







# Network models of primary melanoma microenvironments identify key melanoma regulators underlying prognosis

Won-Min Song <sup>1,2,3,8</sup>, Praveen Agrawal<sup>4,5,8</sup>, Richard Von Itter<sup>4,5</sup>, Barbara Fontanals-Cirera<sup>4,5</sup>, Minghui Wang <sup>1,2,3</sup>, Xianxiao Zhou <sup>1,2,3</sup>, Lara K. Mahal <sup>6</sup>, Eva Hernando <sup>4,5</sup>✉ & Bin Zhang <sup>1,2,3,7</sup>✉

Melanoma is the most lethal skin malignancy, driven by genetic and epigenetic alterations in the complex tumour microenvironment. While large-scale molecular profiling of melanoma has identified molecular signatures associated with melanoma progression, comprehensive systems-level modeling remains elusive. This study builds up predictive gene network models of molecular alterations in primary melanoma by integrating large-scale bulk-based multi-omic and single-cell transcriptomic data. Incorporating clinical, epigenetic, and proteomic data into these networks reveals key subnetworks, cell types, and regulators underlying melanoma progression. Tumors with high immune infiltrates are found to be associated with good prognosis, presumably due to induced CD8+ T-cell cytotoxicity, via *MYO1F*-mediated M1-polarization of macrophages. Seventeen key drivers of the gene subnetworks associated with poor prognosis, including the transcription factor *ZNF180*, are tested for their pro-tumorigenic effects in vitro. The anti-tumor effect of silencing *ZNF180* is further validated using in vivo xenografts. Experimentally validated targets of *ZNF180* are enriched in the *ZNF180* centered network and the known pathways such as melanoma cell maintenance and immune cell infiltration. The transcriptional networks and their critical regulators provide insights into the molecular mechanisms of melanomagenesis and pave the way for developing therapeutic strategies for melanoma.

<sup>1</sup>Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY, USA. <sup>2</sup>Mount Sinai Center for Transformative Disease Modeling, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY, USA. <sup>3</sup>Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY, USA. <sup>4</sup>Department of Pathology, NYU Grossman School of Medicine, New York, NY, USA. <sup>5</sup>Interdisciplinary Melanoma Cooperative Group, Perlmutter Cancer Center at NYU Langone Health, New York, NY, USA. <sup>6</sup>Department of Chemistry, University of Alberta, Edmonton, Canada. <sup>7</sup>Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY, USA. <sup>8</sup>These authors contributed equally: Won-Min Song, Praveen Agrawal. ✉email: [Eva.Hernando-Monge@nyulangone.org](mailto:Eva.Hernando-Monge@nyulangone.org); [bin.zhang@mssm.edu](mailto:bin.zhang@mssm.edu)

**M**elanoma is a highly aggressive tumor that accounts for less than 5% of all skin cancers, but 80% of skin cancer-related deaths<sup>1</sup>. Melanoma is a malignancy of pigment-producing cells, i.e., melanocytes, located predominantly in the skin. These pigments play protective roles against ultraviolet (UV) radiation in healthy individuals, safeguarding cells against DNA damage. Accumulation of UV exposure in melanocytes contributes to high mutational burden<sup>1,2</sup>, conferring mitogenic signals via defective DNA repair and replicative mechanisms and inducing oncogenic mutations such as those observed in *BRAF* and *NRAS*<sup>3</sup>. Mutations in *BRAF*, *NRAS*, *NF1*, and *KIT* have been all implicated as melanoma drivers<sup>1,2,4,5</sup>.

Primary melanomas are heterogeneous in terms of molecular and clinical features, as well as their tumor microenvironment. Interaction with diverse stromal cells can activate pro-invasive programs such as an epithelial-to-mesenchymal transition in some primary tumor cells, leading to metastatic dissemination<sup>6,7</sup>. Dynamic changes in the tumor microenvironment can also lead to immune surveillance escape<sup>8,9</sup> and are predictive of melanoma prognosis<sup>10</sup>. Within the tumor microenvironment, tumor-reactive T lymphocytes play a central role in suppressing tumor growth by infiltrating into malignant lesions and selectively killing cancer cells<sup>11,12</sup>. However, some subclones of melanoma tumors evade the immunosurveillance by intra-tumoral expression of programmed cell death ligand 1 (PD-L1), which binds to the co-inhibitory checkpoint receptor, programmed cell death protein 1 (PD-1)<sup>13,14</sup>.

While understanding the heterogeneity is critical for patient treatment, several factors have hampered a comprehensive molecular characterization of primary melanoma tumors. First, most large-scale, multi-omics studies focus on metastatic tumors or combine analysis of primary and metastatic tumors<sup>2</sup>. Further, these large-scale studies are often bulk-based and confounded by the diversity of cell-types within cancer. This complicates the identification of cell type-specific signaling circuits within the microenvironment. Although machine learning methods such as CIBERSORT<sup>15</sup> and ESTIMATE<sup>16</sup> can estimate relative cell compositions in bulk samples to some degree, they cannot replace the high-resolution analysis of cell-type-specific interactions from scRNA-seq.

Systems biology, especially network biology approaches, have proven effective for integrating diverse, large-scale datasets in complex human diseases<sup>17–39</sup>. Here, we applied an integrative multi-scale gene network analysis framework to jointly analyze the primary melanoma bulk RNA-sequencing data from The Cancer Genome Atlas (denote as pSKCM) and a published single-cell transcriptomic dataset<sup>40</sup>. We hypothesized that co-expressed gene modules associated with the patients' prognosis capture dysregulated pathways in primary melanoma etiology. By generating prognosis gene signatures from the TCGA data, we were able to intersect these signatures with gene modules and identify the enriched modules, subnetworks, and network drivers as pro-tumorigenic regulators of primary melanoma. Similarly, gene signatures associated with (epi-)genomic alterations were utilized to inform gene modules affected by these alterations. This integrative approach has proven effective in identifying causal molecular alterations in complex diseases such as Alzheimer's disease<sup>41,42</sup>, asthma<sup>43</sup>, breast cancer<sup>44</sup>, and gastric cancer<sup>45</sup>.

Our study revealed key immune cell types and signaling pathways, and predicted their regulators underlying primary tumors with varying degrees of tumor infiltration by jointly analyzing bulk and single-cell data. Key findings were replicated in independent bulk<sup>46</sup> and scRNA-seq datasets<sup>40,47</sup>. Further, the predicted pro-tumorigenic regulators of melanoma were validated via *siRNA* screening in vitro and in vivo xenografts.

## Results

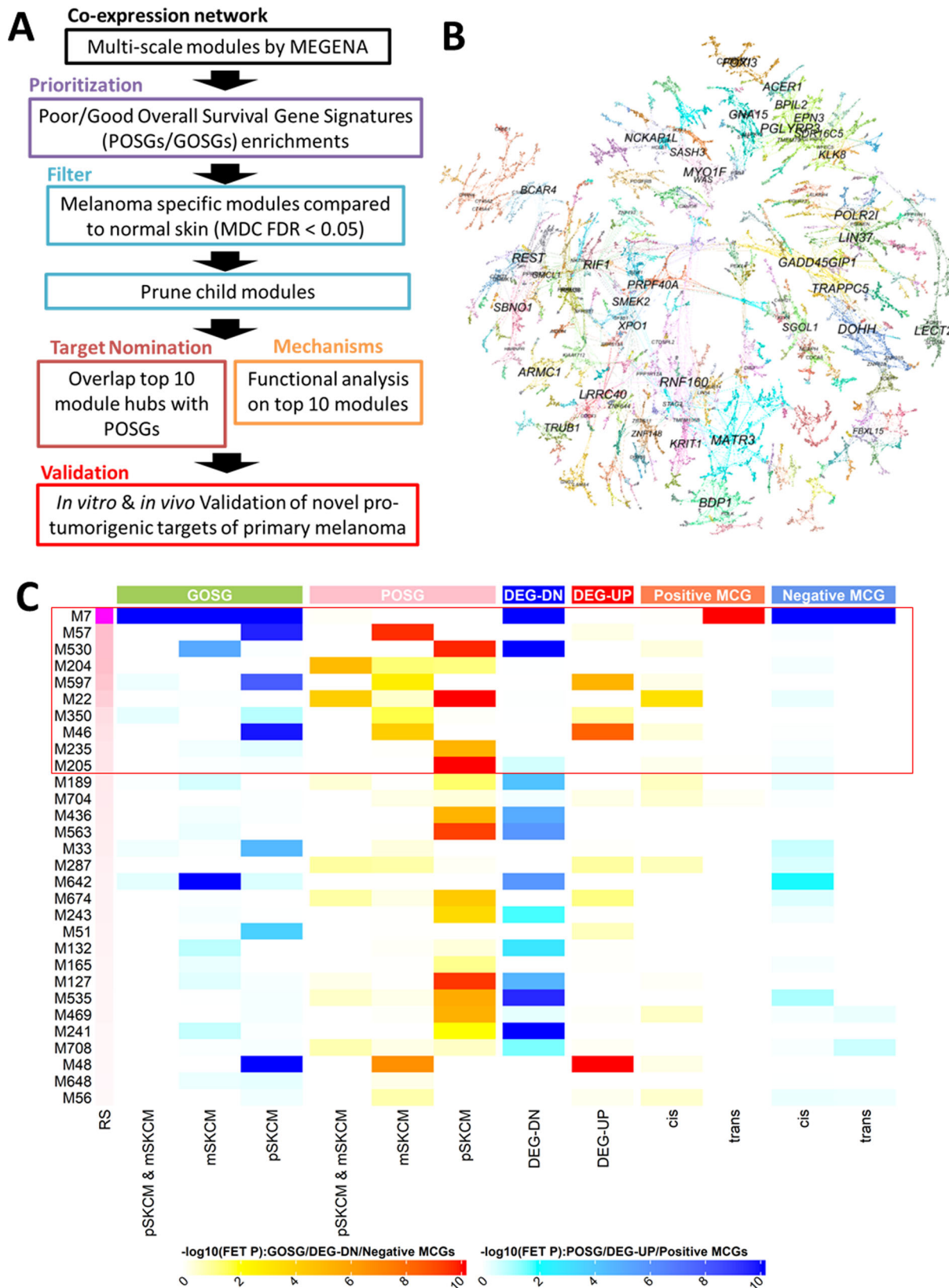
**Integrative network biology analysis of primary melanoma.** We constructed a gene co-expression network from the bulk-based primary skin cutaneous melanoma (pSKCM) RNA-seq data in The Cancer Genome Atlas (TCGA) to identify co-expressed gene modules (subnetworks) and their key regulators by multiscale embedded gene co-expression network analysis (MEGENA) (Fig. 1, Supplemental Fig. 2A, B)<sup>48</sup>. A total of 221 gene modules were prioritized by enrichment for the genes associated with overall survival and known primary melanoma-specific pathways (see “Methods”).

Among the top ten ranked modules, five (M7, M597, M57, M46, and M350) were enriched for the genes whose upregulation was associated with increased overall survival (denoted GOSG; Supplementary Fig. 2B) while the others (M530, M22, M205, M204, and M235; Supplementary Fig. 2A) were enriched for the genes whose upregulation was associated with poor prognosis (denoted POSG). In general, not many POSG enriched modules are associated with known pathways or functions. Only one POSG enriched module M22 significantly overlaps with generic transcription pathways curated by REACTOME (corrected Fisher's Exact Test  $p$  value (cFET  $p$ ) = 3.45E – 54, 13 fold-enrichment(FE)), due to co-expressions of zinc finger proteins.

In contrast, the top-ranked GOSG enriched module, M7 (Fig. 1C; cFET  $p$  = 5.76E – 193, 13.2 FE) is characterized by immune response pathways, including the hallmark IFN- $\gamma$  response (cFET  $p$  = 4.95E – 95, 9.29 FE), the *CTLA4* pathway (cFET  $p$  = 6.20E – 8, 10.6 FE), TNF $\alpha$  signaling via NF $\kappa$ B (cFET  $p$  = 8.36E – 16, 3.6 FE) and the T-cell receptor alpha pathway (cFET  $p$  = 2.56E – 11, 13.1 FE). Hyper-methylation in cis-CpG sites of several genes in M7 was associated with their down-regulation (cis-MCG: cFET  $p$  = 4.01E – 10, 2.43 FE). As altered gene expression and methylation affect each other directly and indirectly, we dissected the causal relationships between these genes and the altered CpG sites. We identified *PTPN6* and *PTPRCAP* as potential modulators of epigenetic regulation of the leukocyte activation pathway in M7 (see Epigenetic silencing of T-cell activation leads to poor prognosis in Supplementary Methods; epigenetic regulatory network in Supplementary Data 6).

**Upregulated immune response subnetwork associated with good overall survival.** The module most strongly associated with a good prognosis, M7, is comprised of sub-modules<sup>48</sup> reflecting distinct immune response pathways, including leukocyte activation (M110) and antigen presentation (M112, Supplementary Fig. 3B). M112 captured the PD1/PD-L1 signaling pathway, including PD-L1 and several key regulators of the PD-L1 transcription pathway as its hubs, such as *STAT1* and *IRF1*<sup>49</sup>. The immune regulatory role of M112 was further supported by its enrichment of the genes upregulated by PD1 ligation in T cells<sup>50</sup> (cFET  $p$  = 1.62E – 21, 16.6 FE). The hubs of M7 were also associated with the T-cell activation pathway (cFET  $p$  = 3.86E – 10, 40.2 FE; Supplementary Data 3).

We dissected the immune cell populations involved in M7 and its sub-modules by integrating the scRNA-seq data comprised of 4645 cells from 19 melanoma patients (GEO accession: GSE72056)<sup>40</sup>. The individual cells were annotated with the published and inferred cell types, including melanoma, T cells (CD3+), macrophages, endothelial, and cancer-associated fibroblasts. We further classified T-cell subpopulations into CD8+ cells, CD4+ T cells, resting memory (RM) CD4+ T cells, and M1-polarized macrophages (M1 macrophages) (Supplementary Fig. 4A; see “Methods”). Genes in M7 sub-modules were expressed in distinct cell types, predominantly melanoma cells (expressed in M401, M402, M110, M639, and M642), M1-macrophages (expressed in M400, M401, and



M642), B-cells (expressed in M111), CD8+ T cells (expressed in M399, M401, M402, and M403) and CD4+ T cells (expressed in M399, M401, and M402, Supplementary Fig. 4C).

To understand gene interactions within each cell population, we constructed a co-expression network for each cell cluster (labeled as CLS1, CLS2, etc.), termed as sc-networks (see

Supplementary Fig. 4B). Then, we evaluated the concordance between the pSKCM network and the sc-networks to identify cell populations involved in the bulk-based gene interactions. For example, the interferon- $\gamma$  (IFN- $\gamma$ ) response modules M401 and M402 significantly overlap the macrophage-enriched modules in the CLS3 sc-network (cFET  $p = 9.48E - 20$ ). This makes sense as

**Fig. 1 Analytic flow of the co-expression network analysis of the primary skin cutaneous melanoma cohort in TCGA.** **A** Overall workflow. **B** Global co-expression network of the primary melanoma samples from TCGA. Gene modules identified with the default resolution parameter  $\alpha = 1$  are shown in distinct colors. Highly connected hub genes are labeled with respective gene symbols. **C** Heatmap representation of molecular features associated with the top 30 gene modules by survival analysis. Heatmap color represents an enrichment of respective gene signatures in the modules. GOSG represents the good prognosis associated genes in pSKCM and/or metastatic melanoma (mSKCM) cohorts. POSG includes the poor prognosis associated genes in pSKCM and/or mSKCM cohorts. DEG-UP and DEG-DN are the upregulated and downregulated gene signatures in tumors in comparison with adjacent normal tissues, respectively. MCG represents the methylation correlated genes (MCG). MCG with a positive (negative) correlation between the gene and a CpG site is called Positive (negative) MCG. MCG is also classified into *cis*- or *trans*-regulation dependent on the distance between the respective CpG site's location and the genomic location of the gene (see the "Methods").

IFN- $\gamma$  is a known activator of macrophages<sup>51</sup>. The *MYO1F* neighborhood from the pSKCM network was significantly enriched in the *MYO1F* neighborhood in the CLS3 network (Fig. 2A; cFET = 2.12E - 6, 211 FE; Supplementary Fig. 4G). In single-cell transcriptomes, *MYO1F* was consistently expressed in M1-macrophages and IFN- $\gamma$ -secreting CD8+ T cells in the discovery cohort and an independent cohort of 33 melanoma tumors<sup>47</sup> (Fig. 2B; Supplementary Fig. 5B; Supplementary Fig. 6). These *MYO1F*<sup>high</sup> M1-macrophages cells expressed PD-L1 (*CD274* in Fig. 2B), and *MYO1F*<sup>high</sup> CD8+ T cells showed high IFN- $\gamma$  expression (*IFNG* in Fig. 2B). Given that M1-macrophages are pro-inflammatory and interact with cytotoxic CD8+ T cells<sup>52</sup>, these results suggest *MYO1F* may play an essential role in IFN- $\gamma$ /PD-L1 signaling in primary melanoma.

The findings from single-cell transcriptomes are consistent with those from the pSKCM bulk RNA-seq data. *MYO1F* was correlated with the inferred M1-macrophage abundance in pSKCM but not with M0-/M2-macrophages (Fig. 2C). We leveraged the published immune subtypes of the pSKCM samples by Thorsson et al.<sup>53</sup> to clarify which immune subtype is associated with *MYO1F*. *MYO1F* showed predominantly high expression in the IFN- $\gamma$  dominant subtype, C2 (Supplementary Fig. 7A), and was correlated with inferred M1/M2 macrophage ratio and CD8+ T-cell abundance within C2 (Supplementary Fig. 7B).

In agreement with these observations, *MYO1F* has been identified as an upstream regulator that hyper-activates macrophage *STAT1*, leading to M1-polarization and IFN- $\gamma$  secretion by stimulating intercellular adhesion<sup>54</sup>. *MYO1F* may also regulate leukocytes. *MYO1F* is associated with the T cell co-inhibitory receptor, PD-L1, in both bulk cohorts (Fig. 2D). PD-L1 upstream genes such as *IRF1* and *STAT1* are more highly expressed in *MYO1F*<sup>high</sup> M1-macrophages than in the other CD19+ monocyte/macrophages (Supplementary Fig. 6).

To validate the IFN- $\gamma$  response subnetwork in macrophages captured in the pSKCM network, we curated the loss-of-function (LoF) signatures of a key IFN- $\gamma$  response pathway regulator, *STAT1*, from bone marrow-derived macrophages (BMM) of *Stat1* knockout (*Stat1*-KO) mice (GSE48970) by querying the CREED database<sup>55</sup>. As expected, the downregulated genes from *Stat1*-KO BMM are significantly enriched in the *STAT1* centered subnetwork in the pSKCM network (Supplementary Fig. 3C; cFET  $p = 2.80E - 31$ , 34.6 FE). Furthermore, *STAT1* expression was correlated with inferred M1-macrophage abundance by CIBERSORT in pSKCM (Spearman  $\rho = 0.59$ ,  $p = 4.28E - 11$ ; Supplementary Fig. 3E). These results support the *STAT1* centered network depicts IFN- $\gamma$  response of macrophages in pSKCM. Altogether, *MYO1F* may modulate IFN- $\gamma$  response in M1-macrophages and cytotoxic CD8+ T cells via the *STAT1/IRF1/PD-L1* axis.

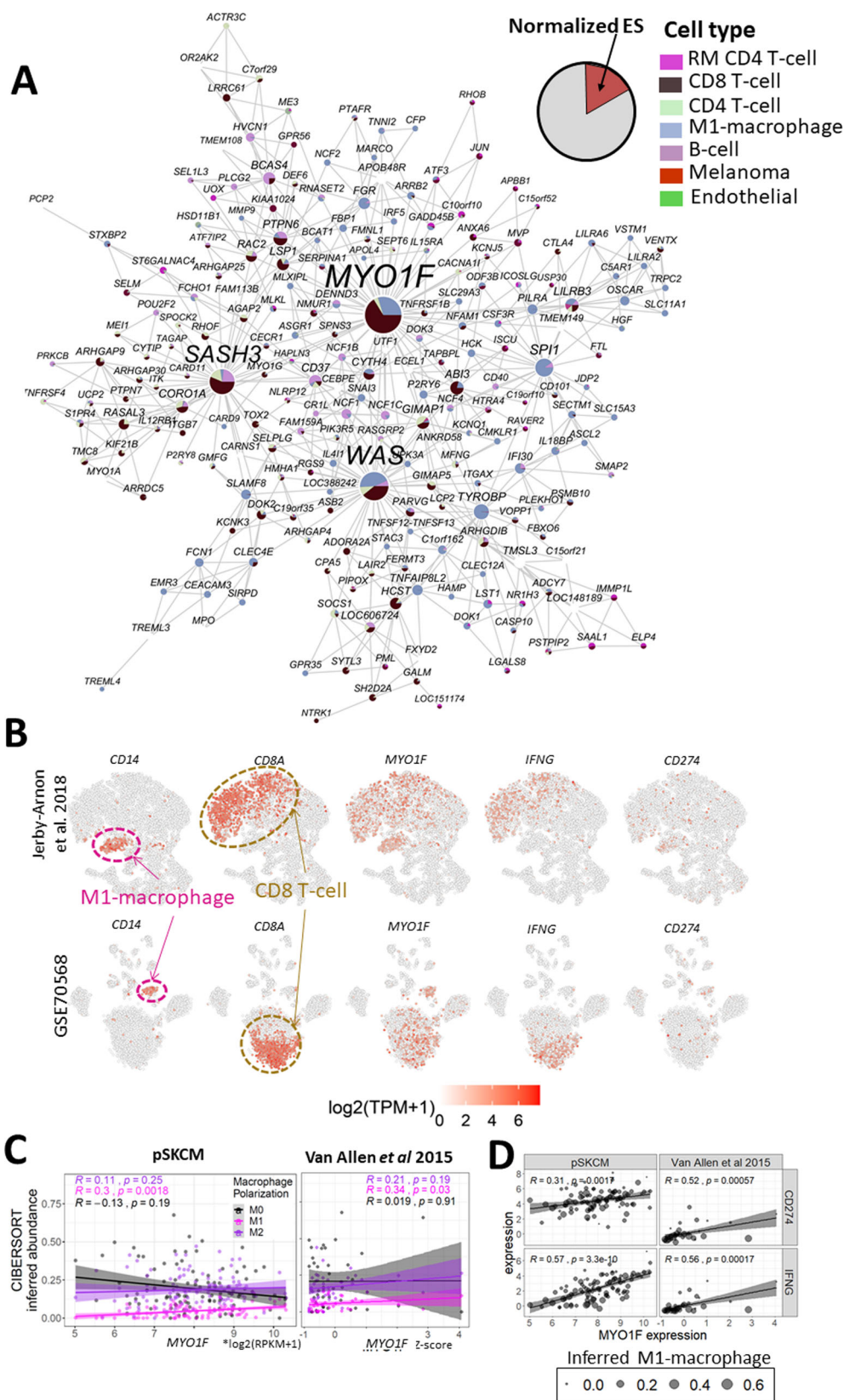
**Upregulation of intra-tumoral DNA repair and mRNA splicing pathways associates with poor prognosis.** We further characterized the top-ranked modules associated with poor prognosis,

including M530, M22, M205, M204, and M235 (Supplementary Fig. 2A). Although these modules are not associated with known pathways and functions in the MSigDB, the hub genes in these modules include known regulators of pre-mRNA splicing such as serine-arginine protein-like factors, *SFRS3*, and *SFRS13A*<sup>56</sup>, which are the hubs of M235, and a spliceosome associated protein, *SR140* as a hub of M530<sup>57</sup>. M205 includes a nuclear pore component (*NUP205*)<sup>58</sup> and a splicing factor (*TNPO3*)<sup>59</sup>. On the other hand, the hubs of M22 include DNA repair regulators such as *FANCM*<sup>60</sup> (a component of the Fanconi anemia core complex) and *ZNF180*, which has a locus associated with aberrant genomic rearrangements<sup>61</sup>. Interestingly, *FANCM* and *ZNF180* gene expressions were correlated with expression of a DNA mismatch repair protein, *MSH2*, from the reverse-phase protein array (RPPA) datasets (Supplementary Data 9) with  $\rho = 0.530$  (adjusted  $p = 4.53E - 2$ ) and 0.529 (adjusted  $p = 4.48E - 2$ ), respectively. In melanoma, high expression of DNA repair genes including *MSH2* is associated with metastasis and may explain their resistance to chemo- and radio-therapies through maintaining the genetic stability<sup>62</sup>. These data suggest DNA repair and mRNA splicing pathways are linked to poor prognosis in primary melanoma.

To identify primary melanoma prognosis regulators, we examined the intersection between the previously identified top ten modules' hub genes and the poor prognosis signatures (Fig. 3A). This scheme nominated 18 regulators, which were more likely to be essential for melanoma cell viability than the poor prognosis signature alone (Wilcox  $p$  value = 3.03E - 4; Supplementary Fig. 10) based on CRISPRi screening data in the Achilles database<sup>63</sup>. These potential regulators' gene expression profiles were correlated with the inferred tumor purity score by ESTIMATE, supporting them as tumor cell-intrinsic genes (Supplementary Fig. 9). Higher expression of these key regulators in tumors was associated with lower CD8+ T cell as inferred by CIBERSORT (Supplementary Fig. 9), suggesting intra-tumoral expression of these genes was associated with low infiltration of cytotoxic T cells. Therefore, these predicted regulator genes capture intra-tumoral pathways essential to tumor viability and confer poor prognosis in primary melanoma.

**Systematic validation of candidate network drivers of poor prognosis.** To validate the 18 candidate regulators of the molecular networks underlying primary melanoma prognosis, we

performed siRNA knockdown experiments in SKmel147 (*NRAS* mutant) and A375 (*BRAF* mutant) cell lines (see siRNA screening of candidate targets in Supplementary Methods). Of the 18 candidates, siRNAs for 17 were readily available for high-throughput screening assays to assess the impact on cell growth and invasion. siRNA-mediated knockdown of five genes (*KRIT1*, *ZNF680*, *SRSF10*, *ZNF180*, and *TMEM160B*) in SKmel147 cells significantly reduced cell growth ( $p$  value < 0.05), relative to non-targeting controls (NTC) (top, Fig. 3B). Knockdown of eight key regulators (*PPP1R2*, *G2E3*, *MKLN1*, *ZNF225*, *ZNF180*, *ZNF347*, *KRIT1*, and *U2SURP* (*SR140*)) in SKmel147 cells significantly



reduced invasion (middle, Fig. 3B). A similar siRNA invasion screening on A375 cells yielded significant differences for five genes: *TBC1D23*, *PPP1R2*, *ZNF180*, *MYNN*, and *ZNF347* (bottom, Fig. 3B). Altogether, we validated the pro-tumorigenic effects of 13 (76.4%) of 17 candidate regulators tested in SKmel147 and A375 cells (Table 1).

**ZNF180 silencing antagonizes melanoma cell proliferation and invasion in vitro and in vivo.** Among the 13 successfully validated functional candidate genes, *PPP1R2* (a hub of M530), *ZNF180*, and *ZNF347* (hubs of M22) showed robust effects in multiple siRNA functional assays. Transient depletion of *PPP1R2*, *ZNF180*, and *ZNF347* significantly reduced the invasive capacity

**Fig. 2 MYO1F as a potential regulator of IFN- $\gamma$  response in primary melanoma microenvironment. A** The network of a POSG-enriched module, M401. Pie size in a node piechart is proportional to the specificity of the respective gene in a cell type, defined as an enrichment score,  $-\log_{10}(\text{FET FDR})$ , which is derived from the enrichment test for the cells expressing the respective gene and the cells that are predicted to belong to the cell type. Enrichment scores for each gene are normalized by the sum of the enrichment scores for all the seven cell types. Cell types with FET FDR < 0.05 are shown in the piecharts. Pie color indicates different cell types: resting memory (RM) CD4 T cell (magenta), CD8 T cell (dark brown), CD4 T cell (palegreen), M1-macrophage (skyblue), B-cell (dark plum), Melanoma (dark red), Endothelial (green). **B** *MYO1F* is expressed in M1-macrophages and CD8 T cells along with PD-L1 (CD274). tSNE plots of the scRNA-seq datasets from GSE72056 and Jerby-Arnon et al.<sup>47</sup> show the cellwise expressions of macrophage marker (*CD14*) and CD8 T-cell marker (*CD8A*) along with *MYO1F*, *INF- $\gamma$*  (*IFNG*), and PD-L1 (*CD274*). The scale for expression levels by  $\log_2(\text{TPM} + 1)$  values are shown at the bottom. **C** Scatter plots between *MYO1F* and CIBERSORT inferred macrophage populations (M0–M2) in the pSKCM and Van Allen et al.<sup>46</sup> cohorts. 95% confidence intervals for Spearman correlations between *MYO1F* and macrophages are (–0.29, 0.33) for M0, (0.036, 0.59) for M1 and (–0.11, 0.49) for M2. **D** Spearman correlation between *MYO1F* expression and PD-L1(*CD274*)/*INF- $\gamma$*  (*IFNG*) in the pSKCM and Van Allen et al. 2015 cohorts. Point size is proportional to inferred M1-macrophage abundance. The 95% confidence intervals of the correlation coefficients for *CD274* and *IFNG* in pSKCM are (0.12, 0.47) and (0.42, 0.69), respectively. The confidence intervals for *CD274* and *IFNG* in the Van Allen et al. 2015 cohort are (0.25, 0.72) and (0.30, 0.74), respectively.

of SKmel147 and A375 cells in trans-well invasion assays (middle, Fig. 3B). Of these candidates, only the knock-down of *ZNF180* showed significant anti-proliferative effects in SKmel147.

Therefore, we further investigated the anti-proliferative effect of *ZNF180* silencing in vivo. Lower *ZNF180* transcript and protein levels confirmed effective knock-down in sh*ZNF180*–SKmel147 cells relative to shNTC-transduced control cells (Supplementary Fig. 14). SKmel147 cells stably infected with control (shNTC) or sh*ZNF180* expressing lentivirus were injected into the flanks of NOD/Shi-scid/IL-2Rgamma null (NSG) mice ( $n = 12/\text{group}$ ). Mice injected with sh*ZNF180* transduced cells showed significantly reduced tumor growth ( $p < 0.0001$ ) and weight at termination ( $p < 0.0001$ ) compared to mice injected with cells carrying a non-targeting shRNA (Fig. 3D–G). This supports the role of *ZNF180* in melanoma growth in vivo. Together, our data indicate that *ZNF180* is essential for melanoma growth both in vitro and in vivo.

RNA-sequencing revealed transcriptional changes elicited by *ZNF180* silencing in SKmel147 cells (see RNA-sequencing in Supplementary Methods). Genes differentially expressed in response to silencing *ZNF180* (termed si*ZNF180*-DEGs; Supplementary Data 8) were significantly enriched in the subnetwork of *ZNF180* in the pSKCM network using threshold-free rank–rank hypergeometric overlap test (see Analysis of RNA-sequencing data of siRNA knock-down SKmel147 cells in Supplementary Methods; Supplementary Fig. 11E). Moreover, the genes down-regulated by si*ZNF180* were enriched in the four-layer network neighborhood of *ZNF180* in the pSKCM network (Supplementary Fig. 12A; cFET  $p = 1.41\text{E} - 83$ , 3.69 FE), thus validating the topological structure of pSKCM gene network. Such findings were replicated in the co-expression network analyses of the metastasis SKCM (mSKCM) and Van Allen et al.<sup>46</sup> cohorts (Supplementary Fig. 12B).

***ZNF180* is a multifunctional driver of primary melanoma etiology.** The si*ZNF180*-DEGs were associated with several aberrant signaling pathways in cancer and pSKCM prognosis. The downregulated genes (si*ZNF180*-DN) significantly overlap with pSKCM-POSG (cFET  $p = 1.07\text{E} - 172$ , 3.00 FE). The intersection between these two signatures was associated with the cell cycle pathway (cFET  $p = 2.15\text{E} - 6$ , 3.60 FE). On the contrary, the upregulated genes (si*ZNF180*-UP) significantly overlapped the good prognosis signatures in pSKCM (FET FDR = 7.32E – 54, 2.03 FE), and their intersection was associated with the generic immune system pathway (cFET  $p = 6.37\text{E} - 3$ , 2.64 FE).

We further examined a potential correlation between *ZNF180* and protein expressions from the RPPA to identify candidate interacting proteins. Protein expressions of mismatch repair protein, MSH2, and plasminogen activator inhibitor, PAI-1, were

significantly correlated with *ZNF180* expression with  $\rho = 0.54$  (adjusted  $p = 4.48\text{E} - 2$ ) and  $-0.44$  (adjusted  $p = 4.83\text{E} - 2$ ), respectively. Particularly, MSH2 closely interacted with the genes down-regulated by si*ZNF180* in the protein-protein interaction network (Fig. 3H). Given the role of MSH2 and DNA repair in melanoma cells to maintain tumor cells' genetic stability, the data suggests *ZNF180* may play a role in regulating these pro-tumorigenic pathways.

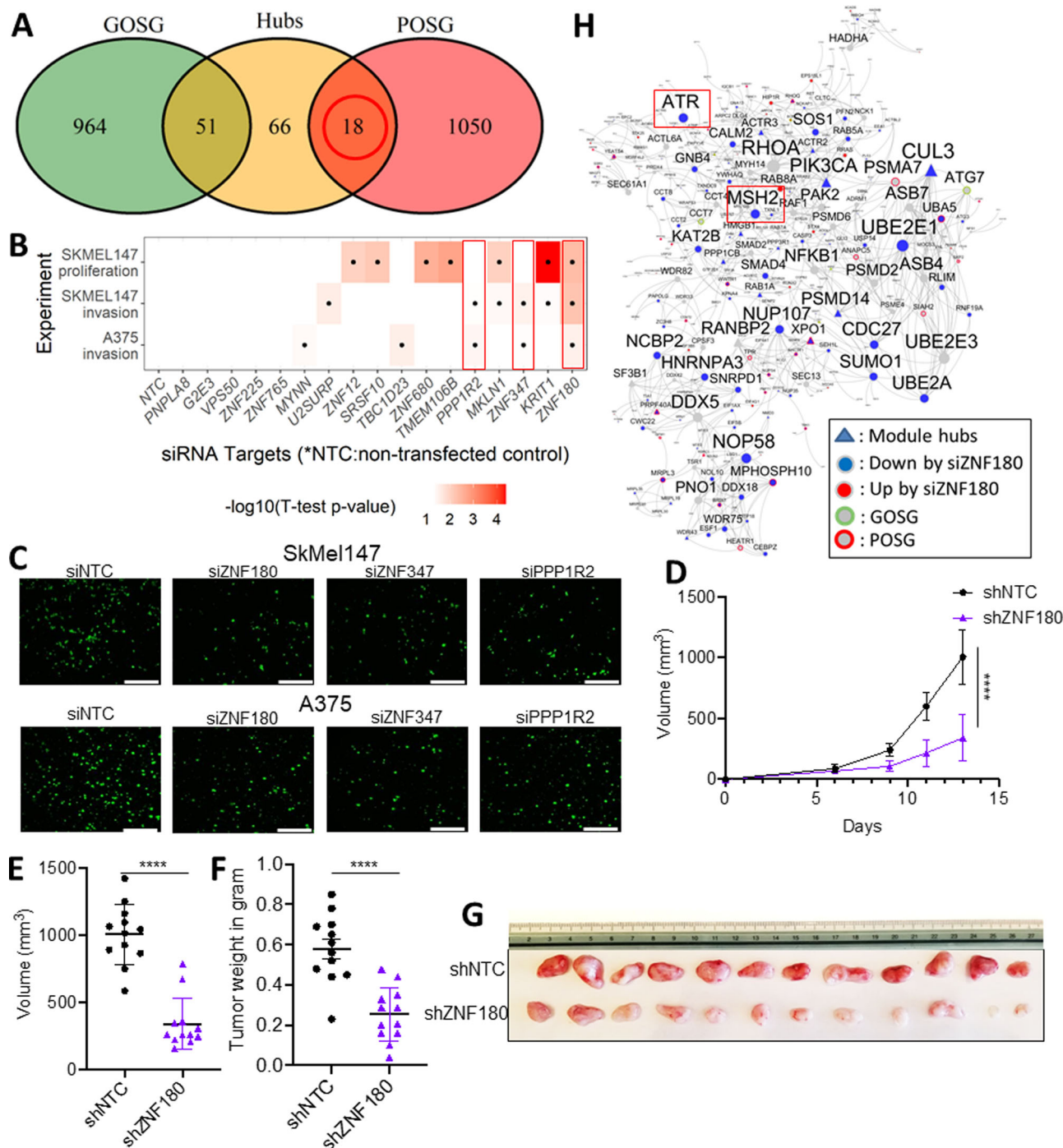
The gene modules enriched for the si*ZNF180*-DEGs more comprehensively captured the signaling pathways controlled by *ZNF180*. The modules enriched in the genes upregulated by si*ZNF180* included tumor suppressors and negative cell cycle regulators (M597; Supplementary Table 1). In contrast, those enriched for the genes downregulated by si*ZNF180* were associated with DNA repair, epithelial–mesenchymal transition (M25), oncogenesis (M24, M25, and M257), pre-mRNA splicing (M25 and M30), protein modification/degradation and chromatin modification/remodeling (Fig. 4D; see Differentially expressed pathways by *ZNF180* suppression in SKMEL147 cells in vitro in Supplementary Methods).

Notably, M25 emerged as a core module regulated by *ZNF180* as it was enriched for not only the down-regulated genes by si*ZNF180* but also the poor prognosis signatures from pSKCM (Fig. 4D). M25 was associated with oncogenic signaling pathways, including MYC targets (Hallmark MYC targets V1: FET FDR = 9.43E – 4, 2.58 FE). M25 also constituted a coherent protein interaction network, comprised of high confidence interactions in the STRING database (>70% confidence; Fig. 3D)<sup>64</sup>. The protein interaction network within M25 included established oncogenes such as *ROCK1*, *PIK3CA*, *RHOA*<sup>65,66</sup>, and the key DNA repair regulators, *MSH2*, and *ATR* (Fig. 3D). These results indicated close interactions between the DNA repair signaling and oncogenic signaling pathways in primary melanoma.

To understand the cell populations regulated by *ZNF180*, we deconvoluted the pSKCM bulk samples into de novo cell types from the melanoma single-cell transcriptome studies using the cellular population mapping (CPM) algorithm implemented in “scBio” R package<sup>67</sup>. *ZNF180* expression was consistently correlated with increased cancer-associated fibroblasts and decreased M1-macrophage abundance in two studies, GSE72056<sup>40</sup> (Fig. 4C), and Jerby-Arnon et al.<sup>47</sup> (Supplementary Fig. 5D). Overall, our results show that *ZNF180* is a multifunctional driver of oncogenic and DNA repair pathways that may also modulate tumor infiltration by M1 macrophages and cancer-associated fibroblasts.

## Discussion

This study constructed a compendium of multi-scale gene networks from bulk and single-cell RNA-seq data to dissect complex



**Fig. 3 Systematic identification and validation of *ZNF180* as a promising target in primary melanoma.** **A** Intersections between the poor survival gene signature from pSKCM (pSKCM-POSG), and the good survival gene signature (pSKCM-GOSG), and the hubs of the top ten modules. Eighteen hub genes which are also pSKCM-POSG were nominated for siRNA screenings. **B** siRNA screening of the nominated targets for proliferation in (top) and invasion (middle) in SKmel147, and invasion in A375 cells (bottom). The heatmap shows  $-\log_{10}(t \text{ test } p \text{ value})$  to remark the significance of the siRNA screening results, and significant results are marked by dots. Red rectangles highlight consistently validated genes (*ZNF180*, *ZNF347* and *PPP1R2*). **C** Trans-well matrigel invasion assays on SKmel147 and A375 cells transfected with siNTC or siZNF180, siZNF347, and siPPP1R2. Invading cells were quantified by counting the number of SKmel147 and A375 cells that invaded into the basal side of Matrigel-coated trans-well inserts after 8 and 6 h respectively,  $n = 5$  fields per replicate; 3 replicates per condition, representative images are shown. Scale bar corresponds to 100  $\mu\text{m}$ . **D** Growth curve representing average tumor volume over time in mice injected with shNTC or shZNF180-transduced SKmel147 cells ( $n = 12$  per group). Tumor volume (**E**) and tumor mass (**F**), and (**G**) images of resected tumors taken 13 days post injection. Data are presented as mean  $\pm$  SD in (**D-F**). \*\*\*\* denotes significance of the comparison by two-tailed unpaired  $t$  test by  $p < 0.0001$ . **H** M25 captures protein interactions curated in the STRING database with a confidence score  $>70\%$ . Node colors and border colors denote differentially expressed genes by siZNF180 (siZNF180-DEG) and pSKCM survival signatures as shown in the legend.

**Table 1** List of top nominated key genes associated with prognosis in primary melanoma tumors.

| Hub gene                                       | Connectivity | Cox P    | log-rank P | Module |
|--|--------------|----------|------------|--------|
| <i>PPP1R2</i> <sup>a,b</sup>                   | 17           | 0.048877 | 0.04945458 | M530   |
| <i>TBC1D23</i> <sup>b</sup>                    | 17           | 0.05934  | 0.04945458 | M530   |
| <i>MYNN</i> <sup>b</sup>                       | 16           | 0.002709 | 0.00286385 | M530   |
| <i>SR140</i><br>( <i>U2SURP</i> ) <sup>a</sup> | 16           | 0.002702 | 0.00142604 | M530   |
| <i>G2E3</i> <sup>a</sup>                       | 25           | 0.018412 | 0.04492035 | M22    |
| <i>ZNF180</i> <sup>c,a,b</sup>                 | 24           | 0.004064 | 0.00422203 | M22    |
| <i>ZNF225</i> <sup>a</sup>                     | 23           | 0.206873 | 0.02719113 | M22    |
| <i>ZNF765</i>                                  | 19           | 0.002705 | 0.01459658 | M22    |
| <i>ZNF347</i> <sup>a,b</sup>                   | 17           | 0.013813 | 0.01459658 | M22    |
| <i>SFRS13A</i>                                 | 25           | 0.00736  | 0.00562611 | M235   |
| ( <i>SRSF10</i> ) <sup>a</sup>                 |              |          |            |        |
| <i>KRIT1</i> <sup>a,c</sup>                    | 47           | 0.017451 | 0.00616876 | M205   |
| <i>TMEM106B</i> <sup>c</sup>                   | 33           | 0.016207 | 0.00063893 | M205   |
| <i>PNPLA8</i>                                  | 25           | 0.034177 | 0.00142604 | M205   |
| <i>CCDC132</i><br>( <i>VPS50</i> )             | 22           | 0.18037  | 0.021213   | M205   |
| <i>ZNF680</i> <sup>c</sup>                     | 20           | 0.02294  | 0.00142604 | M205   |
| <i>ZNF12</i>                                   | 18           | 0.001513 | 0.00286385 | M205   |
| <i>MKLN1</i> <sup>a</sup>                      | 18           | 0.082066 | 0.03678236 | M205   |

<sup>a</sup>Validated genes from invasion screening in SKmel147 cells.

<sup>b</sup>Validated genes from invasion screening in A375 cells.

<sup>c</sup>Validated genes from proliferation screening in SKmel147 cells.

Connectivity of a node (i.e., a gene) is the number of links incident to the node in the co-expression network of TCGA-pSKCM. Cox P (i.e., Cox *p* value) was obtained from a two-sided Cox proportional hazard model for the respective gene expressions. Log-rank *p* (i.e., log-rank *p* value) was obtained by grouping patients by median expressions of the respective genes via a two-sided log-rank test. The module shows the gene module membership of the respective genes. Genes highlighted by bold fonts were chosen for comprehensive transcriptome analysis by RNA-sequencing of SKmel147 cells transfected with the respective siRNAs.

molecular interactions and co-regulation underlying primary melanoma pathogenesis. Our analysis revealed a number of pathways and regulators underlying primary melanoma and its microenvironment. The upregulation of the immune response module, M7, is a strong positive prognostic factor and correlated with increases in CD8+ T cells and M1-macrophages. In contrast, the regulators of the subnetworks predictive of poor prognosis highlighted intra-tumoral DNA repair pathways. We systematically tested the anti-tumoral effects of 17 candidates via siRNA knockdown experiments in vitro. *ZNF180* had the most consistent anti-tumor effects and was further validated in vivo.

Network drivers of key cancer modules as identified by MEGENA are potential regulators of many other genes, while networks provide functional context for how those regulators operate. While the traditional prognosis analysis offers a list of prognostic genes to identify enriched known pathways, such simple analyses are constrained on pre-collected knowledge-bases, limiting the discovery of de novo disease pathways. On the contrary, network-based findings offer mechanistic insights such as gene modules (i.e., de novo pathways) and hubs (i.e., potential regulators) that complement traditional approaches. Identification of *MYO1F* and *ZNF180* and their intrinsic network models supports the advantages of the integrative network approach adopted in this study.

Notably, *MYO1F* emerged as a potential regulator in macrophages. *MYO1F* is an influential hub in the bulk-based pSKCM network (i.e., a top hub in the M401 module) and the network from the pro-inflammatory M1 macrophage cell cluster in the single-cell transcriptomic data (i.e., a hub in M1-macrophage enriched cell cluster). We confirmed high *MYO1F* expression in M1 macrophage populations in the melanoma single-cell transcriptomic data (Fig. 2B).

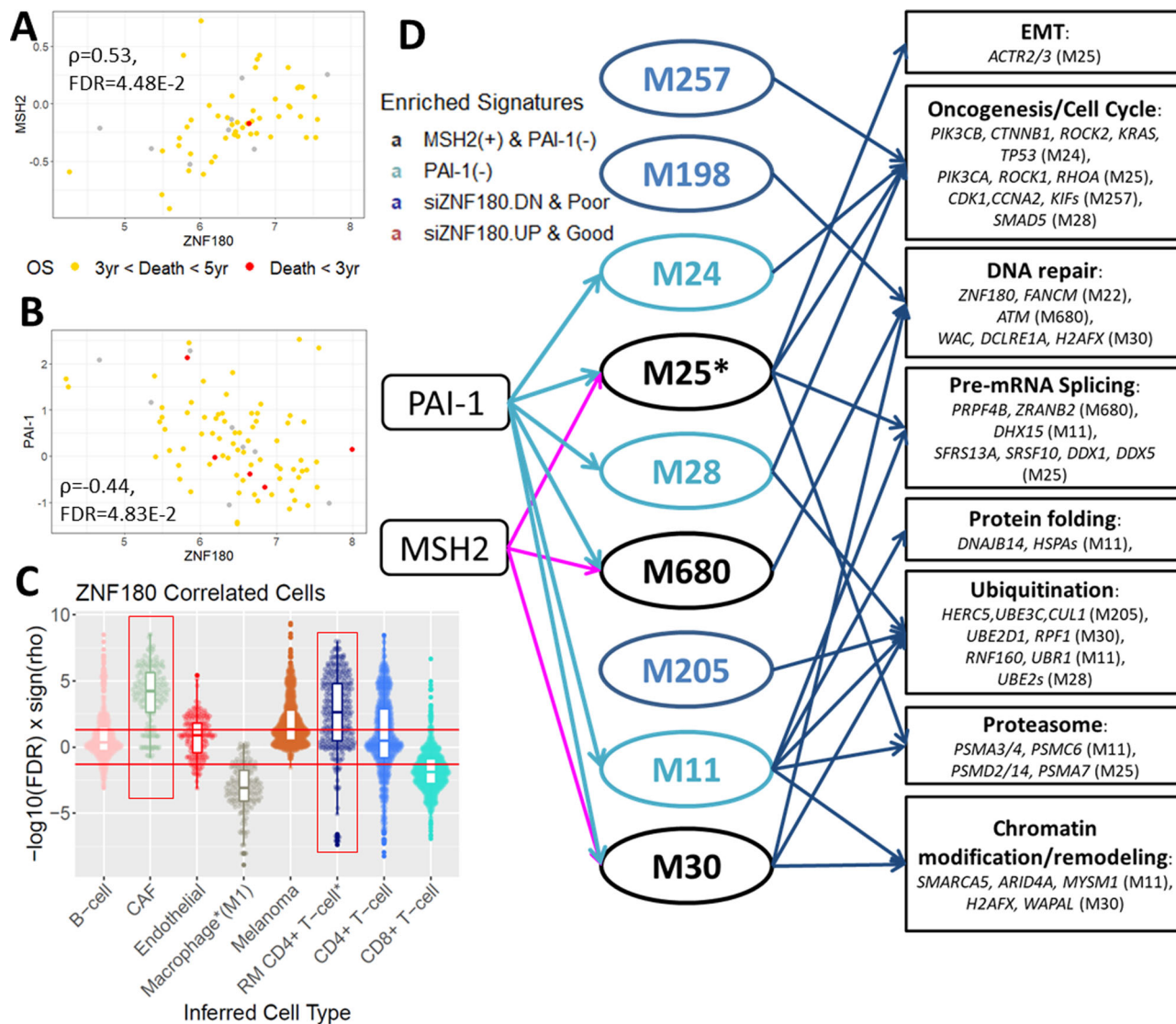
The data indicate *MYO1F* plays a role in the INF- $\gamma$  pathway, mediated by *STAT1* activation in M1-macrophages. *MYO1F* regulates M1-polarization in macrophages within intestinal mucosa by stimulating intercellular adhesion and promotes INF- $\gamma$  secretion via hyper-activating *STAT1*, a hub gene of the same immune response module, M76. High INF- $\gamma$  secretion in this pathway is consistent with the significant upregulation of *MYO1F* in an immune-subtype in the INF- $\gamma$  dominant microenvironment by Thorsson et al.<sup>53</sup>. To establish the role of *STAT1* in this pathway within melanoma, we developed a *STAT1*-centered network model from pSKCM, and intersected it with *STAT1* perturbation signature in the *Stat1*-KO mouse model from bone-marrow-derived macrophages (BMM). We observed a significant overlap between the *STAT1*-centered network and *STAT1* perturbation signature from BMM (Supplementary Fig. 3C). *STAT1* expression was correlated with inferred M1-macrophage abundance by CIBERSORT in pSKCM (Spearman  $\rho = 0.59$ ,  $p = 4.28E - 11$ ; Supplementary Fig. 3E). As there is only one tumor with somatic *STAT1* mutation in pSKCM, we do not have the power to determine the relationship between *STAT1* mutation and M1 macrophage abundance. Thus, the *STAT1*-centered network presents the detailed gene interactions in the *STAT1* pathways in macrophages. Given the close interaction between *STAT1* and *MYO1F* (Pearson  $\rho = 0.32$ ,  $p = 9.88E - 4$ ), these observations also support the regulatory roles of *MYO1F* in M1-polarization and *STAT1* activation in macrophages, leading to increased INF- $\gamma$  secretion in the melanoma microenvironment.

The data also suggest that *MYO1F* may play a role in mediating interactions between CD8+ T cells and M1-macrophages underlying activated adaptive immune systems in the primary melanoma microenvironment. *MYO1F* mRNA expression was significantly correlated with INF- $\gamma$  (i.e., *IFNG*) in an independent melanoma bulk transcriptomic dataset (Fig. 2D). In the single-cell transcriptomic data, INF- $\gamma$  was expressed in CD8+ T cells, while *MYO1F* was expressed in CD8+ T cells and M1-macrophages (Fig. 2B). Indeed, co-localization of M1-macrophages and CD8+ T cells alone can induce INF- $\gamma$  production by CD8+ T cells<sup>68</sup>, enhancing their motility and cytotoxicity<sup>69</sup>. Taken together with *MYO1F*'s regulatory role in driving M1-polarization, the correlation between INF- $\gamma$  and *MYO1F* in the bulk RNA-seq datasets may reflect such enhanced cytotoxicity due to *MYO1F*-mediated M1-polarization, which in turn increases the abundance of co-localized M1-macrophages and CD8+ T cells.

High expression of a co-inhibitory T-cell checkpoint, PD-L1, was also found to be associated with a good prognosis in pSKCM (log-rank  $p = 1.32E - 2$ ). Although an increase in PD-L1 was implicated in cytotoxic T-cell inhibition and thereby promoted immune-surveillance escape in solid tumors<sup>49,70</sup>, the prognostic relevance of high PD-L1 expression has been somehow controversial. In a broad spectrum of solid tumors, several studies reported high PD-L1 was associated with worse outcomes<sup>70-72</sup>, whereas high PD-L1 in metastatic melanoma was found to be correlated with higher infiltrating T-cell content and better prognosis<sup>73</sup>. In our study, high PD-L1 may reflect the abundance of *MYO1F*<sup>high</sup> M1-macrophages modulating an INF- $\gamma$  dominant microenvironment and is independent of T-cell suppression. We observed high PD-L1 expression mostly within the *MYO1F*<sup>high</sup> M1-macrophages, indicating they were predominant sources of PD-L1 (Fig. 2B). Interestingly, the correlation between *CD274* (i.e., PD-L1 expressions) and *IFNG* (i.e., INF- $\gamma$  expressions) (Fig. 2D) also suggests high PD-L1 as a marker for the INF- $\gamma$  dominant microenvironment.

Together, the *MYO1F*-*PD-L1*/*PD1*-*IFN* $\gamma$  axis was identified as a prognostic pathway involving overall INF- $\gamma$  production and antitumoral activities of cytotoxic T lymphocytes (Fig. 5A).



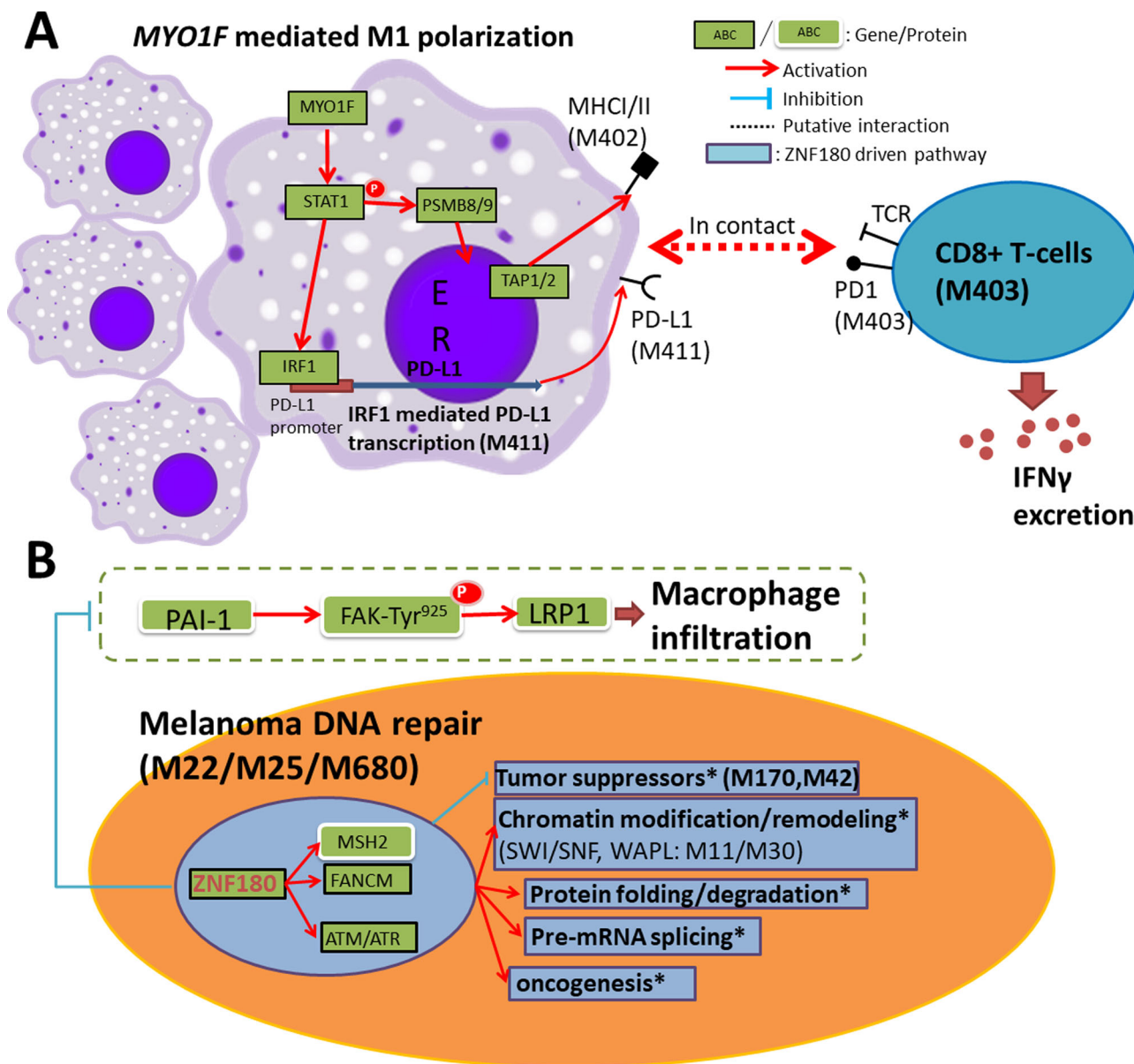


**Fig. 4** Dissection of *ZNF180*-regulated pathways in melanoma. **A, B** Correlations between *ZNF180* expressions (x-axis) and protein expressions (y-axis; **A**: MSH2, **B**: PAI-1) by RPPA. Dot color indicates patients' prognosis in overall survival (death within 3 years of follow-up: red, otherwise yellow). **C** Cell abundance is correlated with *ZNF180* expressions in pSKCM. The relative abundance of 4645 cells with annotated cell types (in x-axis) were inferred across 103 pSKCM samples by cell population mapping (CPM) algorithm in the scBio package. The inferred abundance was then tested for correlation with *ZNF180* expressions in pSKCM. Each dot represents the correlation between the relative abundance of one cell from the scRNA-seq with *ZNF180* expressions in pSKCM. The x-axis is cell types of the respective cells, and the y-axis is  $-\log_{10}(\text{FDR}) \times \text{sign}(\rho)$ . The red horizontal lines mark  $\text{FDR} = 0.05$  threshold, and red rectangles highlight the most distinctively correlated cell types. The border of a boxplot show the lower quantile, median, or upper quantiles while a whisker spans towards the minima and the maxima. **D** Summary of the pathways associated with the modules potentially regulated by *ZNF180*. Enriched signatures in the modules are color-coded.

Among 17 nominated targets from the pSKCM network, *ZNF180* showed the most consistent pro-tumorigenic effect (Fig. 4D). Consistent with the suppression of DNA repair pathways in melanoma cells by the knockdown of *ZNF180*, the module primarily regulated by siZNF180, M25, harbors interactions between DNA repair pathways involving MSH2, and oncogenic pathways involving *PIK3CA*, Rho kinases, and GTPases. Indeed, overexpression of MSH2 in primary melanoma was associated with poor prognosis, likely by enabling tumor cells to proliferate and metastasize as reported<sup>62</sup>. In contrast, several modules upregulated by silencing of *ZNF180* include tumor suppressors such as *BRMS1* (Breast cancer metastasis suppressor 1; within M170)<sup>74</sup> and negative regulator of cell cycle *GADD45-GIP1*<sup>75</sup> as hub genes. In pSKCM, PAI-1 protein expression is negatively correlated with *ZNF180* expression. PAI-1 has been

shown to promote macrophage infiltration in melanoma via phosphorylation of FAK-Tyr<sup>925</sup><sup>76</sup> (Fig. 5B), suggesting *ZNF180* may also regulate immune cell infiltration in primary melanoma.

In conclusion, this study has identified multi-faceted axes of tumor progression and immune evasion of primary melanoma including enhanced CD8+ T-cell cytotoxicity likely via *MYO1F* mediated macrophage M1-polarization and *ZNF180* as a molecular regulator of DNA repair and immune cell infiltration in melanoma cells. The network models have revealed additional regulators that have not been fully characterized in this study. More mechanistic studies (e.g., genetic and epigenetic regulation) using immune-competent models are needed to understand the functional impact of the prioritized regulators of melanoma (e.g., *MYO1F* and *ZNF180*) on tumor progression and immune modulation.



**Fig. 5** Key signaling pathways in primary melanoma captured by the bulk pSKCM and single-cell gene networks. **A** MYO1F mediated M1-polarization of macrophages is associated with a good prognosis. Contact with CD8+ T cells leads to increased IFN $\gamma$  excretion in the tumor microenvironment. **B** ZNF180 overexpression drives DNA repair and downstream melanoma pathways and suppresses macrophage infiltration via decreased PAI-1 protein expression.

## Methods

### Bioinformatics methods

**Gene expression data processing.** Illumina HiSeq RNA Sequencing data, processed by reads per kilobase per million (RPKM) method from TCGA (i.e. RNASeqV2), has been downloaded and comprehensive data quality control has been performed. Primary and metastatic tumor samples were collectively processed by  $\log_2(\text{RPKM} + 1)$  transform, followed by quantile-normalization. From then, we split the data into primary-/metastatic-specific gene expression matrix, corrected for batch effects by the center, platform, and tissue source site (transcription starting site (TSS)) ids from TCGA sample barcodes, and corrected for confounding factors including race, age, and gender by capturing residuals with intercepts from linear regression model by  $\text{lm}()$  function from R software (version 3.4.2). This resulted in 103/353 annotated primary/metastatic tumor tissue samples across 19,047 genes.

**Correlation analysis between methylation and gene expression.** We firstly processed the level 3 DNA methylation data from The Cancer Genome Atlas (TCGA), corresponding to  $M$ -values from Illumina Infinium Human DNA Methylation 450k platform. Upon quantile-normalization, the  $M$ -values were further adjusted for batch variables, including TSS, plate and center ids, and gender and age at diagnosis by taking residuals with intercepts from the linear model, gene expression  $\sim \alpha_1 \text{TSS} + \alpha_2 \text{plate} + \alpha_3 \text{center id} + \alpha_4 \text{gender} + \alpha_5 \text{age} + \alpha_0$ . We identified gene

expressions associated with DNA methylation within the metastatic tumor or primary tumor samples by pairwise Spearman's correlation analysis. *Cis*-regulated genes were defined by applying 1.5 kbps from TSS.

**Extraction of prognostic gene signatures.** We performed a Wald test to systematically evaluate the prognostic significance of stratifying patients by individual gene expressions by the medians. We used the function `survdiff()` from the survival 2.41-3R package<sup>77</sup> to calculate log-rank  $p$  value for overall survival outcome within primary and metastatic tumors separately and applied log-rank  $p$  value  $< 0.05$ . The common prognostic gene signature across primary and metastatic tumors was then defined as the intersection of the respective significant gene signatures and was utilized for ranking the gene modules through enrichment test.

**Co-expression network analysis.** Gene co-expression networks were constructed by multiscale embedded gene co-expression network analysis (MEGENA)<sup>48,78,79</sup>. MEGENA first selects gene pairs with significant correlations ( $\text{FDR} < 0.05$ ). The significant gene pairs are then sorted by absolute Pearson's correlation coefficients and then embedded sequentially by running down the sorted gene pairs to test if each pair can be drawn on the three-dimensional topological sphere without crossing other edges (i.e., planarity test). This edge embedding process is terminated if any of the following three conditions are met: (1) no more edges can be

added, (2) the number of embedded edges reaches the maximal number of edges for the planar network, i.e.,  $3(|V| - 2)$  ( $|V|$  = the number of genes), and (3) the significant list of gene pairs has been tested fully for planarity<sup>80</sup>. The resulting co-expression network belongs to a class of geometrical networks called “planar filtered networks (PFNs)” which can be drawn on the surface of the sphere without any link intersections<sup>80</sup>.

PFN then goes through unsupervised clustering to identify network clusters (i.e., gene modules) at various compactness resolutions by multi-scale clustering analysis (MCA). MCA splits the parent module into child modules by searching for a partition optimizing for Newman’s modularity ( $Q$ )<sup>81</sup>. Then, the compactness of each child module is evaluated by the compactness measure  $\nu$ , defined as below

$$\nu = \frac{\overline{\text{SPD}}}{\log(|V|)^\alpha} \quad (1)$$

where  $V$  is the set of genes in the network,  $\overline{\text{SPD}}$  is the average of shortest path distances of all node pairs, and  $\alpha$  is the resolution parameter. Given that the denominator  $\log(|V|)^\alpha$  is the hallmark of the small-world property represented by the scaling relation  $\overline{\text{SPD}} \sim \log(|V|)$  when  $\alpha = 1$ ,  $\nu$  measures the coherence of a network’s topology<sup>82</sup>. Therefore, a smaller  $\alpha$  identifies more compact clusters. For a given cluster (network), MCA searches through a range of  $\alpha$  values for a resolution scale that leads to more compact clusters than the parent cluster. The resulting gene modules are organized in a hierarchy that represents a multiscale organization of gene modules with different degrees of compactness. The hierarchy captures a series of higher-order relationships (i.e., parent) modules possessing children modules residing within these parent modules. MEGENA identifies more compact children modules within the parent modules<sup>48</sup>.

Candidate key drivers of gene modules are further identified as the nodes with significantly ( $p < 0.05$ ) higher network connectivity than the randomly permuted planar networks<sup>48</sup>. Gene modules are then annotated by the enriched MSigDB signatures and associated with outcomes through the enrichment test of the previously identified gene signatures from the survival and differential expression analyses.

**Integrative network analyses of pSKCM cohort.** To handle large-scale multi-faceted Omics data in primary skin cutaneous melanoma (pSKCM), we developed an integrative network analysis framework to identify and prioritize altered molecular subnetworks in primary melanoma etiology. The approach is anchored in constructing a multiscale co-expression network by MEGENA<sup>48</sup>.

We next generated a set of gene signatures reflecting genomic and epigenomic alterations (Supplementary Fig. 1, IV; co-expression network analysis in Supplementary Methods). The functional impact of epigenetic alterations was determined by gene signatures correlated with CpG sites annotated in the Illumina Infinium Human DNA Methylation 450k platform. If the significant correlation is observed between expressions of gene  $x$  and methylation at a proximal CpG site (less than 1.5 kbps within TSS) with Bonferroni corrected  $p < 0.05$ , the gene is called a *cis*-methylation-correlated gene (*cis*-MCG). If the CpG site is beyond 1.5 kbps from the TSS, the gene is called a *trans*-MCG (Supplementary Data 5). Enrichment of these signatures in a gene module reflects the impact of epigenetic alterations on the module.

From the bulk-based TCGA pSKCM data, MEGENA identified 221 gene modules with at least 50 genes, which were prioritized by enrichment for overall survival associated genes in pSKCM (log-rank  $p < 0.05$ ; extraction of prognostic gene signatures in Supplementary Methods; see Supplementary Data 1 for the gene signatures).

We filtered out the co-expression network modules from MEGENA (see Supplementary Fig. 2A for the scheme) by comparing their connectivity to the normal skin cases from the genotype-tissue expression (GTEx) consortium<sup>83</sup>, before the module ranking process. We hypothesized network connectivity of normal skin driven modules resemble that of a normal skin co-expression network. We tested pSKCM module connectivity against normal skin co-expression network from the GTEx<sup>83</sup> via module differential connectivity analysis<sup>41</sup>. Gene modules with significant gain or loss of overall gene-gene connectivity, compared to normal skin with an overall false-discovery rate (FDR)  $< 0.05$  threshold were retained (see Supplementary Data 2B for the summaries of ranked modules).

Then, the importance of each gene module was determined by enrichment of the prognostic signatures. Specifically, we performed Fisher’s Exact Test to assess overrepresentation of the common prognostic gene signature and utilized  $p$  value to rank order the modules. Given the module hierarchy in the PFN, module ranking by enrichment test results in redundancy in the ranking due to the overlap between each parent module and its child modules. To handle this, we retained only the best-ranked module in each branch of the module hierarchy.

**Analysis of melanoma single-cell RNA sequencing data.** We downloaded the log2 (TPM/10 + 1) gene expression matrix from Gene Expression Omnibus under the accession number, GSE72056<sup>40</sup> (denoted GSE72056). GSE72056 served as a discovery single-cell dataset in combination with the discovery bulk RNA-seq cohort, pSKCM. This dataset was comprised of 1257 malignant cells, 2064 T cells, 525 B cells, 125 macrophages, 65 endothelial cells, 61 cancer associated-fibroblasts and 52 natural killer cells. We characterized these cells in both unsupervised and supervised manners to annotate cell subpopulations.

We utilized the *scan* (1.10.1) R package for the unsupervised clustering analysis of the single-cell RNA-seq data. The unsupervised clustering based on random walk community detection, namely the walk trap algorithm<sup>84</sup>, was applied by constructing the  $k$ -nearest neighbor ( $k$ NN) graph in the first ten PCs. The optimal value  $k$  was determined by detecting the elbow in the entropy curve for  $k$ NN  $\epsilon$   $[2, \sqrt{N_c}]$  ( $N_c$  = number of cells).  $k$ NN = 19 emerged as the optimal result, leading to 15 cell clusters with sizes ranging from 77 cells to 488 cells. Then, each cell cluster’s markers were identified by contrasting each cell cluster with the rest using *limma* (3.36.1) R package with FDR  $< 0.05$  and fold increase  $> 1.2$ .

Cell type inference was performed independently from the clustering results based on the cell type markers identified by Schelker et al. 2017<sup>85</sup>. A decision-tree algorithm implemented in R package *rpart* (4.1-13) was used to train classifiers which were then subsequently applied to the entire set of the cells. Cell type assignment for each cell was further refined by a known cell type most enriched in the  $k$ NN of the cell (see Supplementary Data 4A for cell types in the scRNA-seq data from GSE72056). Compared to the published cell types identified in GSE72056<sup>40</sup>, we achieved an accuracy of 84% and all the cell clusters had distinct cell types. For each cell cluster, we then constructed a cell-cluster specific co-expression network by MEGENA. Lowly expressed genes were filtered out if they were expressed in less than 20 cells and gene pairs with significant Spearman’s correlations (FDR  $< 0.05$ ) were taken as an input for MEGENA<sup>48</sup>.

To validate the findings from GSE72056, we analyzed the scRNA-seq data of 33 melanoma tumors by Jerby-Arnon et al.<sup>47</sup>. We performed the same unsupervised clustering using the  $k$ NN approach, yielding 17 cell clusters (Supplementary Fig. 5A; see Supplementary Data 4B for cell cluster annotation). The cell type of each cell cluster was determined by the marker genes of the key cell types identified through GSE72056 (shown in Supplementary Fig. 5B). The marker genes include *PTPRC* (immune cell marker), *CD3E* (T-cell marker), *CD8A* (CD8+ T-cell marker), *CD19* (B-cell marker), *CD14* (monocyte/macrophage marker), *MITF* (melanoma cell marker), *VWF*, *CDHS* (endothelial cell marker), *COL1A1*, and *FAP* (cancer-associated fibroblasts, CAFs)<sup>85</sup>. In addition, we checked the chemokine markers of macrophage polarization, adapted from Rozer<sup>86</sup> and Duluc et al.<sup>87</sup>. As shown in Supplementary Fig. 6, the cell cluster 17 has the most pronounced expression of M1-macrophage-specific chemokines, hence named as M1-macrophage in Supplementary Fig. 5A.

**Identification of cell subpopulation specificity of a gene.** For a given gene  $i$ , we first identified a list of cells expressing the gene with a threshold TPM <sub>$i$</sub>   $> 0$  and then tested how these cells were enriched for the cells in each inferred cell type (i.e., cell subpopulation) by FET. Multiple-testing corrected FET  $p$  value (i.e., FDR)  $< 0.05$  was used as a threshold to designate a cell type (i.e., cell subpopulation) for the gene.

**Cell composition analysis of pSKCM transcriptome data.** The RPKM data of the pSKCM samples were taken as an input for the CIBERSORT web app at <https://cibersort.stanford.edu><sup>15</sup> with 100 permutations to identify the relative abundance of immune cell populations. LM22 markers derived from purified distinct immune cell populations were used to infer the abundance. The results are provided in Supplementary Data 7A. We also performed ESTIMATE<sup>16</sup> to infer more granular cell compositions including tumor cells, immune and stromal cells to complement CIBERSORT results based on L22 immune markers (provided in Supplementary Data 7B).

However, cell type characteristics in the tumor microenvironment are substantially different from the healthy microenvironment and is disease dependent<sup>67,85</sup>. In this case, leveraging disease matched scRNA-seq has shown improvements in cell type abundance inference<sup>67,85</sup>. To increase the confidence in the inferred results, we used cell type-specific expression from the scRNA-seq data to infer the relative abundance of individual single cells in the bulk pSKCM samples using the CPM algorithm implemented in the scBio R package<sup>67</sup>. As a result, mapping scores were computed between bulk samples and individual cells from the single-cell transcriptomic data (Supplementary Data 7D).

Detailed cell types of some CD3+ T cells and macrophages were not available in the scRNA-seq data. We performed a correlation analysis between the CPM inferred and CIBERSORT inferred abundances as LM22 markers included detailed cell type characteristics to address this issue. Supplementary Fig. 4D (upper panel) shows that CIBERSORT confirmed the detailed cell types correspond to M1-macrophage and RM CD4+ T cells. Indeed, the CD3+ T cells without unassigned T-cell subtypes showed prominent CD4 expression (Supplementary Fig. 4E), supporting that it was a CD4+ T-cell subtype. Henceforth, this cell population is denoted RM CD4+ T cell.

**In-silico validation of gene perturbation signatures.** SKmel147 cells transduced with siRNA against the top three target genes (namely, *ZNF180*, *ZNF347*, and *PPP1R2*) from the screening were selected for RNA-sequencing to identify respective knockdown gene signatures by differential expression analysis. We validated the network structure around a gene subject to a functional perturbation (for instance, gene knockdown or overexpression) by testing enrichment of the respective differentially expressed gene signatures in the network neighborhoods of the perturbed genes. Specifically,  $l$ -layer neighborhood of a gene,  $g$ , is defined as the set of genes whose shortest paths to  $g$  consists of  $l$  edges at most. The network structure is

successfully validated if the BH FDR corrected FET  $p$  value  $< 0.05$  within the  $l$ -layer neighborhood. A network structure with more than 10% of the total number of nodes in the global will not be tested as we are more interested in testing the local structure around a gene.

**Identification of cell populations expressing specific genes.** Per gene  $i$ , we first identified the list of cell types expressing the gene with threshold  $TPM_i > 0$ . Then, the list of expressing cells was tested for enrichments in the inferred cell types by Fisher's Exact Test (FET). FET FDR  $< 0.05$  was used as a threshold to designate a cell type as expressing the gene  $i$ .

**Module conservation analysis between pSKCM and scRNA-seq.** Each gene module in the PFN from the pSKCM was compared against the cell-cluster specific modules from the scRNA-seq data through FET. Two modules are conserved if the Bonferroni corrected FET  $p$  value is smaller than 0.05, the enrichment fold-change is greater than 2, and the overlap includes at least 30% of the genes in each module. Supplementary Data 4C includes detailed information about the module conservation analysis.

## Experimental methods

**Western blotting.** SKMel147 melanoma cells were lysed in ice-cold RIPA buffer supplemented with protease inhibitor cocktail (Roche), centrifuged at 14,000 rpm at 4 °C for 20 min and the supernatant was collected. Protein in the supernatant was estimated by DC protein assay (Bio-rad). Totally, 20  $\mu$ g of protein were resolved by NuPage 4–12% Bis–Tris (Invitrogen) and transferred onto Immobilon-P polyvinylidene fluoride membranes (Millipore). Membranes were washed with distilled water followed by blocking with 5% nonfat dry milk in Tris-buffered saline supplemented with 0.05% Tween-20 (TBST) for 2 h at room temperature. Membranes were washed briefly with TBST then incubated in anti-ZNF180 (1:1000; Sigma SAB1306465), and anti-Tubulin (1:5000; Sigma, T9026) primary antibodies diluted in 5% nonfat dry milk in TBST (0.05% Tween-20) and incubated on a plate shaker for overnight at 4 °C. Membranes were washed for 3 times with TBST followed by incubation with horseradish peroxidase-conjugated anti-mouse or anti-rabbit secondary antibodies (1:5000; Sigma). Membranes were washed three times with TBST then signals were detected using Clarity Western ECL Blotting Substrate (BioRad) and imaged in LICOR Odyssey Fc imaging system.

**Cell lines and cell culture.** Cell line SKMel147 was obtained from Dr. Alan Houghton's laboratory (MSKCC, New York). A375 was purchased from American Type Culture Collection. Cells were cultured in DMEM (Invitrogen) containing 10% (v/v) fetal bovine serum and 1% (v/v) penicillin/streptomycin. Cell lines were maintained in a 5% CO<sub>2</sub> incubator at 37 °C. Cell lines were routinely tested to exclude mycoplasma contamination.

**Reverse transfection of siRNA.** Transfection conditions were optimized for SKMel147 and A375 melanoma cell lines using fluorescein-labeled oligos (Block-IT, Invitrogen). Liposomal transfection complexes with siRNA pools for each of the candidate genes and with NTC (siRNA ON-target plus SMARTpools, Dharmacon, 50 nM) were generated with Lipofectamine 2000 (Invitrogen), following the manufacturer's recommendations. Media was changed after 6 h incubation with liposomal complexes. Forty-eight hours after initiation of transfection, cells were used for RNA extraction, or for proliferation or invasion assays.

**RNA extraction, reverse transcription, and qRT-PCR.** Total RNA was extracted using miRNeasy QIAGEN mini kit according to the manufacturer's protocol. Totally, 500 ng of RNA were reverse transcribed using TaqMan RT reagents (Applied Biosystems) with random hexamers following the manufacturer's recommendations. cDNA was diluted with RNase and DNase free water prior to use in a quantitative real-time polymerase chain reaction (qRT-PCR). GAPDH was used as a housekeeping gene in qRT-PCR. Transcripts were quantified by ABI StepOne Real-Time PCR system (Applied Biosystems) using Power SYBR Green PCR MasterMix and following 2-step cycling parameters: holding stage, 10 min at 95 °C, and cycling stage 40 cycles of 95 °C for 15 s followed by 60 °C for 1 min followed by melting curve stage. The experiment was performed in three technical replicates.

**Cell proliferation assays.** Transfected cells were seeded at  $3 \times 10^3$  cells per well in 96-well plates. The following day (day 0) and every 24 h after (up to 3 days), cells were fixed in 0.1% glutaraldehyde and stored in phosphate-buffered saline (PBS) at 4 °C. Cells were then stained with 0.5% crystal violet, dried, and dissolved with 15% acetic acid. Optical density was read at 590 nm. For normalization and control purposes, cells transfected with NTC (siNTC) were present on each experiment.

**In vitro invasion assay.** Cell invasion was measured using 24-well Fluoroblok inserts (8 mm Becton Dickinson). Optimization was performed for SKMel147 and A375 cell lines to identify assay time length and Matrigel concentration. Briefly, siRNA transfected SKMel147 or A375 cells (40,000 cells per insert) were suspended in the serum-free medium over a Matrigel coating (Becton Dickinson), and a medium supplemented with 10% serum was used as a chemo-attractant. Cells that invaded after 6–8 h were stained in 4  $\mu$ g/ml Calcein AM dye (ThermoFisher) for

1 h and counted in 5 different fields using a fluorescent microscope. For each independent experiment, three replicates per condition were run. The average of cell counts from three replicates per condition was used for plotting results. For control, cells transfected with NTC (siNTC) were present in each experiment. Cell counts for each well were normalized to the mean counts (of replicate wells) for the corresponding condition in the cell input plate to control for cell proliferation effects that may have occurred between initiation of transfection and assay seeding.

**RNA-sequencing of siRNA transfected SKMel147 cells.** Total RNA was extracted 72 h post transfection of siNTC, siZNF180, siZNF347, or siPPP1R2 in SKMel147 cell line in three replicates. RNA was extracted using miRNeasy QIAGEN mini kit. RNA was then processed with Ribo-Zero rRNA Removal Kit (Illumina) to remove rRNA, and further processed into sequencing libraries using Illumina TruSeq Stranded Total RNA Library Prep, following manufacturer's protocol. All libraries were sequenced on Illumina HiSeq2500 (150 M, 50 bp paired-end).

**In vivo validation of protumorigenic effects of ZNF180 in SKMEL147 xenografts.** SKMel147 cells were infected with doxycycline (DOX)-inducible shNTC or shZNF180 lentivirus and selected with 2  $\mu$ g/ml puromycin for 48 h. Cells were induced with 2  $\mu$ g/ml DOX for 3 days, trypsinized, washed with PBS, then suspended in sterile PBS at a concentration of  $2 \times 10^6$  cells per 150  $\mu$ l, and maintained on ice until injection. Immediately before injection, cell aliquots were mixed with Matrigel (Becton Dickinson). Totally, 150  $\mu$ l of cell/matrigel (1:1) suspensions were injected subcutaneously in the right flank of NOD/Shi-scid/IL-2Rgamma null (NSG, Jackson labs #005557) 20-weeks-old male mice ( $n = 12$  per group). Mice were fed DOX-containing food (200 mg/kg weight). When primary tumors were palpable (6 days post injection), length ( $l$ ) and width ( $w$ ) were measured with calipers, 3 times weekly over a period of 13 days. Tumor volume was calculated using the formula  $(l \times w^2)/2$ . Tumor weight was measured at endpoints. Animal experiments were in accordance with guidelines set forth by the Institutional Animal Care and Use Committee (IACUC) of NYU (protocol # S16-00051).

**Viral production.** Totally,  $4 \times 10^6$  HEK293T cells were seeded per 10 cm tissue culture plate and incubated overnight at 37 °C and 5% CO<sub>2</sub>. After seeding, HEK293T were co-transfected with PLKO-Tet-On-shNTC or PLKO-tet-on-shZNF180 (shRNA sequence: CCGGTGTCCTTGTGTGCATCAAAGCTCG AGCTTTGATGCACAACAAGGACATTTTTG) lentiviral expression constructs (12  $\mu$ g), viral packaging plasmid (psPAX2, 8  $\mu$ g), and viral envelope plasmid (pMD2.G, 4  $\mu$ g) using Lipofectamine 2000 (Invitrogen) following manufacturer's recommendations. Viral supernatant was collected and 0.45 mm filtered at 48 h post transfection and stored at 4 °C for short-term use (1–5 days) or –20 °C for long-term storage (5–30 days).

**In vivo xenograft.** SKMel147 cells were infected with DOX-inducible shNTC or shZNF180 lentivirus and selected with 2  $\mu$ g/ml puromycin for 48 h. Cells were induced with 2  $\mu$ g/ml DOX for 3 days, trypsinized, washed with PBS, then suspended in sterile PBS at a concentration of  $2 \times 10^6$  cells per 150  $\mu$ l, and maintained on ice until injection. Immediately before injection, cell aliquots were mixed with Matrigel (Becton Dickinson). Totally, 150  $\mu$ l of cell/matrigel (1:1) suspensions were injected subcutaneously in the right flank of NOD/Shi-scid/IL-2Rgamma null (NSG, Jackson labs #005557) 20-weeks-old male mice ( $n = 12$  per group). Mice were fed DOX-containing food (200 mg/kg weight). When primary tumors were palpable (6 days post injection), length ( $l$ ) and width ( $w$ ) were measured with calipers, 3 times weekly over a period of 13 days. Tumor volume was calculated using the formula  $(l \times w^2)/2$ . Tumor weight was measured at endpoints.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The discovery bulk gene expression data, analyzed as primary melanoma (pSKCM), are available from The Cancer Genome Atlas (TCGA) data portal (<https://portal.gdc.cancer.gov/>). The discovery single-cell RNA sequencing data is available from Gene Expression Omnibus (GEO) under accession number [GSE72056](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE72056), and we downloaded the normalized TPM values (file name: [GSE72056\\_melanoma\\_single\\_cell\\_revised\\_v2.txt.gz](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE72056)). The validation single-cell RNA sequencing data from Jerby-Arnon et al. (2018) was downloaded from [https://singlecell.broadinstitute.org/single\\_cell/study/SCP109/melanoma-immunotherapy-resistance](https://singlecell.broadinstitute.org/single_cell/study/SCP109/melanoma-immunotherapy-resistance). The clinical and gene expression data ( $z$ -score transformed from log-transformed FPKM values) for the validation bulk cohort<sup>46</sup> were downloaded from cbiportal website, [https://www.cbiportal.org/study/summary?id=skcm\\_dfci\\_2015](https://www.cbiportal.org/study/summary?id=skcm_dfci_2015). The RNA-sequencing data from siRNA silencing of ZNF180, ZNF347 and PPP1R2 are available in GEO under accession number [GSE161385](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE161385). In addition, the processed data are available from the corresponding author upon reasonable request. Source data are provided with this paper.

## Code availability

MEGENA software (version 1.3.7) is publicly available as an R package in The Comprehensive R Archive Network (CRAN) (<https://cran.r-project.org/web/packages/>)

[MEGENA/index.html](#)). The development version of MEGENA is available from GitHub repository (<https://github.com/songw01/MEGENA>, <https://doi.org/10.5281/zenodo.4276493>). We utilized 'cit' package for the Causal Inference Test (CIT) of methylation changes to the transcriptome. Additional codes for detailed implementation of MEGENA, CIT, in silico gene-gene interaction validation and other codes are available at [https://github.com/songw01/melanoma\\_network\\_codes](https://github.com/songw01/melanoma_network_codes).

Received: 5 September 2019; Accepted: 21 January 2021;

Published online: 22 February 2021

## References

- Bertolotto, C. Melanoma: from melanocyte to genetic alterations and clinical options. *Science* **2013**, 635203 (2013).
- Cancer Genome Atlas, N. Genomic classification of cutaneous melanoma. *Cell* **161**, 1681–1696 (2015).
- Lo, J. A. & Fisher, D. E. The melanoma revolution: from UV carcinogenesis to a new era in therapeutics. *Science* **346**, 945–949 (2014).
- Kiuru, M. & Busam, K. J. The NF1 gene in tumor syndromes and melanoma. *Lab. Investig.* **97**, 146–157 (2017).
- Willmore-Payne, C., Holden, J. A., Hirschowitz, S. & Layfield, L. J. BRAF and c-kit gene copy number in mutation-positive malignant melanoma. *Hum. Pathol.* **37**, 520–527 (2006).
- Fares, J., Fares, M. Y., Khachfe, H. H., Salhab, H. A. & Fares, Y. Molecular principles of metastasis: a hallmark of cancer revisited. *Signal Transduct. Target Ther.* **5**, 28 (2020).
- Kirstein, J. M., Hague, M. N., McGowan, P. M., Tuck, A. B. & Chambers, A. F. Primary melanoma tumor inhibits metastasis through alterations in systemic hemostasis. *J. Mol. Med.* **94**, 899–910 (2016).
- Nirschl, C. J. et al. IFN $\gamma$ -dependent tissue-immune homeostasis is co-opted in the tumor microenvironment. *Cell* **170**, 127–141 e115 (2017).
- Neagu, M., Constantin, C. & Tanase, C. Immune-related biomarkers for diagnosis/prognosis and therapy monitoring of cutaneous melanoma. *Expert Rev. Mol. Diagn.* **10**, 897–919 (2010).
- Imredi, E. et al. Aquaporin-1 protein expression of the primary tumor may predict cerebral progression of cutaneous melanoma. *Pathol. Oncol. Res.* <https://doi.org/10.1007/s12253-018-0513-6> (2018).
- Mahmoud, F. et al. Immune surveillance in melanoma: from immune attack to melanoma escape and even counterattack. *Cancer Biol. Ther.* **18**, 451–469 (2017).
- Ostroumov, D., Fekete-Drimusz, N., Saborowski, M., Kuhnel, F. & Woller, N. CD4 and CD8 T lymphocyte interplay in controlling tumor growth. *Cell Mol. Life Sci.* **75**, 689–713 (2018).
- Kashani-Sabet, M. Tumor progression by immune evasion in melanoma: role of the programmed cell death-1/programmed cell death-1 ligand 1 interaction. *Cancer* **116**, 1623–1625 (2010).
- Lowney, J. K., Boucher, L. D., Swanson, P. E. & Doherty, G. M. Interferon regulatory factor-1 and -2 expression in human melanoma specimens. *Ann. Surg. Oncol.* **6**, 604–608 (1999).
- Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).
- Yoshihara, K. et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612 (2013).
- Albert, R., Jeong, H. & Barabasi, A. L. Error and attack tolerance of complex networks. *Nature* **406**, 378–382 (2000).
- Barabasi, A. L. & Oltvai, Z. N. Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* **5**, 101–113 (2004).
- Boone, C., Bussey, H. & Andrews, B. J. Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* **8**, 437–449 (2007).
- Butte, A. J. & Kohane, I. S. Mutual information relevance networks: functional clustering using pairwise entropy measurements. In *Proc. Pac. Symp. Biocomput. 2000* (eds Altman, R. B., Dunker, A. K., Hunter, L., Lauderdale, K. & Klein, T. E.) 418–429 (World Scientific, Singapore, 1999).
- Chen, Y. et al. Variations in DNA elucidate molecular networks that cause disease. *Nature* **452**, 429–435 (2008).
- Emilsson, V. et al. Genetics of gene expression and its effect on disease. *Nature* **452**, 423–428 (2008).
- Gardner, T. S., di Bernardo, D., Lorenz, D. & Collins, J. J. Inferring genetic networks and identifying compound mode of action via expression profiling. *Science* **301**, 102–105 (2003).
- Gargalovic, P. S. et al. Identification of inflammatory gene modules based on variations of human endothelial cell responses to oxidized lipids. *Proc. Natl Acad. Sci. USA* **103**, 12741–12746 (2006).
- Gavin, A. C. et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
- Horvath, S. et al. Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc. Natl Acad. Sci. USA* **103**, 17402–17407 (2006).
- Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. & Barabasi, A. L. The large-scale organization of metabolic networks. *Nature* **407**, 651–654 (2000).
- Lee, I., Date, S. V., Adai, A. T. & Marcotte, E. M. A probabilistic functional network of yeast genes. *Science* **306**, 1555–1558 (2004).
- Schadt, E. E. Molecular networks as sensors and drivers of common human diseases. *Nature* **461**, 218–223 (2009).
- Schadt, E. E. et al. An integrative genomics approach to infer causal associations between gene expression and disease. *Nat. Genet.* **37**, 710–717 (2005).
- Schadt, E. E. et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6**, e107 (2008).
- Stuart, J. M., Segal, E., Koller, D. & Kim, S. K. A gene-coexpression network for global discovery of conserved genetic modules. *Science* **302**, 249–255 (2003).
- Tran, L. M. et al. Inferring causal genomic alterations in breast cancer using gene expression data. *BMC Syst. Biol.* **5**, 121 (2011).
- Tu, Z. et al. Integrating siRNA and protein-protein interaction data to identify an expanded insulin signaling network. *Genome Res.* **19**, 1057–1067 (2009).
- Wang, I. M. et al. Systems analysis of eleven rodent disease models reveals an inflammatory signature and key drivers. *Mol. Syst. Biol.* **8**, 594 (2012).
- Yang, X. et al. Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver. *Genome Res* **20**, 1020–1036 (2010).
- Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, Article17 (2005).
- Zhu, J. et al. Stitching together multiple data dimensions reveals interacting metabolomic and transcriptomic networks that modulate cell regulation. *PLoS Biol.* **10**, e1001301 (2012).
- Zhu, J. et al. Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. *Nat. Genet.* **40**, 854–861 (2008).
- Tirosh, I. et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendrogloma. *Nature* **539**, 309–313 (2016).
- Zhang, B. et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* **153**, 707–720 (2013).
- Wang, M. et al. Transformative network modeling of multi-omics data reveals detailed circuits, key regulators, and potential therapeutics for Alzheimer's disease. *Neuron* **S0896-6273**, 8 (2020).
- Choi, H., Song, W. M., Wang, M., Sram, R. J. & Zhang, B. Benzo[a]pyrene is associated with dysregulated myelo-lymphoid hematopoiesis in asthmatic children. *Environ. Int.* **128**, 218–232 (2019).
- Katsyv, I. et al. EPRS is a critical regulator of cell proliferation and estrogen signaling in ER+ breast cancer. *Oncotarget* **7**, 69592–69605 (2016).
- Song, W. M. et al. Multiscale network analysis reveals molecular mechanisms and key regulators of the tumor microenvironment in gastric cancer. *Int. J. Cancer* <https://doi.org/10.1002/ijc.32643> (2019).
- Van Allen, E. M. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).
- Jerby-Arnon, L. et al. A cancer cell program promotes T cell exclusion and resistance to checkpoint blockade. *Cell* **175**, 984–997 e924 (2018).
- Song, W. M. & Zhang, B. Multiscale embedded gene co-expression network analysis. *PLoS Comput. Biol.* **11**, e1004574 (2015).
- Garcia-Diaz, A. et al. Interferon receptor signaling pathways regulating PD-L1 and PD-L2 expression. *Cell Rep.* **19**, 1189–1201 (2017).
- Quigley, M. et al. Transcriptional analysis of HIV-specific CD8+ T cells shows that PD-1 inhibits T cell function by upregulating BATF. *Nat. Med.* **16**, 1147–1151 (2010).
- Su, X. et al. Interferon-gamma regulates cellular metabolism and mRNA translation to potentiate macrophage activation. *Nat. Immunol.* **16**, 838–849 (2015).
- Qian, J. et al. The IFN-gamma/PD-L1 axis between T cells and tumor microenvironment: hints for glioma anti-PD-1/PD-L1 therapy. *J. Neuroinflamm.* **15**, 290 (2018).
- Thorsson, V. et al. The immune landscape of cancer. *Immunity* **48**, 812–830 e814 (2018).
- Piedra-Quintero, Z. L. et al. Myosin 1F regulates M1-polarization by stimulating intercellular adhesion in macrophages. *Front. Immunol.* **9**, 3118 (2018).
- Wang, Z. et al. Extraction and analysis of signatures from the Gene Expression Omnibus by the crowd. *Nat. Commun.* **7**, 12846 (2016).
- Cowper, A. E., Caceres, J. F., Mayeda, A. & Sreaton, G. R. Serine-arginine (SR) protein-like factors that antagonize authentic SR proteins and regulate alternative splicing. *J. Biol. Chem.* **276**, 48908–48914 (2001).
- Rappsilber, J., Ryder, U., Lamond, A. I. & Mann, M. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **12**, 1231–1245 (2002).

58. Grandi, P. et al. Nup93, a vertebrate homologue of yeast Nic96p, forms a complex with a novel 205-kDa protein and is required for correct nuclear pore assembly. *Mol. Biol. Cell* **8**, 2017–2038 (1997).
59. Kataoka, N., Bachorik, J. L. & Dreyfuss, G. Transportin-SR, a nuclear import receptor for SR proteins. *J. Cell Biol.* **145**, 1145–1152 (1999).
60. Xue, Y., Li, Y., Guo, R., Ling, C. & Wang, W. FANCM of the Fanconi anemia core complex is required for both monoubiquitination and DNA repair. *Hum. Mol. Genet.* **17**, 1641–1652 (2008).
61. Wang, R., Cukerman, E., Heng, H. H. & Liew, C. C. Identification of a locus of zinc finger genes in human chromosome 19q13.1-q13.3 region by fluorescence in situ hybridization. *Somat Cell Mol. Genet.* **22**, 245–248 (1996).
62. Kauffmann, A. et al. High expression of DNA repair pathways is associated with metastasis in melanoma patients. *Oncogene* **27**, 565–573 (2008).
63. Cowley, G. S. et al. Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies. *Sci. Data* **1**, 140035 (2014).
64. Szklarczyk, D. et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–D452 (2015).
65. Wei, L., Surma, M., Shi, S., Lambert-Cheatham, N. & Shi, J. Novel Insights into the roles of Rho Kinase in cancer. *Arch. Immunol. Ther. Exp.* **64**, 259–278 (2016).
66. Samuels, Y. & Waldman, T. Oncogenic mutations of PIK3CA in human cancers. *Curr. Top. Microbiol. Immunol.* **347**, 21–41 (2010).
67. Frishberg, A. et al. Cell composition analysis of bulk genomics using single-cell data. *Nat. Methods* **16**, 327–332 (2019).
68. Ahmed, F., Ibrahim, A., Cooper, C. L., Kumar, A. & Crawley, A. M. Chronic hepatitis C virus infection impairs M1 macrophage differentiation and contributes to CD8(+) T-cell dysfunction. *Cells* <https://doi.org/10.3390/cells8040374> (2019).
69. Bhat, P., Leggatt, G., Waterhouse, N. & Frazer, I. H. Interferon-gamma derived from cytotoxic lymphocytes directly enhances their motility and cytotoxicity. *Cell Death Dis.* **8**, e2836 (2017).
70. Wang, Q., Liu, F. & Liu, L. Prognostic significance of PD-L1 in solid tumor: an updated meta-analysis. *Medicine* **96**, e6369 (2017).
71. Xiang, X. et al. Prognostic value of PD-L1 expression in patients with primary solid tumors. *Oncotarget* **9**, 5058–5072 (2018).
72. Li, Y. et al. The prognostic and clinicopathological roles of PD-L1 expression in colorectal cancer: a systematic review and meta-analysis. *Front. Pharm.* **10**, 139 (2019).
73. Kluger, H. M. et al. Characterization of PD-L1 expression and associated T-cell infiltrates in metastatic melanoma samples from variable anatomic sites. *Clin. Cancer Res.* **21**, 3052–3060 (2015).
74. Meehan, W. J. et al. Breast cancer metastasis suppressor 1 (BRMS1) forms complexes with retinoblastoma-binding protein 1 (RBP1) and the mSin3 histone deacetylase complex and represses transcription. *J. Biol. Chem.* **279**, 1562–1569 (2004).
75. Nakayama, K., Nakayama, N., Wang, T. L. & Shih Ie, M. NAC-1 controls cell growth and survival by repressing transcription of Gadd45/GIP1, a candidate tumor suppressor. *Cancer Res.* **67**, 8058–8064 (2007).
76. Thapa, B., Koo, B. H., Kim, Y. H., Kwon, H. J. & Kim, D. S. Plasminogen activator inhibitor-1 regulates infiltration of macrophages into melanoma via phosphorylation of FAK-Tyr(9)(2)(5). *Biochem. Biophys. Res. Commun.* **450**, 1696–1701 (2014).
77. A Package for Survival Analysis in S v. 2.38 (2015).
78. Song, W.-M., Di Matteo, T. & Aste, T. Building complex networks with Platonic solids. *Phys. Rev. E* **85**, 046115 (2012).
79. Song, W.-M., Di Matteo, T. & Aste, T. Nested hierarchies in planar graphs. *Discret. Appl. Math.* **159**, 2135–2146 (2011).
80. Tumminello, M., Aste, T., Di Matteo, T. & Mantegna, R. N. A tool for filtering information in complex systems. *Proc. Natl Acad. Sci. USA* **102**, 10421–10426 (2005).
81. Newman, M. E. Modularity and community structure in networks. *Proc. Natl Acad. Sci. USA* **103**, 8577–8582 (2006).
82. Song, W. M., Di Matteo, T. & Aste, T. Building complex networks with Platonic solids. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* **85**, 046115 (2012).
83. Consortium, G. T. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
84. Pons, P. & Latapy, M. Computing communities in large networks using random walks. *Lect. Notes Comput. Sci.* **3733**, 284–293 (2005).
85. Schelker, M. et al. Estimation of immune cell content in tumour tissue using single-cell RNA-seq data. *Nat. Commun.* **8**, 2032 (2017).
86. Roszer, T. Understanding the mysterious M2 macrophage through activation markers and effector mechanisms. *Mediators Inflamm.* **2015**, 816460 (2015).
87. Duluc, D. et al. Tumor-associated leukemia inhibitory factor and IL-6 skew monocyte differentiation into tumor-associated macrophage-like cells. *Blood* **110**, 4319–4330 (2007).

## Acknowledgements

This research was supported by the Mount Sinai seed fund and NIH funds: U01AG046170 (PI Zhang), P01CA206980 (PI Berwick), and 5P50CA225450 (PI Osman).

## Author contributions

E.H. and B.Z. conceived and designed the study. W.M.S. performed the multiscale network analysis. M.W. and X.Z. conducted the data analysis. E.H. and L.M. designed the validation experiments. P.A., R.V.L., and B.F.-C. performed the validation experiments. W.M.S., P.A., E.H., and B.Z. wrote the paper. All the authors reviewed and revised the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-21457-0>.

**Correspondence** and requests for materials should be addressed to E.H. or B.Z.

**Peer review information** *Nature Communications* thanks Rutao Cui, Carla Robles-Espinoza, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021