# HAND1 and BARX1 act as transcriptional and anatomic determinants of malignancy in gastrointestinal stromal tumor

**Matthew L. Hemming**[a,b], **Shannon Coy**[c], **Jia-Ren Lin**[d,e], **Jessica L. Andersen**[b], **Joanna Przybyl**[f], **Emanuele Mazzola**[g], **Amr H. Abdelhamid Ahmed**[a,b], **Matt van de Rijn**[f], **Peter K. Sorger**[d,e,h], **Scott A. Armstrong**[i], **George D. Demetri**[a,b,h], **Sandro Santagata**[c,d,h]

[a]Department of Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts, USA.

[b]Sarcoma Center, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, USA.

[c]Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA.

[d]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, USA.

[e]Laboratory of Systems Pharmacology, Harvard Medical School, Boston, Massachusetts, USA.

[f]Stanford University, Stanford, CA, USA.

[g]Department of Data Science, Dana-Farber Cancer Institute and Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA.

[h]Ludwig Center at Harvard, Boston, Massachusetts, USA.

[i]Department of Pediatric Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts, USA.

## Abstract

**Purpose:** Gastrointestinal stromal tumor (GIST) arises from interstitial cells of Cajal (ICC) or their precursors, which are present throughout the gastrointestinal tract. While gastric GIST is commonly indolent and small intestine GIST more aggressive, a molecular understanding of disease behavior would inform therapy decisions in GIST. Although a core transcription factor (TF) network is conserved across GIST, accessory TFs HAND1 and BARX1 are expressed in a disease state-specific pattern. Here, we characterize two divergent transcriptional programs maintained by HAND1 and BARX1, and evaluate their association with clinical outcomes.

**Experimental Design:** We evaluated RNA-seq and TF chromatin immunoprecipitation with sequencing (ChIP-seq) in GIST samples and cultured cells for transcriptional programs associated with HAND1 and BARX1. Multiplexed tissue-based cyclic immunofluorescence (CyCIF) and

**Corresponding Authors**: Matthew L. Hemming, Dana-Farber Cancer Institute, 450 Brookline Ave, Boston, Massachusetts 02215. Phone: 617-632-5204; mhemming@partners.org. Sandro Santagata, Brigham and Women's Hospital, 60 Fenwood Road, Hale BTM 8002P, Boston, Massachusetts 02115. Phone: 617-525-5686; ssantagata@bics.bwh.harvard.edu.

immunohistochemistry evaluated tissue and cell-level expression of TFs and their association with clinical factors.

**Results:** We show that HAND1 is expressed in aggressive GIST, modulating *KIT* and core TF expression and supporting proliferative cellular programs. In contrast, BARX1 is expressed in indolent and micro-GISTs. HAND1 and BARX1 expression were superior predictors of relapse-free survival, as compared to standard risk stratification, and they predict progression-free survival on imatinib. Reflecting the developmental origins of accessory TF programs, HAND1 was expressed solely in small intestine ICCs, while BARX1 expression was restricted to gastric ICCs.

**Conclusions:** Our results define anatomic and transcriptional determinants of GIST and molecular origins of clinical phenotypes. Assessment of HAND1 and BARX1 expression in GIST may provide prognostic information and improve clinical decisions on the administration of adjuvant therapy.

## INTRODUCTION

Gastrointestinal stromal tumor (GIST) is the most common soft tissue sarcoma, with population-based annual incidence rates of approximately 1 per 100,000 (1,2). GIST arises from transformation of interstitial cells of Cajal (ICC) or their cellular progenitors (3), which normally act as pacemaker cells regulating gut motility. In at least 85% of cases, GIST lesions harbor activating mutations in one of the related type III receptor tyrosine kinases (RTKs) *KIT* or *PDGFRA*, with ligand-independent kinase signaling mediating oncogenesis (4). Other less common genetic variants can also give rise to GIST, such as inactivating mutations in genes encoding components of the succinate dehydrogenase (SDH) complex or *neurofibromatosis type 1* (*NF1*), which have characteristically unique disease presentations and a more indolent clinical course (5,6).

Following the identification of *KIT* mutations as drivers of GIST (7), the tyrosine kinase inhibitor (TKI) imatinib was found to inhibit KIT signaling and effectively treat patients with GIST (8,9). With the advent of imatinib, durable disease control measured in years can be achieved in nearly 80% of patients with metastatic GIST (10). The success of imatinib has led to its regulatory approval for use as adjuvant therapy for primary resected GIST at high risk of recurrence, as well as in neoadjuvant (pre-operative) therapy to shrink tumors for which surgery would lead to unacceptable risk or morbidity (11–14). In primary resected GIST, risk of recurrence is defined by characteristics such as anatomic site of origin, size and mitotic rate, but more reliable predictive markers would be of clinical benefit (15). Aside from GIST mutational subtype, no molecular markers are clinically available to identify those patients most likely to benefit from receiving adjuvant imatinib. Particularly in cases in which tumors are labeled as Intermediate Risk, when there are borderline assessments of mitotic index, or when neoadjuvant imatinib is administered and confounds mitotic analyses of the surgical specimen, novel molecular markers would be very useful in risk stratification. Further, there are few prognostic tools to guide clinical management in the setting of metastatic disease and to support the rational development of new therapeutics. The lack of clinically useful molecular markers is related to a limited understanding of the biological factors that influence GIST oncogenesis and inter-individual variation in the malignant phenotype.

Core transcription factor (TF) circuits establish and maintain normal lineage-restricted cellular development and function throughout biological systems (16,17). Cancer exploits the core TF networks of precursor cells, either through amplifying or modifying lineage-specific transcriptional programs (18–20). We have previously mapped transcriptionally active chromatin and transcription factor DNA occupancy in GIST and characterized the enhancer and transcriptional landscape of GIST tissue samples and cell lines (21). Using this information, we identified a TF network supporting the GIST gene expression program. Further, through comparative analysis of localized and metastatic tumors, we identified two accessory TFs, HAND1 and BARX1, which are expressed in a pattern specific to disease state. Here, we characterize the transcriptional programs driven by these two accessory TFs, describe clinical, mutational and functional phenotypes in which they are engaged, use high-dimensional tissue imaging to identify high-risk cellular populations, and explore the anatomic and developmental origins of the programs that characterize GIST pathogenesis. Our findings suggest a new molecular subclassification of GIST that justifies further study to inform basic, translational and clinical studies.

## MATERIALS AND METHODS

### Cell Culture and Virus Production.

All cell lines tested negative for mycoplasma infection on routine surveillance (MycoAlert, Lonza Bioscience). Human embryonic kidney (HEK) 293FT (Thermo Fisher Scientific Cat# R70007, RRID: CVCL_6911) and the GIST cell line GIST-T1 (Cosmo Bio Cat# PMC-GIST01-COS, RRID:CVCL_4976; KIT mutation in exon 11    560–578) were cultured in Dulbecco's modified Eagle's medium containing 10% FBS, 2 mM L-glutamine, 100 mg/ml penicillin, and 100 mg/ml streptomycin. GIST-T1 cell line identity was confirmed by sequencing *KIT* exons to confirm the expected mutation. Cell lines were thawed from original or derived stocks and used in the described experiments within approximately 3 months. Transfections were performed with X-tremeGene (Roche). Lentiviral production was performed as previously described (22). Briefly, 293FT cells were cotransfected with pMD2.G (Addgene #12259), psPAX2 (Addgene #12260) and the lentiviral expression plasmid. Viral supernatant was collected at approximately 72 h and debris removed by centrifugation at $1,000g$ for 5 min. Cells were transduced with viral supernatant and polybrene at 8 μg/mL by spinoculation at $680g$ for 60 min. For growth over time assays, $15 \times 10^3$ cells were dispensed per well in a 96 well plate and cell count performed approximately twice per week on a Guava easyCyte Flow Cytometer (EMD Millipore).

### Cloning and CRISPR.

Cell lines stably expressing a human codon-optimized *Streptococcus pyogenes* Cas9 (Addgene #73310) were generated by viral transduction. CRISPR single-guide RNAs (sgRNAs) targeting HAND1, ETV1 or control constructs have been previously described (21). The BARX1 lentiviral expression vector was synthesized with codon optimization (Twist Bioscience). The Dependency Map (DepMap) portal data was accessed through depmap.org (23), utilizing the CRISPR (Avana) Public 20Q3 release.

### RNA-seq.

GIST-T1 cells were transduced with the indicated sgRNAs and incubated for 5 days prior to collection. Total RNA was isolated using an RNeasy Plus Kit (Qiagen), and concentration measured by Nanodrop (Thermo Fisher Scientific) and quality by Bioanalyzer (Agilent). Library preparation was performed using the NEBNext Ultra II non-stranded library prep kit (New England Biolabs). Paired-end 150 bp sequencing was performed on a NovaSeq 6000 (Illumina). All GIST tumor sequencing data has been previously published (21,24–26). RNA-seq data were aligned to hg19 using STAR (27) with expression quantification using Cufflinks (RRID:SCR_014597) (28) to generate gene expression values in fragments per kilobase of transcript per million mapped reads (FPKM) units. EdgeR (RRID:SCR_012802) was used for differential expression analysis of tumor sequencing data (29). Merging of ChIP and differential expression data was performed using binding and expression target analysis (BETA, RRID:SCR_005396) (30). Gene set enrichment analysis (GSEA, RRID:SCR_003199) (31) was performed using Hallmark gene lists in the Molecular Signatures Database (software.broadinstitute.org/gsea/).

### ChIP-seq and ATAC-seq.

All Chip-Seq and ATAC-seq data were aligned to the human reference genome assembly hg19 using Bowtie2 (RRID:SCR_005476) (32). Normalized read density was calculated using Bamliquidator (version 1.0) read density calculator. Aligned reads were extended by 200 bp and the density of reads per base pair was calculated. In each region, the density of reads was normalized to the total number of million mapped reads, generating read density in units of reads per million mapped reads per bp (rpm/bp). Individual ChIP-seq track displays were generated using bamplot (github.com/linlabbcm).

### Tumor Samples, Immunohistochemistry and CyCIF.

Formalin fixed and paraffin embedded (FFPE) archival GIST samples were obtained under an institutional review board (IRB) approved protocol at Brigham and Women's Hospital/ Dana-Farber Cancer Institute or Stanford University. The majority of tumor samples were evaluated as components of a tissue microarray. Where available, pathologic and molecular features of tumors and clinical outcomes were compiled. For relapse-free survival assessments, all included patients received care before the regulatory approval of imatinib as a standard adjuvant therapy option. Normal gastrointestinal tissues were retrieved from the archives of Brigham and Women's Hospital with IRB approval as part of a discarded/excess tissue protocol.

FFPE sections were de-paraffinized, dehydrated and endogenous peroxidase activity blocked. Antigen retrieval was performed in a pressure cooker in citrate buffer (S1699, Dako) at 123°C at 15 PSI for 45 s. Slides were incubated with HAND1 antibody (1:300, OriGene Cat# TA502671, RRID:AB_11125431) or BARX1 antibody (1:100, Atlas Antibodies Cat# HPA055858, RRID:AB_2682947) for 45 min, washed, then incubated with either Labeled Polymer-HRP anti-mouse secondary antibody (K4007, Dako) for HAND1 or Post Primary (Leica Novolink) followed by Novolink Polymer Detection System (RE7150-K, Leica) for BARX1. Slides were then incubated with Dako DAB+ solution (K3468, Dako) and counterstained with hematoxylin. Additional antibodies used for IHC include PDGFRA

(1:200, Cell Signaling Technology Cat# 5241, RRID:AB_10692773), DOG-1 (1:50, Leica Biosystems Cat# NCL-L-DOG-1, RRID:AB_10555293), SDHB (1:250, Abcam Cat# ab14714, RRID:AB_301432), c-KIT (1:250, Agilent Cat# A4502, RRID:AB_2335702), PLAGL1 (1:400, Thermo Fisher Scientific Cat# MA5–31853, RRID:AB_2787476) and HOXC10 (1:300, Thermo Fisher Scientific Cat# PA5–31078, RRID:AB_2548552). For scoring cases, we used a visual scoring system including relative intensity of signal (0–3 scale) and the percent of cells positive (in 5% increments). A positive case was defined as 1+ scoring in 5% or more of tumor cells in two of three cores.

c-KIT (brown) / HAND1 (red) double IHC was performed by incubating sections with c-KIT antibody (1:150, A4502, Dako) for 45 min, followed by Labeled Polymer-HRP anti-rabbit secondary antibody (K4011, Dako) and then DAKO DAB+ solution. Antigen retrieval was then performed in a pressure cooker in citrate buffer (S1699, Dako) at 123°C at 15 PSI for 45 s followed by incubation with HAND1 at (1:300, TA502671, Origene) and then incubation with AP Polymer anti-mouse secondary antibody (Ultra Vision LP, TL-125-AP, Thermo Scientific), followed by incubation with Alkaline Phosphatase (Red Substrate Kit, SK-5105, Vector). Slides were counterstained with Hematoxylin. c-KIT (red) / BARX1 (brown) double IHC was performed by incubating sections with c-KIT antibody (1:150, A4512, Dako) for 45 min, followed by AP Polymer anti-mouse secondary antibody (Ultra Vision LP, TL-125-AP, Thermo Scientific) and then incubated with Alkaline Phosphatase (Red Substrate Kit, SK-5105, Vector). Antigen retrieval was then performed in a pressure cooker in citrate buffer (S1699, Dako) at 123°C at 15 PSI for 45 s followed by incubation with Novolink Polymer Detection Protein Block (RE7150-K, Leica), BARX1 at 1:100 (HPA055858, Sigma) for 45 min, followed by Post Primary and then Novolink Polymer Detection System (RE7150-K, Leica) and developed with Dako DAB+ solution (K3468, Dako) and counterstained with hematoxylin. Differentiation of KIT-positive ICC from mast cells in normal gastrointestinal tissues was supported by cell morphology and tissue localization, with ICC favored in spindled cells with ovoid nuclei and prominent cytoplasmic processes, which are readily identifiable along the myenteric plexus in most specimens, while mast cells exhibit round nuclei and cell bodies with no distinct cytoplasmic processes.

Tissue-based cyclic immunofluorescence (t-CyCIF) consisted of iterative cycles of antibody incubation, imaging, and fluorophore inactivation, and was performed on FFPE specimens as previously described (33,34) using antibodies listed (Table S2). c-KIT signal using this protocol was non-specific and inadequate to detect differences between GIST subsets. Slides were initially prepared as described above for HAND1 and BARX1 IHC. Image acquisition was performed with a RareCyte CyteFinder Slide Scanning Fluorescence Microscope. Photobleaching was performed with a solution of 4.5% $H_2O_2$ and 20mM NaOH in 1X PBS and incubation under a light emitting diode (LED) for 2 h at room temperature. Image processing was performed as previously described to define cell states and identify mesoscale neighborhood information.

## Immunoblotting.

Cells were lysed in RIPA buffer containing protease inhibitor cocktail (Roche) and centrifuged at 14,000*g* for 10 min to remove genomic DNA and debris. Protein concentrations were determined using a bicinchoninic acid-based assay (Pierce Biotechnology). Protein samples were subjected to SDS-PAGE and Western blotting with the following antibodies: BARX1 (1:500, Thermo Fisher Scientific Cat# PA5–68362, RRID:AB_2688478) and ERK (1:2,000; Cell Signaling Technology Cat# 9107, RRID:AB_10695739). Western blots were probed with anti-mouse or anti-rabbit secondary antibodies and detected using the Odyssey CLx infrared imaging system (LI-COR Biosciences). Immunoblots shown are representative of at least three independent experiments.

## Statistical analysis.

Center values, error bars, *P*-value cutoffs, number of replicates and statistical tests are identified in the corresponding figure legends. Replicates represent separate tumors where indicated, or in experiments using the same cell line the replicates represent separate treatments. Error bars are shown for all data points with replicates as a measure of variation within each group. Box plot elements include the center line representing the median, the box representing the upper and lower quartiles, and the whiskers plotted using the Tukey method. Samples sizes were not predetermined. The investigator scoring IHC cases was blinded to clinical information and outcome assessment during data acquisition. Kaplan-Meier analysis of relapse-free survival was calculated from the date of initial diagnosis to the date of first metastasis, relapse, last follow-up or death. Progression-free survival was calculated from the date of imatinib initiation as first line therapy for metastatic GIST until documented disease progression. Hazard ratios and *P*-values were calculated using the logrank test.

## Data and Materials Availability.

Novel RNA-seq data is available through the GEO Publication Reference ID GSE151323. Additional RNA-seq and ChIP-seq data sets analyzed in this study include GSE71119 (24), GSE95864 (21), GSE107447 (25) and PRJNA521803 (26).

# RESULTS

## HAND1 and BARX1 expression are associated with distinct transcriptional programs.

Previously, we used multiple methods to identify and characterize a core group of TFs in GIST tissue samples and cell lines; these methods included chromatin immunoprecipitation with sequencing (ChIP-seq), assay for transposase-accessible chromatin using sequencing (ATAC-seq) and RNA-seq. These core TFs cooperatively interact in a regulatory network that is shared across GIST mutational subtypes and clinical phenotypes (21). Accessory TFs HAND1 and BARX1, however, were expressed in disease-state specific patterns predictive of metastasis-free survival. To better understand the distinct transcriptional programs driven by these two accessory TFs and their relationship to clinical factors, we characterized gene expression datasets from 35 GIST clinical samples for which clinical annotation was

available. We used differential expression analysis of two clinically distinct subgroups to compare transcription in RTK-driven localized and never-recurrent samples (the 'No Recurrence' group) and localized and later-recurrent or metastatic GIST (the 'Metastatic/ Recurrent' group) (Fig. 1A, Table S1). We found that the overall gene expression program was highly similar across all tumors regardless of clinical status, with <5% of all expressed genes ($n$=10,000) exhibiting differential expression between the two groups. Among all differentially expressed transcripts, only the TFs BARX1 and PLAGL1 were enriched in 'No Recurrence' GIST, whereas HAND1 and HOXC10 were the only TFs enriched in 'Metastatic/Recurrent' GIST. BARX1 and HAND1 showed the greatest differences in expression between groups, with >100-fold enrichment of each transcript in their respective group. Moreover, HAND1 expression was commonly and exclusively found in KIT-mutant tumors in the 'Metastatic/Recurrent' group. In contrast, BARX1 was expressed at high levels in KIT-mutant GIST in the 'No Recurrence' localized group as well as in PDGFRA-mutant GIST and SDH-deficient GIST (Fig. 1B).

Recently, an independent cohort of RNA-seq data from 75 GIST specimens was reported (35). Combining this publically available dataset with our clinically annotated samples, we found a strong anti-correlation between HAND1 and BARX1 expression values, with high levels of expression of one or the other TF observed in all but a subset of samples (total $n$=110; $n$=15 with FPKM <50 for both TFs) (Fig. 1C). Dimensionality reduction using principal component analysis (PCA) of the public RNA-seq data from all samples that express HAND1 ($n$=21) or BARX1 ($n$=39) revealed that tumor clustering was correlated with accessory TF expression (Fig. 1D). We then used gene set enrichment analysis (GSEA) to identify the gene expression programs associated with these disease state-specific accessory TFs (i.e., analyzing 'No Recurrence' versus 'Metastatic/Recurrent' groups in the clinically annotated cohort, and BARX1+ versus HAND1+ GIST in the independent published cohort). The 'Metastatic/Recurrent' and the HAND1+ groups showed significant enrichment in Hallmark signatures including G2M Checkpoint and MTORC1 Signaling gene sets, while the 'No Recurrence' and BARX1+ groups both showed enrichment in the Epithelial Mesenchymal Transition (EMT) gene set (Fig. 1E). The G2M Checkpoint was among the most enriched gene sets in the HAND1-expressing groups, suggestive of higher proliferation and a more aggressive phenotype (Fig. 1F–G). Taken together, data from these independent cohorts indicate that unique gene expression programs are associated with accessory TF expression, and that HAND1-positive tumors engage cell growth-associated signal transduction pathways and proliferation programs.

### HAND1 regulates the core GIST TF network and supports *KIT* gene expression.

To explore the mechanism behind the association of HAND1 with more aggressive disease, we used a CRISPR-based system to inactivate HAND1, or the core GIST TF ETV1 (36) as comparator, in the KIT-dependent GIST cell line GIST-T1. Loss of either HAND1 (sgHAND1) or ETV1 (sgETV1) decreased GIST cell proliferation in a 21-day assay as compared to a luciferase guide RNA control (sgLuc) (Fig. 2A). We performed RNA-seq on cells collected prior to the onset of proliferative effects (i.e., 5 days of sgRNA exposure). Compared to control (sgLuc), cells treated with sgHAND1 and sgETV1 showed global changes in gene expression (Fig. 2B, Fig. S1A). Following acute loss of either HAND1 or

ETV1, Hallmark gene sets were enriched for the EMT signature; in comparison, cells transduced with sgLuc were enriched for MTORC1 Signaling (Fig. 2C–E). No significant differences in the G2M Checkpoint gene set were noted at this early time point. To combine data on genes differentially expressed due to loss of HAND1 with known genomic regions of HAND1 binding identified by ChIP-seq, we utilized binding and expression target analysis (BETA) (21,30). BETA showed that HAND1 loss of function was responsible both for significant downregulation or upregulation of HAND1-associated genes (Fig. 2F). A waterfall plot of the significantly altered genes following sgHAND1 treatment showed enrichment for members of the core GIST TF network among both negatively and positively regulated genes, as well as other GIST-associated genes including KIT and negative regulators of KIT signaling (Fig. 2G). BETA analysis of sgETV1 transduced cells revealed similar global dysregulation of ETV1-regulated genes, though changes were less pronounced as compared to sgHAND1 (Fig. S1B).

Several genes exhibited significantly different changes in gene expression with sgETV1 or sgHAND1 exposure. As expected, levels of ETV1 mRNA in the sgETV1 condition and HAND1 mRNA in the sgHAND1 condition showed the greatest decreases, which serves as evidence of CRISPR-mediated on-target nonsense-mediated decay (Fig. 2H). We also found that HOXC10, a non-core TF associated with metastatic GIST (Fig. 1A), was decreased only in the sgHAND1 condition, suggesting co-regulation of these metastasis-associated TFs. Similarly, FOXF1 (37) and HOXA11 expression levels fell only with HAND1 loss. Both sgHAND1 and sgETV1 exposure led to a significant increase in PITX1 (21) expression, suggesting co-repression of this core GIST TF. Notably, the expression of *KIT*, the primary oncogene in these cells, was significantly decreased by loss of HAND1 but not ETV1, while negative regulators of KIT signaling SPRY4 and DUSP6 showed significantly decreased expression with loss of either ETV1 or HAND1 (Fig. 2I). Nonetheless, both TFs showed similar binding at disease relevant loci by ChIP-seq (Fig. 2J, Fig. S1C), with HAND1 loss consistently having a greater effect on mRNA levels at target genes (Fig. S1D). Forced expression of BARX1 in these cells resulted in no observable proliferative phenotype (Fig. S1E–F). No other cell lines profiled by the DepMap project (23) exhibit dependencies upon either of these accessory TFs, suggesting their unique relevance in GIST (Fig. S1G). Taken together, these data support a role for HAND1 in driving an aggressive disease phenotype by regulating the core GIST TF network, maintaining *KIT* gene expression and supporting proliferation-associated signal transduction.

## Selective HAND1 expression in metastatic or High Risk KIT-mutant GIST.

To further characterize the expression of HAND1, BARX1, and other disease-state associated TFs (i.e., HOXC10 and PLAGL1) in GIST tissues resected from patients, we developed and optimized immunohistochemistry (IHC)-based assays for each of these proteins and characterized their expression in a cohort of 87 tumors. We also characterized these tumors using IHC for DOG1, KIT, PDGFRA and SDHB. We found that HAND1 and HOXC10 expression were restricted to KIT-mutant GIST (Fig. 3A). BARX1 and PLAGL1 were expressed in KIT-mutant GIST, albeit in a smaller fraction of the tumors, but both were additionally expressed in PDGFRA-mutant and SDH-deficient GIST, which characteristically exhibit a peri-nuclear Golgi-like staining pattern for PDGFRA (38) and

loss of SDHB (39), respectively. Among KIT-mutant GISTs, only the metastatic gastric tumors – as opposed to the localized gastric GIST – expressed HAND1 and HOXC10, whereas the majority of all small bowel-derived GIST (i.e., localized or metastatic) expressed both high-risk markers (Fig. 3B). In contrast, BARX1 was most commonly expressed in localized gastric GIST. Subdividing tumors by activating KIT mutations, both exon 9 and 11 subtypes exhibited expression of all accessory TFs, while metastatic tumors bearing exon 13 or compound mutations exhibited enrichment for HAND1 and HOXC10 with an associated loss of BARX1 (Fig. 3C).

As is expected for TFs, HAND1, HOXC10, BARX1 and PLAGL1 localized to the nucleus (Fig. 3D, Fig. S2A). There was clear co-expression of HAND1 and HOXC10, and separately of BARX1 and PLAGL1, with mutual exclusion of expression of these accessory TFs between a majority of tumors (Fig. S2B). In tumors expressing either HAND1 or BARX1, >80% of all tumor cells expressed either protein (Fig. S2C). HAND1 positivity had a strong positive predictive value for the mutational subtype of tumors bearing KIT mutations, with PDGFRA expression also showing predictive value for PDGFRA mutant tumors and SDHB loss for SDH-deficient GIST, as expected (Fig. 3E). These results demonstrate the feasibility of assessing accessory TF expression in clinical specimens and support the association of accessory TFs with mutational and clinical subtypes of GIST.

### HAND1-positive tumor cells express proliferative markers and can arise within indolent tumors.

To further explore the association of HAND1 and BARX1 with markers of cellular proliferation in this GIST cohort, we performed multiplexed tissue-based cyclic immunofluorescence (CyCIF) (33,34) of 87 tumors to enable marker expression to be compared at a single cell level. We simultaneously assayed the tumors for the expression of HAND1, BARX1, PDGFRA, SDHB, Ki-67 and PCNA as well as other markers informative of cell type or state (Table S2). SDHB signal was lost from SDH-deficient GIST, while PDGFRA expression was significantly elevated in PDGFRA-mutant tumors, as expected (Fig. 4A). There was significantly higher HAND1 and BARX1 signal in KIT-mutant tumors that had scored positive for these markers by IHC (cross-validating the two imaging methods), with correspondingly lower levels of HAND1 and higher levels of BARX1 in PDGFRA-mutant and SDH-deficient tumors (Fig. 4B). We also observed a significant correlation between Ki-67 signal and mitotic index assessed by expert review (Fig. 4C), and also between Ki-67 and HAND1 signal (Fig. 4D). Selected markers of signal transduction and proliferation, including phospho-ERK and p21, were not different between GIST mutational subtypes, whereas the mitochondrial protein COXIV was present at significantly higher levels in SDH-deficient GIST (Fig. S4A), consistent with the increased abundance of mitochondria in this GIST subtype (40). Analysis of several immune cell cluster of differentiation (CD) markers, including CD163, CD68 and CD11b, revealed that only a small number of these immune subsets infiltrated the tumor microenvironment, and that there were no significant differences between GIST mutational subtypes (Fig. S3B). Whereas Ki-67 and PCNA signal were higher in metastatic tumors in comparison to localized tumors, S100A, a previously reported marker of high-risk GIST (41), was

expressed only in small bowel derived GIST, in both localized and metastatic tumors (Fig. S3C).

To study associations in the expression of TFs and proliferative markers at a single-cell level, we segmented CyCIF images and recorded signal intensities in each channel on a per-cell basis. We then used Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction of combined single cell data from all tumors. We focused on cells that expressed HAND1 and/or BARX1, thereby limiting the analysis to neoplastic cells. BARX1-positive cells clustered together with SDHB-deficient and PDGFRA-high cells, whereas HAND1-positive cells exhibited higher levels of PCNA and Ki-67 (Fig. 4E). Cells expressing S100A represented a subset of the HAND1-positive population (Fig. S3D). Ki-67- and PCNA-positive cells were significantly more enriched in the HAND1-positive population (Fig. 4F, Table S3).

To determine if a small population of HAND1-positive cells was present in localized gastric GIST, we analyzed single cell data from the 15 gastric-primary GIST samples within our cohort (Fig. 3B). UMAP showed co-clustering of a small population of HAND1-positive cells (3,273 of the 24,859 tumor cells; 13%); these cells expressed higher levels of Ki-67 and PCNA compared to BARX1-positive cells (Fig. 4G–I). By IHC, five KIT-mutant tumors from this cohort expressed both HAND1 and BARX1, with at least 5% of neoplastic cells being positive for either marker in the same tumor. Single cell analysis of these tumors showed enrichment of Ki-67 and PCNA in HAND1-positive cells in comparison to BARX1-positive cells (Fig. S3E–F). CyCIF also identified an additional population of cells that expressed both HAND1 and BARX1, and this subgroup had the highest Ki-67 signal intensity of all tumor cells (Fig. S3G–H, Table S3). Taken together, these data demonstrate that HAND1 expression is correlated with cellular proliferation, and that small populations of HAND1-positive cells may exist in localized tumors which otherwise appear histologically indolent. These cells may represent sub-populations prone to evolve over time to dominate the tumor population and drive recurrence and metastasis.

## HAND1 and BARX1 expression associate with anatomic location, mutation and clinical outcomes.

We next extended our analysis to a larger cohort of tumors from 437 patients, many with detailed information on tumor genotype (305 of 437), anatomic origin (384 of 437) and clinical course (130 of 437; Table S4 details a cohort summary). BARX1 and HAND1 IHC showed that the majority of GISTs arising from the esophagus and stomach expressed BARX1 (84%), whereas the majority of GISTs arising from the small or large intestine or rectum expressed HAND1 (75%) (Fig. 5A). Stratifying the cohort by mutational subtype, HAND1 was expressed most frequently in KIT-mutant tumors (53%), in agreement with our findings from the primary cohort (Fig. 5B). One of thirty-four PDGFRA-mutant tumors expressed HAND1, as did both NF1-associated GISTs (Fig. 5B). HAND1 and BARX1 positivity were seen in all subtypes of KIT mutations, although HAND1 was present in nearly all of the exon 9 tumors that commonly originate from the small bowel (42) and in nearly all GIST with compound mutations arising in metastatic and TKI-resistant disease (43) (Fig. 5C). While primary GISTs expressed BARX1 and HAND1 at similar rates, the

majority (63/84, 75%) of metastatic or recurrent GIST cases were positive for HAND1 (Fig. 5D). The majority of all tumors expressed either HAND1 or BARX1 (343/437, 78%), and a minority were positive (28/437, 6%) or negative (66/437, 15%) for both markers, consistent with the gene expression profiling data (Fig. 5E, Fig. 1C). In cases that expressed both HAND1 and BARX1, the majority (22/28, 79%) expressed predominantly one marker over the other or had low expression of both markers (Fig. S4A–B). As part of this cohort, we also evaluated seven micro-GISTs, defined as gastric GISTs ≤ 1 cm in size with benign clinical behavior (44); all were positive for BARX1 and negative for HAND1 (Fig. S4C–D). For GIST negative for both HAND1 and BARX1, cases were overrepresented by tumors arising from the colon or rectum (Table S5). Extragastric GIST positive only for BARX1 were overrepresented in tumors arising from the esophagus or gastroesophageal junction (GEJ; Table S6).

For a subset of these tumors (130 of 437, 30%), annotation was available to assess clinical outcomes. For macroscopically resected primary tumors, standard risk stratification into High Risk and Non-High Risk groups using tumor size, anatomic location and mitotic index showed a trend, but no statistically significant difference, in relapse-free survival (Fig. 5F, $P = 0.1171$). In contrast, stratification of patients by the expression of HAND1 or BARX1 showed a significant difference in relapse-free survival ($P = 0.0092$), with an associated hazard ratio of 2.62 (95% CI 1.14–6.00) (Fig. 5G). A subset of these tumors expressed both or neither marker, but the sample size was too small to permit statistical analysis of these groups. Multivariate analysis of HAND1 expression, BARX1 expression and Recurrence Risk level using a Cox proportional hazards model of relapse free survival showed independent but non-significant trends towards shorter relapse-free times (i.e. higher risk of relapse) with HAND1 expression or High Risk level, and longer relapse-free times (i.e. lower risk of relapse) with BARX1 expression; similarly, HAND1 expression and extragastric tumors showed independent non-significant trends towards shorter relapse-free survival (Fig. S5A–B). Finally, in patients diagnosed with metastatic GIST and started on first-line imatinib, progression-free survival was significantly shorter in HAND1-expressing tumors ($P = 0.0127$), with a hazard ratio of 2.322 (95% CI 1.27–4.25) (Fig. 5H); only one of these tumors, in the HAND1-negative group, expressed BARX1, and 46 of 49 tumors had oncogenic *KIT* mutations while the remaining 3 cases had no mutational testing available. These data demonstrate that HAND1 and BARX1 expression correlates with clinical outcome and may exceed the current standard of care tools for predicting relapse-free survival; these new biomarkers may also offer a tool for predicting progression-free survival in the metastatic setting.

### Restricted anatomic and spatial expression of HAND1 and BARX1 in ICCs.

To evaluate accessory TF expression in normal ICCs, which provide insights into the cell-of-origin of GIST, we performed IHC on normal tissues along the gastrointestinal tract. Using two-marker IHC, we identified ICCs with an antibody recognizing KIT and characterized accessory TF expression using antibodies recognizing either HAND1 or BARX1. We assessed several thousand individual ICCs derived from different anatomic regions spanning from the esophagus to the rectum. ICCs from the stomach showed frequent expression of BARX1, but never HAND1; by contrast, the small intestine ICCs showed frequent

expression of HAND1 but not BARX1 (Fig. 6A). The majority of ICCs in these disparate anatomic compartments express HAND1 or BARX1, and in a mutually exclusive pattern (Fig. 6B). In contrast, ICCs of the esophagus, appendix, colon and rectum expressed neither BARX1 nor HAND1. In the small intestine, where intramuscular and myenteric ICCs could be readily distinguished, HAND1 expression was more common in ICCs within the myenteric plexus (Fig. S5C), suggesting transcriptional differences in ICCs at different stations within the same organ; this suggests that specific ICC subsets are at higher risk for malignant transformation. In aggregate, these data outline anatomic and spatial differences in accessory TF expression in ICCs, and suggest that anatomic restriction of accessory TF expression underlies the differential clinical behavior of GIST derived from the stomach versus the small intestine.

## DISCUSSION

A core circuit of TFs is essential for establishing and maintaining a normal transcriptional state of cellular homeostasis, and this state can become dysregulated during the process of neoplastic transformation (16–20). Accessory disease-state specific TFs can modulate biological and molecular phenotypes and thereby shape clinical behavior. Here, we show that BARX1 and HAND1 are the principal accessory TFs in GIST that modify the core transcriptional circuitry and show strong associations with clinical features and disease outcomes. In addition, we reveal a striking anatomic restriction of accessory TF expression that reflects the developmental origins of these tumors and the difference in malignant potential between gastric and small intestine GIST.

BARX1 is a member of the Bar subclass of homeobox transcription factors and has been studied in multiple developmental contexts, including its role in coordinating gastric myogenesis and intestinal rotation (45). HAND1 (Heart And Neural Crest Derivatives Expressed 1) is a basic helix-loop-helix family transcription factor best understood for its role in early embryonic development and cardiac morphogenesis (46). Both BARX1 and HAND1 have previously been studied in cancer contexts, with reports suggesting suppressive or supportive roles in different tumor histologies (47–49). In GIST, there is remarkable homogeneity of the gene expression program between tumors (Fig. 1A), though mutation and clinical subtypes display subtle transcriptional differences as also observed in other studies (35,50,51). Early in the course of disease there is typically limited chromosomal instability in GIST (52,53), including an absence of amplifications involving the *KIT* locus (54). The lack of chromosomal aberrations, together with a recurrent enhancer landscape across disease states (21), suggests that GISTs rely upon native ICC chromatin and transcriptional environments. It appears that chromosomal stability is selected for, at least in part, to support the enhancer-driven expression of *KIT* in the majority of tumors (55). This dependency places accessory TFs HAND1 and BARX1 in a pivotal position to influence disease biology by modifying the basal transcriptional state of GIST cells.

Active signal transduction and cell cycle programs, as observed at the transcriptional level in HAND1-expressing tumors, is consistent with the higher mitotic rate and proliferative features seen in High Risk and metastatic GIST (12,53). The primary influence of HAND1 may be to modify the expression of, or protein interactions within, the core TF network to

support the GIST growth program, most notably *KIT* gene expression. The HAND1-driven program may not only offer prognostic value, but also represent a therapeutic vulnerability either itself or by targeting genes whose expression is regulated by HAND1. For example, HAND1 regulates GPR20 expression (Fig. S1D), and an anti-GPR20 antibody-drug conjugate has now entered clinical trials (NCT04276415). Additional mechanistic studies are needed to better define how HAND1 orchestrates a more malignant phenotype, which may reveal additional therapeutic vulnerabilities. We observed recurrent enrichment of an EMT signature in BARX1-expressing GIST, or following loss of HAND1 function *in vitro*. This may suggest that, in the context of GIST, other biological features of the EMT program are engaged including those involved in organ development, fibrosis or tissue repair (56); this contrasts with epithelial-derived malignancies. Of note, our analysis of GIST transcription focused on biologically and clinically distinct categories of GIST and did not prioritize subtypes by incidence; additionally, full clinical annotation of all samples (i.e. adjuvant therapy use) was unavailable, which represents a weakness of this study.

In our primary tumor cohort, HAND1 expression had a strong positive predictive value for KIT mutation. In all samples with known underlying mutation, HAND1 was expressed in 53% (134/255) of KIT mutant tumors, 3% (1/34) of PDGFRA mutant tumors, 0% (0/14) of SDH-deficient tumors, and 100% (2/2) of NF1-associated tumors. The lone PDGFRA mutant GIST that expressed HAND1 also expressed BARX1 and was found to have an uncommon in-frame deletion (p.D842_H845del); at resection it was classified as a Low Risk gastric GIST. Since NF1-associated GIST have a different underlying etiology and clinical course than more common GIST subtypes (6), and these tumors are derived from small bowel ICCs that natively express HAND1, the contribution of HAND1 expression to oncogenesis in this tumor type is less clear. While the majority of GISTs express either HAND1 or BARX1, 15% are negative and 6% positive for both markers. The numbers of these GIST subsets in our cohort were too small to assess outcomes, and further study is needed to assess BARX1 and HAND1 expression in a larger cohort of these rare tumor subtypes and characterize their clinical course. Further, given the retrospective nature of these studies and limited available clinical data, additional studies are needed to better establish these findings.

Multiplexed tissue imaging (CycIF) makes it possible to study correlations in the expression of HAND1, BARX1 and other markers of cellular proliferation at the single cell level. Consistent with what was observed in bulk tumor RNA-seq data and *in vitro* with loss of HAND1, GIST cells expressing HAND1 have higher expression of proliferative markers. Rare HAND1-positive cells were also observed in localized gastric GIST, which predominantly express BARX1. Conceivably, since stomach ICCs do not natively express HAND1, as malignant clones of gastric-derived GIST evolve over time they may activate HAND1 expression, giving rise to this double-positive tumor subset. With disease progression, the naturally higher proliferative rates of HAND1-positive cells may cause them to dominate the tumor landscape in the recurrent or metastatic setting, and cells may eventually lose expression of BARX1. Additional study of accessory TFs in case series with paired primary and metastatic samples may further define this evolution. Our observation that BARX1 or HAND1 is absent in ICCs of the esophagus, colon, appendix and rectum agrees with clinical observations that these anatomic compartments are less competent to

support transformation into GIST (57). That the majority (85%) of GISTs express BARX1 and/or HAND1, inclusive of tumors from sites whose ICCs natively lack these TFs, suggests that the presence or elaboration of these accessory TFs is commonly required to support malignancy. From *in vitro* experiments, ectopic expression of BARX1 was not deleterious to the HAND1-driven proliferative program (Fig. S1E–F), suggesting that both TFs are permissive of a malignancy program, but that HAND1 is the driver of more aggressive behavior. The prominent anti-correlation of these accessory TFs in ICCs and GIST samples warrants further research on their role in ICC development and mechanisms of cross-regulation.

S100A has previously been reported to be a marker of high-risk GIST (41); however, in our tumor cohort we only identified S100A within tumors derived from the small intestine. Thus, while HAND1 and S100A expression may both be correlated with small intestine origin, S100A appears to be solely associated with this anatomic site while HAND1 appears to be an acquired driver of aggressive behavior in other regions. CyCIF allowed the interrogation of select immune subsets in a diverse array of GIST samples. We found that myeloid-lineage (CD11b) and macrophage (CD68, CD163) markers were present at low levels in tumors, and there was no significant difference between GIST mutational subtypes.

The anatomic and spatially restricted expression of BARX1 and HAND1 in ICCs, and associated intrinsic differences in clinical behavior based on the anatomic origin of GIST, further argues for the role these TFs play in the pathogenesis of GIST. ICCs in different anatomic compartments have previously been shown to have different densities, distributions and gene expression profiles (58,59), though the results presented here are the first to make a mechanistic link between a gene expression program in ICCs and clinical behavior upon transformation into GIST. Additional genetic studies in laboratory models of GIST altering accessory TF expression in ICCs may further clarify their role in directing oncogenesis.

The greatest potential value in assessing HAND1 and BARX1 expression in GIST lies in predicting the eventual course of the disease in patients. Especially for macroscopically resected GIST deemed Intermediate Risk, or where there is discrepancy, lack of expertise, or an inability to determine mitotic index (i.e., in cases confounded by pre-operative administration of imatinib), HAND1 and BARX1 expression may help support clinical decisions on assessing the relative benefits of adjuvant imatinib therapy. As the majority of tumor cells stain positive for either HAND1 or BARX1, immunohistochemistry is a standard and effective means of assessing the presence or absence of these biomarkers. In the metastatic setting, our data show that HAND1 expression is associated with a shorter progression-free survival, which may inform disease assessment intervals or clinical trial eligibility criteria. Further study is warranted to assess the added clinical value of HAND1 and BARX1 assessment, and whether routine assay of these TFs could enhance the current standard of care and improve risk stratification for GIST patients.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

COMPETING INTERESTS

G.D.D. has served as a scientific consultant with sponsored research from Bayer, Pfizer, Novartis, Roche/Genentech, Epizyme, LOXO Oncology, AbbVie, GlaxoSmithKline, Janssen, PharmaMar, ZioPharm, Daiichi-Sankyo, AdaptImmune, Ignyta and Mirati; serves as a scientific consult for GlaxoSmithKline, EMD-Serono, Sanofi, ICON plc, WCG/Arsenal Capital, Polaris Pharmaceuticals, MJ Hennessey/OncLive, MEDSCAPE; is a consultant or scientific advisory board member with minor equity holding in G1 Therapeutics, Caris Life Sciences, Champions Biotechnology, Bessor Pharmaceuticals, Erasca Pharmaceuticals, RELAY Therapeutics, Caprion/HistoGeneX; is a board of Directors member and scientific advisory board consultant with minor equity holding in Blueprint Medicines, Merrimack Pharmaceuticals (ended October 2019) and Translate BIO; holds patents/royalties from Novartis to DFCI for use of imatinib in GIST; and has non-financial interests in McCann Health, Alexandria Summit and is the AACR Science Policy and Government Affairs Committee Chair. S.A.A. has received research support from Janssen, Novartis, and AstraZeneca. P.K.S. is on the scientific advisory board of RareCyte, Inc., whose product was used to acquire this data, and Glencoe Software, Inc., whose product was used to visualize the data; P.K.S. is also on the board of directors or scientific advisory board of Applied BioMath and NanoString. S.S. has consulted for RareCyte, Inc. None of these relationships constitute a conflict of interest for the present work. The remaining authors declare no conflict of interest.

## REFERENCES

1. Ducimetière F, Lurkin A, Ranchère-Vince D, Decouvelaere A-V, Péoc'h M, Istier L, et al. Incidence of sarcoma histotypes and molecular subtypes in a prospective epidemiological study with central pathology review and molecular testing. Najbauer J, editor. PLoS ONE. Public Library of Science; 2011;6:e20294. [PubMed: 21826194]

2. Ma GL, Murphy JD, Martinez ME, Sicklick JK. Epidemiology of Gastrointestinal Stromal Tumors in the Era of Histology Codes: Results of a Population-Based Study. Cancer Epidemiology Biomarkers & Prevention. 2015;24:298–302.

3. Kindblom LG, Remotti HE, Aldenborg F, Meis-Kindblom JM. Gastrointestinal pacemaker cell tumor (GIPACT): gastrointestinal stromal tumors show phenotypic characteristics of the interstitial cells of Cajal. Am J Pathol. 1998;152:1259–69. [PubMed: 9588894]

4. Hemming ML, Heinrich MC, Bauer S, George S. Translational insights into gastrointestinal stromal tumor and current clinical advances. Annals of Oncology. 2018;3:557–9.

5. Boikos SA, Pappo AS, Killian JK, LaQuaglia MP, Weldon CB, George S, et al. Molecular Subtypes of KIT/PDGFRAWild-Type Gastrointestinal Stromal Tumors. JAMA Oncol. 2016;:1–7.

6. Andersson J, Sihto H, Meis-Kindblom JM, Joensuu H, Nupponen N, Kindblom L-G. NF1-associated gastrointestinal stromal tumors have unique clinical, phenotypic, and genotypic characteristics. Am J Surg Pathol. 2005;29:1170–6. [PubMed: 16096406]

7. Hirota S, Isozaki K, Moriyama Y, Hashimoto K, Nishida T, Ishiguro S, et al. Gain-of-function mutations of c-kit in human gastrointestinal stromal tumors. Science. 1998;279:577–80. [PubMed: 9438854]

8. Heinrich MC, Griffith DJ, Druker BJ, Wait CL, Ott KA, Zigler AJ. Inhibition of c-kit receptor tyrosine kinase activity by STI 571, a selective tyrosine kinase inhibitor. Blood. 2000;96:925–32. [PubMed: 10910906]

9. Demetri GD, Mehren von M, Blanke CD, Van den Abbeele AD, Eisenberg B, Roberts PJ, et al. Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. N Engl J Med. 2002;347:472–80. [PubMed: 12181401]

10. Patrikidou A, Dômont J, Chabaud S, Ray-Coquard I, Coindre J-M, Bui-Nguyen B, et al. Long-term outcome of molecular subgroups of GIST patients treated with standard-dose imatinib in the BFR14 trial of the French Sarcoma Group. Eur J Cancer. Elsevier Ltd; 2016;52:173–80. [PubMed: 26687836]

11. Miettinen M, Lasota J. Gastrointestinal stromal tumors: Pathology and prognosis at different sites. Seminars in Diagnostic Pathology. 2006;23:70–83. [PubMed: 17193820]

12. DeMatteo RP, Ballman KV, Antonescu CR, Corless C, Kolesnikova V, Mehren von M, et al. Long-term results of adjuvant imatinib mesylate in localized, high-risk, primary gastrointestinal stromal tumor: ACOSOG Z9000 (Alliance) intergroup phase 2 trial. Annals of Surgery. 2013;258:422–9. [PubMed: 23860199]

13. Joensuu H, Eriksson M, Sundby Hall K, Hartmann JT, Pink D, Schütte J, et al. One vs three years of adjuvant imatinib for operable gastrointestinal stromal tumor: a randomized trial. JAMA. 2012;307:1265–72. [PubMed: 22453568]

14. Raut CP, Espat NJ, Maki RG, Araujo DM, Trent J, Williams TF, et al. Efficacy and Tolerability of 5-Year Adjuvant Imatinib Treatment for Patients With Resected Intermediate- or High-Risk Primary Gastrointestinal Stromal Tumor. JAMA Oncol. 2018;4:e184060–8. [PubMed: 30383140]

15. Joensuu H, Vehtari A, Riihimäki J, Nishida T, Steigen SE, Brabec P, et al. Risk of recurrence of gastrointestinal stromal tumour after surgery: an analysis of pooled population-based cohorts. Lancet Oncol. 2012;13:265–74. [PubMed: 22153892]

16. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 2011;25:2227–41. [PubMed: 22056668]

17. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Molecular Cell. Elsevier Inc; 2010;38:576–89. [PubMed: 20513432]

18. Pomerantz MM, Li F, Takeda DY, Lenci R, Chonkar A, Chabot M, et al. The androgen receptor cistrome is extensively reprogrammed in human prostate tumorigenesis. Nature Publishing Group. Nature Publishing Group; 2015;47:1346–51.

19. Goldstein AS, Huang J, Guo C, Garraway IP, Witte ON. Identification of a cell of origin for human prostate cancer. Science. 2010;329:568–71. [PubMed: 20671189]

20. Lin CY, Lovén J, Rahl PB, Paranal RM, Burge CB, Bradner JE, et al. Transcriptional Amplification in Tumor Cells with Elevated c-Myc. Cell. Elsevier; 2012;151:56–67. [PubMed: 23021215]

21. Hemming ML, Lawlor MA, Zeid R, Lesluyes T, Fletcher JA, Raut CP, et al. Gastrointestinal stromal tumor enhancers support a transcription factor network predictive of clinical outcome. Proc Natl Acad Sci USA. 2018;115:E5746–55. [PubMed: 29866822]

22. Hemming ML, Elias JE, Gygi SP, Selkoe DJ. Proteomic profiling of gamma-secretase substrates and mapping of substrate requirements. PLoS Biol. 2008;6:e2571–15.

23. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. Nature Publishing Group; 2012;483:603–7. [PubMed: 22460905]

24. Lesluyes T, Pérot G, Largeau MR, Brulard C, Lagarde P, Dapremont V, et al. RNA sequencing validation of the Complexity INdex in SARComas prognostic signature. Eur J Cancer. 2016;57:104–11. [PubMed: 26916546]

25. Flavahan WA, Drier Y, Johnstone SE, Hemming ML, Tarjan DR, Hegazi E, et al. Altered chromosomal topology drives oncogenic programs in SDH-deficient GIST. Nature. Springer US; 2019;:1–8.

26. Vitiello GA, Bowler TG, Liu M, Medina BD, Zhang JQ, Param NJ, et al. Differential immune profiles distinguish the mutational subtypes of gastrointestinal stromal tumor. J Clin Invest. 2019;129:1863–77. [PubMed: 30762585]

27. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2012;29:15–21. [PubMed: 23104886]

28. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28:511–5. [PubMed: 20436464]

29. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2009;26:139–40. [PubMed: 19910308]

30. Wang S, Sun H, Ma J, Zang C, Wang C, Wang J, et al. Target analysis by integration of transcriptome and ChIP-seq data with BETA. Nature Protocols. Nature Publishing Group; 2013;8:2502–15. [PubMed: 24263090]

31. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545–50. [PubMed: 16199517]

32. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.1–R25.10. [PubMed: 19261174]

33. Du Z, Lin J-R, Rashid R, Maliga Z, Wang S, Aster JC, et al. Qualifying antibodies for image-based immune profiling and multiplexed tissue imaging. Nature Protocols. 2019;14:2900–30. [PubMed: 31534232]

34. Lin J-R, Izar B, Wang S, Yapp C, Mei S, Shah PM, et al. Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. Elife. 2018;7:545.

35. Vitiello GA, Bowler TG, Liu M, Medina BD, Zhang JQ, Param NJ, et al. Differential immune profiles distinguish the mutational subtypes of gastrointestinal stromal tumor. J Clin Invest. 2019;8:466–16.

36. Chi P, Chen Y, Zhang L, Guo X, Wongvipat J, Shamu T, et al. ETV1 is a lineage survival factor that cooperates with KIT in gastrointestinal stromal tumours. Nature. 2010;467:849–53. [PubMed: 20927104]

37. Ran L, Chen Y, Sher J, Wong EWP, Murphy D, Zhang JQ, et al. FOXF1 Defines the Core-Regulatory Circuitry in Gastrointestinal Stromal Tumor. Cancer Discov. 2018;8:234–51. [PubMed: 29162563]

38. Pauls K, Merkelbach-Bruse S, Thal D, Buttner R, Wardelmann E. PDGFRalpha- and c-kit-mutated gastrointestinal stromal tumours (GISTs) are characterized by distinctive histological and immunohistochemical features. Histopathology. 2005;46:166–75. [PubMed: 15693889]

39. Gill AJ, Chou A, Vilain R, Clarkson A, Lui M, Jin R, et al. Immunohistochemistry for SDHB divides gastrointestinal stromal tumors (GISTs) into 2 distinct types. Am J Surg Pathol. 2010;34:636–44. [PubMed: 20305538]

40. Szarek E, Ball ER, Imperiale A, Tsokos M, Faucz FR, Giubellino A, et al. Carney triad, SDH-deficient tumors, and Sdhb+/− mice share abnormal mitochondria. Endocrine-Related Cancer. 2015;22:345–52. [PubMed: 25808178]

41. Perez D, Demartines N, Meier K, Clavien P-A, Jungbluth A, Jaeger D. Protein S100 as prognostic marker for gastrointestinal stromal tumors: a clinicopathological risk factor analysis. J Invest Surg. 2007;20:181–6. [PubMed: 17613693]

42. Künstlinger H, Huss S, Merkelbach-Bruse S, Binot E, Kleine MA, Loeser H, et al. Gastrointestinal Stromal Tumors With KIT Exon 9 Mutations: Update on Genotype-Phenotype Correlation and Validation of a High-Resolution Melting Assay for Mutational Testing. Am J Surg Pathol. 2013;37:1648–59. [PubMed: 24061512]

43. Antonescu CR, Besmer P, Guo T, Arkun K, Hom G, Koryotowski B, et al. Acquired resistance to imatinib in gastrointestinal stromal tumor occurs through secondary gene mutation. Clin Cancer Res. 2005;11:4182–90. [PubMed: 15930355]

44. Rossi S, Gasparotto D, Toffolatti L, Pastrello C, Gallina G, Marzotto A, et al. Molecular and clinicopathologic characterization of gastrointestinal stromal tumors (GISTs) of small size. Am J Surg Pathol. 2010;34:1480–91. [PubMed: 20861712]

45. Jayewickreme CD, Shivdasani RA. Control of stomach smooth muscle development and intestinal rotation by transcription factor BARX1. Developmental Biology. Elsevier; 2015;405:21–32.

46. Knöfler M, Meinhardt G, Bauer S, Loregger T, Vasicek R, Bloor DJ, et al. Human Hand1 basic helix-loop-helix (bHLH) protein: extra-embryonic expression pattern, interaction partners and

identification of its transcriptional repressor domains. Biochem J. 2002;361:641–51. [PubMed: 11802795]

47. Yan C, Ji Y, Huang T, Yu F, Gao Y, Gu Y, et al. An esophageal adenocarcinoma susceptibility locus at 9q22 also confers risk to esophageal squamous cell carcinoma by regulating the function of BARX1. Cancer Letters. Elsevier Ltd; 2018;421:103–11. [PubMed: 29454095]

48. Wang G, Liu J, Cai Y, Chen J, Xie W, Kong X, et al. Loss of Barx1 promotes hepatocellular carcinoma metastasis through up-regulating MGAT5 and MMP9 expression and indicates poor prognosis. Oncotarget. 2017;8:71867–80. [PubMed: 29069753]

49. Hoyos JM, Ferraro A, Sacchetti S, Keller S, De Martino I, Borbone E, et al. HAND1 gene expression is negatively regulated by the High Mobility Group A1 proteins and is drastically reduced in human thyroid carcinomas. Oncogene. 2008;28:876–85. [PubMed: 19060921]

50. Antonescu CR, Viale A, Sarran L, Tschernyavsky SJ, Gonen M, Segal NH, et al. Gene expression in gastrointestinal stromal tumors is distinguished by KIT genotype and anatomic site. Clin Cancer Res. American Association for Cancer Research; 2004;10:3282–90. [PubMed: 15161681]

51. Allander SV, Nupponen NN, Ringnér M, Hostetter G, Maher GW, Goldberger N, et al. Gastrointestinal stromal tumors with KIT mutations exhibit a remarkably homogeneous gene expression profile. Cancer Research. 2001;61:8624–8. [PubMed: 11751374]

52. Lagarde P, Perot G, Kauffmann A, Brulard C, Dapremont V, Hostein I, et al. Mitotic Checkpoints and Chromosome Instability Are Strong Predictors of Clinical Outcome in Gastrointestinal Stromal Tumors. Clin Cancer Res. 2012;18:826–38. [PubMed: 22167411]

53. Heinrich MC, Patterson J, Beadling C, Wang Y, Debiec-Rychter M, Dewaele B, et al. Genomic aberrations in cell cycle genes predict progression of KIT-mutant gastrointestinal stromal tumors (GISTs). Clin Sarcoma Res. BioMed Central; 2019;:1–15. [PubMed: 30651969]

54. Tabone S, Théou N, Wozniak A, Saffroy R, Deville L, Julié C, et al. KIT overexpression and amplification in gastrointestinal stromal tumors (GISTs). Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease. 2005;1741:165–72. [PubMed: 15869870]

55. Hemming ML, Lawlor MA, Andersen JL, Hagan T, Chipashvili O, Scott TG, et al. Enhancer Domains in Gastrointestinal Stromal Tumor Regulate KIT Expression and are Targetable by BET Bromodomain Inhibition. Cancer Research. 2019;:canres.1888.2018–31.

56. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009;119:1420–8. [PubMed: 19487818]

57. Corless CL, Heinrich MC. Molecular Pathobiology of Gastrointestinal Stromal Sarcomas. Annu Rev Pathol Mech Dis. 2008;3:557–86.

58. Vanderwinden JM, Rumessen JJ. Interstitial cells of Cajal in human gut and gastrointestinal disease. Microsc Res Tech. 1999;47:344–60. [PubMed: 10602294]

59. Lee MY, Ha SE, Park C, Park PJ, Fuchs R, Wei L, et al. Transcriptome of interstitial cells of Cajal reveals unique and selective gene signatures. Shi X-Z, editor. PLoS ONE. 2017;12:e0176031–25. [PubMed: 28426719]

## STATEMENT OF TRANSLATIONAL RELEVANCE

The biological underpinnings of divergent clinical behavior in gastrointestinal stromal tumor (GIST) are incompletely characterized. By profiling subtypes of GIST using RNA-sequencing, chromatin immunoprecipitation with sequencing, spatial multiplexed imaging and histopathology, we show that HAND1 orchestrates a transcriptional program driving aggressive disease, which is developmentally encoded or evolves in a preponderance of malignant GIST. By contrast, BARX1 is expressed in indolent and micro-GISTs, and BARX1 expression is commonly lost in recurrent or metastatic disease. We found that HAND1 and BARX1 expression patterns were superior predictors of relapse-free survival compared to standard risk stratification in GIST, and they predict progression-free survival on imatinib in the metastatic setting. Evaluation of these transcription factors by immunohistochemistry is feasible and may provide clinically actionable prognostic information. These findings warrant further investigation on the predictive and therapeutic value of HAND1 and BARX1 expression in GIST.
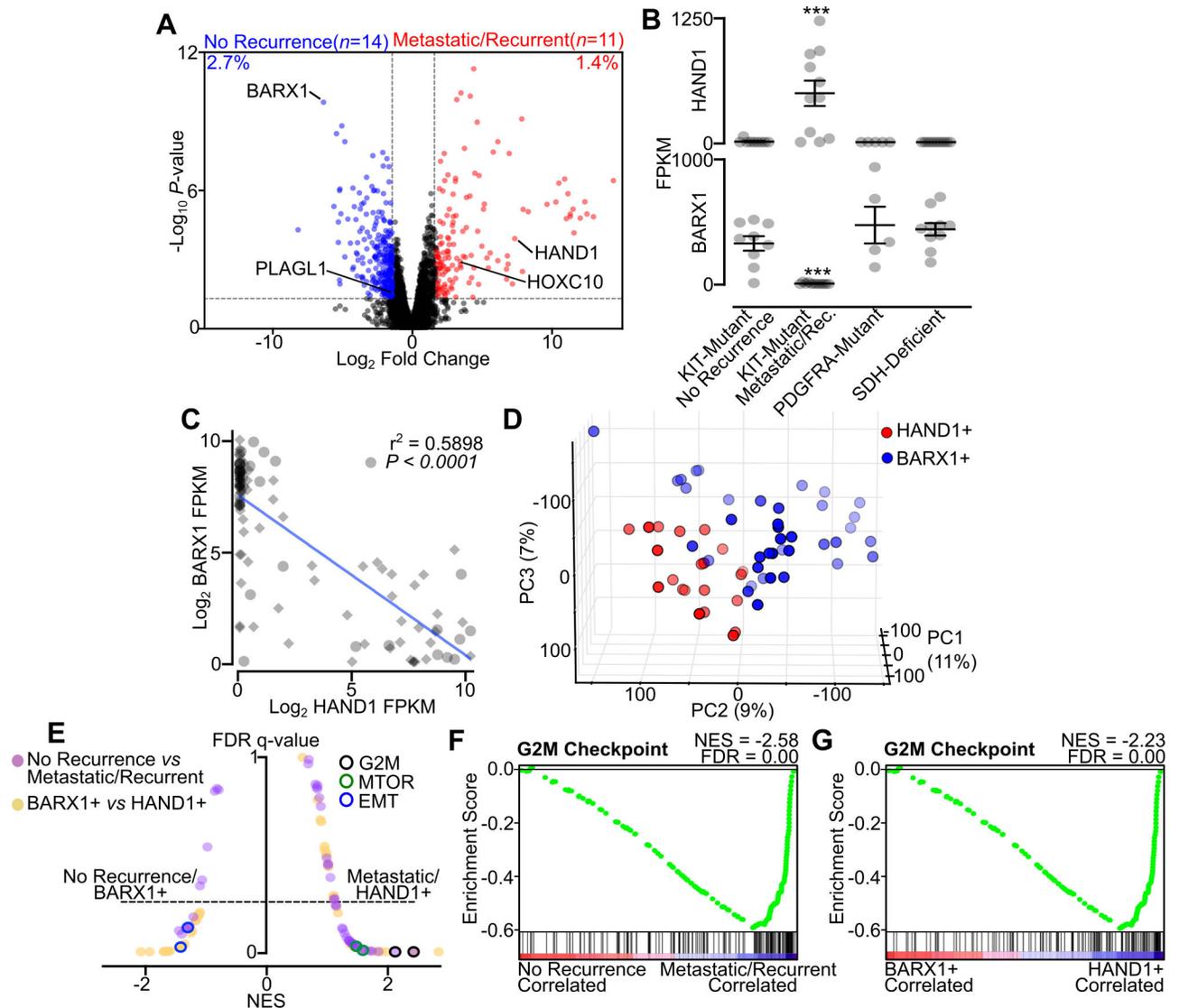
**Figure 1. HAND1 is associated with metastatic GIST and a distinctive transcriptional program.**
**A,** Volcano plot of RNA-seq data from localized and never recurrent GIST ('No Recurrence', $n$=14) or localized and later recurrent or metastatic GIST ('Metastatic/Recurrent', $n$=11). All differentially expressed TFs are labeled. The percent of genes differentially expressed is shown ($n$=10,000 total expressed transcripts). **B,** FPKM of HAND1 and BARX1 in No Recurrence KIT-mutant GIST ($n$=9), Metastatic/Recurrent KIT-mutant GIST ($n$=11), PDGFRA-mutant GIST ($n$=5) or SDH-deficient GIST ($n$=10). Data were analyzed by one-way ANOVA with Tukey's multiple comparison test (compared to KIT-mutant localized GIST; ***,$P$<0.001). **C,** Correlation of BARX1 and HAND1 expression in all GIST samples ($n$=110), with the clinically annotated cohort indicated with circles and validation cohort with diamonds. The Pearson correlation is shown. **D,** PCA of GIST validation cohort RNA-seq (35) stratified by HAND1-positivity ($n$=21) and BARX1-positivity ($n$=39) with threshold for expression of 50 FPKM. **E,** Butterfly plot of all Hallmark gene sets indicating the NES and FDR q-value for the two RNA-seq data cohorts
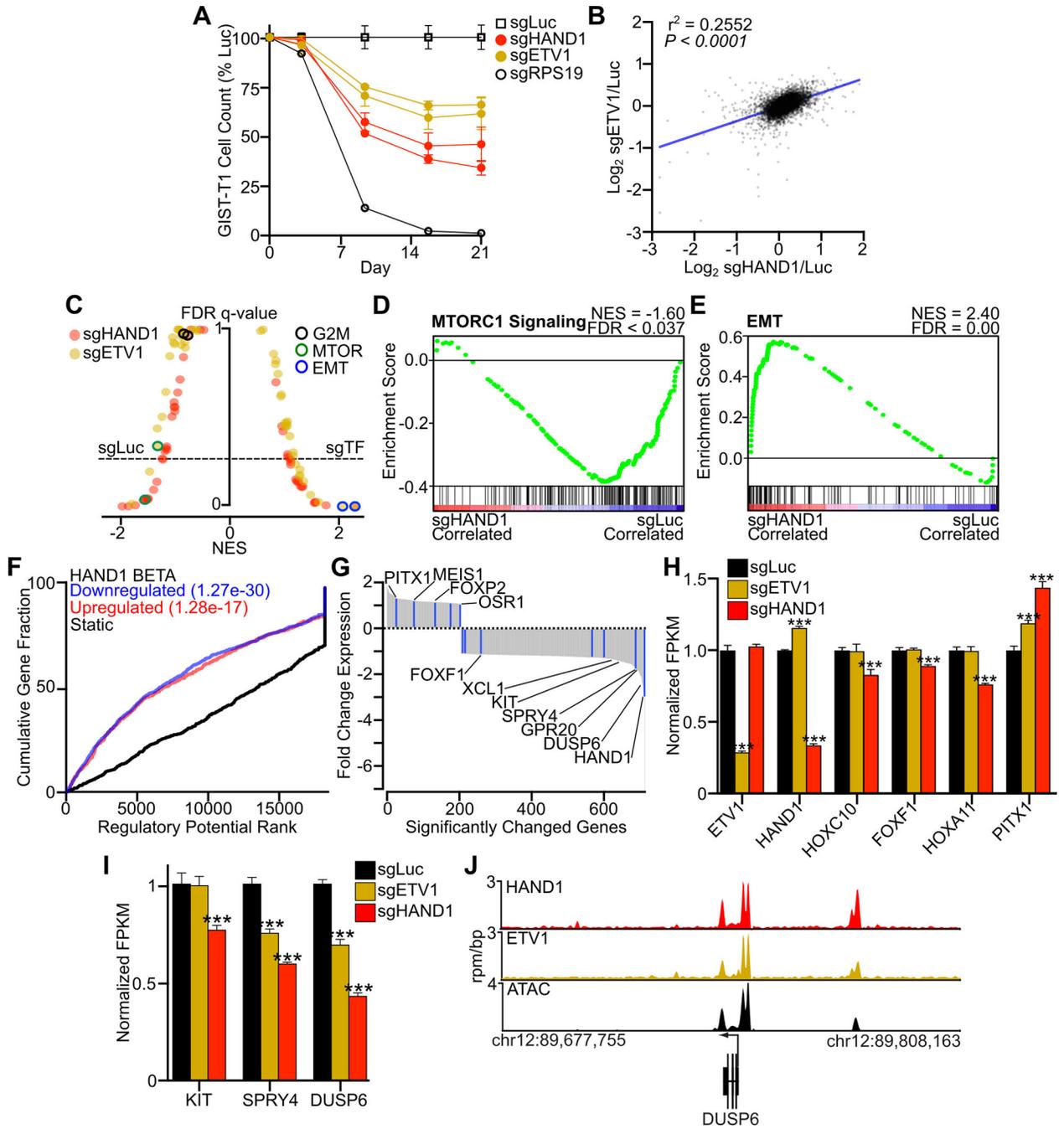
('No Recurrence' and 'Metastatic/Recurrent' GIST in purple, BARX1- or HAND1-positive GIST in yellow). The G2M Checkpoint, MTORC1 Signaling and Epithelial Mesenchymal Transition (EMT) gene sets are indicated for each condition. **F-G,** GSEA showing the Hallmark G2M Checkpoint gene set in the independent GIST RNA-seq data sets.

**Figure 2. HAND1 is necessary for GIST cell growth and transcriptional regulation.**
**A,** Growth over time assay following transduction with two distinct sgRNAs targeting *HAND1* or *ETV1*, or control sgRNAs against luciferase (Luc) and the essential ribosomal gene *RPS19*. Cell count is normalized to sgLuc control (*n*=3 per construct). **B,** Log$_2$ correlation of the ratio of expressed transcripts in sgHAND1 or sgETV1 to sgLuc conditions following 5 days of sgRNA expression. The Pearson correlation is shown. **C,** Butterfly plot of all Hallmark gene sets indicating the NES and FDR q-value comparing sgHAND1 (red) and sgETV1 (yellow) to sgLuc conditions. The G2M Checkpoint, MTORC1 Signaling and

EMT gene sets are indicated for each condition. **D-E,** Hallmark MTORC1 Signaling and EMT gene sets comparing sgLuc and sgHAND1. **F,** Binding and Expression Target Analysis (BETA) integrating HAND1 ChIP-seq and differential gene expression for sgHAND1 and sgLuc comparisons. **G,** Waterfall plot showing the fold change of differentially expressed HAND1-regulated genes resulting from sgHAND1 expression. Core transcription factors are indicated in blue, and select genes are labeled. **H,** Expression of select GIST TFs in cells treated with sgHAND1, sgETV1 or sgLuc as control. **I,** Expression of KIT and negative regulators of KIT signaling in cells treated with sgHAND1, sgETV1 or sgLuc as control. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test ($n$=3; compared to sgLuc; ***,$P$<0.001). **J,** ChIP-seq tracks for HAND1 and ETV1 and ATAC peaks at the *DUSP6* locus.
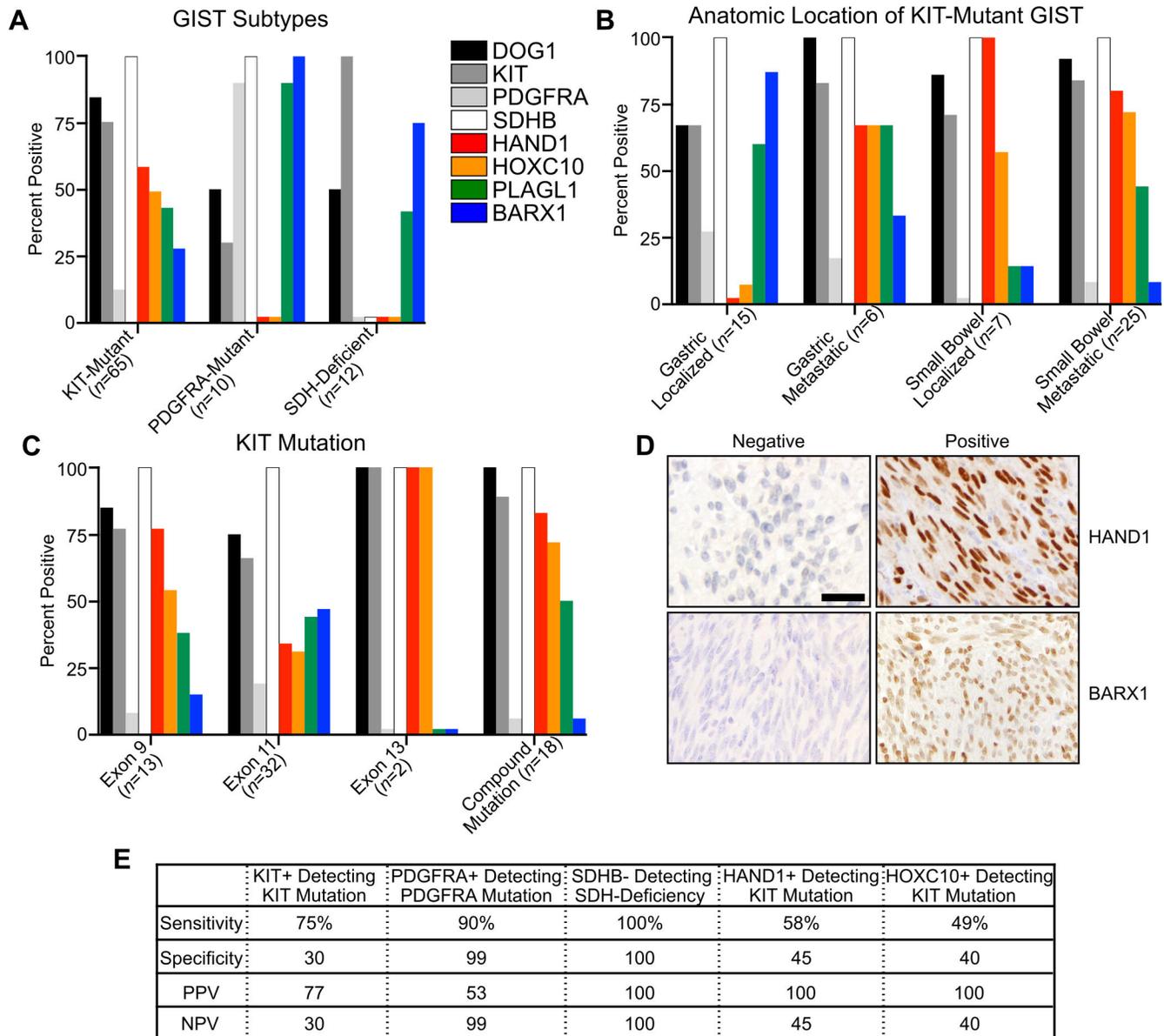
**Figure 3. Immunohistochemistry of GIST with standard and novel biomarkers.**
**A,** Percent positive expression in GIST subtypes including KIT-mutant ($n$=65), PDGFRA-mutant ($n$=10), and SDH-deficient ($n$=12) tumors for the proteins DOG1, KIT, PDGFRA, SDHB, HAND1, HOXC10, PLAGL1 and BARX1. **B,** IHC of KIT-mutant GIST stratified by anatomic location and localized or metastatic disease. **C**, IHC of KIT-mutant GIST stratified by mutational subtype. All exon 13 and compound mutation tumors were derived from metastatic disease. **D,** IHC of GIST with exemplary positive and negative expression of HAND1 and BARX1; scale bar indicates 20 μm. **E,** Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of select proteins for mutational subtype.
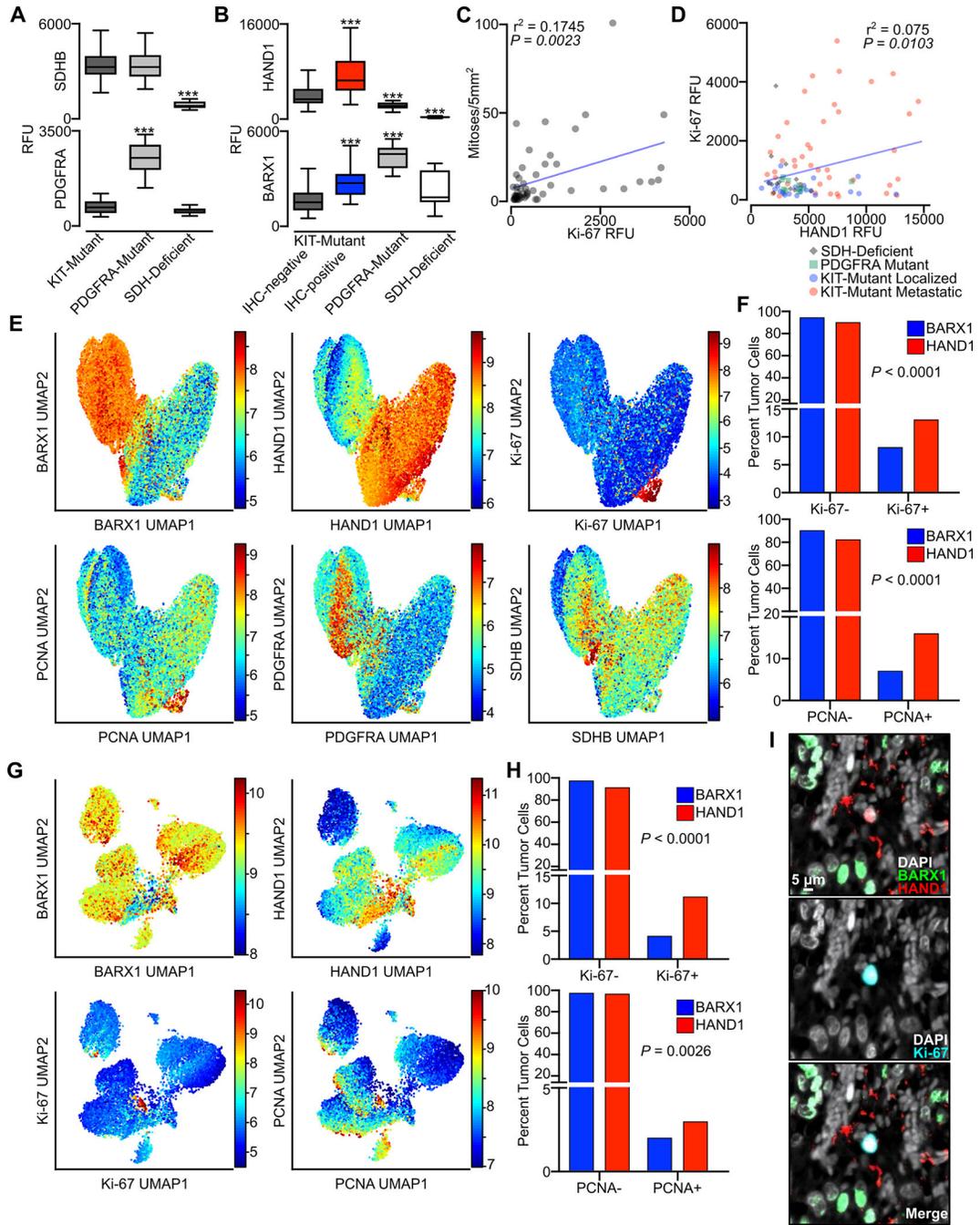
**Figure 4. Tissue level and single cell analysis of accessory TFs and proliferative markers from multiplexed CyCIF imaging.**

**A,** Relative intensity of signal in whole tissue (relative fluorescence units, RFU) of SDHB and PDGFRA in KIT-mutant (*n*=65), PDGFRA-mutant (*n*=10), and SDH-deficient (*n*=12) GIST. Data were analyzed by one-way ANOVA with Tukey's multiple comparison test (compared to KIT-mutant GIST; \*\*\*,*P*<0.001). **B,** Relative intensity of signal in whole tissue for HAND1 or BARX1 in KIT-mutant tumors stratified by IHC-positivity or negativity for each marker, or PDGFRA-mutant and SDH-deficient GIST. **C,** Correlation of

mitotic index and Ki-67 signal intensity for samples with available clinical annotation ($n$=51). **D,** Correlation of Ki-67 and HAND1 signal intensity, with samples distinguished by mutational subtypes ($n$=87). The Pearson correlation for all samples is shown. **E,** UMAP of all tumor samples for single tumor cells (87 tumors, $n$=156,036 cells), defined by the expression of HAND1 and/or BARX1, with heatmaps showing BARX1, HAND1, Ki-67, PCNA, PDGFRA and SDHB $\log_2$ RFU. **F,** Percent of all tumor cells that were negative or positive for Ki-67 (upper panel) or PCNA (lower panel) stratified by BARX1 (blue, $n$=65,340) or HAND1 (red, $n$=81,501) expression. Data were analyzed by Fisher's exact test with $P$ value indicated. **G,** UMAP of all HAND1 and/or BARX1 positive cells from localized gastric GIST (15 tumors, $n$=24,859 cells) showing BARX1, HAND1, Ki-67 and PCNA. **H,** Percent of localized gastric tumor cells that were negative or positive for Ki-67 (upper panel) or PCNA (lower panel) stratified by BARX1 (blue, $n$=21,586) or HAND1 (red, $n$=1,552) expression. **I,** Image showing signal for DAPI (white), BARX1 (green), HAND1 (red) and Ki-67 (blue) in a localized gastric GIST (scale bar, 5 μm).

**A**

| | *n* | HAND1+ | BARX1+ |
|---|---|---|---|
| Esophagus/GEJ | 6 | 2 | 3 |
| Stomach | 211 | 18 | 180 |
| Small Intestine | 140 | 116 | 22 |
| Colon | 7 | 3 | 1 |
| Rectum | 20 | 7 | 1 |
| Unknown | 53 | 31 | 15 |

**B**

| | *n* | HAND1+ | BARX1+ |
|---|---|---|---|
| KIT | 255 | 134 | 95 |
| PDGFRA | 34 | 1 | 34 |
| SDH | 14 | 0 | 10 |
| NF1 | 2 | 2 | 0 |
| Unknown | 132 | 40 | 83 |

**C**

| KIT Mutation | *n* | HAND1+ | BARX1+ |
|---|---|---|---|
| 9 | 30 | 26 | 2 |
| 11 | 193 | 84 | 89 |
| 13,14,17 | 10 | 5 | 3 |
| Compound | 22 | 19 | 1 |

**D**

| KIT mutant tumors | *n* | HAND1+ | BARX1+ |
|---|---|---|---|
| Primary | 171 | 71 | 84 |
| Metastatic/Recurrent | 84 | 63 | 11 |

**E**

| | HAND1+ | HAND1- |
|---|---|---|
| BARX1+ | 28 (6%) | 194 (44) |
| BARX1- | 149 (34) | 66 (15) |

**F** — Relapse Free Survival vs Time (years). HR = 1.84 (0.92-3.68), *P* = 0.1171. Non-High Risk (*n*=28); High Risk (*n*=53).

**G** — Relapse Free Survival vs Time (years). HR = 2.62 (1.14-6.00), *P* = 0.0092. HAND1-/BARX1+ (*n*=40); HAND1+/BARX1- (*n*=21).

**H** — Progression Free Survival vs Time (years). HAND1-negative (*n*=8); HAND1-positive (*n*=41). HR = 2.322 (1.27-4.25), *P* = 0.0127.

**Figure 5. Expression characteristics of HAND1 and BARX1 in GIST and clinical outcomes.**
Expression of HAND1 and BARX1 stratified by anatomic location (**A**), tumor mutation (**B**),
KIT mutation subtype (**C**), and disease status of KIT mutant tumors (**D**) across all samples
(*n*=437). **E,** Frequency of co-expression of BARX1 and HAND1 across all samples (*n*=437).
**F-G,** Kaplan-Meier plots of relapse free survival following primary tumor resection
stratified by risk status (F) or HAND1 and BARX1 expression (G). **H,** Kaplan-Meier plot of
progression free survival on first line imatinib for patients with metastatic GIST stratified by
HAND1 expression.
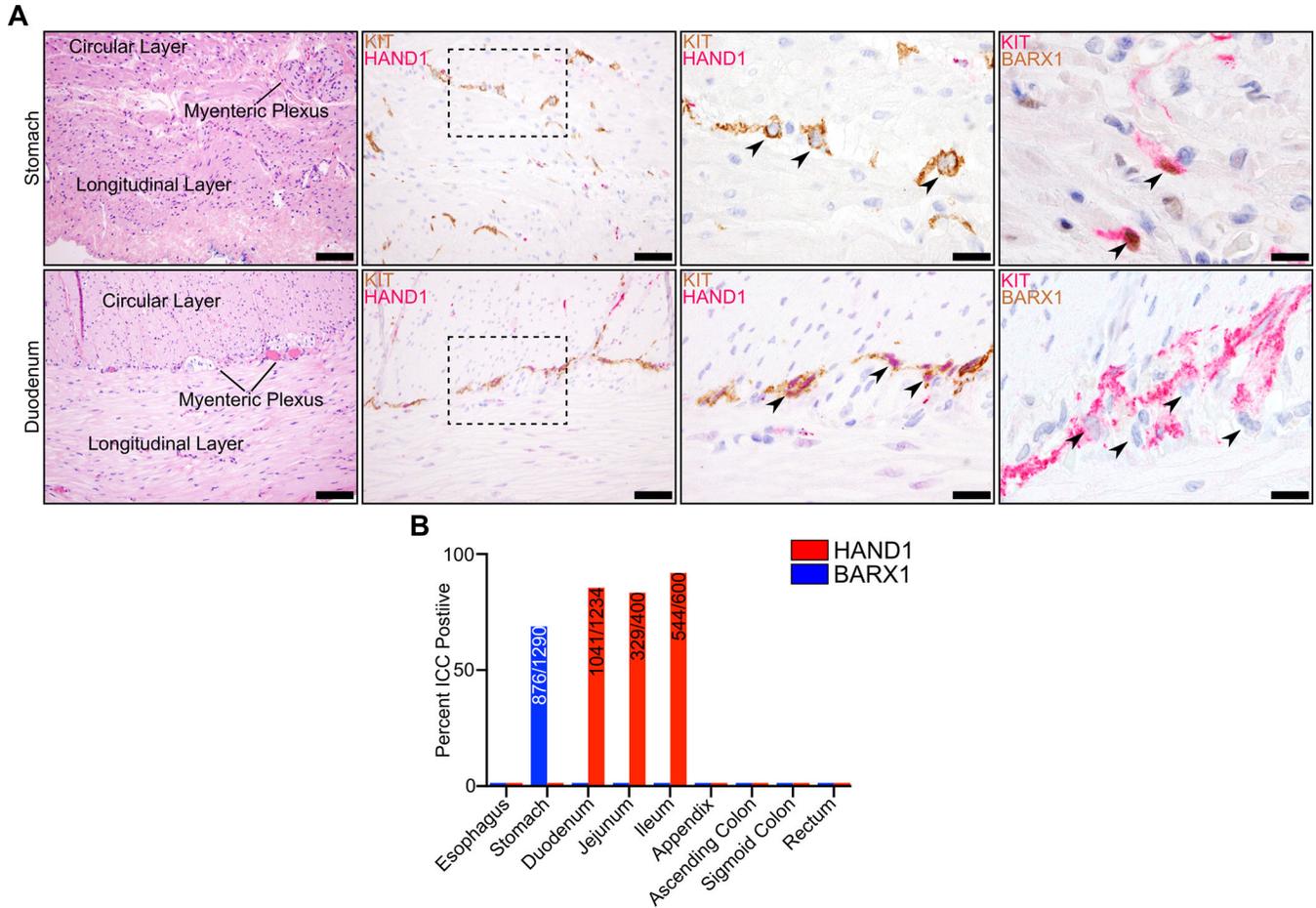
**Figure 6. Anatomic and spatial restriction of HAND1 and BARX1 expression in ICC.**
**A,** H&E staining of normal tissue with muscular layers and myenteric plexus indicated (left, scale bar 100 μm) in stomach (top row) and duodenum (bottom row). KIT (brown) and HAND1 (red) IHC at the myenteric plexus (second from left, scale bar 50 μm) with individual ICC indicated by an arrowhead (third from left, scale bar 20 μm). KIT (red) and BARX1 (brown) IHC at the myenteric plexus (right, scale bar 20 μm). **B,** Percentage of ICC expressing positive for HAND1 or BARX1 in normal gastrointestinal tissues. Inset values indicate the numerator of positively expressing ICC over the denominator of total ICC. For quantification, 100 ICC were evaluated per section, where possible, in at least 7 sections per anatomic location.