10.1038/s41467-024-50527-2
25. Wei Y et al (2022) Single-cell analysis and functional characterization uncover the stem cell hierarchies and developmental origins of rhabdomyosarcoma. Nat Cancer 3:961–975. https://doi.org:10.1038/s43018-022-00414-w
26. Patel AG et al (2022) The myogenesis program drives clonal selection and drug resistance in rhabdomyosarcoma. *Dev Cell* 57, 1226–1240 e1228 https://doi.org:10.1016/j.devcel.2022.04.003
27. Venkadakrishnan VB et al (2024) Lineage-specific canonical and non-canonical activity of EZH2 in advanced prostate cancer subtypes. Nat Commun 15:6779. https://doi.org:10.1038/s41467-024-51156-5
28. Chen Y, Zhu H, Luo Y, Tong S, Liu Y (2024) EZH2: The roles in targeted therapy and mechanisms of resistance in breast cancer. Biomed Pharmacother 175:116624. https://doi.org:10.1016/j.biopha.2024.116624

29. Gounder M et al (2020) Tazemetostat in advanced epithelioid sarcoma with loss of INI1/SMARCB1: an international, open-label, phase 2 basket study. Lancet Oncol 21:1423–1432. https://doi.org:10.1016/S1470-2045(20)30451-4
30. Kawano S et al (2016) Preclinical Evidence of Anti-Tumor Activity Induced by EZH2 Inhibition in Human Models of Synovial Sarcoma. PLoS ONE 11:e0158888. https://doi.org:10.1371/journal.pone.0158888
31. O'Donnell E 3 et al (2025) Genetic and epigenetic characterization of sarcoma stem cells across subtypes identifies EZH2 as a therapeutic target. NPJ Precis Oncol 9:7. https://doi.org:10.1038/s41698-024-00776-7
32. Pejler G (2019) The emerging role of mast cell proteases in asthma. Eur Respir J 54. https://doi.org:10.1183/13993003.00685-2019
33. Ruijtenberg S, van den Heuvel S (2016) Coordinating cell proliferation and differentiation: Antagonism between cell cycle regulators and cell type-specific gene expression. Cell Cycle 15:196–212. https://doi.org:10.1080/15384101.2015.1120925
34. DeMartino J et al (2023) Single-cell transcriptomics reveals immune suppression and cell states predictive of patient outcomes in rhabdomyosarcoma. Nat Commun 14:3074. https://doi.org:10.1038/s41467-023-38886-8
35. Cheng S et al (2021) A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* 184, 792–809 e723 https://doi.org:10.1016/j.cell.2021.01.010
36. Davis KL et al (2022) A Phase I/II Trial of Nivolumab plus Ipilimumab in Children and Young Adults with Relapsed/Refractory Solid Tumors: A Children's Oncology Group Study ADVL1412. Clin Cancer Res 28:5088–5097. https://doi.org:10.1158/1078-0432.CCR-22-2164
37. DeSantes KB et al (2025) Dose Escalation Study of Enoblituzumab in Children and Young Adults with B7-H3-Expressing Relapsed or Refractory Solid Tumors. Cancer Res Commun 5:1574–1583. https://doi.org:10.1158/2767-9764.CRC-25-0293. A Phase I, Open-Label
38. Houreih MA et al (2009) Alveolar rhabdomyosarcoma with neuroendocrine/neuronal differentiation: report of 3 cases. Int J Surg Pathol 17:135–141. https://doi.org:10.1177/1066896908319444
39. Okuno N et al (2024) Primary Gastric Alveolar Rhabdomyosarcoma: A Potent Mimicker of Neuroendocrine Carcinoma. Cureus 16. https://doi.org:10.7759/cureus.69907. e69907
40. Bahrami A, Gown AM, Baird GS, Hicks MJ, Folpe AL (2008) Aberrant expression of epithelial and neuroendocrine markers in alveolar rhabdomyosarcoma: a potentially serious diagnostic pitfall. Mod Pathol 21:795–806. https://doi.org:10.1038/modpathol.2008.86
41. Whitfield HJ et al (2026) High-Risk Rhabdomyosarcomas Feature a Convergent Cell State. Cancer Res. https://doi.org:10.1158/0008-5472.CAN-25-4403
42. Ramaglia M et al (2016) High EZH2 expression is correlated to metastatic disease in pediatric soft tissue sarcomas. Cancer Cell Int 16:59. https://doi.org:10.1186/s12935-016-0338-x
43. Williamson D et al (2010) Fusion gene-negative alveolar rhabdomyosarcoma is clinically and molecularly indistinguishable from embryonal rhabdomyosarcoma. J Clin Oncol 28:2151–2158.

https://doi.org:10.1200/JCO.2009.26.3814
44. Crane JN et al (2022) Clinical group and modified TNM stage for rhabdomyosarcoma: A review from the Children's Oncology Group. Pediatr Blood Cancer 69:e29644. https://doi.org:10.1002/pbc.29644
45. Team RC (2026) R: A Language and Environment for Statistical Computing. https://www.R-project.org
46. Foroutan M et al (2018) Single sample scoring of molecular phenotypes. BMC Bioinformatics 19:404. https://doi.org:10.1186/s12859-018-2435-4
47. Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32:2847–2849. https://doi.org:10.1093/bioinformatics/btw313
48. Charrad M, Ghazzali N, Boiteau V, Niknafs A, NbClust (2014) An R Package for Determining the Relevant Number of Clusters in a Data Set. J Stat Softw 61:1–36. https://doi.org:https://doi.org/10.18637/jss.v061.i06
49. Ritchie ME et al (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47. https://doi.org:10.1093/nar/gkv007
50. Hao Y et al (2024) Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol 42:293–304. https://doi.org:10.1038/s41587-023-01767-y
51. Hafemeister C, Satija R (2019) Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol 20:296. https://doi.org:10.1186/s13059-019-1874-1
52. Korsunsky I et al (2019) Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16:1289–1296. https://doi.org:10.1038/s41592-019-0619-0

Figures

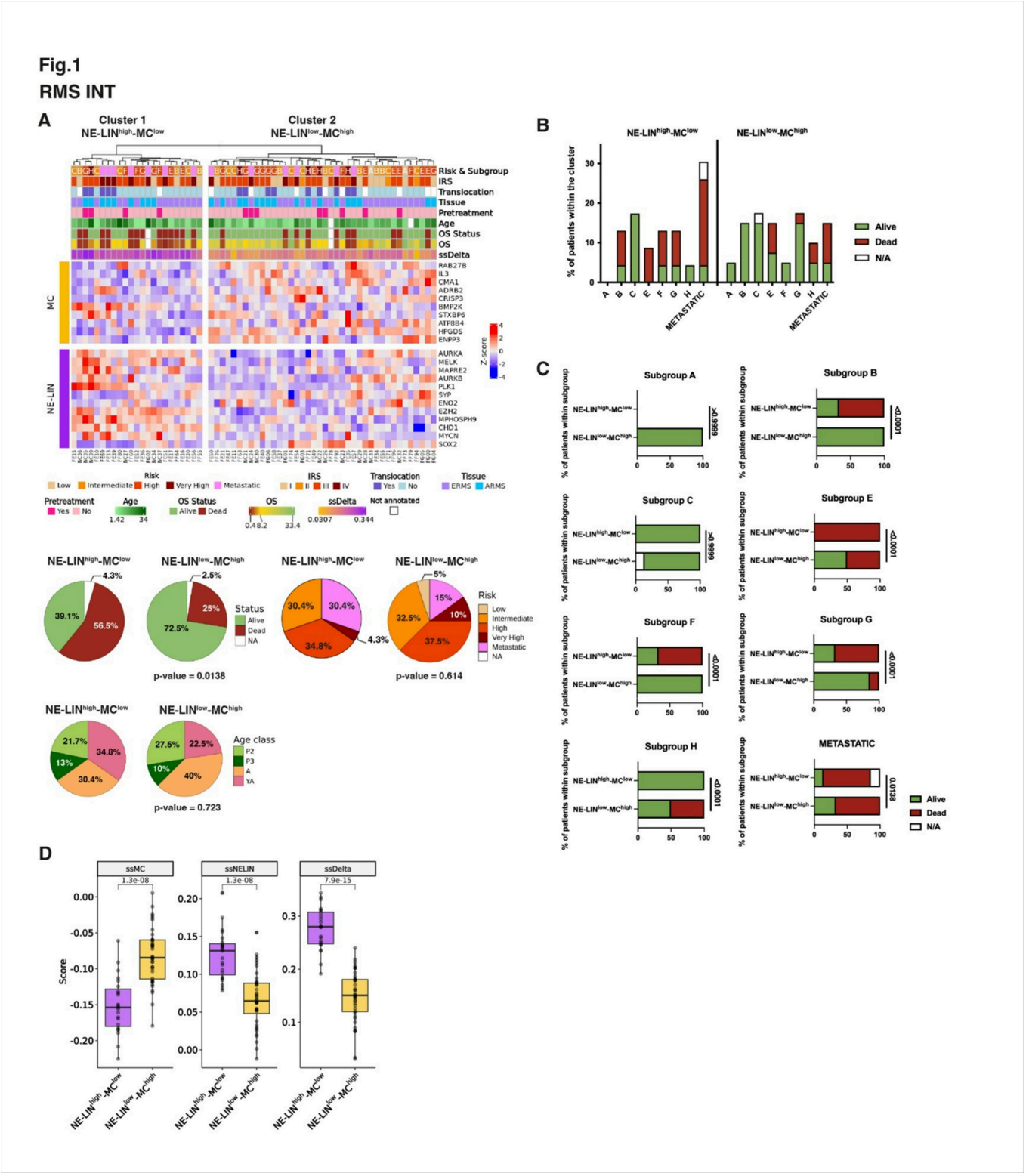


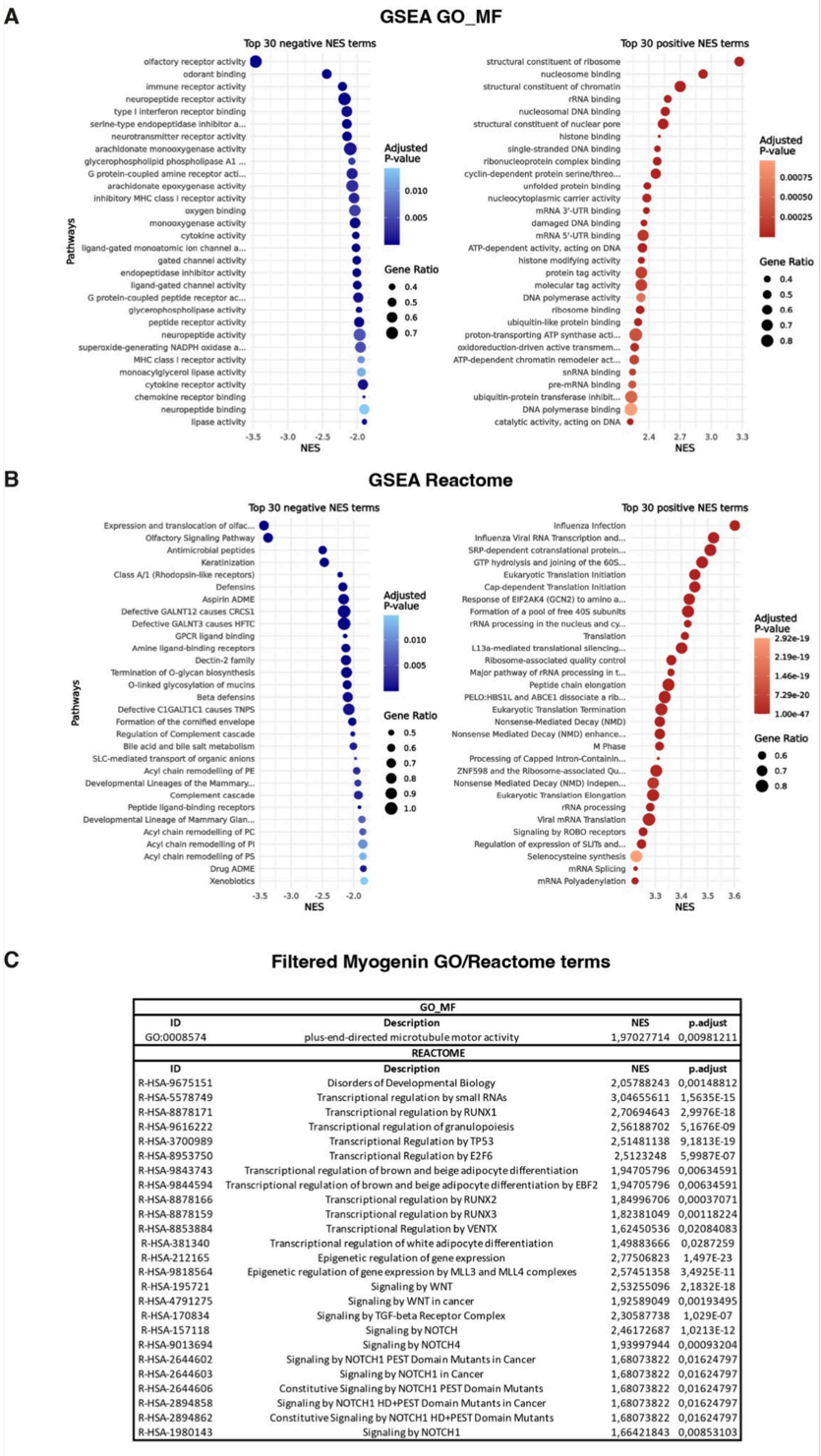
Figure 1

The NE-LIN^{high}-MC^{low} signature defines a new transcriptional state in RMS associated with poor outcome. A. Heatmap shows hierarchical clustering of RMS tumor samples of the INT dataset according to NE-LIN and MC genes. Only samples with ERMS or ARMS histology were analyzed (n=61 tumor samples). Indication for IRS, translocation status, histology (ERMS or ARMS), pretreatment, age, and OS status are reported. Pie charts show relative frequencies of OS status, risk subgroup (according to the

EpSSG localized RMS prognostic nomogram risk group [21], or metastasis), and age among each cluster (NE-LIN^{high}/MC^{low} and NE-LIN^{low}/MC^{high}). Statistical comparisons were performed using Fisher's exact test. Age classes were categorized as follows: Pediatric: P1 £1 year; P2 >1 to £ 6 years; P3 >6 to £12 years; Adolescent: A >12 to £20 years; Young Adult: YA ³20 years. **B.** Bar plot shows the relative percentage of OS status, calculated within each cluster, in patients divided by risk category. **C.** Bar plots show the relative percentage of OS status calculated within each clinical subgroup (EpSSG risk group for localized RMS or metastatic) in the two cluster. Statistical comparisons were performed using Fisher's exact test. **D.** Box plots showing expression of the ssMC, ssNE-LIN and ssDelta scores in the two clusters shown in A. Statistical comparisons were performed using or Wilcoxon test.

Fig.2

RMS INT: NE-LIN^{high}-MC^{low} vs NE-LIN^{low}-MC^{high}



GO_MF			
ID	**Description**	**NES**	**p.adjust**
GO:0008574	plus-end-directed microtubule motor activity	1,97027714	0,00981211
REACTOME			
ID	**Description**	**NES**	**p.adjust**
R-HSA-9675151	Disorders of Developmental Biology	2,05788243	0,00148812
R-HSA-5578749	Transcriptional regulation by small RNAs	3,04655611	1,5635E-15
R-HSA-8878171	Transcriptional regulation by RUNX1	2,70694643	2,9976E-18
R-HSA-9616222	Transcriptional regulation of granulopoiesis	2,56188702	5,1676E-09
R-HSA-3700989	Transcriptional Regulation by TP53	2,51481138	9,1813E-19
R-HSA-8953750	Transcriptional Regulation by E2F6	2,5123248	5,9987E-07
R-HSA-9843743	Transcriptional regulation of brown and beige adipocyte differentiation	1,94705796	0,00634591
R-HSA-9844594	Transcriptional regulation of brown and beige adipocyte differentiation by EBF2	1,94705796	0,00634591
R-HSA-8878166	Transcriptional regulation by RUNX2	1,84996706	0,00037071
R-HSA-8878159	Transcriptional regulation by RUNX3	1,82381049	0,00118224
R-HSA-8853884	Transcriptional Regulation by VENTX	1,62450536	0,02084083
R-HSA-381340	Transcriptional regulation of white adipocyte differentiation	1,49883666	0,0287259
R-HSA-212165	Epigenetic regulation of gene expression	2,77506823	1,497E-23
R-HSA-9818564	Epigenetic regulation of gene expression by MLL3 and MLL4 complexes	2,57451358	3,4925E-11
R-HSA-195721	Signaling by WNT	2,53255096	2,1832E-18
R-HSA-4791275	Signaling by WNT in cancer	1,92589049	0,00193495
R-HSA-170834	Signaling by TGF-beta Receptor Complex	2,30587738	1,029E-07
R-HSA-157118	Signaling by NOTCH	2,46172687	1,0213E-12
R-HSA-9013694	Signaling by NOTCH4	1,93997944	0,00093204
R-HSA-2644602	Signaling by NOTCH1 PEST Domain Mutants in Cancer	1,68073822	0,01624797
R-HSA-2644603	Signaling by NOTCH1 in Cancer	1,68073822	0,01624797
R-HSA-2644606	Constitutive Signaling by NOTCH1 PEST Domain Mutants	1,68073822	0,01624797
R-HSA-2894858	Signaling by NOTCH1 HD+PEST Domain Mutants in Cancer	1,68073822	0,01624797
R-HSA-2894862	Constitutive Signaling by NOTCH1 HD+PEST Domain Mutants	1,68073822	0,01624797
R-HSA-1980143	Signaling by NOTCH1	1,66421843	0,00853103

Figure 2

Pathway enrichment analysis in the INT dataset reveals activation of developmental regulatory networks and reduced immune signaling in PAAS-RMS. A-B. Bubble plots illustrate the top-30 negative (left panels) and top-30 positive (right panels) enriched gene sets identified by GSEA in the NE-LIN^{high}/MC^{low} vs NE-LIN^{low}/MC^{high} comparison, ranked according to their normalized enrichment score (NES); only significantly enriched pathways are shown. **C**. Selected myogenesis-related gene sets

significantly enriched in the NE-LIN^{high}/MC^{low} cluster. The table reports NES and adjusted *P* value for each gene set.

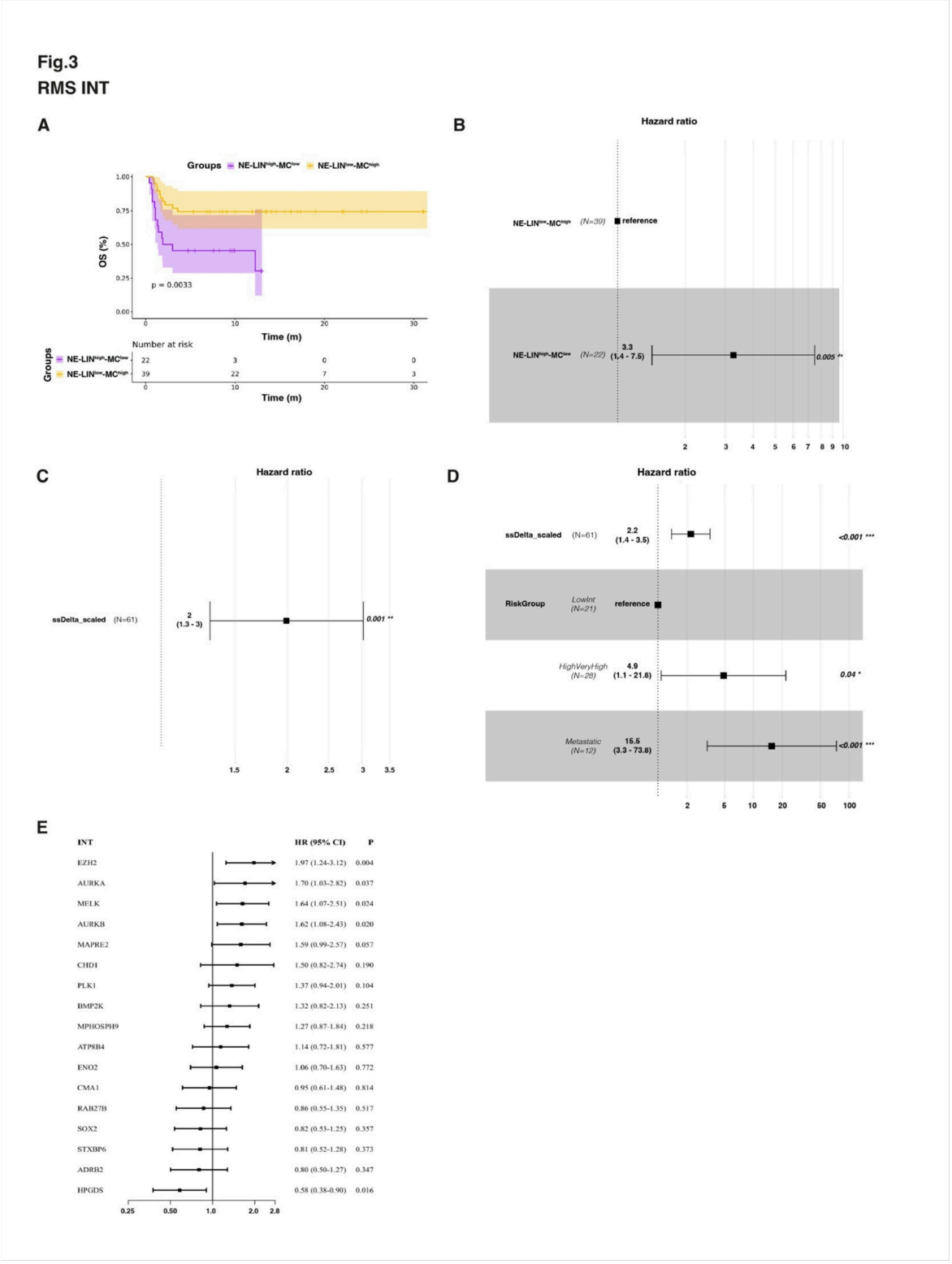


Figure 3

Prognostic association of the PAAS state in RMS. A. Kaplan-Meier curves for OS stratified by clusters in the INT dataset (n=61 patients). Log-rank test was used to evaluate statistical significance, and number

at risk table illustrates the number of patients across the time measure in months (m). **B-E.** Cox proportional hazards analyses for overall survival in patients of the INT cohort. Forest plots display hazard ratios (HRs), 95% confidence intervals (CIs), and corresponding *P* values. Hazard ratios greater than 1 indicate an increased risk of death. **B.** Univariate analysis of the NE-LIN^{high}/MC^{low} vs NE-LIN^{low}/MC^{high} clusters. **C.** Univariate analysis of the ssDelta signature score. **D.** Multivariable analysis of the ssDelta signature score adjusted for clinical variables (EpSSG risk group for localized disease, or metastatic status). **E.** Multivariable analyses of individual genes from the NE-LIN and MC signatures adjusted for clinical variables (EpSSG risk group).

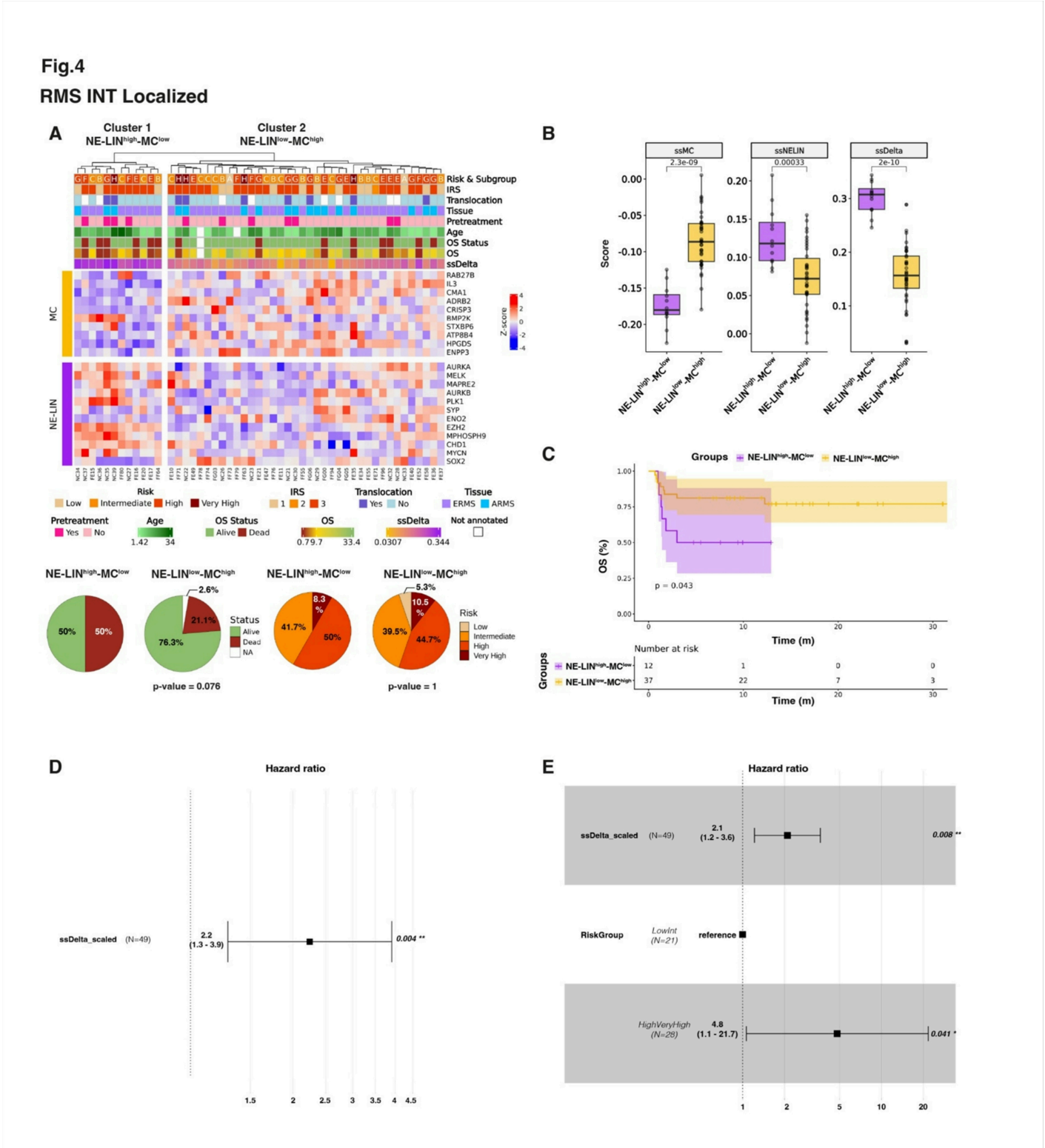


Figure 4

Prognostic association of the PAAS state in localized RMS. A-E. Analyses were limited to localized RMS patients within the INT cohort (n=49) **A.** Heatmap shows hierarchical clustering samples according to NE-LIN and MC signature expression. Indication for IRS, translocation status, histology (ERMS or ARMS), pretreatment, OS status, are reported. Pie charts show relative frequencies of OS status and risk subgroup (according to the EpSSG localized RMS prognostic nomogram risk group, [21]) among each

cluster (NE-LIN^{high}/MC^{low} and NE-LIN^{low}/MC^{high}). Statistical comparisons in pie charts were performed using Fisher's exact test. **B.** Box plots show expression of the ssMC, ssNE-LIN and ssDelta scores in the two clusters shown in A. Statistical comparisons were performed using Wilcoxon test. **C.** Kaplan-Meier curves for OS stratified by clusters in patients with localized RMS of the INT cohort. Log-rank test was used to evaluate statistical significance, and number at risk table illustrates the number of patients across the time measure in months (m). **D-E.** Cox proportional hazards analyses for overall survival in patients with localized RMS of the INT cohort. Forest plots display hazard ratios (HRs), 95% confidence intervals (CIs), and corresponding *P* values. Hazard ratios greater than 1 indicate an increased risk of death. **D.** Univariate analysis of the ssDelta signature score. **E.** Multivariable analysis of the ssDelta signature score adjusted for clinical variables (EpSSG risk group for localized disease, or metastatic disease).

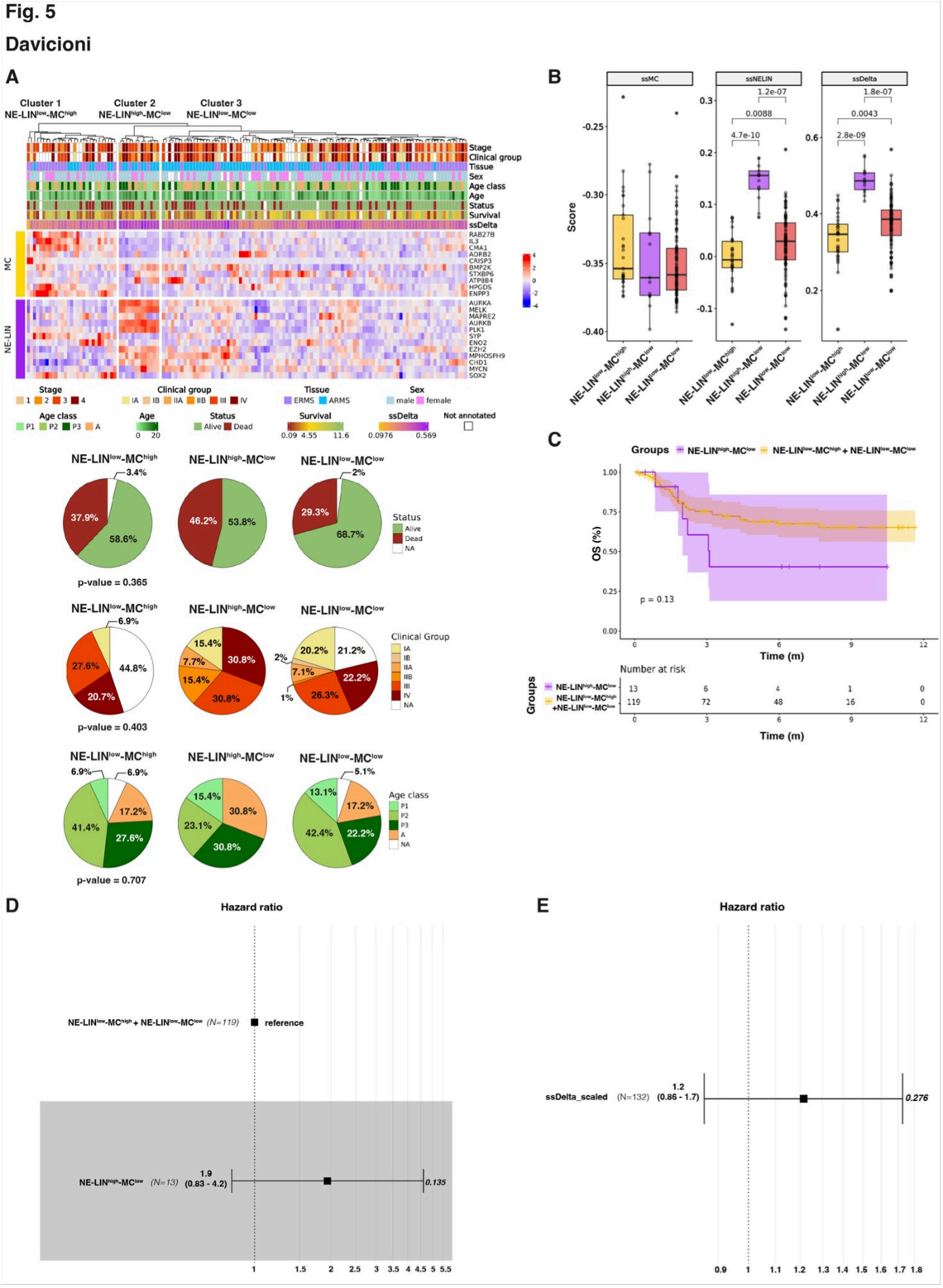


Figure 5

Validation of the PAAS-RMS state its association with survival in the Davicioni cohort. A. Heatmap shows hierarchical clustering of RMS tumor samples of the Davicioni dataset according to NE-LIN and MC genes. Only samples with ERMS or ARMS histology were analyzed (n=141). Indication for histology (ERMS or ARMS), sex, age, and OS status are reported. Pie charts show relative frequencies of OS status, COG/IRS Clinical Group (I–IV) [44], and age among each cluster (NE-LIN^{high}/MC^{low}, NE-

LIN^{low}/MC^{high}, and NE-LIN^{low}/MC^{low}). Statistical comparisons were performed using Fisher's exact test. Age classes were categorized as follows: Pediatric: P1 £1 year; P2 >1 to £ 6 years; P3 >6 to £12 years; Adolescent: A >12 to £20 years; Young Adult: YA ³20 years. **B.** Box plots showing expression of the ssMC, ssNE-LIN and ssDelta scores in the three clusters shown in A. Statistical comparisons were performed using or Wilcoxon test. **C.** Kaplan-Meier curves for OS stratified by clusters in the Davicioni dataset. The NE-LIN^{low}/MC^{high} was compared with the combination of the NE-LIN^{low}/MC^{low}+NE-LIN^{low}/MC^{low} clusters. Log-rank test was used to evaluate statistical significance, and number at risk table illustrates the number of patients across the time measure in months (m). **D-E.** Cox proportional hazards analyses for overall survival in the Davicioni dataset. Forest plots display hazard ratios (HRs), 95% confidence intervals (CIs), and corresponding *P* values. Hazard ratios greater than 1 indicate an increased risk of death. **D.** Univariate analysis of the NE-LIN^{high}/MC^{low} vs NE-LIN^{low}/MC^{high} + NE-LIN^{low}/MC^{low} clusters. **E.** Univariate analysis of the ssDelta signature score.

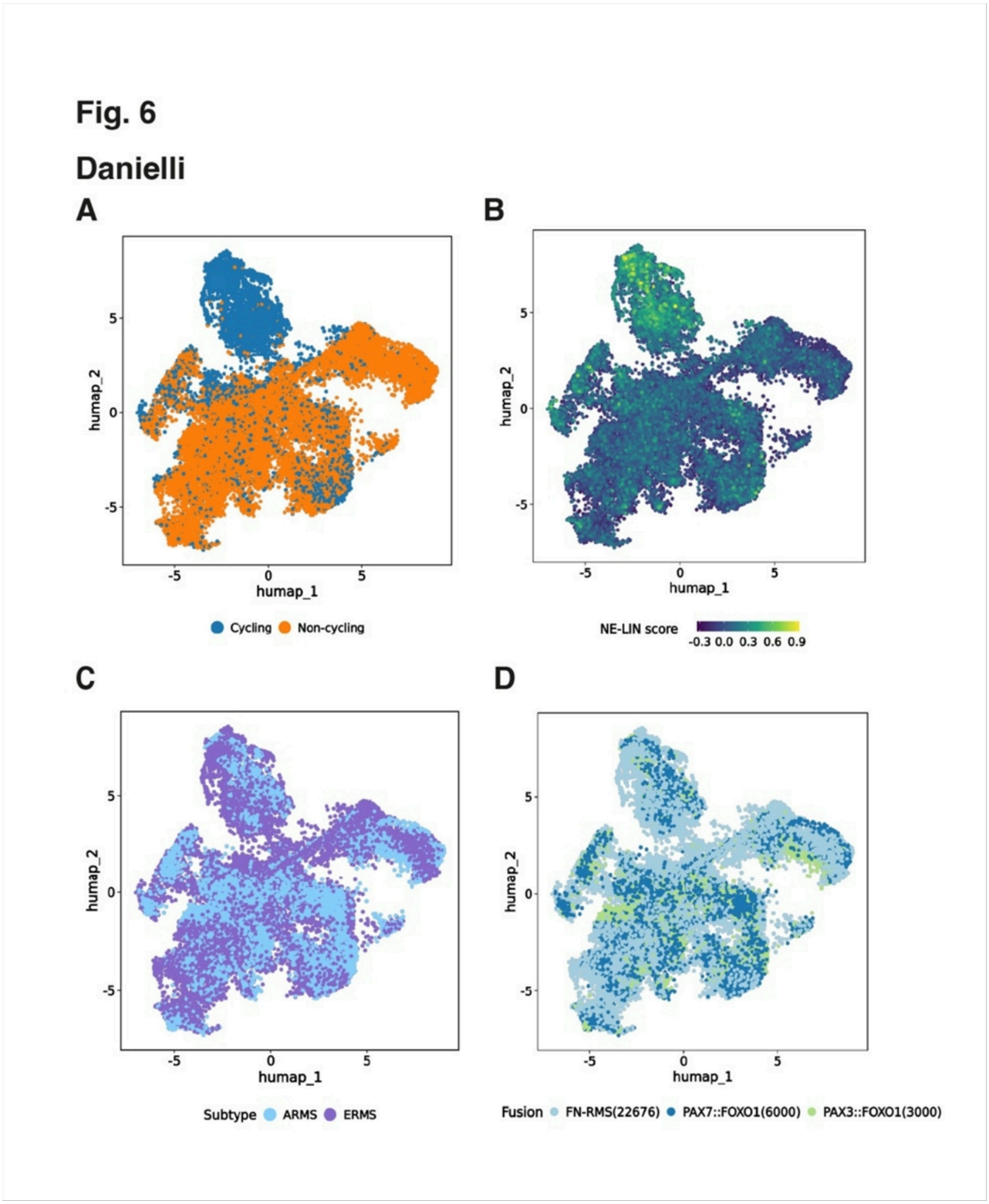


Figure 6

s.c./s.n. RNA-Seq analyses reveal the presence of the NE-LIN^{high} transcriptional status across further independent RMS cohorts. **A-D.** The analysis was performed on the 21 RMS samples (n=31676 cells) from two independent s.c./s.n. RNA-Seq cohorts [25,26] integrated and harmonized in the Danielli dataset [24]. UMAPs show cycling and non-cycling tumor cell clusters identified according to Danielli *et al.* (A), expression of the gssNE-LIN score (B), and visualization of tumor subtype (C) and fusion status (D).

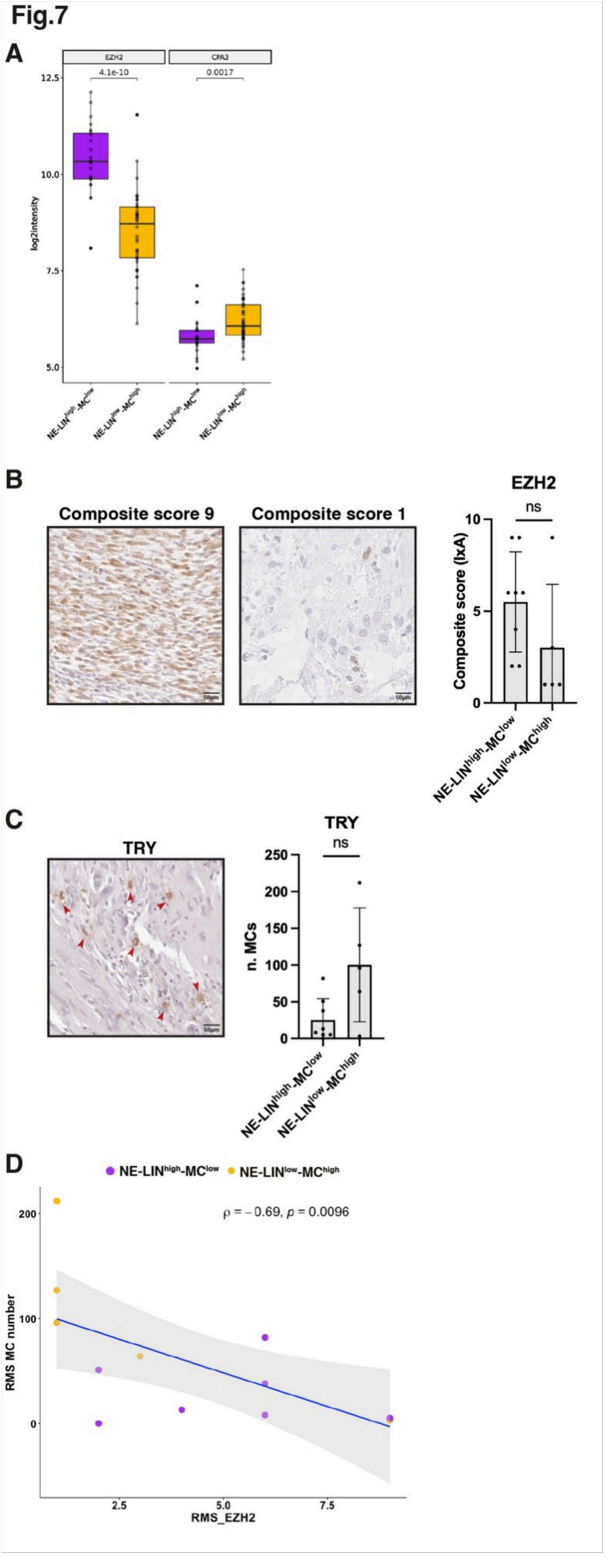


Figure 7

EZH2 expression and MC infiltration are inversely correlated in RMS tumor samples. A. Box plots showing the expression of *EZH2* and *CPA3* in the INT data set (n=61), comparing patients in the NE-LIN^{high}/MC^{low} vs NE-LIN^{low}/MC^{high} clusters. Statistical comparisons were performed using or Wilcoxon test. **B.** Representative images showing staining of the highest and lowest scores (left panel), and histogram showing values of EZH2 composite score in tumor RMS samples from the two clusters (right

panel; n= 5-8). **C.** Representative images of tryptase staining (left panel) and histogram showing MC number in tumor RMS samples (right panel; n=5-8). **B-C.** Histograms report means ± SD of all biological samples, which are represented by dots. Statistical comparisons were performed using Mann-Whitney test. **D.** Scatter plot with Spearman's correlation analysis coefficient (r) and p-value (p), showing the relationship between EZH2 composite score and MC number in and RMS samples analyzed in B-C.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- nrreportingsummaryPapersignaturesRMS.pdf
- SupplementatyTable1Singaturesused.xlsx
- SupplementaryTable2DEGSPAASRMSINTChoort.xlsx
- SupplementaryTableS3.DEGSPAASRMSDavicioniCohort.xlsx
- PaperSignaturesRMSSUPPLEMENTARYInformation.docx