Epstein-Barr Virus Selects for Stem Cell Lineages in Breast Cancer and Normal Tissues

Bernard Friedenson

Dept Biochemistry and Molecular Genetics and Illinois Cancer Center, College of Medicine, University of Illinois Chicago, Chicago, Illinois, *bernief@uic.edu*

ABSTRACT

Epstein-Barr virus (EBV) causes persistent damage to breast cancer genomes, so viral particles need not be continually present to cause malignancy. Because basal type stem cells are present in breast cancer but missing in normal breast, the hypothesis was tested that EBV creates or selects for breast cancer stem cells. Bioinformatic comparisons identified methylation of cis-regulators of gene expression in breast cancer chromosomes that closely matched EBV-associated cancers. On certain chromosomes, most of these methylated regulators governed stem cell related differentiation pathways. The shared methylation sites were found in breast cancers typical of the basal layer, where breast stem cells are thought to originate. A few of the same shared breast cancer methylation marks were detected in non-malignant keratinocytes with a cured EBV infection, suggesting EBV epigenetic changes are difficult to reverse. Control experiments with methylation sites on chromosome 17 in (non-EBV) skin cell carcinoma did not correspond to abnormal methylation in breast cancer. Moreover, randomly generated epigenetic sites showed negligible overlap with EBV cancers, so the observed matches are unlikely in non-EBV cancers and unlikely due to chance. Additionally, breast cancer genomes had variable damage to immune responses typical of EBV infection, so immune variation helps explain why not everyone gets EBV cancer. Coupled with other breast cancer safeguards compromised by EBV, epigenetic changes in stem cell related genes that match EBV cancers make the breast more susceptible to cancer triggers.

Introduction

Saliva, blood, breast milk **[1]**, and semen transmit Epstein-Barr virus (EBV). Transmission is so effective that nearly everyone acquires EBV infection at some point in their lives [2]. The infection provides no known benefit to its host and sometimes causes nasopharyngeal cancer (NPC) and Burkitt's lymphoma (BL). NPC and endemic BL do not occur without EBV infection, so they are unequivocal examples of EBV-cancers. EBV attaches to human chromosomes as circular episomes which replicate along with host DNA, but the episomes can be lost during replication in NPC and BL cells [3, 4].

In previous work, NPC and BL were used as cancer genome models to determine whether there are EBV-related changes in breast cancer genomes. Breast cancer genomes were found to have critical chromosome damage resembling EBV-induced damage [5, 6]. Genomic comparisons found that thousands of breakpoints and other lesions in breast cancer closely matched EBV-associated cancers. Many of these chromosome lesions are pathogenic and would accelerate breast cancer even if the infection clears or EBV episomes are lost. This persistent damage explains why EBV particles need not be consistently detected in breast cancer [5, 6].

Breast cancer [7], NPC and BL [8-10] all arise from cancer stem cells - the rare cells within a tumor that can self-renew and initiate new tumors [11, 12]. This common origin suggests the hypothesis that breast cancer arises because EBV disrupts normal breast stem cell lineages. If this is true, lesions in breast cancer that interfere with stem cell lineages might match EBV cancers [5] . The virus becomes most successful if the cells it infects cannot be correctly replaced so EBV selects for cells that have damaged or lost the ability to restore themselves correctly.

Stem cells were previously believed to become malignant through the same mechanisms as somatic cells. Stem cells were thought to be more vulnerable than somatic cells because of their plasticity, self-renewal, and longevity. However, stem cells are either absent or vanishingly rare in healthy breast tissue [13]. Single cell transcriptomics of over 700,000 normal breast cells found no clear evidence of a stem cell

population with a unique transcription profile[13]. No proliferating cells were detected in basal cells or cell states with stem cell markers, challenging the concept that a basal stem cell fuels epithelial homeostasis. In contrast, the ability to proliferate enabled the identification of progenitors in the luminal compartment. One of these populations of progenitor cells aligned with hormone dependent proliferation [13]. Results from transcriptomics of normal breast cells create a paradox when compared to transcriptomics of breast cancers. In contrast to the absence of canonical stem cells in the basal layer of the normal breast, basal layers in breast cancers show strong evidence for stem cells or their descendants. For instance, the cancer subtype containing basal single cells had the presumptive progenitor cell of origin for basal breast cancer [14]. Stem cells were also present in cells in the cancer microenvironment. One state of cancer associated fibroblasts in the cancer stroma "had features of mesenchymal stem cells with high expression of stem cell markers (ALDH1A1, KLF4, and LEPR)." [14]. Perivascular-like cells in the stroma also expressed markers related to stem cells [14] .

EBV infection might resolve this paradox if it provides a mechanism to produce or select for malignant stem cells. EBV infection actively reshapes host cell genomes[5] and epigenomes to select for cells that facilitate viral replication [15]. A precondition for this reshaping is a mechanism of immune escape that releases control of EBV infection. This requirement contributes to explaining why not everyone with EBV infection gets cancer. In NPC and BL, massive changes in host cell methylation [16, 17] then occur as EBV changes accessibility to chromatin, reprogramming cancer epigenomes [18, 19]. Methylation patterns are strong predictors of gene expression levels [20, 21] and unambiguous correlations have been established between cis-methylation sites and genes expressed in breast cancers [22]. Reprogramming can affect stem cell differentiation pathways and the maintenance and regeneration of tissue damage [23]. If this scenario applies to breast cancer, then some breast cancer methylation sites should correspond to those in chromosomes from EBV-associated cancers. The matching sites should be related to genes that regulate stem cell behavior at steps prior to their terminal differentiation. These changes may cause long-lasting damage to breast cancer genomes so that viral particles need not be continually present to cause malignancy.

The research presented below tested these hypotheses. DNA methylation was observed in regions of breast cancer chromosomes that regulate stem cell differentiation, function, and microenvironment. This methylation does not exist in normal cells, but was found in identical or virtually identical positions in prototype EBV-cancers. The agreement in completely unrelated data was so strong and so consistent with established effects of EBV that it supports the hypothesis that EBV selects for epigenetic modifications in stem cells. This selection prevents replacement of abnormal cells and facilitates viral replication. Genes affected include HOX genes and those encoding for morphogens. Differentiation pathways become blocked or deregulated. At least one large subset of breast cancers in the cohort had diverse immune escape mechanisms [24], so serious that they would allow EBV infection. Malignant transformation also creates a homing microenvironment that facilitates growth and multiplication of cancer stem cell seeds in multiple tissues. This milieu provides nutrition, cytokines, and protection from immune responses [25].

Results

Methylation sites in breast cancer that control critical genes align with methylation sites in EBV cancers.

The minimum distance between abnormal methylation at breast cancer gene control sites and the EBV cancers NPC and BL was calculated for all 22 human autosomes. The most frequent agreement was within 10 base pairs and included the two EBV cancers (Fig. 1, blue and orange bars in 16 of 22 chromosomes). The positions of the shared methyl groups were sometimes exactly the same.

Other calculations also showed that chance occurrence was unlikely and supported similarity of abnormal methylation distributions in breast and EBV cancers. For example, comparing 2319 abnormal methylation positions in NPC to 789 breast cancer positions on chromosome 2(242,696,762 bases) found 15 positions that agreed within 10 base pairs while 0.158 matches would be expected by chance. The probability of observing $k = 15$ or more matches with $\lambda = 0.158$ under a Poisson distribution model gives $p \approx 1.3 \times 10^{-27}$. This value is so low it strongly suggests the observed agreement is not

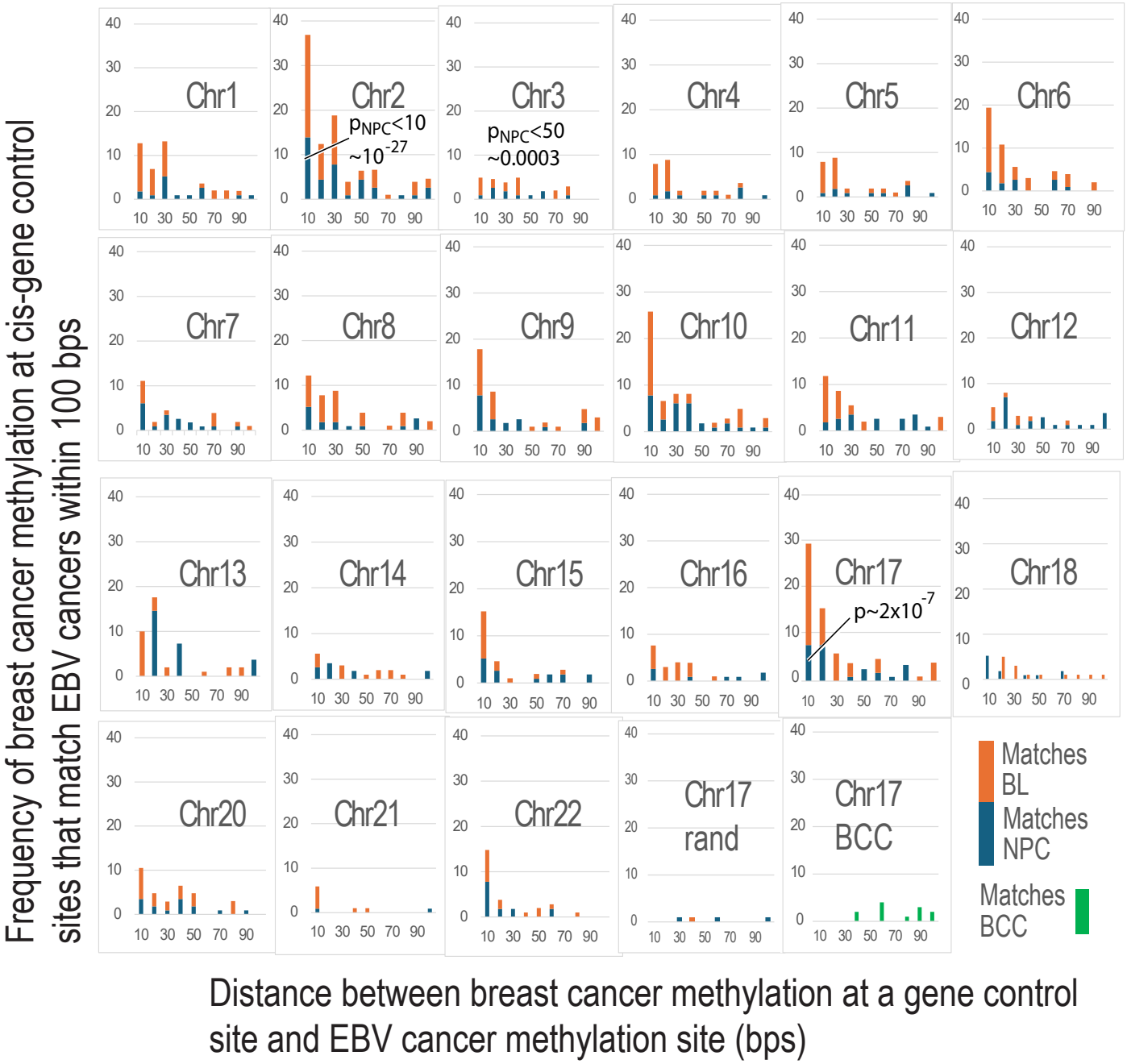

Fig. 1. Minimum distance between abnormal methylation sites in breast cancer and EBV cancers. On most chromosomes, abnormal breast cancer methylation sites are within 10 base pairs of an abnormal methylation site in the EBV cancers NPC and BL. Comparison of abnormal methylation sites on chromosome 17 to a random distribution (lower right) in a sample trial shows almost no matches, while other trials were blank. The non-EBV cancer BCC (green bars) also does not resemble the matches on chr17. Multiple comparisons were used to find values that agreed and P values for the comparisons with NPC methylation sites are shown for chromosomes 2, 3, and 17.

random. The same calculation using 9 matches that were found <10 bases apart on chromosome 17 gave $p \approx 2x10^{-7}$. Chromosome 3 had fewer matches than either chromosome 2 or 17 but 5 matches were <50 bases apart and still statistically significant ($p \approx 0.0003$). In addition Mann-Whitney comparisons of the entire distributions of abnormal methylation in breast cancer vs. NPC on seven chromosomes (2, 7, 9,10-12, and 20) give high P-values, which indicate that there is no significant difference between the distributions of abnormal methylation groups. On chromosomes 9, 10, and 11, the confidence interval for the median difference included zero, further suggesting that the medians were not significantly different. Even when the methylation sites on a chromosome near each other were sparse, Mann-Whitney comparisons still could not exclude the possibility that abnormally methylated DNA base distributions were the same.

Ten trials using random data rather than breast cancer data produced blank or nearly blank results (e.g. Fig. 1, lower right). Another control compared these abnormal methylation sites to 13,766 methylation sites on chromosome 17 [26] in basal cell carcinoma of the skin (BCC). The skin cancers had not been corrected by subtracting methylation in normal cells, so much larger numbers of methylation sites had to be compared to breast cancer. Nonetheless matching sites were much rarer and weaker in comparison to breast cancer vs NPC or BL (Fig. 1, lower right) . Mann-Whitney testing of abnormal methylation sites in breast cancer and BCC gave statistical evidence that the distributions were different (two-sided $p<0.0001$ with a 60.7% probability that a randomly selected value from BCC is greaten than one from breast cancer).

These results show that breast cancer chromosomes have frequent abnormal methylation at virtually the same positions as cancers known to be caused by EBV. The sharing of methylation loci between breast and EBV cancers suggests the abnormal methyl groups are essential to produce all three cancers. To characterize this effect, the genes in breast cancers established as controlled by cis-methylation [22] at loci matching EBV cancers were next determined.

Stem cell related gene controls predominate at sites methylated in the breast cancer positions that match EBV cancers.

The next analyses asked whether there were identifiable functions that were preferentially targeted in the breast cancer methylated gene regulators that matched EBV-cancers. Chromosome 17 was the first chromosome evaluated because it encodes the *BRCA1* and *ERBB2 (HER2)* genes, which both have significant roles in breast cancer. Chromosome 17 in the breast cancer cohort [22] had 405 gene entries with abnormal methylation at their regulatory sites and 85% of sequences that regulated these genes were <=200 bp in length.

Methylation sites on chromosome 17 in breast cancer [22], NPC [18], and BL [19] that were not present in normal cells were compared in detail. The first hint that abnormal methylation affected the same functions in the three cancers was found in regulatory regions of *SOX9* (Fig. 2). *SOX9* counterbalances the effect of the core stem cell gene *SOX2*, which regulates pluripotency. In cancer, *SOX9* and *SOX2* are epigenetic regulators that together determine cancer cell plasticity and metastatic progression [27, 28]. In addition, an intricate signaling channel network governs the self-renewal and lineage determination of mammary stem cells. *SOX9* is in this network and controlled by *NOTCH* signaling [29]. *KRT19*, a control point for *NOTCH* signaling, is methylated at positions within 3 bases in breast cancer and BL. Breast cancer and EBV-related cancers had another link to stem cells through shared methylated clusters that affected homeobox genes *HOXB2* and *HOXB4*. *HOX* genes produce a code that contributes to driving tissue stem cells to differentiate into the correct lineages in order to maintain and repair their respective tissues and organs [30]. *HOX* genes are often regulated by morphogen gradients (e.g. Wnt, BMP, FGF, RA, Shh) and are inhibited by epigenetic marks until functional differentiation begins at gastrulation. The marks are removed on differentiation and *HOX* genes become active to guide development of various body regions [31]. As shown in the first panel in Fig. 2, *HOXB2* and *HOXB4* control sites in breast cancers with abnormal methylation at sites that closely correspond to EBV cancers.

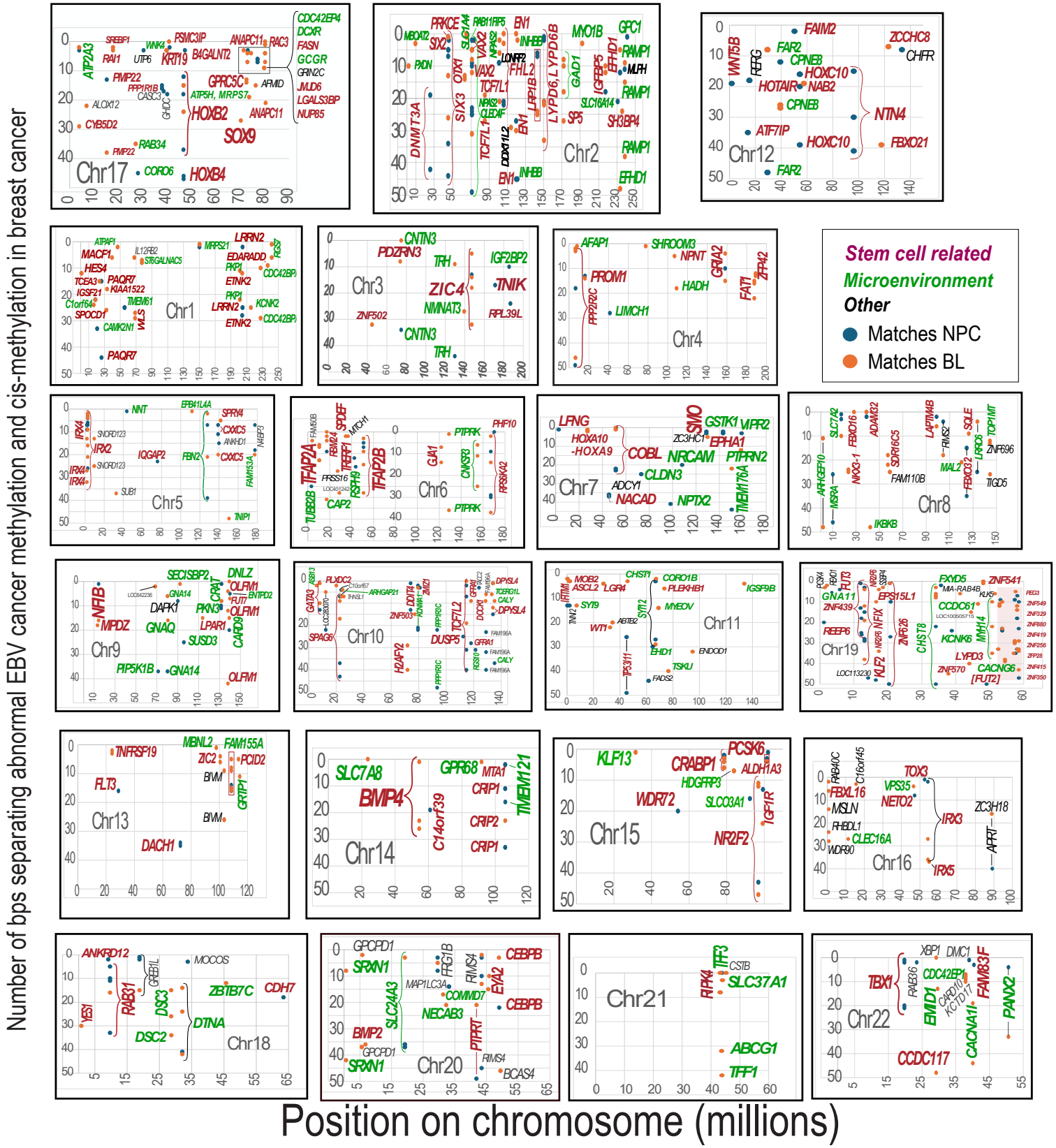

Fig 2. Methylation of DNA controlling breast cancer stem cell genes is different from normal and closely matches the EBV cancers NPC and BL (blue and orange dots, respectively. See the legend in the second row). Genes related to stem cells (red text) or those controlling the environment (green text) are the predominant genes in breast cancer that are close to positions with abnormal methylation in EBV associated cancers. Genes that affect other functions are indicated in black. Distances on the vertical axis represent positions of abnormal cis-methylation in breast cancer and EBV cancers that agree within <50 bp. Many positions are identical or almost identical for all three types of cancers (top of each panel). The three chromosomes in the top row are described in the text. The red shaded area in chromosome 19 for zinc finger proteins infers matches, but the data is incomplete.

As a control, the number of CpG islands in NPC DNA sequences that matched breast cancer sequences were calculated as 100 times the numbers of CpG islands in DNA divided by the length of DNA. All the comparisons at any window size involved dense numbers of CpG islands in the differentially methylated DNA, ranging from 24 to 7.4, that did not depend on the length of DNA. *HOXB2* gave a value of 16.5 and *NUP85* gave a low value of 7.4. Other regions on chromosome 17 gave values as low as 1.31. The lengths of matching DNA sequences were as low as 42 base pairs (*HOXB2*). These results show that relatively short CpG-rich DNA segments were selected for unusual methylation in known EBV mediated cancers and some of these segments closely matched breast cancer cis gene control regions.

Finding *SOX9, KRT19,* and *HOX* genes with abnormal methylation at about the same positions in both breast and EBV cancers raised the possibility that EBV induces methylation that interferes with breast stem cell differentiation. Data in Fig. 1 hint at this possibility. Although young and old stem cells are similar, the numbers of stem cell and their potential cell fates decline with age. The number of abnormal methyl groups in children with BL was almost always higher than in those that matched adults with NPC. Based on this observation, subsequent analyses pursued the possibility that EBV had also targeted stem cells on the other autosomal chromosomes in breast cancers.

Breast cancer stem cell genes are selected at methylation sites that match EBV-cancers.

Essentially identical abnormal methylation positions in breast cancer and EBV cancers extended to other chromosomes with preferential targeting of stem cell related genes or transcription factors that influence cell lineage. Supplementary Table S1 compiles the functions of about 350 of these genes. As shown by the red text in Fig. 2, breast cancer gene promoters and enhancer regions are within 50 bps of methylated positions in EBV cancers. These matching methylated gene regulators affect morphogen responsive *HOX* genes and genes with other associations to stem cells on all twenty-two non-sex chromosomes.

For chromosome 17 (Fig. 3 and supplementary Table S1), some relationship to stem cells was found for about 66% of abnormal methylated positions in breast cancer cis gene control positions corresponding to EBV cancers. Across all 22 chromosomes, 48% of these genes were related to stem cells or to pathways essential for stem cell differentiation. Multiple chromosomes showed methylation controls for genes that disrupt cell fate decisions, self renewal, and tissue homeostasis. A few examples below illustrate these effects.

Chromosome 2 genes with altered controls in breast cancers matching EBV cancers had strong relationships to stem cells (Figs 2,3 and Table S1). Genes that shared methylation loci included *SIX2*, *SIX3*, *VAX2*, and *EN1*. While *HOX* genes are master regulators of stem cell destinations and identity, *SIX, VAX,* and *EN* genes participate in finer regulation in specific tissues and regions during development. All these genes nevertheless activate or suppress responses to morphogen gradients during development [GeneCards] and to maintain tissue homeostasis [32]. A different control exists on chromosome 12 where the *HOTAIR* gene contributes to regulating *HOX* expression. *HOTAIR* is 2200 bases of antisense non-coding RNA that inhibits *HOXD* cluster transcription. *HOTAIR* binds to specific chromatin modification complexes to regulate chromatin. Interactions with *HOTAIR*/*PRC* control *HOXD* gene transcriptional activity on chromosome 2.

Supplementary Table S1 gives 167 instances of abnormal methylation that mark stem cell related functions or differentiation pathways with significant specificity in the breast, nasopharynx, or hematopoietic system. These examples suggest that pluripotent cells in breast cancer may have been infected and modified by EBV.

Differentiation pathways affected by similarities in abnormal methylation

To further investigate the effects of EBV on stem cell pluripotency, the frequency was estimated of differentiation pathways with abnormal EBV-related methylation sites. Six chromosomes were tested based on numerous close (<10 bp) matches to EBV cancers in positions of abnormal methylation and the known distribution of *HOX* genes. As shown

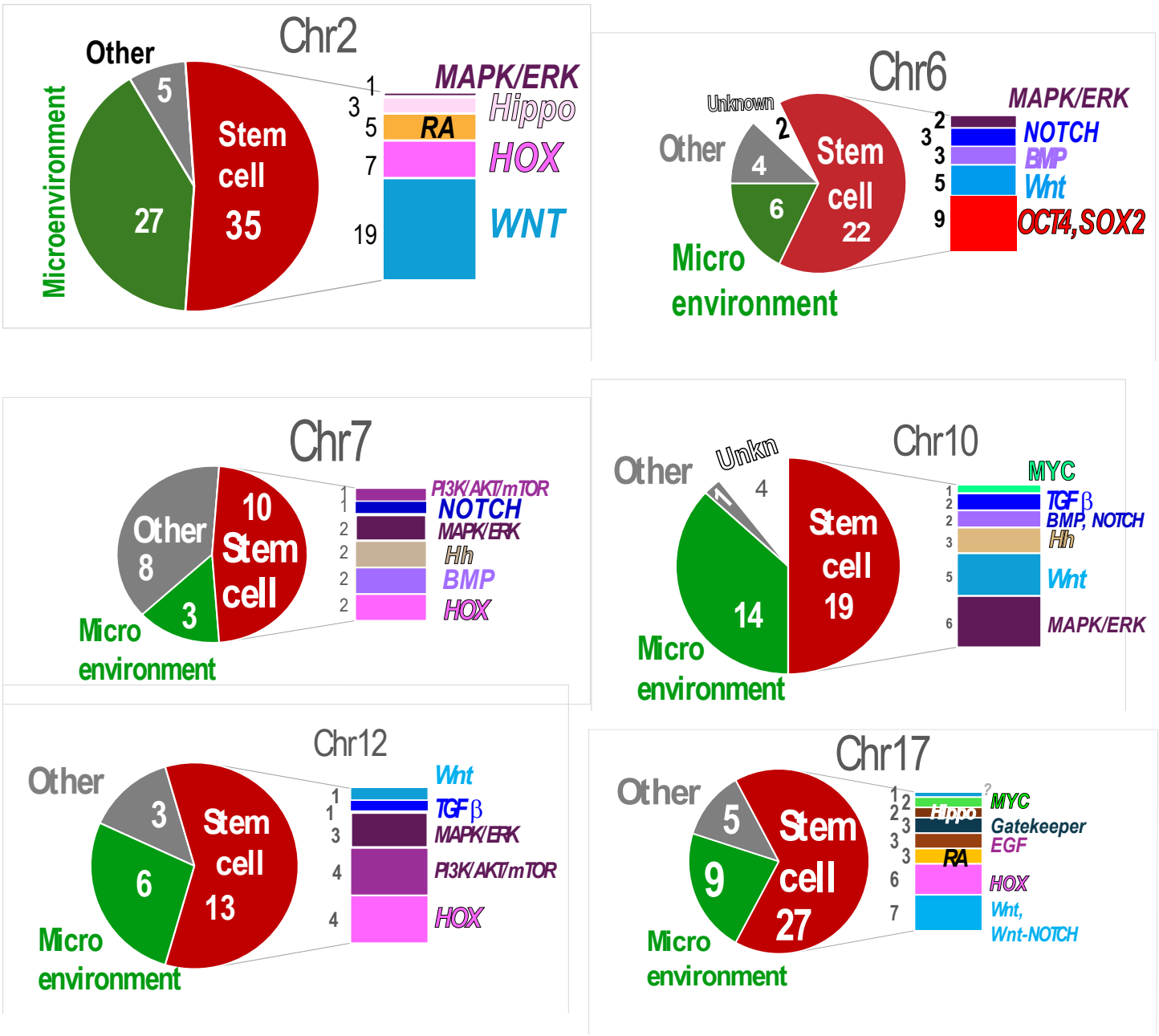

Fig. 3. On some chromosomes, most breast cancer gene control regions with abnormal cis-methylation corresponding to EBV cancers affect stem cell differentiation pathways and morphogenesis. Occurrences of shared abnormal methylation loci impacting stem cells, microenvironment or other properties are indicated on the pie charts. Differentiation pathways most frequently affected by the breast cancer cis-methylation sites that match EBV cancer are shown in bar charts to the right of the pie charts. The number of times a given stem cell differentiation pathway is affected by shared abnormal methylation is given at the left of the bar chart. Abnormal breast cancer methylation on chromosome 6 that matches EBV cancers directly affects the pluripotency genes OCT4 and SOX2.

in Fig. 3, morphogen associated pathways had abnormal methylation, but the sets of pathways differed on different chromosomes. Only the predominant morphogen connected pathways are indicated, but abnormal methylation usually had links to multiple additional pathways in different contexts. On chromosome 6, the stem cell genes *OCT4* and *SOX2* were the most frequent targets. These two gene cooperate and are essential to reprogram somatic cells to maintain and regain pluripotent stem cells. Their expression is under tight control, and their deregulation can collapse the entire pluripotency network [33, 34]. In addition, *HOX* gene controls were abnormally methylated on four chromosomes, where they were frequent targets. Other genes connected to Wnt and other stem cell differentiation pathways are also predominant targets. These results show that abnormal methylation in breast cancers at sites closely related to EBV can have a profound effect on stem cell pluripotency and the ability to regenerate tissue damage.

Effect of methylation on gene expression at breast cancer gene control sites related to EBV.

In the METABRIC cohort, correlations exist between cis-specific (same chromosomes) methylation and expression "at hundreds of promoters and over a thousand distal elements" [22]. Effects of these cis-methylations near those in EBV cancers were tested. Expression levels were available for 20,603 genes from 1980 breast cancer patients in the METABRIC cohort [35] . As shown in Fig. 4, abnormal methylation on established cis gene control regions and related to EBV cancers more often caused over-expression of stem cell related genes, but under expression also occurred frequently. The most highly over-expressed and under-expressed genes (>50 instances) were labeled in Fig. 4 and they include specific links to cancer development in the context of stem cells. Table 1 groups these abnormally expressed genes according to whether they are generally pluripotency factors, morphogens, homeobox genes, or other differentiation regulators in stem cells.

Some over-expressed genes (highlighted in red in Fig. 4) can contribute to creating breast cancer stem cells, either leaving cells in a progenitor-like state so that they do not

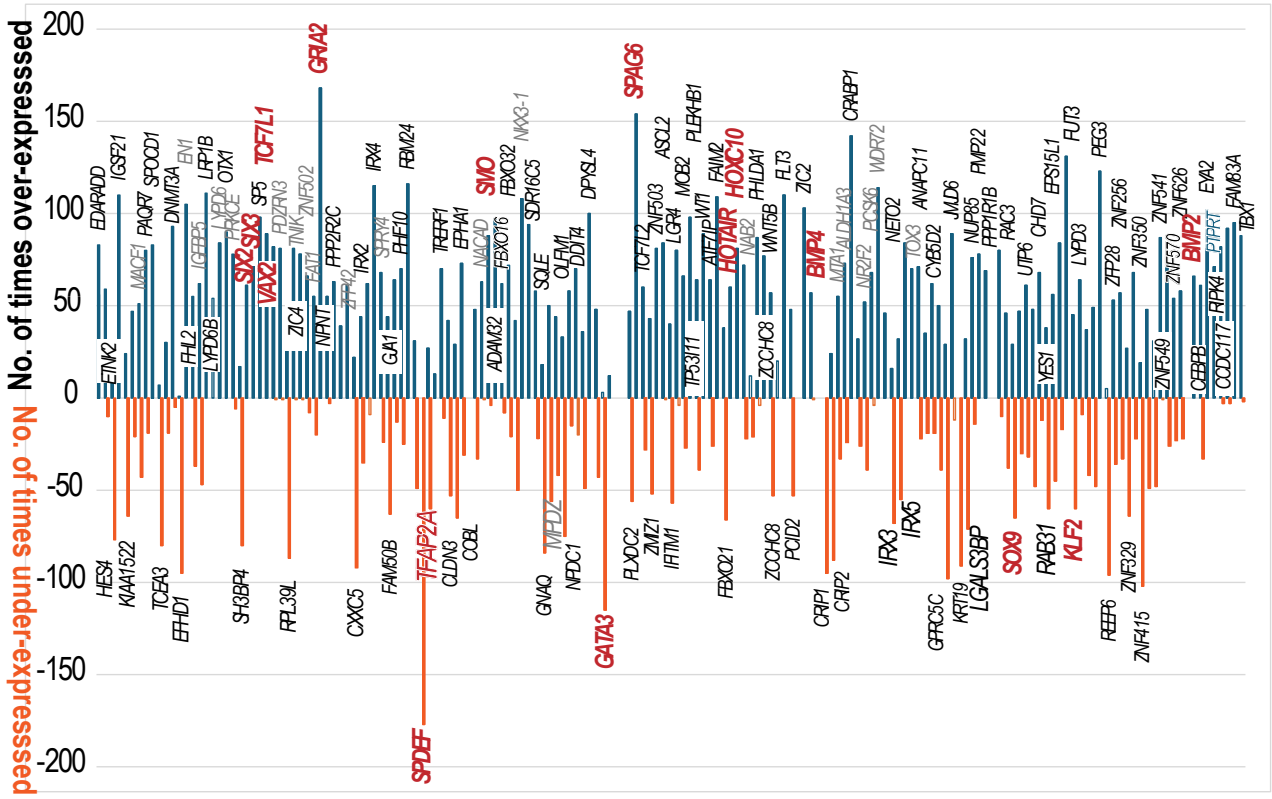

Fig. 4. Genes with abnormal methylation in breast cancers matching EBV cancers (+- 50 base pairs) are over or under expressed. The number of times these genes are over or under expressed in the METABRIC cohort is shown on the vertical axis. Genes over- or under-expressed more than 50 times are labeled on the 22 autosomes. Genes mentioned in the text as examples are highlighted in red.

mature into specialized types or reprogramming them to ectopic expression [36]. The genes labeled as pluripotency regulators include *SOX9* and *KLF2*. *SOX9* opposes the pluripotency factor *SOX2* and pushes cells to differentiate [37], while *KLF2* maintains self renewal and stemness [38]. Overexpressed *TCFL1* in stem-like cancer cells keeps them in a pluripotent state, unable to differentiate by silencing Wnt targets [39].

Other highlighted over-expressed genes include morphogens (e.g. *BMP2* on chromosome 20*, BMP4* on chromosome 14, and morphogen responsive genes (e.g. *SMO* on chromosome 7) that direct differentiation into specific cellular identities. On chromosome 2, *SIX2* [40], *SIX3* and *VAX2* are over-expressed homeobox genes that help translate positional information from *HOX* genes into development.

On chromosome 6, abnormal methylation causes under-expression of *TFAP2A.* Normally TFAP2A combines with OCT4 and SOX2 proteins to initiate exit from pluripotency [41], so methylation of *TFAP2A* regulatory elements locks cells into a self-renewing state. In a different context, *TFAP2A* also cooperates with *GATA2/3* and *TFAP2C* in the "Tetra" network. This network represses pluripotency genes like *OCT4* and couples this repression with initiating differentiation toward the placental support system. Under-expressed *TFAP2A* disrupts this network, preventing proper lineage commitment [42]. *TFAP2A* binds to epigenetically inactive placental genes when their differentiation begins. Without *TFAP2A* contributing to activating these genes, they remain silent, and differentiation stops.

HOXC10 (located on chromosome 12) expression is normally low in most differentiated tissues, but its over-expression can activate *ERK, AKT,* and *p65* oncogenic pathways and epithelial mesenchymal transition-related genes [43]. *HOTAIR* is encoded within the same *HOXC* gene cluster on chromosome12. *HOTAIR* over-expression causes histone H3K27 methylation, which increases breast cancer invasiveness [44].

Further examples are also included and summarized in Table 1. Over-expressed and under-expressed genes at EBV methylation sites on any of multiple chromosomes can

reprogram transcriptional networks, affecting hundreds of downstream genes involved in differentiation and malignancy.

Control Experiment showing both ER-Negative and ER-positive breast Cancer Cells have methylation in positions similar to EBV-cancers.

Many estrogen receptor (ER)-negative breast cancer cells are believed to originate from the basal layer of mammary epithelium. Basal-like breast cancers typically lack expression of ER, PR (progesterone receptor), and HER2, and are often referred to as triple-negative breast cancers. They are highly heterogeneous and notoriously hard to treat. These basal cells are located in the outer layer of the mammary ducts and are distinct from luminal cells, which typically give rise to ER-positive cancers. The basal layer contains stem and progenitor cells as candidates for the origin of aggressive subtypes of basal-like breast cancer.

Conventional stem cells are absent from the basal layer of normal adult breast tissue [13], yet breast cancers exhibit stem-like cells that are the presumptive origin of basal breast cancers [14]. Both normal and cancer datasets agree that there are stem like cells in the luminal layer. If EBV selects or produces basal stem cell phenotypes, then ER negative and ER positive breast cancers should have abnormal methylation in stem cell genes that matches methylation sites in EBV cancers. As shown in Fig. 5, ER-negative breast cancers have from 10 to 52% of methylation near methylation sites in EBV cancers on various chromosomes. Thus EBV methylation sites may select or produce cancer stem-like cells that originate breast cancer in both basal type and luminal layers.

Non-malignant cells from people who have recovered from an EBV infection exhibit abnormal DNA chromosome methylation in a few of the same loci as breast cancer cells.

To what extent do the results thus far apply to non-malignant cells with an active or a resolved past EBV infection? Are gene control sites methylated by viral infection in non-malignant cells related to those in EBV cancer? DNA methylation in EBV-infected oral keratinocytes differs from uninfected oral keratinocytes and produces DNA scars after the virus is cleared [45]. To determine if these infection related methylation scars were related

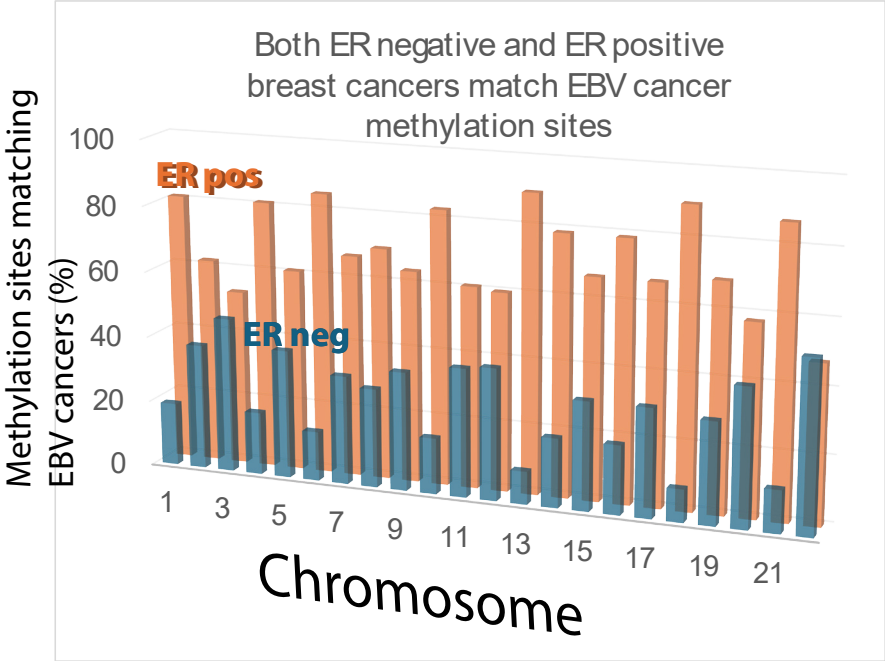

Fig. 5. Both ER negative (typically basal type) and ER positive (typically luminal type) breast cancers have abnormal cis-methylation in sites that closely match cancers associated with EBV. This is true for all 22 autosomes and the matching sites are predominantly stem cell related genes.

to breast cancers, the distributions of methyl groups in breast cancer cells were compared to cured, but previously infected oral keratinocytes.

Linear regression analysis revealed a correlation between epigenetic marks in these cell types (r=0.93, r2=0.87, p<0.0001) with a power for 5% significance of 99.99%. Kolmogorov-Smirnov tests (p=0.89) and Spearman's rank correlation (p=0.88, CI=0.72-0.95) gave consistent results. These results show that methylation in previously infected non-malignant cells correlates with methylation in breast cancer cells. Again these results are unlikely due to pure chance. For example, chromosome 22 has 299 methylation sites in oral keratinocytes. Methylation that controls the RAB36 gene is 7 bases from one of 136 methylation sites in breast cancer (Fig. 6) and other genes that align are closer. The number of sites that agree expected by pure chance is 15X136x299/51324926=0.0119 which gives a p value of 0.012, assuming a Poisson distribution. The p values for chromosome 22 show that pure chance cannot explain the relatively rare agreement between non-malignant keratinocytes with a past EBV infection and breast cancers.

Next, the exact genome coordinates were compared. As shown in Fig. 6 for cells from cured EBV infections, 14/22 autosomal chromosomes had methylation scars close to methylated loci that affected stem cell gene control in breast cancer. A few examples illustrate clear similarities. For chromosome 2, 20 of the 27 genes with abnormal methylation also had at least some relation to stem cells. On chromosome 12, the *HOXC* gene and *HOTAIR*, the *HOXD* controller are abnormally methylated, showing a clear relationship to the breast cancer stem cell data in Fig. 2. The *NFIX* gene on chromosome 19 had abnormal methylation in breast cancer, EBV cancer, and EBV infection. *NFIX* inhibition participates in both cancer and infection and may be essential for EBV pathogenesis. *NFIX* suppresses breast cancer cell proliferation by delaying mitosis through downregulation of *CDK1* expression. Eight chromosomes had new methylation loci that were still related to breast cancer stem cell gene controls (Supplementary Table S2).

These results show that EBV methylates non-malignant infected keratinocytes in much the same ways as it does in malignant breast cells even if the infection clears.

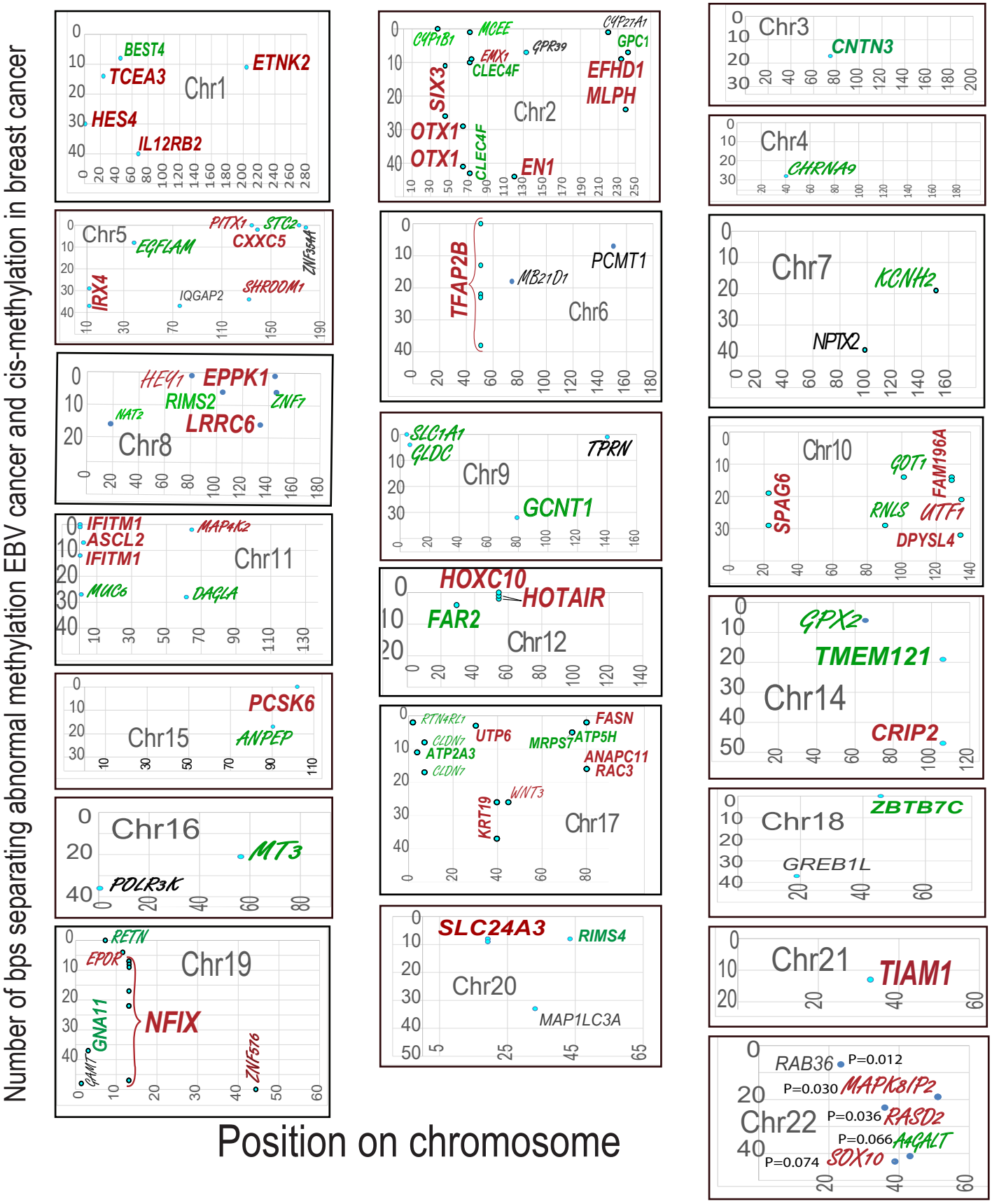

Fig. 6 Some DNA gene control elements have abnormal cis-methylation in breast cancer that matches non-malignant oral keratinocytes that have recovered from an EBV infection. Positions that match either NPC or BL cancers were compared to abnormal methylation in breast cancers. The abnormal methylation sites focus on stem cell gene controls (red) and genes that affect the microenvironment (green). Other genes with matching methylated controls are indicated in black. These types of genes that are methylated in non-malignant oral keratinocytes but not in breast cancers are indicated by the script font. For chr22, p values show that the matching values include those that are unlikely due to pure chance.

Persistent methylation marks on stem cell gene regulators in non-malignant keratinocytes suggest that EBV associated methylation scars persist and are not easy to completely reverse.

Control experiment shows abnormal methylation at sites essential to produce basal cell carcinoma of skin keratinocytes are not associated with EBV effects on non-malignant oral keratinocytes.

Skin basal cell carcinoma (BCC) has no known relationship to EBV and was used as a control to test the association between methylation marks and EBV cancers in keratinocytes. Nearly 7000 hypermethylation sites in BCC [26] were compared to abnormal methylation sites in oral keratinocytes with a cured EBV infection. Fig. 7 shows that numbers of methylation sites in BCC (unrelated to EBV) near EBV- associated normal oral keratinocytes generally agreed less often than expected by chance. Using a 10 base pair window gave 1 match with 1.32 expected by chance alone. Assuming a Poisson distribution gives a p value of 0.73 (a 73% chance of seeing at least one match by chance alone). The p value rose to 0.27 at a 50 base pair window, which remained statistically non-significant.

Breast cancer mechanisms of immune escape release EBV restraints to help explain why not everyone gets EBV cancers.

EBV infects over 90% of people worldwide, but only a small fraction develop EBV-associated cancers. Variability in the activity of immune related genes may contribute to why only some people develop breast cancer. To test this idea, modification and expression of the genes related to immune responses were compared among the data for breast cancer patients.

As shown in Fig. 8A, 358 genes were near methylation sites in EBV cancers and 84 (23%) of the genes had functions related to the immune response. The 84 genes that impacted the immune system were mainly divided between those related to stem cells (40/84 genes) and those related to stem cell microenvironment (32/84 genes).

To further assess the influence of cis-methylation on immune-related genes, Batra's 4143 autosomal genes controlled by cis-methylation [22] were compared to 1480

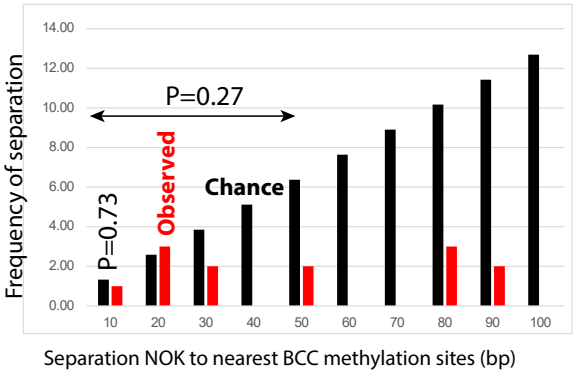

Fig 7. The frequency of methylated genes in normal oral keratinocytes that have recovered from an EBV infection (NOK)s within 50 bp of methylation sites in basal cell carcinoma of the skin is generally less than expected by chance.

immunity genes [24] using METABRIC data from an ER+/PR-/HER2- subset of breast cancers. From the 4143 cis-methylation controlled breast cancer genes, 154 were related to the immune response in the subgroup and these are labeled in Fig 8B. Many of them are in tight clusters with 22 immune related genes near cis-methylation sites shared with EBV cancer chromosomes. If additional members of the same stem cell differentiation pathway are included based on a compendium of oncogenic pathways [46], the total rises to at least 50 immune genes. Of the original 22 immune genes, 19 were related to stem cell functions or conditioning the microenvironment.

To characterize potential consequences of these methylation anomalies, immune escape mechanisms were compared between ER+/PR−/HER2− subgroups in the METABRIC cohort and EBV-associated cancers [24]. Table 2 shows four obvious immune escape mechanisms in breast cancers that are also associated with EBV infections.

In contrast 161 genes expressed in immune cells have been associated with breast cancer. These associations provide insight into tissue specific effects of gene expression[47]. The 161 immune related genes matched 97 METABRIC genes with expression values available and Fig 8B shows wide variation in their expression in different tumors. For example, as shown in Fig 8C, *APOBEC3A* and *APOBEC3B* are significantly overexpressed in a few breast cancers. The consequences of this overexpression have been linked to breast cancer tumor evolution, therapy resistance, and increased tumor mutational burden [48, 49]. Both *APOBEC3* genes are overexpressed in EBV infection and *APOBEC3A* is also overexpressed in malaria [50]. The two enzymes deaminate cytosine to uracil in single-stranded DNA, especially during DNA replication or repair. This leads to C-to-T and C-to-G mutations, especially in TCA and TCT motifs. Signatures correlating with ABOBEC overactivity are common mutations in breast cancer genomes [49].

These results show that methylation like that in EBV-related cancers is significant in shaping the immune response. The overexpression of *APOBEC3* genes is one force in a few breast cancers that extends the effects into oncogenic pathways. Immune imbalances are reiterated from EBV, further increasing chances that additional triggers will result in cancer. For example, chronic malaria exposure triggers BL in EBV-infected children. As shown in Fig. 8D, the genetic backgrounds for all these effects can vary widely. The approximately 20,603 genes measured in the cohort are over or under

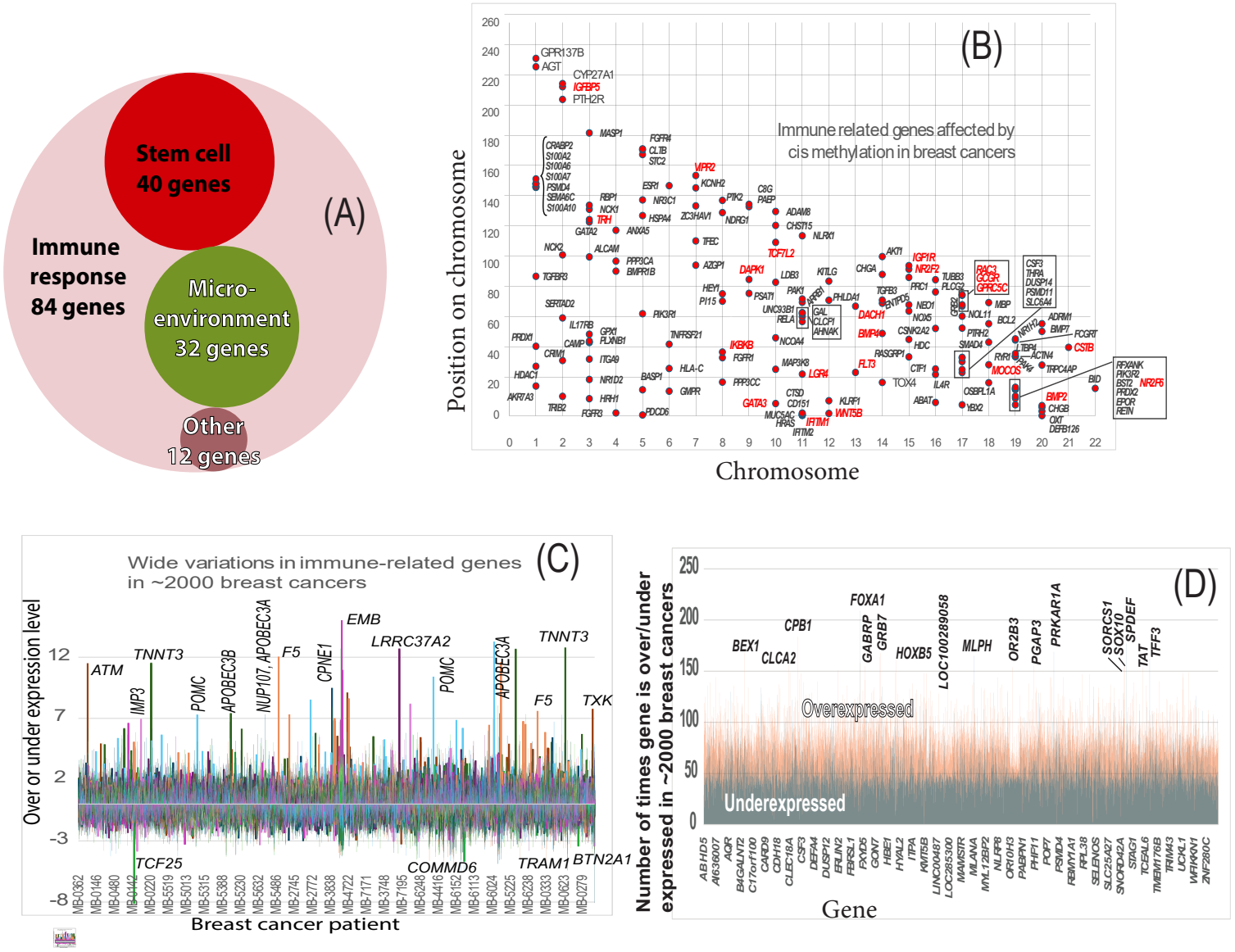

Fig. 8A. Control of genes in breast cancer by cis-methylation related to stem cells, the microenvironment, and other functions overlaps immune responses. This overlap occurs in about 23% of the cis methylated breast cancer gene controls near abnormal EBV methylation sites. Additional mechanisms of immune compromise are also characteristic of EBV infection. Fig. 8B. Immune response genes are affected by cis methylation in breast cancer. 4143 breast cancer genes under cis-methylation control include immune related genes on every chromosome that often cluster together. Fig 8C. Expression of immune-related genes varies widely in breast cancers APOBEC3B and 3A over-expression is common among breast cancers and both genes are significantly over-expressed in a few breast cancers. APOBEC3A and 3B are linked not only to breast cancer but also to EBV infection. Fig. 8D. Wide variation in gene expression in breast cancers is typical of all the breast cancers in the cohort. Only small percentages of the total 20603 genes are over or under expressed. This creates highly variable backgrounds for the development of breast cancers and contributes to explaining why not everyone with EBV infection gets breast cancer.

expressed in only low percentages of the ~2000 breast cancers. These varying environments create different breast cancer backgrounds and contribute to explaining why not everyone with EBV infection gets breast cancers.

The abundant EBV-like sequences in human chromosomes represent a small section of EBV DNA.

To study the feasibility of a future vaccine against EBV, human chromosomes 3, 8, 11, and 17 DNA were arbitrarily selected for comparison to all known and available viral sequences. As shown in Fig. 9, many positions along the entire lengths of the four chromosomes resemble EBV and retroviruses, but the distributions of virus-like sequences differ. Chromosome 8 had the strongest resemblance to EBV with 6014 matching sequences (EBV homology scores over 500, indicating 355-500 identical bases). Because EBV-like sequences are so abundant on human chromosomes, EBV can mimic human sequences. The following five EBV-like human sequences have oncogenic properties in NPC [51]:

- *LMP2A* and *LMP2B* regulate the host aryl hydrocarbon pathway [52] and maintain latent infection.
- *A73* correlates with NPC occurrence [53]
- *BALF4* is a virion envelope glycoprotein in spikes that attaches to host cell surface proteoglycans. The viral and cellular membranes then fuse, enabling virus entry. *BALF4* dramatically enhances human cell infections [54]
- *BALF3* is an endonuclease that mediates mature virion production and packaging during the EBV lytic cycle[55]. The BALF3 protein causes double-strand breaks and micronuclei [56].
- *LMP-1* reprograms infected cells, preventing differentiation and inducing proliferation and inflammation [57]

This result shows that thousands of full or partial copies of EBV-like sequences relate to human genome instability (Figs. 9A, B, C) and emphasizes how dangerous EBV infection can become. Fortunately, human DNA only matches about 2500 bases within the 173,000 bases in EBV genomes. Only a few of all possible EBV genetic regions

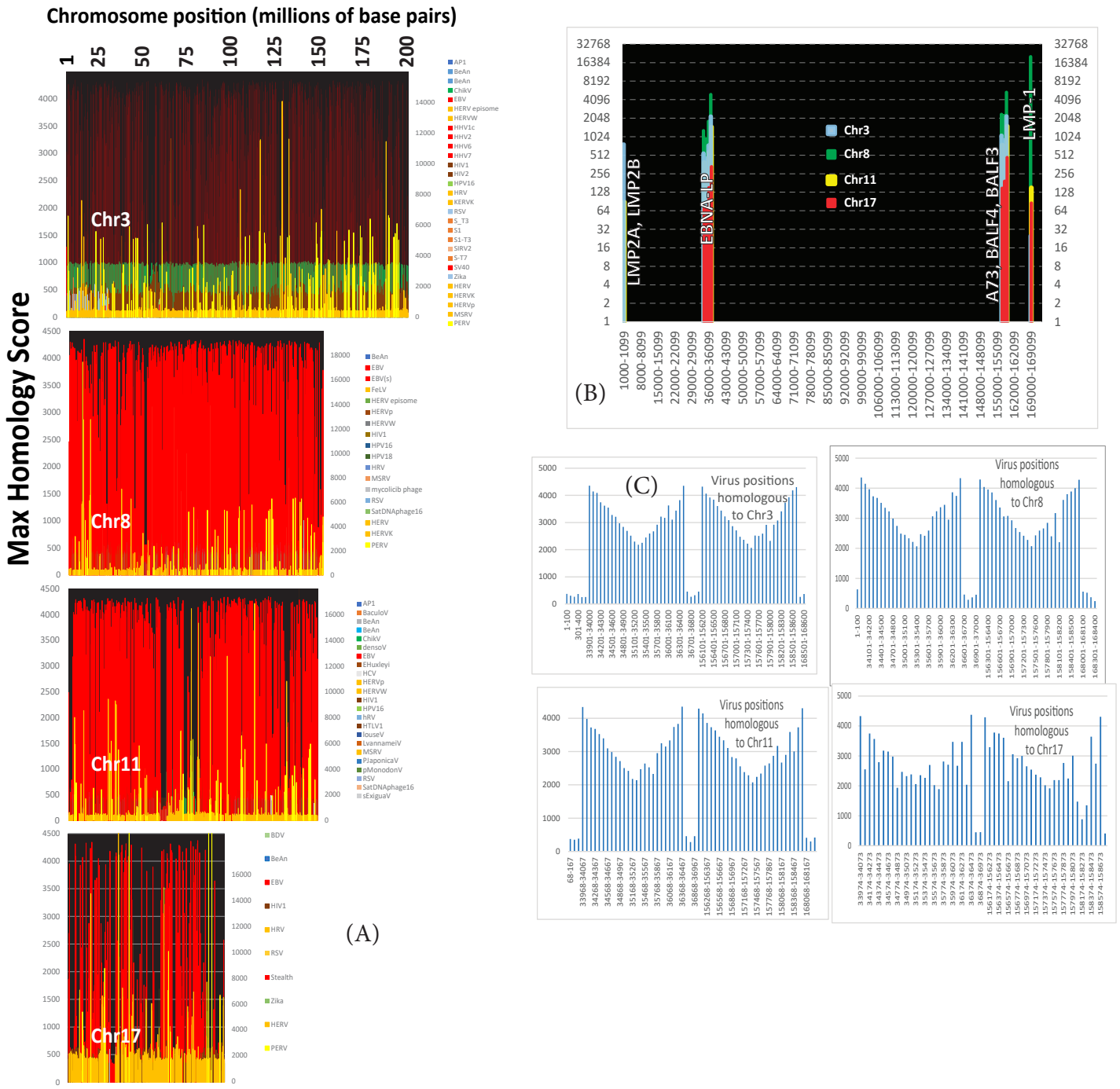

Fig 9A. Positions on chromosomes 3, 8, 11, and 17 with significant DNA sequence homology to any known virus. Fig 9B. Genes on the EBV chromosome that correspond to the positions of homology in Fig. 9A. Fig. 9C. Number of homologous sequences at each position in the EBV genome.

probably exist in the human genome suggesting that the products of these regions should be excluded from vaccines or treatment protocols.

Discussion

Methylation of breast cancers focuses on the same stem cell gene control elements that are targeted in EBV-cancers and even in cleared EBV infections. EBV changes the accessibility and expression of genes encoding critical transcription factors activated across the entire human genome. For instance, abnormal methylation shared with EBV cancers was common for *HOX* genes and their interactions with morphogen and morphogen regulated pathways. These HOX interacting routes included Wnt-beta catenin, *BMP, TGF, FGF*, Retinoic acid, *Hedgehog*, *NOTCH* and *Hippo*. Abnormal methylation of *HOX* genes can promote cancer not only by confusing tissue organization, but also by increasing their oncogenic activity and interfering with their roles as tumor suppressors. At least one *HOX* gene has been correlated with breast cancer outcome [58]. Genome safeguards disabled by EBV are characteristic evidence of prior EBV infection, regardless of whether the infection is still active [5]. The epigenetic changes that match EBV infected cells in the current work might account for some of the similarities in chromatin breaks found between breast cancers and EBV cancers [5].

The current results probably explain why stem cells or their lineage descendants are absent in normal adult human breast basal cells but apparent in breast cancers. The virus causes persistent changes in stem cell gene control elements and this produces or selects for cancer stem cells. These stem cells are friendly to viral replication and labile to malignancy.

Direct model experiments indicating that EBV infection affects breast stem cells have already been published [59]. The initial experiments engineered mammary epithelial cell lines (MECs) with hTERT, SV40 large T-antigen, and Ras V12 oncogene. EBV marked with the gene for green fluorescent protein cooperated with the Ras oncogene to multiply early MEC progenitors that had a stem cell phenotype and a differentiation block. The ability to form

mammospheres increased up to ten-fold and was used to count stem cells. The infected cell lines were then introduced into immunodeficient (NOD / SCID) mice. The emergence of triple negative mammary tumors greatly accelerated in the immunodeficient mice if the cd21 receptor was present [59]. Although there are potential objections to the methodology, organoid culture assays, genetic lineage tracing in human cancers, and in vivo microscopy of mouse breast tumors have all supported the initial xenotransplantation experiments [60].

Additional published laboratory results also support the current findings. Single cell transcriptomic profiling of breast cancer found a differentiation depending on signals from the microenvironment and diversity in mesenchymal differentiation states, including stem cell markers [14]. The EBV protein LMP2A is capable of inducing stem like cell populations in NPC [61]. EBV is also known to infect hematopoietic stem cells giving rise to a wide variety of different infected lymphoid and myeloid cell types [62]. EBV protein LMP-1 binds the *CEBPA* enhancer and induces dedifferentiation [10]. *CEBPB* is the *CEBPA* binding partner in a heterodimer. *CEBPB* control regions on chromosome 20 were differentially methylated in breast cancer and corresponded to EBV cancers. *CEBPB* is essential for differentiation in hematopoietic stem cells, muscle and skeletal stem cells.

EBV reprogramming of its host's epigenetic system has been observed in NPC where there are characteristic and massive changes in methylation [16, 17]. NPC has hypermethylation in at least 80% of cases [18]. In the remaining 20%, subtle changes in specificity and activity of methyltransferases DNMT3A and DNMT3B may account for methylation deficits [63]. EBV suppresses the removal of methyl groups by demethylation pathways, which also alters chromatin exposure [18]. Like NPC, BL causes epigenetic reprogramming by silencing at least one critical transcription factor [64]. Both NPC and BL have forms in which EBV is undetectable, compatible with loss of virus as the cancer progresses.

Epidemiologic studies reported EBV infection was about five-fold more likely in breast cancer tissues than in non-malignant controls [65], but the virus can disappear [66]. However, EBV still leaves evidence of infection behind even in non-malignant cells. EBV-

infected oral keratinocytes that expressed telomerase have hypermethylated CpG islands as scars that normal cells do not have. The methyl groups remain even if the virus is lost during repeated host cell divisions. These epigenetic alterations change gene expression, interfere with differentiation, and can lead to an invasive phenotype [45, 67]. Breast cancer methylation was observed near the same positions. This persistent methylation interferes with differentiation in keratinocytes and does not occur in never-infected cells [45]. More broadly, the implication is that a virus infection can change the human genome, even if the virus does not integrate its DNA. Many of these changes are deleterious and support wider use of vaccination against virus disease. Basal cell carcinoma which is unrelated to EBV is also unrelated to persistent methylation in keratinocytes.

Gene therapy or genetic manipulation of stem cells is a promising general strategy for treating many human disorders. The results here show that stem cells are susceptible to damage from infection and require at least an epigenetic quality assessment before use.

The significant homology between the human and EBV genomes complicates the development of a vaccine or treatment protocol. The results presented here and previously [5] predict that vaccination against near-universal EBV infection could significantly lower the incidence of breast cancer.

Methods

Breast cancer patient data.

All breast cancer data came from female patients with primary breast cancers. Original publications [35] and Kaggle are freely available sources used to obtain clinical and genome data from the METABRIC cohort of 1904 patients. Briefly, 1459 patients in the cohort had tumors classified as ER positive breast tumors, and 445 as ER negative. Grades 1, 2 and 3 were in 165, 740, and 927 patients respectively. Patient ages ranged from about 22 to 96, with most (1200/1904) between ages 47 and 72. The cancers were mostly ductal (1454 patients). The original publications also described tissue collection procedures, clinical data, gene expression profiles, copy number aberrations, and point mutations [35,

68]. The publications also provided stage, grade, tumor size, and the status of lymph nodes, ER data from in situ hybridization, PAM50 gene signature [69], and breast cancer driver genes [22].

Breast cancer methylation data.

Using samples from the METABRIC cohort, published breast cancer differential DNA methylation positions had been derived [22] from whole genome reduced representation bisulfite DNA sequencing of 1538 breast tumors and 244 adjacent normal samples. Identification of methylation sites in breast cancers that correlated with gene expression through promoters and enhancers on the same chromosome was taken from published results [22]. Briefly the published identification procedure screened breast cancer genomes for Pearson correlation of promoters on the same chromosome as methylation sites. Large methylation ("trans") effects that could effect gene controls on multiple chromosomes were eliminated. If a known gene promoter had the strongest, most negative correlation with its corresponding gene, it was taken as a specific ("in cis") effect. Correlations were all based on >50 samples with FDR's limited to <0.05). Non-promoter cis-regulation candidate searches did not depend on a one to one correspondence between methylation sites and regulatory targets. Since a locus can correlate with multiple genes on the same chromosome, correlation values for every locus within 500 kb of any of the genes were ranked. The genomic distance (in bp) was then calculated between each methylated locus and the coordinates assigned to its expression profile. A locus was defined as paired with a gene control region if the transcription start site of its most strongly correlated gene expression was within 500 kb. Shuffled data was used to estimate the probability that the pairing occurred by chance. The ratio between pairing events in real vs shuffled data was then taken as the false discovery rate [22]. This protocol identified 4143 genes on the 22 human autosomal chromosomes as regulated by sequences on the same chromosome as the regulated gene [22].

NPC patients and methylation data.

NPC comparisons to breast cancer were based on unrelated and independently published data [18] for NPC biopsies from 15 sporadic NPC patients and 9 non-tumor adjacent tissues. Their cancers were all primary NPC's at stages II (3 patients), III (7 [patients) and IV (5 patients), with 11 patients surviving. Ages ranged from 38 to 82, composed of 13 males and two females. Differential methylation in NPC was obtained from published whole genome bisulfite sequencing data reported as the average methylation difference between NPC and normal tissue. The results showed 16,910 differentially methylated stretches of DNA had average methylation that differed from the same region in normal DNA by an absolute value of at least 0.2. From these 16,910 regions, 6036 were <500 bp long with average differential methylation of -0.336 to 0.463. Three patient samples were globally hypomethylated and 12 were globally hypermethylated. An independent patient cohort of 48 NPC patients had been used to validate the whole genome bisulfite sequence results [18]. For analysis technical duplicates minimum difference calculations and were not considered.

BL patients and methylation data

BL comparisons to breast cancer came from published data for 13 BL patients [19], ranging in age from 3 to 18. Eleven patients were male and two were female. Tumors in all the patients had breaks in the *MYC* gene and *IG-MYC* translocations. All the tumors had differential hypermethylation that correlated with gene expression. Over two-thirds of the differentially methylated regions were <500 base pairs in length but the lengths ranged from 8-5160. Reference DNA came from the non-neoplastic germinal center B-cells of non-neoplastic tonsils removed from four patients, ages 13-30. The data used in the current work began from 42,381 differentially hypermethylated regions on the 22 autosomal chromosomes. The average methylation in germinal center B-cell DNA segments was subtracted from the average rate in BL and the values were limited to a difference of at least 20% across DNA region that were differentially methylated. The genomic positions of all these regions were compared to genomic coordinates of cis differential hypermethylation sites that affected breast cancer gene controls.

EBV infected normal cell data and transiently infected normal cells

The published data [45] used came from a clonal population originating from normal oral keratinocytes (NOK) that had been immortalized by human telomerase (hTERT) [45]. Briefly, the original publication had infected NOKs by coculture for 24 h with anti-IgG-induced Akata BL cells. After removing B-cells, infected cells were selected for antibiotic resistance on 10 passages. Clones with transient infection were produced by removing antibiotic selection pressure for ten additional passages, after which flow cytometric single-cell cloning was performed. EBV encoded RNA (*EBER*) in situ hybridization confirmed the EBV positivity of infected and transiently infected cells. Uninfected parental cells and plasmid-transfected controls were cultured under identical conditions; infected cells were subjected to single-cell sorting. DNA fingerprinting analysis authenticated the identity of cell clones in relation to uninfected parental controls. Methylation was determined by reduced representation bisulfite sequencing for six samples: uninfected parental cells, control cells transfected with the plasmid vector used, one EBV-positive clone, and three clones transiently infected but EBV-negative [45]. Although the age of the donor was not specified, normal oral keratinocytes typically come from adult donors. For example, a well-known hTERT-immortalized oral keratinocyte line is OKF6/TERT-1, derived from the oral cavity of a male donor, age 57 [70].

Skin basal cell carcinoma methylation data

Basal cell carcinoma data (BCC) came from 16 BCC biopsy specimens from 11 women and 5 men, ages 55-89. The samples were 8 nodular BCCs, and 8 sclerodermiform BCCs [26]. Tumor sites included face, shoulder, head, ear, lip, nose, and pectoral areas. Data had been mapped against the human reference genome (i.e., hg19). Whole genome methylation profiling was with Infinium MethylationEPIC BeadChips.

Calculations of matching abnormally methylated positions.

Distance between abnormally methylated positions in breast and EBV cancers were calculated as the minimum value obtained by subtracting start, end, and mid-points of differential methylation in NPC [18] or BL [71] from these values for cis hypermethylation

[22] in breast cancer. The same piece of DNA could be abnormally methylated at different positions. A strict cutoff of 20% difference from normal cells was used over the entire length of DNA segments being compared. An example of a typical formula used to calculate minimum distances is as follows. =MIN(ABS(D2-$K$2:$M$1532),ABS(E2-$K$2:$M$1532),ABS(F2-$K$2:$M$1532)), where values in column D through F contain breast cancer cis hypermethylation positions (start, end, mid-point) and values in columns K through M contain differential hypermethylation positions (start, end, mid-point) in EBV cancer (either NPC or BL). The Poisson distribution was used as the standard model for estimating the probability of a given number of rare, independent events occurring in a fixed space or time. The p-value was calculated from Poisson distributions as follows: $P(X \geq N) = 1 - P(X < N) = 1 - \sum_{k=0}^{k=N-1} \frac{e^{-\lambda}\lambda^k}{k!}$. Where λ = the number of matches expected from random events and N=the maximum number of matches found. Poisson distributions were evaluated using a Python 3.13 script in a Jupyter notebook within the Anaconda3 platform.

For calculating the control breast cancer vs basal cell skin carcinoma beta values of >=0.5 were used. Beta is defined as methylation sites divided by the total of methylation and non methylation sites.

To test for problems caused by including excessively long DNA sequences, values of differential methylation were cut off in controls at 1000 bps, making only minor differences in the results. DNA sequence positions all came from reads that had all been positioned according to Human Genome Assembly GRCH37 (UCSC Release hg19). These alignments were used directly from the original publications and not corrected or manipulated in any way.

Calculation of frequency of close agreement

Frequencies of agreement between breast cancer and EBV cancer were calculated after normalizing the numbers of NPC differential methylation sites to the same total numbers of cancers as the BL data (15 NPC vs. 13 BL). Frequency calculations were performed in Excel.

Relationships to stem cells and stem cell microenvironments.

Stem cell associated genes were defined as those essential for self-renewal, asymmetric cell division, terminal differentiation from a pluripotent state, or regulators of these processes. In traditional usage, stem cell genes refer to transcription factors and other regulatory molecules—such as *OCT4, NANOG,* or *SOX2*—that maintain pluripotency and self-renewal. In the current work, morphogens and other transcription factors that define cell fate and lineage were also included. Mere expression in stem cells was not considered sufficient for this classification.

The functions of about 350 genes controlled by abnormal methylation at <50 bps separation distances between breast and EBV cancer or EBV infected non-malignant cells were determined by manual curation. Artificial intelligence programs such as Microsoft Copilot and Google provided valuable assistance; however, both occasionally produced errors. Accordingly, all outputs were carefully reviewed, verified, and validated with literature references. Each breast cancer abnormally methylated gene control site within <50 bp of an EBV related methylation site was initially queried for a relationship to stem cells, to regulating stem cell gene expression, differentiation fate, or use as a stem cell marker. Stem cell databases were included and routinely reviewed [25-28]. The primary criterion for identifying a gene as a stem cell related gene was a credible reference documenting this classification. At this stage of the work, embryonic, pluripotent, and multipotent stem cells were not considered separately.

Accompanying searches for some relationship to the stem cell microenvironment were also conducted for matching breast cancer-EBV cancer methylation loci. Genes that influence the microenvironment were taken as those that code for proteins involved in extracellular matrix remodeling, transport factors, structure surrounding the cell, or other modulatory factors. Changes in internal cellular conditions also can directly determine how cells sense and respond to their microenvironment. In contrast some genes were characterized as “Other” (not stem cell genes) because their roles in cell structure or metabolism only indirectly affected stem cell gene products or the stem cell microenvironment.

Breast cancer gene expression data.

Breast cancer gene expression data came from the METABRIC study [35]. The study design had two cohorts: Discovery (997 patient samples) and Validation (995 patient samples). An Illumina HT12 v3 microarray was used to determine gene expression levels. Expression values were presented as Log2(transformed intensity values). To permit cross-sample comparisons, Log2 values had been normalized to z-scores. Z-scores for gene expression that were at least two standard deviations from the mean give a p value of <0.05. Expression of genes that had these z-scores were then compared to stem cell genes with abnormal methylation positions that were <50 bp from methylation positions in EBV cancers.

Calculation of overlap in immune related methylation positions in breast cancers vs methylation position in EBV cancers

An Excel formula was created to check whether the value in cell **K2** (representing an EBV cancer methylation position) falls within a reported range from start to end for any of three sets of methylation positions in CGI regions for breast cancer immune prognosis genes (in rows 2, 3, and 4 across columns AB to AF).

```
=IF(OR( AND(K2 >= MIN($AB$2,$AC$2,$AD$2,$AE$2,$AF$2) - 50, K2 <=
MAX($AB$2,$AC$2,$AD$2,$AE$2,$AF$2) + 50), AND(K2 >=
MIN($AB$3,$AC$3,$AD$3,$AE$3,$AF$3) - 50, K2 <= MAX($AB$3,$AC$3,$AD$3,$AE$3,$AF$3) +
50), AND(K2 >= MIN($AB$4,$AC$4,$AD$4,$AE$4,$AF$4) - 50, K2 <=
MAX($AB$4,$AC$4,$AD$4,$AE$4,$AF$4) + 50) ),K2, "")
```

Comparisons of distributions calculations

Tests for departure from normal distributions and other statistical calculations were made using the StatsDirect statistical program. These tests routinely showed that hypermethylation and breakpoint positions on any chromosomes were unlikely to be normally distributed. Therefore, Mann-Whitney nonparametric tests were used to compare the distributions of differentially methylated positions. Correlation between epigenetic modifications on breast vs EBV cancers was by linear regression analyses with

a power for 5% significance of 99.99%. Kolmogorov-Smirnov and Spearman's tests were also used. Frequencies were plotted to verify the distributions of coinciding areas.

Window sizes to compare methylation sites in breast and EBV cancers

Window sizes of <50 were used to compare differential methylation sites in breast cancer and EBV cancers. This value is consistent with a minimum typical length for enhancer DNA (50-1500 bp). The 50-base comparison at start, end, and middle regulatory sites fully covers <200 bps and thus finds only the minimum number of identities from longer differentially hypermethylated regions. The stringency of the <50 base window size was checked by determining how methylation sites were distributed with respect to transcription start sites. The frequencies of methylated sites at -500,000 to +500,000 distance from the published [22] paired gene transcription start sites were calculated from the pooled data for human chromosomes 1-22. Calculation of the frequencies of separation between differential hypermethylation and transcription start sites for breast cancer chromosomes was done to assess the rigor of a window size of 0 to <50 bps. This test revealed 927/4143 methylation sites within 1000 bases, 261 were within 500 bases, one within 200 bases, and none within fifty bases. Based on this result, the 50 base pair comparison window was judged to be rigorous and was therefore used.

DNA sequence homology.

Homology between human and viral DNA was determined using the NCBI BLASTn program (MegaBLAST). E values less than 1e-10 was the criterion indicating significant homology.

Conclusions

These findings illustrate a selective pressure from a viral infection that acts to facilitate malignancy in cells with stem like properties.

1. **EBV's Role in Breast Cancer**: Epstein-Barr Virus (EBV) infection produces or selects for stem cells that permit viral replication and allow malignancy. Abnormal methylation sites in breast cancer closely match those in EBV-associated cancers, suggesting a common mechanism driven by EBV. These findings explains why stem cells are absent in normal breast basal layer but present in breast cancer.
2. **Impact on Stem Cells**: EBV targets and disrupts the regulation of stem cell-related genes. Reshaping the host cell genome and epigenome creates a cell friendly to viral replication and susceptible to malignancy.
3. **Persistent Epigenetic Changes**: Even after EBV infection clears, the virus leaves lasting methylation scars on non-malignant cells. These scars affect some stem cell control sites, occur in cancers, and may push cells toward malignancy.
4. **Vaccine Potential**: The significant homology between human and EBV genomes complicates vaccine development, but the study identifies specific EBV sequences to exclude from immunogens. A childhood vaccine against EBV could help preserve normal stem cell functions, potentially reducing the risk of breast cancer.
5. **Therapeutic Implications**: Understanding EBV's impact on stem cell genes is crucial for protecting stem cells and for quality control of regenerative medicine.

Data Availability

Datasets used are well-documented and freely available from the original sources or the corresponding author on reasonable request.

Competing Interests

None declared.

References

1. Daud, II, Coleman CB, Smith NA, Ogolla S, Simbiri K, Bukusi EA, et al. Breast Milk as a Potential Source of Epstein-Barr Virus Transmission Among Infants Living in a Malaria-Endemic Region of Kenya. The Journal of infectious diseases. 2015;212(11):1735–42. Epub 20150517. doi: 10.1093/infdis/jiv290. PubMed PMID: 25985902; PubMed Central PMCID: PMCPMC4633760.
2. Kerr JR. Epstein-Barr virus (EBV) reactivation and therapeutic inhibitors. J Clin Pathol. 2019;72(10):651–8. Epub 20190717. doi: 10.1136/jclinpath-2019-205822. PubMed PMID: 31315893.
3. Dittmer DP, Hilscher CJ, Gulley ML, Yang EV, Chen M, Glaser R. Multiple pathways for Epstein-Barr virus episome loss from nasopharyngeal carcinoma. International journal of cancer Journal international du cancer. 2008;123(9):2105–12. doi: 10.1002/ijc.23685. PubMed PMID: 18688856; PubMed Central PMCID: PMCPMC2866115.
4. Kolman JL, Kolman CJ, Miller G. Marked variation in the size of genomic plasmids among members of a family of related Epstein-Barr viruses. Proceedings of the National Academy of Sciences of the United States of America. 1992;89(16):7772–6. doi: 10.1073/pnas.89.16.7772. PubMed PMID: 1323852; PubMed Central PMCID: PMCPMC49793.
5. Friedenson B. Identifying Safeguards Disabled by Epstein-Barr Virus Infections in Genomes From Patients With Breast Cancer: Chromosomal Bioinformatics Analysis. JMIRx Med. 2025;6:e50712. Epub 20250129. doi: 10.2196/50712. PubMed PMID: 39885374; PubMed Central PMCID: PMCPMC11796484.
6. Friedenson B. Evidence of lesions from Epstein-Barr virus infection in human breast cancer genomes. MedRxiv. 2024. doi: Evidence of lesions from Epstein-Barr virus infection in human breast cancer genomes.
7. Zhang X, Powell K, Li L. Breast Cancer Stem Cells: Biomarkers, Identification and Isolation Methods, Regulating Mechanisms, Cellular Origin, and Beyond. Cancers. 2020;12(12). Epub 20201214. doi: 10.3390/cancers12123765. PubMed PMID: 33327542; PubMed Central PMCID: PMCPMC7765014.
8. Zhao M, Tang Z, Wang Y, Ding J, Guo Y, Gao T. A direct negative feedback loop of miR-4721/FOXA1/Nanog promotes nasopharyngeal cell stem cell enrichment and metastasis. Journal of translational medicine. 2021;19(1):387. Epub 20210909. doi: 10.1186/s12967-021-03059-y. PubMed PMID: 34503528; PubMed Central PMCID: PMCPMC8428129.
9. Sell S. Infection, stem cells and cancer signals. Curr Pharm Biotechnol. 2011;12(2):182–8. doi: 10.2174/138920111794295675. PubMed PMID: 21044009; PubMed Central PMCID: PMCPMC3611662.
10. Xie J, Wang Z, Fan W, Liu Y, Liu F, Wan X, et al. Targeting cancer cell plasticity by HDAC inhibition to reverse EBV-induced dedifferentiation in nasopharyngeal carcinoma. Signal Transduct Target Ther.

2021;6(1):333. Epub 20210904. doi: 10.1038/s41392-021-00702-4. PubMed PMID: 34482361; PubMed Central PMCID: PMCPMC8418605.
11. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(7):3983–8. Epub 20030310. doi: 10.1073/pnas.0530291100. PubMed PMID: 12629218; PubMed Central PMCID: PMCPMC153034.
12. Smith J, Zyoud A, Allegrucci C. A Case of Identity: HOX Genes in Normal and Cancer Stem Cells. Cancers. 2019;11(4). Epub 20190410. doi: 10.3390/cancers11040512. PubMed PMID: 30974862; PubMed Central PMCID: PMCPMC6521190.
13. Kumar T, Nee K, Wei R, He S, Nguyen QH, Bai S, et al. A spatially resolved single-cell genomic atlas of the adult human breast. Nature. 2023;620(7972):181–91. Epub 20230628. doi: 10.1038/s41586-023-06252-9. PubMed PMID: 37380767; PubMed Central PMCID: PMCPMC11443819.
14. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, et al. A single-cell and spatially resolved atlas of human breast cancers. Nature genetics. 2021;53(9):1334–47. Epub 20210906. doi: 10.1038/s41588-021-00911-1. PubMed PMID: 34493872; PubMed Central PMCID: PMCPMC9044823.
15. Torne AS, Robertson ES. Epigenetic Mechanisms in Latent Epstein-Barr Virus Infection and Associated Cancers. Cancers. 2024;16(5). Epub 20240229. doi: 10.3390/cancers16050991. PubMed PMID: 38473352; PubMed Central PMCID: PMCPMC10931536.
16. Zhang L, Wang R, Xie Z. Correction to: The roles of DNA methylation on the promotor of the Epstein-Barr virus (EBV) gene and the genome in patients with EBV-associated diseases. Appl Microbiol Biotechnol. 2022;106(21):7373. doi: 10.1007/s00253-022-12216-2. PubMed PMID: 36205765; PubMed Central PMCID: PMCPMC9592632.
17. Zhang L, Wang R, Xie Z. The roles of DNA methylation on the promotor of the Epstein-Barr virus (EBV) gene and the genome in patients with EBV-associated diseases. Appl Microbiol Biotechnol. 2022;106(12):4413–26. Epub 20220628. doi: 10.1007/s00253-022-12029-3. PubMed PMID: 35763069; PubMed Central PMCID: PMCPMC9259528.
18. Ka-Yue Chow L, Lai-Shun Chung D, Tao L, Chan KF, Tung SY, Cheong Ngan RK, et al. Epigenomic landscape study reveals molecular subtypes and EBV-associated regulatory epigenome reprogramming in nasopharyngeal carcinoma. EBioMedicine. 2022;86:104357. Epub 20221111. doi: 10.1016/j.ebiom.2022.104357. PubMed PMID: 36371985; PubMed Central PMCID: PMCPMC9663866.
19. Kretzmer H, Bernhart SH, Wang W, Haake A, Weniger MA, Bergmann AK, et al. DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. Nature genetics. 2015;47(11):1316–25. Epub 20151005. doi: 10.1038/ng.3413. PubMed PMID: 26437030; PubMed Central PMCID: PMCPMC5444523.
20. Yan Y, Chai X, Liu J, Wang S, Li W, Huang T. DeepMethyGene: a deep-learning model to predict gene expression using DNA methylations. BMC Bioinformatics. 2025;26(1):99. Epub 20250408. doi: 10.1186/s12859-025-06115-2. PubMed PMID: 40200144; PubMed Central PMCID: PMCPMC11977931.
21. Dhar GA, Saha S, Mitra P, Nag Chaudhuri R. DNA methylation and regulation of gene expression: Guardian of our health. Nucleus (Calcutta). 2021;64(3):259–70. Epub 20210816. doi: 10.1007/s13237-021-00367-y. PubMed PMID: 34421129; PubMed Central PMCID: PMCPMC8366481.
22. Batra RN, Lifshitz A, Vidakovic AT, Chin SF, Sati-Batra A, Sammut SJ, et al. DNA methylation landscapes of 1538 breast cancers reveal a replication-linked clock, epigenomic instability and cis-regulation. Nat Commun. 2021;12(1):5406. Epub 20210913. doi: 10.1038/s41467-021-25661-w. PubMed PMID: 34518533; PubMed Central PMCID: PMCPMC8437946.
23. Toh TB, Lim JJ, Chow EK. Epigenetics in cancer stem cells. Molecular cancer. 2017;16(1):29. Epub 20170201. doi: 10.1186/s12943-017-0596-9. PubMed PMID: 28148257; PubMed Central PMCID: PMCPMC5286794.

24. Xie P, An R, Yu S, He J, Zhang H. A novel immune subtype classification of ER-positive, PR-negative and HER2-negative breast cancer based on the genomic and transcriptomic landscape. Journal of translational medicine. 2021;19(1):398. Epub 20210920. doi: 10.1186/s12967-021-03076-x. PubMed PMID: 34544424; PubMed Central PMCID: PMCPMC8454077.
25. Tiwari N, Mishra J, Singh N, Mishra SK, Trisal M, Verma S. Epigenetics and cancer stem cells-The world of cancer genesis: A review. J Cancer Res Ther. 2025;21(1):5–9. Epub 20250402. doi: 10.4103/jcrt.jcrt_388_24. PubMed PMID: 40214347.
26. Sand M, Bromba A, Sand D, Gambichler T, Hessam S, Becker JC, et al. Dicer Sequencing, Whole Genome Methylation Profiling, mRNA and smallRNA Sequencing Analysis in Basal Cell Carcinoma. Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology. 2019;53(5):760–73. doi: 10.33594/000000171. PubMed PMID: 31647206.
27. Correction: Epigenetic Switch between SOX2 and SOX9 Regulates Cancer Cell Plasticity. Cancer research. 2017;77(13):3720. doi: 10.1158/0008-5472.CAN-17-1444. PubMed PMID: 28674205.
28. Lin SC, Chou YT, Jiang SS, Chang JL, Chung CH, Kao YR, et al. Epigenetic Switch between SOX2 and SOX9 Regulates Cancer Cell Plasticity. Cancer research. 2016;76(23):7036–48. Epub 20161007. doi: 10.1158/0008-5472.CAN-15-3178. PubMed PMID: 27758880.
29. Chhabra SN, Booth BW. Asymmetric cell division of mammary stem cells. Cell Div. 2021;16(1):5. Epub 20210929. doi: 10.1186/s13008-021-00073-w. PubMed PMID: 34587981; PubMed Central PMCID: PMCPMC8482671.
30. Bhatlekar S, Fields JZ, Boman BM. Role of HOX Genes in Stem Cell Differentiation and Cancer. Stem Cells Int. 2018;2018:3569493. Epub 20180722. doi: 10.1155/2018/3569493. PubMed PMID: 30154863; PubMed Central PMCID: PMCPMC6081605.
31. Steens J, Klein D. HOX genes in stem cells: Maintaining cellular identity and regulation of differentiation. Front Cell Dev Biol. 2022;10:1002909. Epub 20220913. doi: 10.3389/fcell.2022.1002909. PubMed PMID: 36176275; PubMed Central PMCID: PMCPMC9514042.
32. Tahsin S, Sane NS, Cernyar B, Jiang L, Zohar Y, Lee BR, et al. Correction: AR loss in prostate cancer stroma mediated by NF-kappaB and p38-MAPK signaling disrupts stromal morphogen production. Oncogene. 2025;44(1):50. doi: 10.1038/s41388-024-03190-2. PubMed PMID: 39511413.
33. Shi G, Jin Y. Role of Oct4 in maintaining and regaining stem cell pluripotency. Stem Cell Res Ther. 2010;1(5):39. Epub 20101214. doi: 10.1186/scrt39. PubMed PMID: 21156086; PubMed Central PMCID: PMCPMC3025441.
34. Hou Y, Nie Z, Jiang Q, Velychko S, Heising S, Bedzhov I, et al. Emerging cooperativity between Oct4 and Sox2 governs the pluripotency network in early mouse embryos. Elife. 2025;13. Epub 20250227. doi: 10.7554/eLife.100735. PubMed PMID: 40014376; PubMed Central PMCID: PMCPMC11867617.
35. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486(7403):346–52. Epub 20120418. doi: 10.1038/nature10983. PubMed PMID: 22522925; PubMed Central PMCID: PMCPMC3440846.
36. Chen Y, Li M, Wu Y. The occurrence and development of induced pluripotent stem cells. Front Genet. 2024;15:1389558. Epub 20240418. doi: 10.3389/fgene.2024.1389558. PubMed PMID: 38699229; PubMed Central PMCID: PMCPMC11063328.
37. Yamamizu K, Schlessinger D, Ko MS. SOX9 accelerates ESC differentiation to three germ layer lineages by repressing SOX2 expression through P21 (WAF1/CIP1). Development. 2014;141(22):4254–66. doi: 10.1242/dev.115436. PubMed PMID: 25371362; PubMed Central PMCID: PMCPMC4302912.

38. Bourillot PY, Savatier P. Kruppel-like transcription factors and control of pluripotency. BMC Biol. 2010;8:125. Epub 20100927. doi: 10.1186/1741-7007-8-125. PubMed PMID: 20875146; PubMed Central PMCID: PMCPMC2946285.
39. Sierra RA, Hoverter NP, Ramirez RN, Vuong LM, Mortazavi A, Merrill BJ, et al. TCF7L1 suppresses primitive streak gene expression to support human embryonic stem cell pluripotency. Development. 2018;145(4). Epub 20180223. doi: 10.1242/dev.161075. PubMed PMID: 29361574; PubMed Central PMCID: PMCPMC5869011.
40. Kutejova E, Engist B, Self M, Oliver G, Kirilenko P, Bobola N. Six2 functions redundantly immediately downstream of Hoxa2. Development. 2008;135(8):1463–70. Epub 20080305. doi: 10.1242/dev.017624. PubMed PMID: 18321982.
41. Hovland AS, Bhattacharya D, Azambuja AP, Pramio D, Copeland J, Rothstein M, et al. Pluripotency factors are repurposed to shape the epigenomic landscape of neural crest cells. Developmental cell. 2022;57(19):2257–72 e5. Epub 20220930. doi: 10.1016/j.devcel.2022.09.006. PubMed PMID: 36182685; PubMed Central PMCID: PMCPMC9743141.
42. Krendl C, Shaposhnikov D, Rishko V, Ori C, Ziegenhain C, Sass S, et al. GATA2/3-TFAP2A/C transcription factor network couples human pluripotent stem cell differentiation to trophectoderm with repression of pluripotency. Proceedings of the National Academy of Sciences of the United States of America. 2017;114(45):E9579–E88. Epub 20171025. doi: 10.1073/pnas.1708341114. PubMed PMID: 29078328; PubMed Central PMCID: PMCPMC5692555.
43. Fang J, Wang J, Yu L, Xu W. Role of HOXC10 in Cancer. Front Oncol. 2021;11:684021. Epub 20210525. doi: 10.3389/fonc.2021.684021. PubMed PMID: 34113572; PubMed Central PMCID: PMCPMC8185354.
44. Nazari M, Babakhanzadeh E, Mollazadeh A, Ahmadzade M, Mohammadi Soleimani E, Hajimaqsoudi E. HOTAIR in cancer: diagnostic, prognostic, and therapeutic perspectives. Cancer Cell Int. 2024;24(1):415. Epub 20241219. doi: 10.1186/s12935-024-03612-x. PubMed PMID: 39702144; PubMed Central PMCID: PMCPMC11660992.
45. Birdwell CE, Queen KJ, Kilgore PC, Rollyson P, Trutschl M, Cvek U, et al. Genome-wide DNA methylation as an epigenetic consequence of Epstein-Barr virus infection of immortalized keratinocytes. Journal of virology. 2014;88(19):11442–58. Epub 20140723. doi: 10.1128/JVI.00972-14. PubMed PMID: 25056883; PubMed Central PMCID: PMCPMC4178815.
46. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell. 2018;173(2):321–37 e10. doi: 10.1016/j.cell.2018.03.035. PubMed PMID: 29625050; PubMed Central PMCID: PMCPMC6070353.
47. Liu J, Sun W, Li N, Li H, Wu L, Yi H, et al. Uncovering immune cell-associated genes in breast cancer: based on summary data-based Mendelian randomized analysis and colocalization study. Breast cancer research : BCR. 2024;26(1):172. Epub 20241129. doi: 10.1186/s13058-024-01928-0. PubMed PMID: 39614330; PubMed Central PMCID: PMCPMC11606077.
48. Carpenter MA, Temiz NA, Ibrahim MA, Jarvis MC, Brown MR, Argyris PP, et al. Mutational impact of APOBEC3A and APOBEC3B in a human cell line and comparisons to breast cancer. PLoS genetics. 2023;19(11):e1011043. Epub 20231130. doi: 10.1371/journal.pgen.1011043. PubMed PMID: 38033156; PubMed Central PMCID: PMCPMC10715669.
49. Lu H, Lu Z, Wang Y, Chen M, Li G, Wang X. APOBEC in breast cancer: a dual player in tumor evolution and therapeutic response. Front Mol Biosci. 2025;12:1604313. Epub 20250428. doi: 10.3389/fmolb.2025.1604313. PubMed PMID: 40356722; PubMed Central PMCID: PMCPMC12066316.
50. Summerauer AM, Jaggi V, Ogwang R, Traxel S, Colombo L, Amundsen E, et al. Epstein-Barr virus and malaria upregulate AID and APOBEC3 enzymes, but only AID seems to play a major mutagenic role in Burkitt lymphoma. European journal of immunology. 2022;52(8):1273–84. Epub 20220512. doi: 10.1002/eji.202249820. PubMed PMID: 35503749; PubMed Central PMCID: PMCPMC7613445.

51. Tsang CM, Lui VWY, Bruce JP, Pugh TJ, Lo KW. Translational genomics of nasopharyngeal cancer. Semin Cancer Biol. 2020;61:84–100. Epub 20190912. doi: 10.1016/j.semcancer.2019.09.006. PubMed PMID: 31521748.
52. Jiang Y, Xiao H, Sun L, Zhang Y, Liu S, Luo B. LMP2A suppresses the role of AHR pathway through ERK signal pathway in EBV-associated gastric cancer. Virus research. 2021;297:198399. Epub 20210319. doi: 10.1016/j.virusres.2021.198399. PubMed PMID: 33753181.
53. Shen JJ, Niu WN, Zhou M, Zhou F, Zhang HY, Wang L. Association of Epstein Barr virus A73 gene polymorphism with nasopharyngeal carcinoma. Genet Test Mol Biomarkers. 2015;19(4):187–90. Epub 20141002. doi: 10.1089/gtmb.2014.0149. PubMed PMID: 25275219.
54. Neuhierl B, Feederle R, Hammerschmidt W, Delecluse HJ. Glycoprotein gp110 of Epstein-Barr virus determines viral tropism and efficiency of infection. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(23):15036–41. Epub 20021030. doi: 10.1073/pnas.232381299. PubMed PMID: 12409611; PubMed Central PMCID: PMCPMC137540.
55. Chiu SH, Wu MC, Wu CC, Chen YC, Lin SF, Hsu JT, et al. Epstein-Barr virus BALF3 has nuclease activity and mediates mature virion production during the lytic cycle. Journal of virology. 2014;88(9):4962–75. Epub 20140219. doi: 10.1128/JVI.00063-14. PubMed PMID: 24554665; PubMed Central PMCID: PMCPMC3993834.
56. Liu X, Deng Y, Huang Y, Ye J, Xie S, He Q, et al. Nasopharyngeal Carcinoma Progression: Accumulating Genomic Instability and Persistent Epstein-Barr Virus Infection. Curr Oncol. 2022;29(9):6035–52. Epub 20220823. doi: 10.3390/curroncol29090475. PubMed PMID: 36135044; PubMed Central PMCID: PMCPMC9498130.
57. Singh DR, Nelson SE, Pawelski AS, Kansra AS, Fogarty SA, Bristol JA, et al. Epstein-Barr virus LMP1 protein promotes proliferation and inhibits differentiation of epithelial cells via activation of YAP and TAZ. Proceedings of the National Academy of Sciences of the United States of America. 2023;120(20):e2219755120. Epub 20230508. doi: 10.1073/pnas.2219755120. PubMed PMID: 37155846; PubMed Central PMCID: PMCPMC10193989.
58. De Palma FDE, Pol JG, Carbonnier V, Scuderi SA, Mannino D, Montegut L, et al. Epigenetic regulation of HOXA2 expression affects tumor progression and predicts breast cancer patient survival. Cell death and differentiation. 2025;32(4):730–44. Epub 20250120. doi: 10.1038/s41418-024-01430-2. PubMed PMID: 39833374; PubMed Central PMCID: PMCPMC11982354.
59. Hu H, Luo ML, Desmedt C, Nabavi S, Yadegarynia S, Hong A, et al. Epstein-Barr Virus Infection of Mammary Epithelial Cells Promotes Malignant Transformation. EBioMedicine. 2016;9:148–60. Epub 20160521. doi: 10.1016/j.ebiom.2016.05.025. PubMed PMID: 27333046; PubMed Central PMCID: PMCPMC4972522.
60. Clarke MF. Clinical and Therapeutic Implications of Cancer Stem Cells. The New England journal of medicine. 2019;380(23):2237–45. doi: 10.1056/NEJMra1804280. PubMed PMID: 31167052.
61. Kong QL, Hu LJ, Cao JY, Huang YJ, Xu LH, Liang Y, et al. Epstein-Barr virus-encoded LMP2A induces an epithelial-mesenchymal transition and increases the number of side population stem-like cancer cells in nasopharyngeal carcinoma. PLoS pathogens. 2010;6(6):e1000940. Epub 20100603. doi: 10.1371/journal.ppat.1000940. PubMed PMID: 20532215; PubMed Central PMCID: PMCPMC2880580.
62. Wang J, Su M, Wei N, Yan H, Zhang J, Gong Y, et al. Chronic active Epstein-Barr virus disease originates from infected hematopoietic stem cells. Blood. 2024;143(1):32–41. doi: 10.1182/blood.2023021074. PubMed PMID: 37824804.
63. Gao L, Emperle M, Guo Y, Grimm SA, Ren W, Adam S, et al. Comprehensive structure-function characterization of DNMT3B and DNMT3A reveals distinctive de novo DNA methylation mechanisms. Nat Commun. 2020;11(1):3355. Epub 20200703. doi: 10.1038/s41467-020-17109-4. PubMed PMID: 32620778; PubMed Central PMCID: PMCPMC7335073.

64. Manara F, Jay A, Odongo GA, Mure F, Maroui MA, Diederichs A, et al. Epigenetic Alteration of the Cancer-Related Gene TGFBI in B Cells Infected with Epstein-Barr Virus and Exposed to Aflatoxin B1: Potential Role in Burkitt Lymphoma Development. Cancers. 2022;14(5). Epub 20220302. doi: 10.3390/cancers14051284. PubMed PMID: 35267594; PubMed Central PMCID: PMCPMC8909323.
65. Lawson JS, Glenn WK. Catching viral breast cancer. Infectious agents and cancer. 2021;16(1):37. Epub 20210610. doi: 10.1186/s13027-021-00366-3. PubMed PMID: 34108009; PubMed Central PMCID: PMCPMC8191131.
66. Joshi D, Buehring GC. Are viruses associated with human breast cancer? Scrutinizing the molecular evidence. Breast cancer research and treatment. 2012;135(1):1–15. doi: 10.1007/s10549-011-1921-4. PubMed PMID: 22274134.
67. Birdwell CE, Prasai K, Dykes S, Jia Y, Munroe TGC, Bienkowska-Haba M, et al. Epstein-Barr virus stably confers an invasive phenotype to epithelial cells through reprogramming of the WNT pathway. Oncotarget. 2018;9(12):10417–35. Epub 20180102. doi: 10.18632/oncotarget.23824. PubMed PMID: 29535816; PubMed Central PMCID: PMCPMC5828208.
68. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun. 2016;7:11479. Epub 20160510. doi: 10.1038/ncomms11479. PubMed PMID: 27161491; PubMed Central PMCID: PMCPMC4866047.
69. Kensler KH, Sankar VN, Wang J, Zhang X, Rubadue CA, Baker GM, et al. PAM50 Molecular Intrinsic Subtypes in the Nurses' Health Study Cohorts. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2019;28(4):798–806. Epub 20181227. doi: 10.1158/1055-9965.EPI-18-0863. PubMed PMID: 30591591; PubMed Central PMCID: PMCPMC6449178.
70. Dickson MA, Hahn WC, Ino Y, Ronfard V, Wu JY, Weinberg RA, et al. Human keratinocytes that express hTERT and also bypass a p16(INK4a)-enforced mechanism that limits life span become immortal yet retain normal growth and differentiation characteristics. Molecular and cellular biology. 2000;20(4):1436–47. doi: 10.1128/MCB.20.4.1436-1447.2000. PubMed PMID: 10648628; PubMed Central PMCID: PMCPMC85304.
71. Lopez C, Kleinheinz K, Aukema SM, Rohde M, Bernhart SH, Hubschmann D, et al. Genomic and transcriptomic changes complement each other in the pathogenesis of sporadic Burkitt lymphoma. Nat Commun. 2019;10(1):1459. Epub 20190329. doi: 10.1038/s41467-019-08578-3. PubMed PMID: 30926794; PubMed Central PMCID: PMCPMC6440956.
72. Lin YL, Persaud SD, Nhieu J, Wei LN. Cellular Retinoic Acid-Binding Protein 1 Modulates Stem Cell Proliferation to Affect Learning and Memory in Male Mice. Endocrinology. 2017;158(9):3004–14. doi: 10.1210/en.2017-00353. PubMed PMID: 28911165; PubMed Central PMCID: PMCPMC5659671.
73. Khawaja RR, Agarwal A, Fukaya M, Jeong HK, Gross S, Gonzalez-Fernandez E, et al. GluA2 overexpression in oligodendrocyte progenitors promotes postinjury oligodendrocyte regeneration. Cell reports. 2021;35(7):109147. doi: 10.1016/j.celrep.2021.109147. PubMed PMID: 34010640; PubMed Central PMCID: PMCPMC8185898.
74. Wang H, Zhou Y, Yu D, Zhu H. Klf2 contributes to the stemness and self-renewal of human bone marrow stromal cells. Cytotechnology. 2016;68(4):839–48. Epub 20141231. doi: 10.1007/s10616-014-9837-6. PubMed PMID: 25550041; PubMed Central PMCID: PMCPMC4960135.
75. Guo Q, Zhou Y, Ni H, Niu M, Xu S, Zheng L, et al. The SIX2/PFN2 feedback loop promotes the stemness of gastric cancer cells. Journal of translational medicine. 2024;22(1):832. Epub 20240910. doi: 10.1186/s12967-024-05618-5. PubMed PMID: 39256760; PubMed Central PMCID: PMCPMC11389068.
76. Oliphant MUJ, Vincent MY, Galbraith MD, Pandey A, Zaberezhnyy V, Rudra P, et al. SIX2 Mediates Late-Stage Metastasis via Direct Regulation of SOX2 and Induction of a Cancer Stem Cell

Program. Cancer research. 2019;79(4):720–34. Epub 20190103. doi: 10.1158/0008-5472.CAN-18-1791. PubMed PMID: 30606720; PubMed Central PMCID: PMCPMC6586234.
77. Ma TL, Zhu P, Chen JX, Hu YH, Xie J. SIX3 function in cancer: progression and comprehensive analysis. Cancer Gene Ther. 2022;29(11):1542–9. Epub 20220628. doi: 10.1038/s41417-022-00488-9. PubMed PMID: 35764712.
78. Hu X, Yan R, Cheng X, Song L, Zhang W, Li K, et al. The function of sperm-associated antigen 6 in neuronal proliferation and differentiation. J Mol Histol. 2016;47(6):531–40. Epub 20160826. doi: 10.1007/s10735-016-9694-z. PubMed PMID: 27566703.
79. Chen WY, Tsai YC, Yeh HL, Suau F, Jiang KC, Shao AN, et al. Loss of SPDEF and gain of TGFBI activity after androgen deprivation therapy promote EMT and bone metastasis of prostate cancer. Science signaling. 2017;10(492). Epub 20170815. doi: 10.1126/scisignal.aam6826. PubMed PMID: 28811384.
80. Jing S, Chen L, Han S, Liu N, Han M, Yang Y, et al. Expression of TCF7L2 in Glioma and Its Relationship With Clinicopathological Characteristics and Patient Overall Survival. Front Neurol. 2021;12:627431. Epub 20210708. doi: 10.3389/fneur.2021.627431. PubMed PMID: 34305772; PubMed Central PMCID: PMCPMC8296806.
81. Brombin A, Simpson DJ, Travnickova J, Brunsdon H, Zeng Z, Lu Y, et al. Tfap2b specifies an embryonic melanocyte stem cell that retains adult multifate potential. Cell reports. 2022;38(2):110234. doi: 10.1016/j.celrep.2021.110234. PubMed PMID: 35021087; PubMed Central PMCID: PMCPMC8764619.
82. Jin C, Luo Y, Liang Z, Li X, Kolat D, Zhao L, et al. Crucial role of the transcription factors family activator protein 2 in cancer: current clue and views. Journal of translational medicine. 2023;21(1):371. Epub 20230608. doi: 10.1186/s12967-023-04189-1. PubMed PMID: 37291585; PubMed Central PMCID: PMCPMC10249218.
83. Lin X, Kang K, Chen P, Zeng Z, Li G, Xiong W, et al. Regulatory mechanisms of PD-1/PD-L1 in cancers. Molecular cancer. 2024;23(1):108. Epub 20240518. doi: 10.1186/s12943-024-02023-w. PubMed PMID: 38762484; PubMed Central PMCID: PMCPMC11102195.
84. Galassi C, Chan TA, Vitale I, Galluzzi L. The hallmarks of cancer immune evasion. Cancer cell. 2024;42(11):1825–63. Epub 20241010. doi: 10.1016/j.ccell.2024.09.010. PubMed PMID: 39393356.
85. Zuo J, Currin A, Griffin BD, Shannon-Lowe C, Thomas WA, Ressing ME, et al. The Epstein-Barr virus G-protein-coupled receptor contributes to immune evasion by targeting MHC class I molecules for degradation. PLoS pathogens. 2009;5(1):e1000255. Epub 20090102. doi: 10.1371/journal.ppat.1000255. PubMed PMID: 19119421; PubMed Central PMCID: PMCPMC2603334.
86. Lin JH, Lin JY, Chou YC, Chen MR, Yeh TH, Lin CW, et al. Epstein-Barr virus LMP2A suppresses MHC class II expression by regulating the B-cell transcription factors E47 and PU.1. Blood. 2015;125(14):2228–38. Epub 20150128. doi: 10.1182/blood-2014-08-594689. PubMed PMID: 25631773.
87. Sausen DG, Poirier MC, Spiers LM, Smith EN. Mechanisms of T cell evasion by Epstein-Barr virus and implications for tumor survival. Frontiers in immunology. 2023;14:1289313. Epub 20231221. doi: 10.3389/fimmu.2023.1289313. PubMed PMID: 38179040; PubMed Central PMCID: PMCPMC10764432.
88. Xu J, Ahmad A, Jones JF, Dolcetti R, Vaccher E, Prasad U, et al. Elevated serum transforming growth factor beta1 levels in Epstein-Barr virus-associated diseases and their correlation with virus-specific immunoglobulin A (IgA) and IgM. Journal of virology. 2000;74(5):2443–6. doi: 10.1128/jvi.74.5.2443-2446.2000. PubMed PMID: 10666277; PubMed Central PMCID: PMCPMC111728.

Table1 Effects on stem cells by over- and under-expressed genes in breast cancers. *Abnormal methylation of these gene control sites is closely related to abnormal methylation sites in EBV cancers*

Gene	Chr	General description of function	Expression Status in breast cancers	Context and reference
BMP2	20	Morphogen	Overexpressed	Drives differentiation
BMP4	14	Morphogen	Overexpressed	Drives differentiation
CRABP1	15	Morphogen	Overexpressed	Powerful regulator of stem cell fate via retinoic acid signaling, diminishes neural and embryonic stem cell expansion [72].
GRIA2	4	Differentiation regulator	Overexpressed	Neurodevelopment gene. Steers progenitors toward specific cell fates after an injury [73]
HOTAIR	12	Homeobox gene regulator	Overexpressed	Epigenetic regulator Causes H3K27 methylation; increases breast cancer invasiveness
HOXC10	12	Homeobox gene	Overexpressed	Activates *ERK, AKT, p65,* and EMT-related genes
KLF2	19	Pluripotency regulator [38].	Overexpressed	Increases expression of stem cell markers like OCT4 and NANOG, and suppresses differentiation markers to maintain stemness and self-renewal [74]
SIX2	2	Homeobox gene	Overexpressed	Enhances expression of stem cell markers *SOX2* and *Nanog* [75, 76] Activates Wnt/β-catenin and MAPK pathways;
SIX3	2	Homeobox gene	Overexpressed	Modulates Wnt signaling; enhances proliferation and migration in gastric cancer [77]
SMO	7	Morphogen-responsive gene	Overexpressed	Guides differentiation into specific identities
SOX9	17	Pluripotency regulator	Overexpressed	Antagonizes and opposes pluripotency factor *SOX2* [37]
SPAG6	10	Differentiation regulator	Overexpressed	*SPAG6* is a cilia-related gene. Its overexpression biases stem cells toward specific lineages, and inhibits proliferation of neural progenitors [78].

SPDEF	6	Differentiation regulator (epithelial)	Under-expressed (most significantly in Fig. 4)	Crucial in epithelial cell differentiation. *SPDEF* under-expression during androgen deprivation therapy increased TGF-beta signaling activity to promote epithelial mesenchymal transition (which is strongly linked to stem cell properties) [79]
TCF7L1	2	Pluripotency regulator vs. self renewal	Overexpressed	Silences Wnt targets; maintains pluripotent state and prevents differentiation [39]
TCF7L2	2	Differentiation regulator	Overexpressed	In many types of tumors, *TCF7L2* promotes tumor invasion and metastasis, often via epithelial mesenchymal transition [80]. Can either activate or repress Wnt signaling
TFAP2A	6	Pluripotency exit regulator	Under-expressed	Methylation locks cells in self-renewal; disrupts placental lineage commitment [42].
TFAP2B	6	Differentiation regulator	Under-expressed	Determines differentiation outcome and cell fate in an embryonic melanocyte stem cell [81]. Suppresses Wnt/β-catenin pathways in certain cancers [82].Not as well studied as *TFAP2A*
VAX2	2	Homeobox gene	Overexpressed	Translates positional info from *HOX* genes

Table 2 Immune escape mechanisms in ER positive breast cancers shared with EBV cancers

Mechanism in ER-positive breast cancer [83, 84]	**EBV Association**
MHC downregulation	EBV downregulates MHC class I and II molecules to evade cytotoxic T cells [85, 86]
PD-1/PD-L1 upregulation	EBV-associated tumors often upregulate PD-L1, contributing to T cell anergy [87]
Immune exhaustion	Chronic EBV infection leads to T cell exhaustion, especially in latent phases [87]
TGF-β signaling	EBV can induce TGF-β signaling, promoting immunosuppression, EBV replication and tumor progression [88].

Supplementary Table S1. Matching abnormal methylation of gene controls in breast cancers and EBV cancers affects genes associated with stem cells and the microenvironment. S=Stem cell related, M=microenvironment, Other=not directly related to stem cells or microenvironment.

Chrom-osome	CLASS	Gene with abnormal methylation in breast cancer corresponding to EBV cancers	Function related to stem cell self renewal, differentiation, microenvironment, or other
1	S	EDARADD	Guides development of structures that originate from ectodermal stem cells. Critical for interactions between ectoderm and mesoderm during development
1	S	ETNK2	Essential for stem cell pluripotency Synthesis of critical membrane component of both mitochondrial and plasma cell membranes
1	S	HES4	Regulates development especially in the immune system, bone and retina [1]. Responds to NOTCH signaling. Regulated by Wnt and Hedgehog signaling.
1	S	IGSF21	Expressed in stem cell lines. Regulates inhibitory synapse formation in brain [2]
1	S	KIAA1522	Stem cell pluripotency [3]
1	S	LRRN2	Implicated in regulating pluripotency and differentiation [4, 5]. Cell adhesion and signal transduction
1	S	MACF1	Required for beta catenin signaling in nucleus and important for embryonic development. Binds to growing microtubule ends in cytoskeleton,
1	S	PAQR7	Membrane progesterone receptor linked to stem cell ability to regenerate nerve cells [6]
1	S	SPOCD1	Human stem cell fate determination [7]. Directs piRNAs mediated DNA methylation [8]
1	S	TCEA3	Regulates stem cell differentiation potential. Highly expressed in embryonic stem cells [9]. Enhances skeletal muscle differentiation
1	S	WLS	Critical mediator in Wnt pathway insuring their correct trafficking and secretion. Essential for pattering and body axis development
1	M	ATPAF1	Assembly factor for mitochondrial ATP synthase complex in electron transport chain
1	M	CAMK2N1	Mitochondrial stress response [10]
1	M	C1orf64	Androgen receptor target
1	M	CDC42BPA	a kinase regulated by the small GTPase CDC42, involved in actin–myosin cytoskeleton dynamics, which are crucial for cell shape, migration, and adhesion
1	M	KCNK2	Regulates membrane potential found in stem cells. Important for stem cell function but not considered a stem cell gene
1	M	MRPS21	Mitochondrial ribosomal protein
1	M	PKP1	Cell adhesion and communication
1	M	RGS7	Accelerates G protein inactivation, dampening GPCR signaling which mediate responses to extracellular signals. In this way, RGS7 influences cellular perception and response to environmental stimuli.
1	M	SLC2A1	Glucose transporter. One of the transporters that can contribute to the Warburg effect
1	M	ST6GALNAC5	Modifies cell surface interactions i.e. cell-cell, cell-extracellular matrix interactions. Sialyltransferase. Mediates breast cancer brain metastasis [11]
1	M	TMEM61	Regulates actin filopodia formation as in primary cilia, determining how cells move, adhere, and interact with surrounding tissues.

1	*Other*	*IL12RB2*	Crucial to activate T-helper cells after they are already committed . Supports stem cell self renewal [12]
2	***S***	***DNMT3A***	Stem cell activation and function, found in hematopoietic and embryonic stem cells. Enzyme for DNA methylation of host genes. Critical to control T-cells and regulated by EBV in gastric carcinoma [13, 14]
2	***S***	***EFHD1***	Regulates Hippo pathway, which controls stem cell properties . May contribute to stem cell development [15]. Mitochondrial calcium sensor, membrane trafficking. Immune cell activation
2	***S***	***EN1***	Homeobox gene. Regulates stem cells, cell growth and differentiation. Controls pattern formation in nervous system development.
2	***S***	***FHL2***	Stem cell regulator including hematopoietic and mesenchymal cells. Strongest developmental pathway relationship is to Wnt signaling [16]Essential to maintain self renewal during replication stress. Regulates ovarian cancer tumor progression Weakens gastrointestinal tumor expression [17, 18].
2	***S***	***IGFBP5***	Regulates stem cell differentiation and proliferation
2	***S***	***LRP1B***	Regulates neural stem cells. Frequently mutated in cancer by epigenetic or genetic mechanism. Putative tumor suppressor [19].
2	***S***	***LYPD6B***	Expressed in stem cells. Hypermethylation associated with invasive phenotype in melanoma [20]
2	***S***	***LYPD6***	Enhances Wnt beta catenin signals [21]interacts with LRP6 a Wnt coreceptor
2	*S*	***OTX1***	Stem cell development
2	***S***	***PRKCE***	Regulates stem cells
2	***S***	***SH3BP4***	Stem cell regulation
2	***S***	***SIX2***	Direct downstream target of HOXa2 [22] Regulates stem cells
2	***S***	***SIX3***	Homeobox gene. Essential for neuronal differentiation, regulates Wnt pathway [23]Cancer suppressor in most cases but can also be a promoter depending on whether pathways inhibit or promote cancer [24]
2	***S***	***SP5***	Maintains pluripotent stem cells
2	***S***	***TCF7L1***	Regulates pluripotency and lineage specification [25]Embryonic stem cell signature gene
2	***S***	***VAX2***	Homeobox gene, controls basic eye development processes.
2	***M***	***CLEC4F***	Binds galactose and N-acetyl galactosamine which allows Kupffer cells to recognize and respond to these products from damaged cells or pathogens. Marker for macrophages recruited to mitochondrial damage signals [26].Destroys platelets that have been desialylated by bacterial infection [27].
2	***M***	***GAD1***	Response to inhibiting mitochondrial oxidative phosphorylation [28]. Produces inhibitory neurotransmitter gamma amino butyric acid from glutamic acid. Overexpression a poor prognostic indicator in NPC [29]
2	***M***	***GPC1***	Regulates signaling pathways that influence stem cell behavior, including proliferation, differentiation, and migration Involved in EBV infection and immune cell infiltration [30]
2	***M***	***INHBB***	Shapes cancer microenvironment [31]. Regulates cell cycle, apoptosis and synthesis of estradiol and progesterone. Correlates with immune infiltration in gastric cancer [32]
2	***M***	***MBOAT2***	Shapes membrane fluidity and composition, interactions with ECM. Remodels lecithins, Activates quiescent neural stem cells [33]
2	***M***	***MYO1B***	Actin binding motor protein that couples the actin cortex to the plasma membrane, enabling cells to change shape and form protrusions like filopodia. These protrusions are essential for cell migration, especially during development or wound healing.
2	***M***	***NPAS2***	Controls glucose metabolism Upregulates glycolytic genes [34]. Circadian clock gene

2	***M***	***PXDN***	Extracellular matrix and basement membrane formation. Mammary gland development [35]
2	***M***	***RAB11FIP5***	Directs proteins to specific locations in the cell surface, influencing how cells respond to external signals. In salivary epithelial cells, RAB11FIP5 mediates the V-ATPase transport to plasma membrane so that cells can respond to acidic conditions. Essential for calcium-stimulated exocytosis in neuroendocrine cells, affecting release of signaling molecules into surroundings. Also affects mitochondrial position and shape [36]NK cell function
2	***M***	***RAMP1***	Partners with receptors to influence cell response to their surroundings Stress stimulated hematopoiesis [37]
2	***M***	***SLC16A14***	Orphan gene but close relatives are Monocarboxylate transporter that maintains stem cell niche microenvironment, Pyruvate lactate axis, energy metabolism, and gluconeogenesis
2	***M***	***SLC1A4***	Transporter for neutral amino acids Ala, Ser, Cys, Thr, Pro, and OH-Pro
2	*Other*	*DDX11L2*	Originally classified as a pseudogene but may contain a functional promoter.
2	*Other*	*LONRF2*	Protein quality control, Stem cell quality
2	*Other*	*MLPH*	Transports intracellular organelles. Facilitates movement along actin filaments. Development of normal pigmentation in skin, hair, eyes. Indirect and context dependent influence on cell microenvironment.
3	***S***	***PDZRN3***	Development and stability of vascular plexuses, Wnt planar cell polarity [38]. Negative regulator of Wnt pathway [39]. Ubiquitin ligase that marks proteins for destruction.
3	***S***	***RPL39L***	Ribosomal protein paralog highly expressed in pluripotent cells supports pluripotency and differentiation [40], contributes to differentiation in multiple lineages
3	***S***	***TNIK***	serine/threonine kinase that activates Wnt signaling pathway, TNIK interacts with transcription factors including TCF4, adjusting gene response to Wnt signals [41]
3	***S***	***ZIC4***	Transcription factor with important role in development [42]. Target of Wnt pathway
3	***S***	***ZNF502***	Insufficient information but probably a stem cell gene because it belongs to a family with strong stem cell links and is expressed in germ line cells May be involved in stem cell pluripotency affects cell differentiation
3	***M***	***CNTN3***	Cell-cell interactions to organize the extracellular environment. Cell adhesion and directional guidance during differentiation of neural stem cells.
3	***M***	***IGF2BP2***	Regulates glycolysis [43]
3	***M***	***NMNAT3***	NMNAT3 enzyme catalyzes conversion of NMN to NAD^+, primarily within mitochondria. This conversion is essential for redox balance, energy metabolism, and sirtuin activation, especially Sirt3.
3	***M***	***TRH***	Integrator of energy metabolism [44, 45]. Transcription factor produced in the hypothalamus
4	***S***	***FAT1***	Loss promotes cancer initiation stem like cells [46], Essential in development for cell polarity, proliferation and survival. Loss is lethal to mice [47]
4	***S***	***GRIA2***	Cell communication neuron to neuron signaling during brain development. Shapes neuronal circuits [48]. Also expressed in pituitary.
4	***S***	***NPNT***	Regulates sox2 expressed in bulge stem cells other stem cell processes [49],
4	***S***	***PPP2R2C***	Interacts with signaling pathways that direct cell fate [50] Regulatory factor that influences development and stem cell behavior. Encodes PP2A enzyme which affects pluripotency and differentiation.
4	***S***	***PROM1***	Well-known stem cell marker on several types of stem cells [51]

4	***S***	***ZFP42***	Maintains undifferentiated stem cells and regulates their differentiation [52]
4	***M***	***AFAP1***	Cell adhesion, migration. Actin filaments in cytoskeleton, cell signaling
4	***M***	***HADH***	Encodes part of a protein essential for mitochondrial fatty acid oxidation Hydroxyacyl-CoA hydrogenase
4	***M***	***LMCH1***	Cell motility, cytoskeleton, and adhesion. Muscle stress fibers
4	***M***	***SHROOM3***	Regulates cell shape through actin binding. Implicated in neural tube defects
5	***S***	***CXXC5***	An epigenetic regulator that recruits Tet2 demethylase to regulate interferon responses [53]. Modulates the balance between self-renewal and differentiation to influence cell fate decisions [54]
5	***S***	***IQGAP2***	May contribute to differentiation of stem cells. Participates in MAPK/ERK and PI3K/AKT/mTOR pathways, which are options for cell fate determination. Upregulated in EBV-infected cells. Knockdown of IQGAP2 affects the clumping of infected cells. Varied functions related to its binding partner. Modulates the cytoskeleton [55].
5	***S***	***IRX2***	Methylation determines differentiation pathway of induced pluripotent stem cells [56]
5	***S***	***IRX4***	Multi-potent progenitor cell marker
5	***S***	***SPRY4***	Gatekeeper of MAPK/ERK responses, one of the pathways that regulate stem cell fate [57]. SPRY4 activity has been detected in embryonic stem cells
5	***M***	***EPB41L4A***	Interactions between the cytoskeleton and the plasma membrane. Contributes to microenvironmental interactions and to beta catenin signaling in cell fate decisions
5	***M***	***FAM153A***	Influences cell adhesion and migration by affecting cell shape [58]
5	***M***	***FBN2***	Extracellular matrix formation can influence stem cell differentiation
5	***M***	***NNT***	Mitochondrial redox regulation, antioxidant defense
5	***M***	***TNIP1***	Inflammation regulator, tissue homeostasis, autoimmunity
5	*Other*	*ANKHD1*	Cell cycle regulation, proliferation and survival
5	*Other*	*SNORD123*	Guides methylation of ribosomal RNA, indirectly affects stem cells and microenvironment
5	*Other*	*SUB1*	Safeguards genome stability.
6	*S*	***GJA1***	Essential for primed stem cell functions to allow communication between cells [59]. Negative regulator of SOX2 expression [60]
6	***S***	***PHF10***	Essential for chromatin remodeling and hematopoietic stem cell maintenance [61], PHF10 is a key component of the npBAF chromatin remodeling complex, which is essential for the self-renewal and proliferation of multipotent neural stem cells.
6	***S***	***RBM24***	Regulates embryonic lineage differentiation and cell homeostasis [62]Related to BMP signaling [63]. Driver of lineage commitment to cardiac cells
6	***S***	***RPS6KA2***	Enhances stem cell differentiation [64]. Promotes cell growth and differentiation. Downstream effector of MAPK/ERK pathway
6	***S***	***SPDEF***	Influences differentiation and maturation of somatic goblet cells. Regulated by NOTCH signals. Enhances cancer stem cell like properties [65]
6	***S***	***TFAP2A***	Causes embryonic stem cells to differentiate to neural crest cells by combining with OCT4-SOX2 complex [66]. Partners with GATA2/3 and TFAP2C in a transcriptional circuit named the TEtra network. This network represses pluripotency genes like OCT4 to differentiates toward trophectoderm (placental lineage). Inhibiting TFAP2A disrupts this balance, preventing proper lineage commitment [67]. TFAP2A binds to epigenetically inactive placental genes at the onset of differentiation, helping activate them. Without TFAP2A, these genes remain silent, stalling differentiation".
6	***S***	***TFAP2B***	Specifies an embryonic melanocyte stem cell that retains adult potential for multiple fates [68] Suppression of Epithelial-to-mesenchymal transition and Wnt/β-catenin pathways in certain cancers [69]

6	***S***	***TRERF1***	Induces preleukemic stem cell expansion [70]. Transcriptional-regulating factor 1 is a zinc-finger transcriptional regulating protein which interacts with CBP/p300 to regulate the human gene CYP11A1.
6	***M***	***CAP2***	actin cytoskeleton regulation, which affects cell shape, movement, and signaling. Interacts with pathways which control cell adhesion.
6	***M***	***CNKSR3***	Transepithelial sodium transport [GeneCards] Aldosterone mediated sodium transport
6	***M***	***PACSIN1***	Endocytosis, receptor trafficking
6	***M***	***PTPRK***	Protein tyrosine phosphatase, controls adhesion molecules that determine stem cell niche interactions [71]targeting removes stem cell function from colon cancer cells [72] Dephosphorylates beta catenin in Wnt pathway [73]. Promotes differentiation as a key modulator and suppresses proliferation.
6	***M***	***RSPH9***	Cilia function and motility
6	***M***	***TUBB2B***	Microtubule structure, controls how the cell interacts with its environment.
6	*Other*	*MTCH1*	Decides if cell should undergo apoptosis. Guides proteins into place in the membrane. Mitochondrial outer membrane protein
6	*Other*	*PCMT1*	Protein quality control, cell survival
6	*Other*	*PRSS16*	Serine protease shapes T-cell abilities to recognize antigens
6	*Unkn*	FAM50B	Identified in gene signatures in stem cells. Development of germ cells [74]. Regulates immune microenvironment Differential methylation relies on histone modifications in embryonic stem cells. Gene is imprinted in humans and associated with retrotransposon insertion sites [70]
6	*Unkn*	LOC401242	Too little information available
7	***S***	***COBL***	Cell polarity and development
7	***S***	***EPHA1***	Cancer stem cell marker [75]
7	***S***	***HOXA10-HOXA9***	Both HOXA10 and HOXA9 activate stem cell specific genes [76]. Transcription read through associates with viral infections [77]
7	***S***	***LFNG***	Differentiation and cell fate [78]
7	***S***	***NACAD***	Regulates fate of intestinal stem cells [79]. Controls metabolism that stem cell differentiation pathways depend on.
7	***S***	***SMO***	Regulates stem cell self renewal and development Part of hedgehog pathway [80]
7	***M***	***CLDN3***	Controls interact with microenvironment and shapes cell architecture. Contributes to cancer stem cell phenotype and highly expressed in cancer stem cells [81]
7	***M***	***GSTK1***	Glutathione S-transferase, detoxification enzyme
7	***M***	***NPTX2***	Anchors AMPA receptors at synapses by binding to the extracellular section. This creates a more structured extracellular matrix around the synapse [82].
7	***M***	***NRCAM***	Sculptor that shapes stem cell niche microenvironment. Regulates stem cells in the nervous system [83].
7	***M***	***PTPRN2***	Phosphatidylinositol phosphatase that alters phosphoinositide levels in the plasma membrane. These levels affect membrane curvature, vesicle formation, exocytosis, and the actin cytoskeleton. PTPRN2 indirectly releases membrane cofilin which cuts actin filaments. Severing these filaments increases cell relocation, invasion, and adaptation to forces acting on the microenvironment. Vesicle-mediated exocytosis, controls signaling from the extracellular matrix, communication between cells, hormone and neurotransmitter release [84]
7	***M***	***TMEM176A***	Cellular microenvironment in immune cell contexts
7	***M***	***VIPR2***	Water and ion flux.

7	*Other*	*ADCY1*	Adenyl cyclase, may be important to maintain stem cells
7	*Other*	*ZC3HC1*	Stabilizes FANCD2, a critical member of the FA-BRCA pathway for DNA repair.
8	***S***	***ADAM32***	RT-PCR results showed ADAM32 affects expression of genes associated with stemness [85], has a role in development, membrane protein.
8	***S***	***FBXO16***	Marks proteins for destruction by ubiquitylation. Disrupts beta catenin signaling in breast cancer [86]
8	***S***	***FBXO32***	Regulates stem cell differentiation and renewal [87]
8	***S***	***LAPTM4B***	Associated with stemness especially in populations of cancer cells [88]. This differs from its role in normal cells where LAPTM4B is associated with lysosomes and endosomal traffic.
8	***S***	***NKX3-1***	Determines cell fate of pluripotent stem cells in specific contexts i.e. fate-determining factor, steering pluripotent stem cells toward a specialized lineage [89]
8	***S***	***SDR16C5***	A short-chain dehydrogenase/reductase (SDR) superfamily member that catalyzes the oxidation of retinol (vitamin A) into retinaldehyde—a key early step in the retinoic acid biosynthesis pathway. This enzymatic role is essential for mediating retinoic acid signaling, which influences numerous developmental and metabolic processes, but SDR16C5 does this as part of a metabolic pathway rather than by acting directly on stem cell properties or by actively remodeling the microenvironment.
8	***S***	***SQLE***	Key enzyme in cholesterol biosynthesis. Upregulates markers for stem cell properties [90]
8	***M***	***ARHGEF10***	Regulates rho GTPases involved in cell signaling, cytoskeletal dynamics, vesicle trafficking, and neural developmental programs. Interacts with other small GTPases like rad6 and rab8 which control cell shape, polarity and movement [91].
8	***M***	***IKBKB***	Regulates inflammation and immune response activates NF-KB to produce cytokines.
8	***M***	***LRRC6***	Development and function of cilia [92]
8	***M***	***MAL2***	Modulates membrane traffic to control cell response to microenvironment. Vesicular trafficking, formation of lipid rafts, and signal transduction in T-cells and myelin-producing cells
8	***M***	***MSRA***	Protects cells from oxidative stress. A special truncated form protects mouse embryonic stem cells from oxidation [93]. Used as a marker to isolate stem cell populations
8	***M***	***SLC7A2***	Regulates the intracellular levels of cationic amino acids such as Arginine, affects cellular metabolism and energy homeostasis. Alterations directly influence the metabolic fitness of nearby immune cells which depend on adequate levels of L-arginine for an immune response. Helps create an immunosuppressive niche environment. Expressed in embryonic and induced pluripotent stem cells.
8	***M***	***TOP1MT***	Relieves supercoiling of mitochondrial DNA. Helps tumors adapt to environments with limited resources [94]
8	*Other*	*FAM110B*	Mitosis control, locates at centrosomes and spindle poles. Might influence stem cell niche remodeling.
8	*Other*	*RIMS2*	Scaffold protein that regulates release of neurotransmitters and hormones Expressed in stem cells especially during neuronal development [95] Regulates exocytosis. Coordinates vesicle treaffic
8	*Other*	*TIGD5*	Potential influence on chromatin architecture. DNA transposon expressed in neural stem cells and primordial germ cells. Frequently coamplifies with MYC [96]. May be a tumor suppressor.
8	*Other*	*ZNF696*	Regulates transcription mediated by RNA polymerase II, Not extensively studied.
9	***S***	***GNAQ***	Contributes to an Rspo3/Wnt3a-Lgr4-Gnaq pathway, enhances β-catenin signaling in MLL leukemic stem cells [97] . Also interacts with Hippo pathway
9	***S***	***MPDZ***	Inhibits hippo signaling pathway involved in differentiation of stem cells [98]

9	***S***	***NFIB***	Regulates stem cell differentiation and activity in stem cell niche [99]. Associated with NFKB1 and important for NFKB signaling in both B-cells and cytotoxic T-cells [100].
9	***S***	***NPDC1***	Regulates neuronal cell differentiation and proliferation and influences stem cell behavior
9	***S***	***OLFM1***	Development and organization of the nervous system, hematopoiesis [101]
9	***M***	***CARD9***	Bridges pattern receptors and downstream immune signaling, triggers release of inflammatory cytokines. Balances gut microbes and immune responses [102]
9	***M***	***CRAT***	Balances fat and glucose metabolism and energy homeostasis in mitochondria
9	***M***	***DNLZ***	Helps cells respond to environmental stresses like elevated temperature. Preserves mitochondrial functions during stress. Chaperone mediated acquisition of protein 3-dimensional structure.
9	***M***	***ENTPD2***	Architect of the environment around cells. Released in exosomes to hydrolyze extracellular trinucleotides in the external microenvironment. Overexpression prevents myeloid derived monocyte suppressor cells from differentiating [103] Influence stem cell behavior in taste buds [104]
9	***M***	***FUT7***	Adhesive interactions of bone marrow cells to regulate distribution, growth, and development in hematopoietic system [105]. Can also affect microenvironment by control of leukocyte trafficking under conditions of inflammation
9	***M***	***GCNT1***	Glycosylation essential in mucins, and selectins, affecting cell communications
9	***M***	***GNA14***	Shapes the tumor microenvironment [106]
9	***M***	***LPAR1***	G protein coupled receptor that triggers cell signaling in response to the environment controls cell migration, interactions, survival and differentiation. Enhances differentiation of neural stem cells, links to multiple sclerosis, an EBV mediated disease [107]. Stress response in stem cells, transplantation survival, also affects microenvironment [108].
9	***M***	***PIP5K1B***	Actin filament organization determining cell shape, adhesion and mobility. Contributes to endocytosis and exocytosis. Calcium signaling lipid kinase essential for mitochondrial functions as well as cell signaling
9	***M***	***PKN3***	Rho GTPase effector which regulates cytoskeletal structure involved in cell micgration and interactions [109]
9	***M***	***SECISBP2***	Control of selenoprotein translation. Essential to synthesize proteins that respond to oxidative stress in the cell microenvironment.
9	***M***	***SUSD3***	Regulates how cell-cell and cell environment interactions, influences spread and invasion of tumors. Promotes estrogen dependent cell multiplication and cell interaction and migration. Expressed in stem cells
9	*Other*	*DAPK1*	Regulates autophagy and apoptosis. Affects cel fate decisions under stress. Involved in cancer stem cell formation also affects tumor stromal cells [110] Downregulation promotes stem cell properties of cancer cells Cell death associated [111]
10	***S***	***DDIT4***	Regulates stem cell fate links HIF and mTOR pathways [112]
10	***S***	***DOCK1***	"Exclusively abundant in hematopoietic stem cells" [113]
10	***S***	***DPYSL4***	Regulates morphogenesis in tooth germ cells [114]. Semaphorin stimulates RAS/MEK/ERK pathway
10	***S***	***DUSP5***	Regulates MAPK signaling to control cell signaling and differentiation, growth and survival. Phosphatase that regulates signaling pathways, promotes osteogenic differentiation of bone marrow stromal cells [115]. Gatekeeper of developmental and immune processes [116]
10	***S***	***GATA3***	Regulates self-renewal and cell fate in hematopoietic stem cells [117]. Controlled by BMP signaling and Hh pathways in developing neural crest cells in zebrafish [118]
10	***S***	***GFRA1***	Neuron survival and differentiation, essential for self-renewal [119]

10	***S***	***H2AFY2***	Critical for the differentiation of hematopoietic stem cells [120] Epigenetic regulator that modifies chromatin structure.
10	*Other*	*LOC283070*	Long non-coding RNA involved in converting prostate cancer cells to androgen receptor independence. Might affect pathways in stem cells that do not require androgen. Probably influences chromatin structure
10	***S***	***PLXDC2***	Stem cell plasticity, Hematopoietic stem cell marker [121]
10	***S***	***SPAG6***	Affects TGF-beta [122]and PI3K/AKT [123] pathways Hypermethylated during differentiation of human embryonic stem cells [124]. Identified as a breast cancer biomarker [125]
10	***S***	***TCF7L2***	Part of Wnt signaling pathway to maintain normal somatic stem cell homeostasis, also involved in embryonic development [126, 127]
10	***S***	***ZMIZ1***	Modulates activity of stem cell pathways through, NOTCH1, SAD3/4, and beta catenin [128]Transcriptional coactivator that is critical for T-cell development
10	***S***	***ZNF503***	Essential for embryonic development. Activates MYC, negative regulator of GATA3 potential role in breast cancer stem cell like behavior [129]
10	***M***	***ARHGAP21***	Negative regulator of Rho GTPase mediated adhesion and migration, essential for hematopoietic stem cell microenvironment [130]
10	***M***	***ASB13***	Ubiquitin ligase that tags proteins for degradation, may mediate cytokine signaling Knockout increases cell travels and disrupts polymerization of F-actin [131]. ASB13 is part of an axis that regulates how cells respond to the microenvironment.
10	***M***	***CALY***	Calcium signaling a key part of the stem cell environment at various stem cell differentiation stages.
10	***M***	***KCNMA1***	Affects stem cell behavior by encoding a conductance channel to regulate nerve and muscle functions. Affects cell microenvironment
10	***M***	***PPP1R3C***	Regulates gluconeogenesis and glycogen synthesis. Part of a phosphatase complex
10	***M***	***RGS10***	Regulates G-protein signaling
10	*Other*	*FAM196A / INSYN2A*	May be involved in chromosome segregation and DNA repair. Regulates inhibitory neurotransmission, scaffolding component at synapses.
10	*Other*	*TACC2*	Organizes microtubules and couples nucleus to centrosome. Androgen responsive cell cycle regulator [132]
10	*Other*	*THNSL1*	Threonine biosynthesis
10	*Unkn*	*C10orf67*	Probably encodes a mitochondrial protein of unknown function
11	***S***	***ASCL2***	Transcription factor that drives cell fate decisions, highly expressed in stem cells and essential to maintain them [133]
11	***S***	***IFITM1***	Highly expressed in human embryonic stem cells to maintain pluripotency and to protect pluripotent cells from viral infection [134, 135] Encodes protein that restricts cellular entry by diverse viral pathogens [20, 21]. .IFITMs are part of the innate immune system
11	***S***	***LGR4***	Regulator that affects stem cell lineage decisions and behavior. Receptor that modulates the Wnt/beta catenin pathway. Loss of LGR4 interferes with the ability of mammary stem cells to repopulate by affecting SOX2 [136]
11	***S***	***MOB2***	A component of the Hippo signaling pathway that plays a crucial role in stem cell self renewal and differentiation [137].
11	***S***	***PLEKHB1***	Membrane associated signaling functions that might indirect regulate cell differentiation pathways. Could control differentiation in specific cell types. Other genes in the same pathway have stronger links to differentiation decisions

11	***S***	***TP53I11***	Suppresses extracellular matrix independent survival and mesenchymal transitions in mammary epithelial cells
11	***S***	***WT1***	Regulates differentiation of progenitor cells, highly expressed in stem cells [138]
11	***M***	***CHST1***	Carbohydrate sialyltransferase that affects the structure of components of the extracellular matrix which is critical to form the stem cell niche.
11	***M***	***CORO1B***	Cytoskeleton organization, cell motility. Cell structures critical for cell microenvironment, endothelial networks and tissue regeneration [139]
11	***M***	***EHD1***	Endocytosis
11	***M***	***IGSF9B***	Cell-cell adhesion, associated with stem cells
11	***M***	***MYEOV***	Recruitment and behavior of immune infiltration
11	***M***	***SYT12***	Helps cells respond to environmental stress Helps cells interpret and respond to surroundings. Correlates with immune cell infiltration in gastric cancer [140]
11	***M***	***SYT9***	Cellular response to environmental signals. Vesicle mediated transport, clathrin endocytosis
11	***M***	***TSKU***	Key signaling pathways in the cell microenvironment
11	*Other*	*ABTB2*	Cell response to toxins
11	*Other*	*ENDOD1*	Nucleic acid processing and genomic integrity
11	*Other*	*FADS2*	Encodes rate limiting enzyme for biosynthesis of polyunsaturated fatty acids
11	*Other*	*TNNI2*	Protects telomeres, regulates their length
12	***S***	***ATF7IP***	Crucial to regulate hematopoietic stem and progenitor cells by interacting with a histone methyltransferase [141]
12	***S***	***FAIM2***	Contributes to embryonic stem cells differentiating to specific lineage [142] Marker of immune infiltration [143]
12	***S***	***FBXO21***	Regulates maintenance and differentiation of hematopoietic stem cells [144]. Regulates epithelial mesenchymal transition [145]
12	***S***	***HOTAIR***	Epigenetic regulator, that can trigger EMT and regulate stemness of cancer cells [146] Master regulator of chromatin dynamics [147], regulates WNT pathway [148]
12	***S***	***HOXC10***	In mesenchymal stem cells used as a marker, regulator of body plan [149]
12	***S***	***NAB2***	Female germline stem cells, brain development [150]
12	***S***	***NTN4***	Promotes and regulates some stem cell populations [151]
12	***S***	***WNT5B***	Expressed in stem cells and regulates behavior, often linked to cancer stem cells [152]
12	***S***	***ZCCHC8***	Regulatory factor in embryonic stem cells [153]
12	***M***	***CPNE8***	Encoded protein binds to cell membrane in response to calcium signals. Enhances fibroblast and immune cell infiltration of gastric tumor environment.
12	***M***	***FAR2***	Fatty acid metabolism Defects in FAR2 can cause mitochondrial dysfunction [154]. Affects membrane fluidity and interactions with the cell microenvironment
12	*Other*	*CHFR*	Cell cycle checkpoint. Has some connection to stem cells because of its role in early germ cell maintenance.
12	*Other*	*RERG*	Cell growth gene regulated by estrogen
13	***S***	***DACH1***	Regulates self renewal and differentiation [155]. Enhances cell cycle arrest by binding p53. Prevents breast cancer invasion and metastasis. DACH1 is a cytokine hijacked by a prototype foamy virus [156].
13	***S***	***FLT3***	Essential for normal stem cell development and immunity, Hematopoiesis [157].
13	***S***	***PCID2***	Suppresses developmental genes to maintain pluripotency in embryonic stem cells. Prevents premature differentiation [158].

13	S	TNFSRF19	Actively contributes to stem cell maintenance and differentiation Developmental signaling, modulation of apoptosis, and interactions with other signaling molecules such as TRAF family members also called "TROY" gene [159].
13	S	ZIC2	Required for embryonic stem cell specification [160]
13	M	FAM155A	Direct structural role in ion channels that control how neurons respond to the microenvironment. Expressed in embryonic stem cells, expression altered as cells differentiate.
13	M	GRTP1	Mitochondrial function and interaction, glucose metabolism, cell growth and differentiation. Connected to insul/TOR/S6K pathway
13	M	MBNL2	Shapes the immune microenvironment of tumors [161].
13	Other	BIVM	DNA endonuclease activity, Probably involved in DNA repair
14	S	BMP4	Encodes critical molecule that guides stem cell differentiation commits multi-potent stem cells to specific differentiation pathway [162]
14	S	C14orf39/S6OS1	Development and function of neural stem cells. Spindle checkpoint essential for sperm and ovarian function [163]
14	S	CRIP1	Myometrial stem progenitor cell population [164]
14	S	CRIP2	Stem cell gene that affects vascular development and development of various tissues [165]
14	S	MTA1	Ability of stem cells to propagate [166]. Stem cell fate [167]. Deacetylates histone H3K27
14	M	GPR68	Receptor that senses extracellular hydrogen ions.
14	M	SLC7A8	Transporter for neutral amino acids, may influence metabolism during differentiation Essential for oxidative phosphorylation, key supplier of amino acids for innate immunity cells [168]
14	M	TMEM121	Trans membrane protein that responds to and regulates cell behavior that affects surrounding environment, especially in cancer.
15	S	ALDH1A3	Essential for retinoic acid production (pathway to guide HOX gene directed differentiation) Cancer stem cell marker
15	S	CRABP1	Growth factor sensitivity and stemness [169]
15	S	IGF1R	Crucial to establish stem cell pluripotency and self-renewal. Modulates key pathways like PI3K/AKT, RAS/MAPK, and JAK/STAT in stem cells. Niche remodeling [170]
15	S	NR2F2	Regulates differentiation and immune control in stem cells [171]. A long non-coding RNA downregulated in EBV infection [172]
15	S	PCSK6	Protease needed to activate stem cell factors BMP and Nodal [173]
15	S	WDR72	Enhances stem cell properties of lung cancer cells. Activates AKT/HIF signaling [174]
15	M	HDGFRP3	Extracellular signaling that remodels the extracellular environment [175]
15	M	KLF13	Releases chemokines that shape the local environment
15	M	SLCO3A1	Moves organic anions across cell membranes. Part of a shared gene signature across multiple stem cell lines [176]
16	S	FBXL16	Embryonic stem cell differentiation [177]
16	S	IRX3	Hematopoietic differentiation, Neurogenesis in hypothalamus [178], stem cell behavior
16	S	IRX5	Hematopoietic differentiation, Neurogenesis in hypothalamus , stem cell behavior
16	S	NETO2	Significant role in regulating multiple stem cell pathways Wnt, JAK-STAT, MAPK/AKT and TGF-beta [179] although not a core stem cell gene.
16	S	TOX3	Regulates neural stem cell identity [180]
16	M	C16orf45	Cell migration, cell shape, protein-protein interactions

16	*M*	*CLEC16A*	Autophagy, mitophagy, regulates normal glucose responsive insulin release, inflammation, endosomal trafficking
16	*M*	*MSLN*	Membrane antigen, cell adhesion and recognition.
16	*M*	*RAB40C*	GTPase involved in protein degradation**,** membrane trafficking**, and** cell migration
16	M	*VPS35*	Sorting and trafficking proteins, essential for mitochondrial function and fusion
16	M	*WDR90*	Centrosome and cilium structure sensing microenvironment
16	Other	*APRT*	Purine salvage pathway
16	Other	*RHBDL1*	Protein degradation and signaling
16	*Other*	*Z3H18*	RNA processing and turnover
17	*S*	*ANAPC11*	Key regulator of adult keratinocyte stem cell fate and keratinocyte homeostasis [181]. Important in early stage development. Linked to endoderm stem cells Crucial to maintain hematopoietic stem cells. Essential for cell division
17	*S*	*B4GALNT2*	Stem cell regeneration Infection susceptibility [182]
17	*S*	*CYB5D2*	Enhances repair of damaged heart cells by pluripotent stem cells after infarction [183]. Expressed in neural stem cells and used as a stem cell marker Predicted to be involved in nervous system development. Likely affects multiple differentiation pathways.
17	*S*	*FASN*	A metabolic gatekeeper that is essential for the activity of multiple stem cell pathways such as Wnt, NOTCH, and EGF [184]. Influences cell fate decisions thorough metabolic reprogramming. Controls adult neural stem cell activity [185]Cell energy production fatty acid synthase
17	*S*	*GPRC5C*	Regulates stem cell quiescence of hematopoietic stem cells [186] induced by retinoic acid [GeneCards]
17	*S*	*JMJD6*	Demethylase that influences chromatin structure, hydroxylates lysine on histones and non-histones. Promotes stem cell self renewal and ability to differentiate into multiple cell types [187]
17	*S*	*KRT19*	Regulates NOTCH signaling pathway and stem cell properties [188]. Balances Wnt NOTCH signaling [189]
17	*S*	*LGALS3BP*	Regulates positions of neural progenitors [190]Related to EGF signaling [191]
	S	*NUP85*	Nuclear architecture in stem cells
17	*S*	*PMP22*	Regulates stem cell property of self renewal in GC. [192] TEAD1 and YAP/TAZ key hippo effector bind to enhancers upstream of PMP22 [193]
17	*S*	*PPP1R1B*	Stem cell development
17	*S*	*PSMC31P*	Germline male stem cell development [Gene Cards]
17	*S*	*RAC3*	Required to maintain pluripotency of normal stem cells [194]
17	*S*	*RAI1*	Upregulated in response to retinoic acid signaling for differentiation. Brain development, circadian rhythm,
17	*S*	*SOX9*	Stem cell factor critical in embryonic development. Immune cell infiltration into tumors [195].
17	*S*	*SREBF1*	Regulates hematopoietic stem cell fate, function, and survival [196]
17	*S*	*UTP6*	Hypermethylation associated with stem-cell like properties in colorectal cancer stem cells Secondary control of Wnt pathway via FOXK2 [197]. May be involved in transcriptional regulation of pluripotency regulatory factors Low expression increases stemness.
17	*M*	*ATP2A3*	Significant role in controlling cellular microenvironment, and calcium homeostasis and cell signaling. Encodes an ATP-ase enzyme for calcium transport and sequestration. Expressed in embryonic stem cells and in hematopoietic and epithelial cell lineages.
17	*M*	*ATP5H*	Mitochondrial ATP synthase
17	*M*	*CDC42EP4*	Organizes cell shape, movement, and pseudopodia formation, Interacts with GTPases, senses environment

17	*M*	*CORO6*	Actin dynamics, cytoskeleton, cell growth
17	*M*	*DCXR*	Promotes aerobic glycolysis [198]
17	*M*	*GCGR*	Glucagon receptor, cell microenvironment, glucose metabolism, stem cell maintenance [199]
17	*M*	*MRPS7*	Mitochondrial ribosomes
17	*M*	*RAB34*	Ciliogenesis, sensing microenvironment
17	*M*	*WNK4*	Adjusts mitochondrial energetics [200]Regulates renal salt handling [201]
17	*Other*	*AFMID*	Tryptophan metabolism
17	*Other*	*ALOX12*	Lipoxygenase that regulates platelets, inflammation
17	*Other*	*CASC3*	Nonsense mediated mRNA removal. May have some developmental functions [202]
17	*Other*	*GHDC*	Amino acid ligase activity, lactation, Innate immunity
17	*Other*	*GRIN2C*	NMDA receptor subunit
18	*S*	*ANKRD12*	Inhibits nuclear receptors via histone deacetylases [GeneCards]. ANKRD12 silencing promotes invasion [203].
18	*S*	*CHD7*	Chromatin remodeling Regulates endodermal and mesodermal development, colocalizes with canonical stem cell factors like OCT4 [204]
18	*S*	*RAB31*	Fine tunes cell fate decision. Interacts with GLI1, an effector in the hedgehog developmental pathway [205]. Regulates glucose transporter [206]. Membrane traffic, exosome biogenesis, found in a variety of stem cells.
18	*S*	*YES1*	Hippo pathway effector involved in development, growth, repair and homeostasis
18	*M*	*CYB5A*	Cytochrome related to electron transport chain in mitochondrial energy production
18	*M*	*DSC2*	Organization of tissues, stability and even cell signaling in the microenvironment
18	*M*	*DSC3*	Desmosome proteins that control cell-cell adhesion interactions and might affect stem cell niche. Encode a core part of desmosomes that inherently influence the cell microenvironment
18	*M*	*DTNA*	Cell membrane stability in muscle and neural tissues. Mechanical stability of cell membrane to facilitate interactions with the extracellular matrix
18	*M*	*GREB1L*	Glycosylation of estrogen receptor, and promotion of immune cell invasion [207, 208]. Intertwined with cell environment regulation
18	*M*	*ZBTB7C*	Linked to tumor microenvironment [209]. Regulates metabolism and glucose levels
18	*Other*	*MOCOS*	Purine and aldehyde metabolism indirect connection to energy production.
19	*S*	*EPS15L1*	Stem cell development and differentiation, endocytosis [210]
19	*S*	*FUT3*	Fucosyltransferase 3 forms Lewis antigens essential for embryonic development and differentiation pathways. Adds fucose to glycans linked to stem cell characteristics and affects how they differentiate, may enhance cell fate plasticity. Expression reduced in cancer stem cells vs normal stem cells [211]. Fut2 deficits reduce stemness in intestinal stem cells. Lewis antigens serve as binding sites for some pathogens.
19	*S*	*KLF2*	Regulates embryonic stem cell-specific gene expression [212]
19	*S*	*LYPD3*	Drives tumor stemness and immune evasion. Maintains tumor stem cell properties Cell adhesion and migration binds to extracellular matrix
19	*S*	*NFIX*	Regulates production of intermediate progenitor cells from embryonic stem cells in adults regulates the timing of neural differentiation [213] Involved in the oxidative stress response and cell fate decisions [203]. Binds viral and cellular promoters.
19	*S*	*NR2F6*	Regulates mouse stem cell hematopoiesis and myelopoiesis [214] Crucial role in glucose metabolism and energy production in brown fat.
19	*S*	*PEG3*	Parental imprinted gene in stem cells in all tissues tested. Stem cell marker [215], Regulates adipogenesis.
19	*S*	*REEP6*	Photoreceptor development [216] Retinal homeostasis Downstream effector that fine tunes stem cell fate decision. A specialized developmental gene.

19	S	ZFP28	Maintains pluripotency in embryonic stem cells [217]
19	S	ZNF256	Represses gene expression during development, fine tunes gene activity at specific stages. Cortical specification (Cortecon repository)
19	S	ZNF329	Neural differentiation, cortical specification, deep layers (Cortecon repository) Links to developmental disorders that are not well-documented [MayMyGenome.com].
19	S	ZNF350	Recruits KAP1 a corepressor that modifies chromatin to determine stem cell identity and lineage [218].
19	S	ZNF415	Contains a KRAB A box, gene is expressed in stem cell niches,
19	S	ZNF419	Related to tumor stemness according to methylation and characteristics of mRBA expression [219]
19	S	ZNF439	Neural differentiation, cortical specification, deep layers (Cortecon repository)
19	S	ZNF541	Meiosis regulator in germ line male cells, likely stem cell gene for sperm cell fate
19	S	ZNF549	Expressed in primordial germ cells, cortical plate, suggesting a role in stem cell differentiation
19	S	ZNF570	Transcription regulation, potential impact on stem cells but inadequate information
19	S	ZNF626	Neural differentiation, cortical specification, upper layers (Cortecon repository)
19	S	ZNF880	Neural differentiation, cortical specification, upper layers (Cortecon repository)
19	M	CACNG6	Transmembrane regulatory receptor voltage dependent calcium channel. Ion homeostasis.
19	M	CCDC61	In epithelial cells with multi cilia, helps maintain interface between cell and its environment. Centrosome location and assembly [220], essential for cell division Proper construction of basal bodies
19	M	CHST8	Affects how cells interact with their surroundings, hormone and immune responses by catalyzing the sulfation of N-acetyl galactosamine on glycoproteins and glycolipids. Neural stem cell proliferation (results from CRISPR screens). Activated by TGF beta [221]. Loss weakens breast cancer tumor growth [222]
19	M	FXYD5	Subunit of sodium-potassium pump, ion homeostasis an membrane potential Confers properties of stem cells [223]
19	M	GNA11	Signal transduction and calcium regulation. Melanocyte development and melanoma [224]
19	M	KCNK6	Potassium channel
19	M	MYH14	Cytokinesis, cell motility and cell polarity. Interacts with ac tin filaments to control how cells respond to their microenvironment
19	M	PCSK4	Helps specify early development environment by processing growth factors and hormones. Specifies cell fates in placental development
19	Other	KLK5	Skin maintenance after terminal differentiation of keratinocytes. Keratinocyte differentiation and desquamation
19	Other	LOC113230	Non-coding RNA that affects transcription and gene regulation Inhibits arginine synthesis, controlled by TGF-beta induction
19	Other	MIA-RAB4B	Sequesters miRNA as a miRNA sponge to regulate gene expression
19	Other	REXO1	RNA processing
19	Other	SSBP4	Single strand DNA binding, DNA repair
19	Unkn	LOC100505715	unknown
20	S	BMP2	Guides stem cell differentiation fate, “shaping embryonic development, maintaining tissue homeostasis, and influencing disease progression” [225].
20	S	CEBPB	Maintains stem cell self renewal and quiescence, regulates genes associated with cell cycle arrest [226].

20	S	EYA2	Self-renewal [227]
20	S	PTPRT	Hematopoietic stem cell self renewal [228]
20	M	COMMD7	Modulates NF-KB pathway for inflammation. Maintains cell surface levels of receptors, directly affecting cell responses to the microenvironment.
20	M	GPCPD1	Controls metabolism to influence cell microenvironment Glycerophospholipid metabolism
20	M	MAP1L3CA	Microtubule interactions, autophagy
20	M	NECAB3	Regulates beta amyloid production. Interacts with dopamine receptor Activates HIF to use glycolysis at normal levels of oxygen [229]
20	M	SLC24A3	Listed as a cancer stem cell gene [230]. Calcium and sodium balance inside cells exchanges with extracellular environment
20	M	SRXN1	Oxidative stress response
20	Other	BCAS4	Associated with intermediate filaments
20	Other	FRG1B	Pseudogene that may play a role in chromatin remodeling, Elements near FRG1B bind to OCT4 and NANOG. [GeneCards]
20	Other	RIMS4	Neuronal activity and synaptic regulation
21	S	RIPK4	RIPK4 activates IRF6, to downregulate proliferative signals (e.g. p63) and promote keratinocyte differentiation. RIPK4 promotes epidermal differentiation, initiating hippo signaling cascade by phosphorylating LATS1/2, inhibiting YAP/TAZ. RIPK4 also enhances differentiation signaling, recruiting LATS1/2 into liquid condensates [231, 232].
21	M	ABCG1	Mitochondrial function, lipid homeostasis, Removes excess cholesterol from tissues influences stem cell behavior
21	M	SLC37A1	Phosphate regulator, glucose homeostasis
21	M	TFF1	Secreted by gastric mucous cells to stabilize and protect the mucus layer and protect epithelial surfaces.
21	M	TFF3	Mitochondrial function. Protects intestinal barrier
21	Other	CSTB	Inhibits cathepsins to regulate protein turnover within cells
22	S	CCDC117	Stem cell fate determination, cardiac morphogenesis [233, 234]
22	S	FAM83A	Associated with stem cell traits in cancer, activates stem cell pathways such as Wnt-beta-catenin and TGF-beta.
22	S	TBX1	Regulates stem cell quiescence vs. proliferation [235], acts by controlling the stem cell gene SOX2.
22	M	CACNA11	Calcium flux
22	M	CDC42EP1	Effector for CDC42 which controls the actin cytoskeleton
22	M	EMID1	Modifies extracellular environment, extracellular matrix interactions
22	M	KCTD17	Formation of primary cilia on cells which sense and respond to the microenvironment
22	M	PANX2	Regulates cellular processes within the stem cell niche, including cell communication and signaling pathways. Facilitate release of extracellular messengers to modulate extracellular environment
22	M	RAB36	A GTPase required for stem cell maintenance and cell migration in epithelium of digestive tract [236]
22	Other	CARD10	Modulates cytokine and chemokine expression
22	Other	DMC1	DNA meiotic recombinase gene expressed in cells undergoing meiosis, including stem cells
22	Other	XBP1	Cell stress response, misfolded and unfolded protein metabolism, protein quality control. Under some conditions SBP1 can influence stem cell behavior

1. El Yakoubi W, Borday C, Hamdache J, Parain K, Tran HT, Vleminckx K, et al. Hes4 controls proliferative properties of neural stem cells during retinal ontogenesis. Stem Cells. 2012 Dec;30(12):2784–95. PMID: 22969013. doi: 10.1002/stem.1231.
2. Tanabe Y, Naito Y, Vasuta C, Lee AK, Soumounou Y, Linhoff MW, et al. IgSF21 promotes differentiation of inhibitory synapses via binding to neurexin2alpha. Nat Commun. 2017 Sep 1;8(1):408. PMID: 28864826. doi: 10.1038/s41467-017-00333-w.
3. Shi Y, Xiao Q, Huang S, Pan Q, Jia H, Zhou Q, et al. Poor Prognostic Biomarker KIAA1522 Is Associated with Immune Infiltrates in Hepatocellular Carcinoma. J Oncol. 2023;2023:3538928. PMID: 36761433. doi: 10.1155/2023/3538928.
4. Liao CH, Wang YH, Chang WW, Yang BC, Wu TJ, Liu WL, et al. Leucine-Rich Repeat Neuronal Protein 1 Regulates Differentiation of Embryonic Stem Cells by Post-Translational Modifications of Pluripotency Factors. Stem Cells. 2018 Oct;36(10):1514–24. PMID: 29893054. doi: 10.1002/stem.2862.
5. Johnson KR, Mallon BS, Fann YC, Chen KG. Multivariate meta-analysis reveals global transcriptomic signatures underlying distinct human naive-like pluripotent states. PLoS One. 2021;16(5):e0251461. PMID: 33984026. doi: 10.1371/journal.pone.0251461.
6. Castelnovo LF, Thomas P. Membrane progesterone receptor alpha (mPRalpha/PAQR7) promotes migration, proliferation and BDNF release in human Schwann cell-like differentiated adipose stem cells. Mol Cell Endocrinol. 2021 Jul 1;531:111298. PMID: 33930460. doi: 10.1016/j.mce.2021.111298.
7. Zhou D, Zhu F, Huang ZH, Zhang H, Fan LQ, Fan JY. SPOC domain-containing protein 1 regulates the proliferation and apoptosis of human spermatogonial stem cells through adenylate kinase 4. World J Stem Cells. 2022 Dec 26;14(12):822–38. PMID: 36619695. doi: 10.4252/wjsc.v14.i12.822.
8. Zoch A, Auchynnikava T, Berrens RV, Kabayama Y, Schopp T, Heep M, et al. SPOCD1 is an essential executor of piRNA-directed de novo DNA methylation. Nature. 2020 Aug;584(7822):635–9. PMID: 32674113. doi: 10.1038/s41586-020-2557-5.
9. Park KS, Cha Y, Kim CH, Ahn HJ, Kim D, Ko S, et al. Transcription elongation factor Tcea3 regulates the pluripotent differentiation potential of mouse embryonic stem cells via the Lefty1-Nodal-Smad2 pathway. Stem Cells. 2013 Feb;31(2):282–92. PMID: 23169579. doi: 10.1002/stem.1284.
10. Joiner ML, Koval OM, Li J, He BJ, Allamargot C, Gao Z, et al. CaMKII determines mitochondrial stress responses in heart. Nature. 2012 Nov 8;491(7423):269–73. PMID: 23051746. doi: 10.1038/nature11444.
11. Drolez A, Vandenhaute E, Delannoy CP, Dewald JH, Gosselet F, Cecchelli R, et al. ST6GALNAC5 Expression Decreases the Interactions between Breast Cancer Cells and the Human Blood-Brain Barrier. Int J Mol Sci. 2016 Aug 11;17(8). PMID: 27529215. doi: 10.3390/ijms17081309.
12. Zhang S, Morita M, Wang Z, Ooehara J, Zhang S, Xie M, et al. Interleukin-12 supports in vitro self-renewal of long-term hematopoietic stem cells. Blood Sci. 2019 Aug;1(1):92–101. PMID: 35402790. doi: 10.1097/BS9.0000000000000002.

13. Ladle BH, Li KP, Phillips MJ, Pucsek AB, Haile A, Powell JD, et al. De novo DNA methylation by DNA methyltransferase 3a controls early effector CD8+ T-cell fate decisions following activation. Proc Natl Acad Sci U S A. 2016 Sep 20;113(38):10631–6. PMID: 27582468. doi: 10.1073/pnas.1524490113.
14. Song H, Zhang Y, Liu J, Liu W, Luo B. Activation of DNA methyltransferase 3a by Epstein-Barr nuclear antigen 1 in gastric carcinoma. Dig Liver Dis. 2022 Jul;54(7):973–83. PMID: 34215536. doi: 10.1016/j.dld.2021.06.004.
15. Meng K, Hu Y, Wang D, Li Y, Shi F, Lu J, et al. EFHD1, a novel mitochondrial regulator of tumor metastasis in clear cell renal cell carcinoma. Cancer Sci. 2023 May;114(5):2029–40. PMID: 36747492. doi: 10.1111/cas.15749.
16. Brun J, Dieudonne FX, Marty C, Muller J, Schule R, Patino-Garcia A, et al. FHL2 silencing reduces Wnt signaling and osteosarcoma tumorigenesis in vitro and in vivo. PLoS One. 2013;8(1):e55034. PMID: 23383046. doi: 10.1371/journal.pone.0055034.
17. Zhang S, Zhang L, Zhang D, Guo Y, Gao Y, Jiang Z, et al. Four and a half LIM domains 2 (FHL2) attenuates tumorigenesis of gastrointestinal stromal tumors (GISTs) by negatively regulating KIT signaling. Mol Carcinog. 2024 Jul;63(7):1334–48. PMID: 38629424. doi: 10.1002/mc.23727.
18. Hua G, He C, Lv X, Fan L, Wang C, Remmenga SW, et al. The four and a half LIM domains 2 (FHL2) regulates ovarian granulosa cell tumor progression via controlling AKT1 transcription. Cell Death Dis. 2016 Jul 14;7(7):e2297. PMID: 27415427. doi: 10.1038/cddis.2016.207.
19. Principe C, Dionisio de Sousa IJ, Prazeres H, Soares P, Lima RT. LRP1B: A Giant Lost in Cancer Translation. Pharmaceuticals (Basel). 2021 Aug 24;14(9). PMID: 34577535. doi: 10.3390/ph14090836.
20. Koroknai V, Szasz I, Hernandez-Vargas H, Fernandez-Jimenez N, Cuenin C, Herceg Z, et al. DNA hypermethylation is associated with invasive phenotype of malignant melanoma. Exp Dermatol. 2020 Jan;29(1):39–50. PMID: 31602702. doi: 10.1111/exd.14047.
21. Ozhan G, Sezgin E, Wehner D, Pfister AS, Kuhl SJ, Kagermeier-Schenk B, et al. Lypd6 enhances Wnt/beta-catenin signaling by promoting Lrp6 phosphorylation in raft plasma membrane domains. Dev Cell. 2013 Aug 26;26(4):331–45. PMID: 23987510. doi: 10.1016/j.devcel.2013.07.020.
22. Kutejova E, Engist B, Self M, Oliver G, Kirilenko P, Bobola N. Six2 functions redundantly immediately downstream of Hoxa2. Development. 2008 Apr;135(8):1463–70. PMID: 18321982. doi: 10.1242/dev.017624.
23. Shrestha P, Jaganathan A, Huilgol D, Ballon C, Hwangbo Y, Mills AA. Chd5 Regulates the Transcription Factor Six3 to Promote Neuronal Differentiation. Stem Cells. 2023 Mar 17;41(3):242–51. PMID: 36636025. doi: 10.1093/stmcls/sxad002.
24. Ma TL, Zhu P, Chen JX, Hu YH, Xie J. SIX3 function in cancer: progression and comprehensive analysis. Cancer Gene Ther. 2022 Nov;29(11):1542–9. PMID: 35764712. doi: 10.1038/s41417-022-00488-9.
25. Liang R, Liu Y. Tcf7l1 directly regulates cardiomyocyte differentiation in embryonic stem cells. Stem Cell Res Ther. 2018 Oct 11;9(1):267. PMID: 30326964. doi: 10.1186/s13287-018-1015-x.
26. An P, Wei LL, Zhao S, Sverdlov DY, Vaid KA, Miyamoto M, et al. Hepatocyte mitochondria-derived danger signals directly activate hepatic stellate cells and drive progression of liver fibrosis. Nat Commun. 2020 May 12;11(1):2362. PMID: 32398673. doi: 10.1038/s41467-020-16092-0.
27. Jiang Y, Tang Y, Hoover C, Kondo Y, Huang D, Restagno D, et al. Kupffer cell receptor CLEC4F is important for the destruction of desialylated platelets in mice. Cell Death Differ. 2021 Nov;28(11):3009–21. PMID: 33993195. doi: 10.1038/s41418-021-00797-w.

28. Myers MA, Georgiou HM, Byron S, Esposti MD. Inhibition of mitochondrial oxidative phosphorylation induces hyper-expression of glutamic acid decarboxylase in pancreatic islet cells. Autoimmunity. 1999;30(1):43–51. PMID: 10433094. doi: 10.3109/08916939908994759.
29. Lee YY, Chao TB, Sheu MJ, Tian YF, Chen TJ, Lee SW, et al. Glutamate Decarboxylase 1 Overexpression as a Poor Prognostic Factor in Patients with Nasopharyngeal Carcinoma. J Cancer. 2016;7(12):1716–23. PMID: 27698909. doi: 10.7150/jca.15667.
30. Xiong J, Dai YT, Wang WF, Zhang H, Wang CF, Yin T, et al. GPCR signaling contributes to immune characteristics of microenvironment and process of EBV-induced lymphomagenesis. Sci Bull (Beijing). 2023 Nov 15;68(21):2607–19. PMID: 37798178. doi: 10.1016/j.scib.2023.09.029.
31. Jin Y, Cai Q, Wang L, Ji J, Sun Y, Jiang J, et al. Paracrine activin B-NF-kappaB signaling shapes an inflammatory tumor microenvironment in gastric cancer via fibroblast reprogramming. J Exp Clin Cancer Res. 2023 Oct 19;42(1):269. PMID: 37858201. doi: 10.1186/s13046-023-02861-4.
32. Yu W, He G, Zhang W, Ye Z, Zhong Z, Huang S. INHBB is a novel prognostic biomarker and correlated with immune infiltrates in gastric cancer. Front Genet. 2022;13:933862. PMID: 36118865. doi: 10.3389/fgene.2022.933862.
33. Xiaoai Zhao XY, Kévin Contrepois, Francesco Vallania, Mathew Ellenberger, Chloe M. Kashiwagi, Stephanie D. Gagnon, Cynthia J. Siebrand, Matias Cabruja, Gavin M. Traber, Andrew McKay, Daniel Hornburg, Purvesh Khatri, Michael P. Snyder, Richard N. Zare, Anne Brunet. Lipidomic profiling reveals age-dependent changes in complex plasma membrane lipids that regulate neural stem cell aging. Biorxiv. 2022.
34. Bao X, Zhang J, Huang G, Yan J, Xu C, Dou Z, et al. The crosstalk between HIFs and mitochondrial dysfunctions in cancer development. Cell Death Dis. 2021 Feb 26;12(2):215. PMID: 33637686. doi: 10.1038/s41419-021-03505-1.
35. Hao YY, Xiao WQ, Zhang HN, Yu NN, Park G, Han YH, et al. Peroxiredoxin 1 modulates oxidative stress resistance and cell apoptosis through stemness in liver cancer under non-thermal plasma treatment. Biochem Biophys Res Commun. 2024 Dec 17;738:150522. PMID: 39154551. doi: 10.1016/j.bbrc.2024.150522.
36. Landry MC, Champagne C, Boulanger MC, Jette A, Fuchs M, Dziengelewski C, et al. A functional interplay between the small GTPase Rab11a and mitochondria-shaping proteins regulates mitochondrial positioning and polarization of the actin cytoskeleton downstream of Src family kinases. J Biol Chem. 2014 Jan 24;289(4):2230–49. PMID: 24302731. doi: 10.1074/jbc.M113.516351.
37. Suekane A, Saito Y, Nakahata S, Ichikawa T, Ogoh H, Tsujikawa K, et al. CGRP-CRLR/RAMP1 signal is important for stress-induced hematopoiesis. Sci Rep. 2019 Jan 23;9(1):429. PMID: 30674976. doi: 10.1038/s41598-018-36796-0.
38. Sewduth RN, Jaspard-Vinassa B, Peghaire C, Guillabert A, Franzl N, Larrieu-Lahargue F, et al. The ubiquitin ligase PDZRN3 is required for vascular morphogenesis through Wnt/planar cell polarity signalling. Nat Commun. 2014 Sep 8;5:4832. PMID: 25198863. doi: 10.1038/ncomms5832.
39. Konopelski Snavely SE, Susman MW, Kunz RC, Tan J, Srinivasan S, Cohen MD, et al. Proteomic analysis identifies the E3 ubiquitin ligase Pdzrn3 as a regulatory target of Wnt5a-Ror signaling. Proc Natl Acad Sci U S A. 2021 Jun 22;118(25). PMID: 34135125. doi: 10.1073/pnas.2104944118.
40. Wong QW, Li J, Ng SR, Lim SG, Yang H, Vardy LA. RPL39L is an example of a recently evolved ribosomal protein paralog that shows highly specific tissue expression patterns and is upregulated in ESCs and HCC tumors. RNA Biol. 2014;11(1):33–41. PMID: 24452241. doi: 10.4161/rna.27427.

41. Wu X, Zhang Z, Qiu Z, Wu X, Chen J, Liu L, et al. TNIK in disease: from molecular insights to therapeutic prospects. Apoptosis. 2024 Oct;29(9-10):1361–76. PMID: 38853204. doi: 10.1007/s10495-024-01987-w.
42. Chen W, Tang D, Tang D, Dai Y. Epigenetic silencing of ZIC4 contributes to cancer progression in hepatocellular carcinoma. Cell Death Dis. 2020 Oct 23;11(10):906. PMID: 33097694. doi: 10.1038/s41419-020-03109-1.
43. Zhou Y, Yan J, Huang H, Liu L, Ren L, Hu J, et al. The m(6)A reader IGF2BP2 regulates glycolytic metabolism and mediates histone lactylation to enhance hepatic stellate cell activation and liver fibrosis. Cell Death Dis. 2024 Mar 5;15(3):189. PMID: 38443347. doi: 10.1038/s41419-024-06509-9.
44. Lechan RM, Fekete C. The TRH neuron: a hypothalamic integrator of energy metabolism. Prog Brain Res. 2006;153:209–35. PMID: 16876577. doi: 10.1016/S0079-6123(06)53012-2.
45. Nillni EA. Regulation of the hypothalamic thyrotropin releasing hormone (TRH) neuron by neuronal and peripheral inputs. Front Neuroendocrinol. 2010 Apr;31(2):134–56. PMID: 20074584. doi: 10.1016/j.yfrne.2010.01.001.
46. Chen ZG, Saba NF, Teng Y. The diverse functions of FAT1 in cancer progression: good, bad, or ugly? J Exp Clin Cancer Res. 2022 Aug 15;41(1):248. PMID: 35965328. doi: 10.1186/s13046-022-02461-8.
47. Liebig S, Neumann M, Silva P, Ortiz-Tanchez J, Schulze V, Isaakidis K, et al. FAT1 expression in T-cell acute lymphoblastic leukemia (T-ALL) modulates proliferation and WNT signaling. Sci Rep. 2023 Jan 18;13(1):972. PMID: 36653435. doi: 10.1038/s41598-023-27792-0.
48. Coombs ID, Ziobro J, Krotov V, Surtees TL, Cull-Candy SG, Farrant M. A gain-of-function GRIA2 variant associated with neurodevelopmental delay and seizures: Functional characterization and targeted treatment. Epilepsia. 2022 Dec;63(12):e156–e63. PMID: 36161652. doi: 10.1111/epi.17419.
49. Fujiwara H, Ferreira M, Donati G, Marciano DK, Linton JM, Sato Y, et al. The basement membrane of hair follicle stem cells is a muscle cell niche. Cell. 2011 Feb 18;144(4):577–89. PMID: 21335239. doi: 10.1016/j.cell.2011.01.014.
50. Iannotta L, Emanuele M, Favetta G, Tombesi G, Vandewynckel L, Lara Ordonez AJ, et al. PAK6-mediated phosphorylation of PPP2R2C regulates LRRK2-PP2A complex formation. Front Mol Neurosci. 2023;16:1269387. PMID: 38169846. doi: 10.3389/fnmol.2023.1269387.
51. Snippert HJ, van Es JH, van den Born M, Begthel H, Stange DE, Barker N, et al. Prominin-1/CD133 marks stem cells and early progenitors in mouse small intestine. Gastroenterology. 2009 Jun;136(7):2187–94 e1. PMID: 19324043. doi: 10.1053/j.gastro.2009.03.002.
52. Scotland KB, Chen S, Sylvester R, Gudas LJ. Analysis of Rex1 (zfp42) function in embryonic stem cell differentiation. Dev Dyn. 2009 Aug;238(8):1863–77. PMID: 19618472. doi: 10.1002/dvdy.22037.
53. Ma S, Wan X, Deng Z, Shi L, Hao C, Zhou Z, et al. Epigenetic regulator CXXC5 recruits DNA demethylase Tet2 to regulate TLR7/9-elicited IFN response in pDCs. The Journal of experimental medicine. 2017 May 1;214(5):1471–91. PMID: 28416650. doi: 10.1084/jem.20161149.
54. Joshi HR, Hill HR, Zhou Z, He X, Voelkerding KV, Kumanovics A. Frontline Science: Cxxc5 expression alters cell cycle and myeloid differentiation of mouse hematopoietic stem and progenitor cells. J Leukoc Biol. 2020 Aug;108(2):469–84. PMID: 32083332. doi: 10.1002/JLB.1HI0120-169R.
55. Lin KM, Weng LF, Chen SJ, Lin SJ, Tsai CH. Upregulation of IQGAP2 by EBV transactivator Rta and its influence on EBV life cycle. Journal of virology. 2023 Aug 31;97(8):e0054023. PMID: 37504571. doi: 10.1128/jvi.00540-23.

56. Sekiya A, Takasawa K, Arai Y, Horike SI, Akutsu H, Umezawa A, et al. Variation of DNA methylation on the IRX1/2 genes is responsible for the neural differentiation propensity in human induced pluripotent stem cells. Regen Ther. 2022 Dec;21:620–30. PMID: 36514370. doi: 10.1016/j.reth.2022.11.007.
57. Pan H, Xu R, Zhang Y. Role of SPRY4 in health and disease. Front Oncol. 2024;14:1376873. PMID: 38686189. doi: 10.3389/fonc.2024.1376873.
58. Siesser PF, Motolese M, Walker MP, Goldfarb D, Gewain K, Yan F, et al. FAM123A binds to microtubules and inhibits the guanine nucleotide exchange factor ARHGEF2 to decrease actomyosin contractility. Sci Signal. 2012 Sep 4;5(240):ra64. PMID: 22949735. doi: 10.1126/scisignal.2002871.
59. Esseltine JL, Brooks CR, Edwards NA, Subasri M, Sampson J, Seguin C, et al. Dynamic regulation of connexins in stem cell pluripotency. Stem Cells. 2020 Jan;38(1):52–66. PMID: 31646713. doi: 10.1002/stem.3092.
60. Hatabi K, Hirohara Y, Kushida Y, Kuroda Y, Wakao S, Trosko J, et al. Inhibition of Gap Junctional Intercellular Communication Upregulates Pluripotency Gene Expression in Endogenous Pluripotent Muse Cells. Cells. 2022 Aug 30;11(17). PMID: 36078111. doi: 10.3390/cells11172701.
61. Krasteva V, Crabtree GR, Lessard JA. The BAF45a/PHF10 subunit of SWI/SNF-like chromatin remodeling complexes is essential for hematopoietic stem cell maintenance. Exp Hematol. 2017 Apr;48:58–71 e15. PMID: 27931852. doi: 10.1016/j.exphem.2016.11.008.
62. Grifone R, Shao M, Saquet A, Shi DL. RNA-Binding Protein Rbm24 as a Multifaceted Post-Transcriptional Regulator of Embryonic Lineage Differentiation and Cellular Homeostasis. Cells. 2020 Aug 12;9(8). PMID: 32806768. doi: 10.3390/cells9081891.
63. Shi DL, Grifone R, Zhang X, Li H. Rbm24-mediated post-transcriptional regulation of skeletal and cardiac muscle development, function and regeneration. J Muscle Res Cell Motil. 2025 Mar;46(1):53–65. PMID: 39614020. doi: 10.1007/s10974-024-09685-5.
64. Zhang J, Liao JQ, Wen LR, Padhiar AA, Li Z, He ZY, et al. Rps6ka2 enhances iMSC chondrogenic differentiation to attenuate knee osteoarthritis through articular cartilage regeneration in mice. Biochem Biophys Res Commun. 2023 Jun 30;663:61–70. PMID: 37119767. doi: 10.1016/j.bbrc.2023.04.049.
65. Li J, Wan X, Xie D, Yuan H, Pei Q, Luo Y, et al. SPDEF enhances cancer stem cell-like properties and tumorigenesis through directly promoting GALNT7 transcription in luminal breast cancer. Cell Death Dis. 2023 Aug 26;14(8):569. PMID: 37633945. doi: 10.1038/s41419-023-06098-z.
66. Hovland AS, Bhattacharya D, Azambuja AP, Pramio D, Copeland J, Rothstein M, et al. Pluripotency factors are repurposed to shape the epigenomic landscape of neural crest cells. Dev Cell. 2022 Oct 10;57(19):2257–72 e5. PMID: 36182685. doi: 10.1016/j.devcel.2022.09.006.
67. Krendl C, Shaposhnikov D, Rishko V, Ori C, Ziegenhain C, Sass S, et al. GATA2/3-TFAP2A/C transcription factor network couples human pluripotent stem cell differentiation to trophectoderm with repression of pluripotency. Proc Natl Acad Sci U S A. 2017 Nov 7;114(45):E9579–E88. PMID: 29078328. doi: 10.1073/pnas.1708341114.
68. Brombin A, Simpson DJ, Travnickova J, Brunsdon H, Zeng Z, Lu Y, et al. Tfap2b specifies an embryonic melanocyte stem cell that retains adult multifate potential. Cell Rep. 2022 Jan 11;38(2):110234. PMID: 35021087. doi: 10.1016/j.celrep.2021.110234.
69. Jin C, Luo Y, Liang Z, Li X, Kolat D, Zhao L, et al. Crucial role of the transcription factors family activator protein 2 in cancer: current clue and views. J Transl Med. 2023 Jun 8;21(1):371. PMID: 37291585. doi: 10.1186/s12967-023-04189-1.

70. Du W, Amarachintha S, Wilson A, Pang Q. The immune receptor Trem1 cooperates with diminished DNA damage response to induce preleukemic stem cell expansion. Leukemia. 2017 Feb;31(2):423–33. PMID: 27568523. doi: 10.1038/leu.2016.242.
71. Fearnley GW, Young KA, Edgar JR, Antrobus R, Hay IM, Liang WC, et al. The homophilic receptor PTPRK selectively dephosphorylates multiple junctional regulators to promote cell-cell adhesion. Elife. 2019 Mar 29;8. PMID: 30924770. doi: 10.7554/eLife.44597.
72. Storm EE, Durinck S, de Sousa e Melo F, Tremayne J, Kljavin N, Tan C, et al. Targeting PTPRK-RSPO3 colon tumours promotes differentiation and loss of stem-cell function. Nature. 2016 Jan 7;529(7584):97–100. PMID: 26700806. doi: 10.1038/nature16466.
73. Zheng C, Liu T, Wang AQ, Chen XA, Zhang RZ, Wang XC, et al. Protein tyrosine phosphatase receptor type kappa (PTPRK) revisited: evolving insights into structure, function, and pathology. Journal of translational medicine. 2025 May 12;23(1):534. PMID: 40355891. doi: 10.1186/s12967-025-06496-1.
74. Bar S, Vershkov D, Keshet G, Lezmi E, Meller N, Yilmaz A, et al. Identifying regulators of parental imprinting by CRISPR/Cas9 screening in haploid human embryonic stem cells. Nat Commun. 2021 Nov 18;12(1):6718. PMID: 34795250. doi: 10.1038/s41467-021-26949-7.
75. Nagare RP, Sneha S, Sidhanth C, Roopa S, Murhekar K, Shirley S, et al. Expression of cancer stem cell markers CD24, EPHA1 and CD9 and their correlation with clinical outcome in epithelial ovarian tumours. Cancer Biomark. 2020;28(3):397–408. PMID: 32224528. doi: 10.3233/CBM-201463.
76. Ferrell CM, Dorsam ST, Ohta H, Humphries RK, Derynck MK, Haqq C, et al. Activation of stem-cell specific genes by HOXA9 and HOXA10 homeodomain proteins in CD34+ human cord blood cells. Stem Cells. 2005 May;23(5):644–55. PMID: 15849172. doi: 10.1634/stemcells.2004-0198.
77. Caldas P, Luz M, Baseggio S, Andrade R, Sobral D, Grosso AR. Transcription readthrough is prevalent in healthy human tissues and associated with inherent genomic features. Commun Biol. 2024 Jan 15;7(1):100. PMID: 38225287. doi: 10.1038/s42003-024-05779-5.
78. Semerci F, Choi WT, Bajic A, Thakkar A, Encinas JM, Depreux F, et al. Lunatic fringe-mediated Notch signaling regulates adult hippocampal neural stem cell maintenance. Elife. 2017 Jul 12;6. PMID: 28699891. doi: 10.7554/eLife.24660.
79. Ramalho S, Alkan F, Prekovic S, Jastrzebski K, Barbera EP, Hoekman L, et al. NAC regulates metabolism and cell fate in intestinal stem cells. Sci Adv. 2025 Jan 10;11(2):eadn9750. PMID: 39772672. doi: 10.1126/sciadv.adn9750.
80. Wu F, Gao G, Pan T, Yang Z, Xu G, Abbas N, et al. Generation of a SMO homozygous knockout human embryonic stem cell line WAe001-A-16 by CRISPR/Cas9 editing. Stem Cell Res. 2018 Mar;27:5–9. PMID: 29278761. doi: 10.1016/j.scr.2017.12.002.
81. Ma L, Yin W, Ma H, Elshoura I, Wang L. Targeting claudin-3 suppresses stem cell-like phenotype in nonsquamous non-small-cell lung carcinoma. Lung Cancer Manag. 2019 Feb;8(1):LMT04. PMID: 31044015. doi: 10.2217/lmt-2018-0010.
82. Rennich BJ, Luth ES, Moores S, Juo P. Regulation of AMPA receptor trafficking by secreted protein factors. Front Cell Neurosci. 2023;17:1271169. PMID: 38089145. doi: 10.3389/fncel.2023.1271169.
83. Moore A, Chinnaiya K, Kim DW, Brown S, Stewart I, Robins S, et al. Loss of Function of the Neural Cell Adhesion Molecule NrCAM Regulates Differentiation, Proliferation and Neurogenesis in Early Postnatal Hypothalamic Tanycytes. Front Neurosci. 2022;16:832961. PMID: 35464310. doi: 10.3389/fnins.2022.832961.

84. Stojilkovic SS, Sokanovic SJ, Constantin S. What is known and unknown about the role of neuroendocrine genes Ptprn and Ptprn2. Front Endocrinol (Lausanne). 2025;16:1531723. PMID: 39926347. doi: 10.3389/fendo.2025.1531723.
85. Fukazawa T, Tanimoto K, Yamaoka E, Kojima M, Kanawa M, Hirohashi N, et al. Oncogenic Role of ADAM32 in Hepatoblastoma: A Potential Molecular Target for Therapy. Cancers (Basel). 2022 Sep 28;14(19). PMID: 36230656. doi: 10.3390/cancers14194732.
86. Paul D, , Rapole,T, Santra, M. . A New Player in an Old Story: FBXO16 Prevents Breast Cancer Tumorigenesis through Disrupting Cellular Function of Nuclear β-Catenin. J Esperimental Pathology. 2020;1(1). doi: DOI: https://doi.org/10.33696/pathology.1.004.
87. Cai EY, Kufeld MN, Schuster S, Arora S, Larkin M, Germanos AA, et al. Selective Translation of Cell Fate Regulators Mediates Tolerance to Broad Oncogenic Stress. Cell Stem Cell. 2020 Aug 6;27(2):270–83 e7. PMID: 32516567. doi: 10.1016/j.stem.2020.05.007.
88. Liao J, Wang J, Xu Y, Wu Y, Wang M, Zhao Q, et al. LAPTM4B-YAP loop feedback amplification enhances the stemness of hepatocellular carcinoma. iScience. 2023 Jun 16;26(6):106754. PMID: 37213231. doi: 10.1016/j.isci.2023.106754.
89. Lee U, Zhang Y, Zhu Y, Luo AC, Gong L, Tremmel DM, et al. Robust differentiation of human pluripotent stem cells into mural progenitor cells via transient activation of NKX3.1. Nat Commun. 2024 Sep 30;15(1):8392. PMID: 39349465. doi: 10.1038/s41467-024-52678-8.
90. Hou R, Sun X, Cao S, Wang Y, Jiang L. Stabilization of SQLE mRNA by WTAP/FTO/IGF2BP3-dependent manner in HGSOC: implications for metabolism, stemness, and progression. Cell Death Dis. 2024 Dec 1;15(12):872. PMID: 39617776. doi: 10.1038/s41419-024-07257-6.
91. Shibata S, Teshima Y, Niimi K, Inagaki S. Involvement of ARHGEF10, GEF for RhoA, in Rab6/Rab8-mediating membrane traffic. Small GTPases. 2019 May;10(3):169–77. PMID: 28448737. doi: 10.1080/21541248.2017.1302550.
92. Kim DY, Sub YJ, Kim HY, Cho KJ, Choi WI, Choi YJ, et al. LRRC6 regulates biogenesis of motile cilia by aiding FOXJ1 translocation into the nucleus. Cell Commun Signal. 2023 Jun 16;21(1):142. PMID: 37328841. doi: 10.1186/s12964-023-01135-y.
93. Zhang C, Jia P, Jia Y, Weissbach H, Webster KA, Huang X, et al. Methionine sulfoxide reductase A (MsrA) protects cultured mouse embryonic stem cells from H2O2-mediated oxidative stress. J Cell Biochem. 2010 Sep 1;111(1):94–103. PMID: 20506347. doi: 10.1002/jcb.22666.
94. Baechler SA, Factor VM, Dalla Rosa I, Ravji A, Becker D, Khiati S, et al. The mitochondrial type IB topoisomerase drives mitochondrial translation and carcinogenesis. Nat Commun. 2019 Jan 8;10(1):83. PMID: 30622257. doi: 10.1038/s41467-018-07922-3.
95. Mechaussier S, Almoallem B, Zeitz C, Van Schil K, Jeddawi L, Van Dorpe J, et al. Loss of Function of RIMS2 Causes a Syndromic Congenital Cone-Rod Synaptic Disease with Neurodevelopmental and Pancreatic Involvement. Am J Hum Genet. 2020 Sep 3;107(3):580. PMID: 32888510. doi: 10.1016/j.ajhg.2020.08.004.
96. Dai Y, Kawaguchi T, Nishio M, Otani J, Tashiro H, Terai Y, et al. The TIGD5 gene located in 8q24 and frequently amplified in ovarian cancers is a tumor suppressor. Genes Cells. 2022 Nov;27(11):633–42. PMID: 36054307. doi: 10.1111/gtc.12980.
97. Yi H, Wang, J.,,Kavallaris, Wang,J. gr4-Mediated Potentiation Of Wnt/β-Catenin Signaling Promotes MLL Leukemogenesis Via An Rspo3/Wnt3a-Gnaq Pathway In Leukemic Stem Cells. Blood. 2013;122 (21). doi: https://doi.org/10.1182/blood.V122.21.887.887.
98. Liu W, Huang Y, Wang D, Han F, Chen H, Chen J, et al. MPDZ as a novel epigenetic silenced tumor suppressor inhibits growth and progression of lung cancer through the Hippo-YAP pathway. Oncogene. 2021 Jul;40(26):4468–85. PMID: 34108620. doi: 10.1038/s41388-021-01857-8.

99. Becker-Santos DD, Lonergan KM, Gronostajski RM, Lam WL. Nuclear Factor I/B: A Master Regulator of Cell Differentiation with Paradoxical Roles in Cancer. EBioMedicine. 2017 Aug;22:2–9. PMID: 28596133. doi: 10.1016/j.ebiom.2017.05.027.
100. Moretton A, Kourtis S, Ganez Zapater A, Calabro C, Espinar Calvo ML, Fontaine F, et al. A metabolic map of the DNA damage response identifies PRDX1 in the control of nuclear ROS scavenging and aspartate availability. Molecular systems biology. 2023 Jul 11;19(7):e11267. PMID: 37259925. doi: 10.15252/msb.202211267.
101. Anholt RR. Olfactomedin proteins: central players in development and disease. Front Cell Dev Biol. 2014;2:6. PMID: 25364714. doi: 10.3389/fcell.2014.00006.
102. Hartjes L, Ruland J. CARD9 Signaling in Intestinal Immune Homeostasis and Oncogenesis. Front Immunol. 2019;10:419. PMID: 30906296. doi: 10.3389/fimmu.2019.00419.
103. Lu J, Luo Y, Rao D, Wang T, Lei Z, Chen X, et al. Myeloid-derived suppressor cells in cancer: therapeutic targets to overcome tumor immune evasion. Exp Hematol Oncol. 2024 Apr 12;13(1):39. PMID: 38609997. doi: 10.1186/s40164-024-00505-7.
104. Ohmoto M, Jyotaki M, Yee KK, Matsumoto I. A Transcription Factor Etv1/Er81 Is Involved in the Differentiation of Sweet, Umami, and Sodium Taste Cells. eNeuro. 2023 Apr;10(4). PMID: 37045597. doi: 10.1523/ENEURO.0236-22.2023.
105. Le Marer N, Palcic MM, Clarke JL, Davies D, Skacel PO. Developmental regulation of alpha 1,3-fucosyltransferase expression in CD34 positive progenitors and maturing myeloid cells isolated from normal human bone marrow. Glycobiology. 1997 Apr;7(3):357–65. PMID: 9147044. doi: 10.1093/glycob/7.3.357.
106. Oshima H, Ishikawa T, Yoshida GJ, Naoi K, Maeda Y, Naka K, et al. TNF-alpha/TNFR1 signaling promotes gastric tumorigenesis through induction of Noxo1 and Gna14 in tumor cells. Oncogene. 2014 Jul 17;33(29):3820–9. PMID: 23975421. doi: 10.1038/onc.2013.356.
107. Xiao D, Su X, Gao H, Li X, Qu Y. The Roles of Lpar1 in Central Nervous System Disorders and Diseases. Front Neurosci. 2021;15:710473. PMID: 34385905. doi: 10.3389/fnins.2021.710473.
108. Li M, Lv Y, Chen F, Wang X, Zhu J, Li H, et al. Co-stimulation of LPAR(1) and S1PR(1/3) increases the transplantation efficacy of human mesenchymal stem cells in drug-induced and alcoholic liver diseases. Stem Cell Res Ther. 2018 Jun 14;9(1):161. PMID: 29898789. doi: 10.1186/s13287-018-0860-y.
109. Temme L, Laitinen T, Asquith CRM. PKN3: a target in cancer metastasis. Nat Rev Drug Discov. 2022 Oct;21(10):704. PMID: 36104477. doi: 10.1038/d41573-022-00154-7.
110. You MH. Mechanism of DAPK1 for Regulating Cancer Stem Cells in Thyroid Cancer. Curr Issues Mol Biol. 2024 Jul 5;46(7):7086–96. PMID: 39057063. doi: 10.3390/cimb46070422.
111. Yuan W, Ji J, Shu Y, Chen J, Liu S, Wu L, et al. Downregulation of DAPK1 promotes the stemness of cancer stem cells and EMT process by activating ZEB1 in colorectal cancer. J Mol Med (Berl). 2019 Jan;97(1):89–102. PMID: 30460377. doi: 10.1007/s00109-018-1716-8.
112. Gharibi B, Ghuman M, Hughes FJ. DDIT4 regulates mesenchymal stem cell fate by mediating between HIF1alpha and mTOR signalling. Sci Rep. 2016 Nov 23;6:36889. PMID: 27876894. doi: 10.1038/srep36889.
113. Lee SH, Chiu YC, Li YH, Lin CC, Hou HA, Chou WC, et al. High expression of dedicator of cytokinesis 1 (DOCK1) confers poor prognosis in acute myeloid leukemia. Oncotarget. 2017 Sep 22;8(42):72250–9. PMID: 29069784. doi: 10.18632/oncotarget.19706.
114. Yasukawa M, Ishida K, Yuge Y, Hanaoka M, Minami Y, Ogawa M, et al. Dpysl4 is involved in tooth germ morphogenesis through growth regulation, polarization and differentiation of dental epithelial cells. Int J Biol Sci. 2013;9(4):382–90. PMID: 23630450. doi: 10.7150/ijbs.5510.

115. Liu X, Liu X, Du Y, Hu M, Tian Y, Li Z, et al. DUSP5 promotes osteogenic differentiation through SCP1/2-dependent phosphorylation of SMAD1. Stem Cells. 2021 Oct;39(10):1395–409. PMID: 34169608. doi: 10.1002/stem.3428.
116. Kutty RG, Xin G, Schauder DM, Cossette SM, Bordas M, Cui W, et al. Dual Specificity Phosphatase 5 Is Essential for T Cell Survival. PloS one. 2016;11(12):e0167246. PMID: 27936095. doi: 10.1371/journal.pone.0167246.
117. Ku CJ, Hosoya T, Maillard I, Engel JD. GATA-3 regulates hematopoietic stem cell maintenance and cell-cycle entry. Blood. 2012 Mar 8;119(10):2242–51. PMID: 22267605. doi: 10.1182/blood-2011-07-366070.
118. Swartz ME, Lovely CB, Eberhart JK. Variation in phenotypes from a Bmp-Gata3 genetic pathway is modulated by Shh signaling. PLoS Genet. 2021 May;17(5):e1009579. PMID: 34033651. doi: 10.1371/journal.pgen.1009579.
119. He Z, Jiang J, Hofmann MC, Dym M. Gfra1 silencing in mouse spermatogonial stem cells results in their differentiation via the inactivation of RET tyrosine kinase. Biol Reprod. 2007 Oct;77(4):723–33. PMID: 17625109. doi: 10.1095/biolreprod.107.062513.
120. Barrero MJ, Sese B, Marti M, Izpisua Belmonte JC. Macro histone variants are critical for the differentiation of human pluripotent cells. J Biol Chem. 2013 May 31;288(22):16110–6. PMID: 23595991. doi: 10.1074/jbc.M113.466144.
121. Lang L, Chen F, Li Y, Shay C, Yang F, Dan H, et al. Adaptive c-Met-PLXDC2 Signaling Axis Mediates Cancer Stem Cell Plasticity to Confer Radioresistance-associated Aggressiveness in Head and Neck Cancer. Cancer Res Commun. 2023 Apr;3(4):659–71. PMID: 37089864. doi: 10.1158/2767-9764.CRC-22-0289.
122. Zhao B, Yin J, Ding L, Luo J, Luo J, Mu J, et al. SPAG6 regulates cell proliferation and apoptosis via TGF-beta/Smad signal pathway in adult B-cell acute lymphoblastic leukemia. Int J Hematol. 2024 Feb;119(2):119–29. PMID: 38147275. doi: 10.1007/s12185-023-03684-x.
123. Zhang R, Zhu H, Yuan Y, Wang Y, Tian Z. SPAG6 promotes cell proliferation and inhibits apoptosis through the PTEN/PI3K/AKT pathway in Burkitt lymphoma. Oncol Rep. 2020 Nov;44(5):2021–30. PMID: 33000212. doi: 10.3892/or.2020.7776.
124. Zheng DF, Wang Q, Wang JP, Bao ZQ, Wu SW, Ma L, et al. The Emerging Role of Sperm-Associated Antigen 6 Gene in the Microtubule Function of Cells and Cancer. Mol Ther Oncolytics. 2019 Dec 20;15:101–7. PMID: 31660426. doi: 10.1016/j.omto.2019.08.011.
125. Mijnes J, Bringezu S, Berger J, Schalla C, Rose M, von Serenyi S, et al. SPAG6 promotes cell migration and induces epithelial-to-mesenchymal transition in luminal breast cancer cells. bioRxiv. 2022:2022.03.24.485597. doi: 10.1101/2022.03.24.485597.
126. Jin T, Liu L. The Wnt signaling pathway effector TCF7L2 and type 2 diabetes mellitus. Mol Endocrinol. 2008 Nov;22(11):2383–92. PMID: 18599616. doi: 10.1210/me.2008-0135.
127. Del Bosque-Plata L, Hernandez-Cortes EP, Gragnoli C. The broad pathogenetic role of TCF7L2 in human diseases beyond type 2 diabetes. J Cell Physiol. 2022 Jan;237(1):301–12. PMID: 34612510. doi: 10.1002/jcp.30581.
128. Lomeli H. ZMIZ proteins: partners in transcriptional regulation and risk factors for human disease. J Mol Med (Berl). 2022 Jul;100(7):973–83. PMID: 35670836. doi: 10.1007/s00109-022-02216-0.
129. Shahi P, Wang CY, Lawson DA, Slorach EM, Lu A, Yu Y, et al. ZNF503/Zpo2 drives aggressive breast cancer progression by down-regulation of GATA3 expression. Proc Natl Acad Sci U S A. 2017 Mar 21;114(12):3169–74. PMID: 28258171. doi: 10.1073/pnas.1701690114.
130. Xavier-Ferrucio J, Ricon L, Vieira K, Longhini AL, Lazarini M, Bigarella CL, et al. Hematopoietic defects in response to reduced Arhgap21. Stem Cell Res. 2018 Jan;26:17–27. PMID: 29212046. doi: 10.1016/j.scr.2017.11.014.

131. Fan H, Wang X, Li W, Shen M, Wei Y, Zheng H, et al. ASB13 inhibits breast cancer metastasis through promoting SNAI2 degradation and relieving its transcriptional repression of YAP. Genes & development. 2020 Oct 1;34(19-20):1359–72. PMID: 32943576. doi: 10.1101/gad.339796.120.
132. Takayama K, Horie-Inoue K, Suzuki T, Urano T, Ikeda K, Fujimura T, et al. TACC2 is an androgen-responsive cell cycle regulator promoting androgen-mediated and castration-resistant growth of prostate cancer. Mol Endocrinol. 2012 May;26(5):748–61. PMID: 22456197. doi: 10.1210/me.2011-1242.
133. van der Flier LG, van Gijn ME, Hatzis P, Kujala P, Haegebarth A, Stange DE, et al. Transcription factor achaete scute-like 2 controls intestinal stem cell fate. Cell. 2009 Mar 6;136(5):903–12. PMID: 19269367. doi: 10.1016/j.cell.2009.01.031.
134. Fu Y, Zhou Z, Wang H, Gong P, Guo R, Wang J, et al. IFITM1 suppresses expression of human endogenous retroviruses in human embryonic stem cells. FEBS Open Bio. 2017 Aug;7(8):1102–10. PMID: 28781951. doi: 10.1002/2211-5463.12246.
135. Friedlova N, Zavadil Kokas F, Hupp TR, Vojtesek B, Nekulova M. IFITM protein regulation and functions: Far beyond the fight against viruses. Front Immunol. 2022;13:1042368. PMID: 36466909. doi: 10.3389/fimmu.2022.1042368.
136. Wang Y, Dong J, Li D, Lai L, Siwko S, Li Y, et al. Lgr4 regulates mammary gland development and stem cell activity through the pluripotency transcription factor Sox2. Stem Cells. 2013 Sep;31(9):1921–31. PMID: 23712846. doi: 10.1002/stem.1438.
137. Delgado ILS, Carmona B, Nolasco S, Santos D, Leitao A, Soares H. MOB: Pivotal Conserved Proteins in Cytokinesis, Cell Architecture and Tissue Homeostasis. Biology (Basel). 2020 Nov 24;9(12). PMID: 33255245. doi: 10.3390/biology9120413.
138. Martinez-Estrada OM, Lettice LA, Essafi A, Guadix JA, Slight J, Velecela V, et al. Wt1 is required for cardiovascular progenitor cell formation through transcriptional control of Snail and E-cadherin. Nat Genet. 2010 Jan;42(1):89–93. PMID: 20023660. doi: 10.1038/ng.494.
139. Werner AC, Weckbach LT, Salvermoser M, Pitter B, Cao J, Maier-Begandt D, et al. Coronin 1B Controls Endothelial Actin Dynamics at Cell-Cell Junctions and Is Required for Endothelial Network Assembly. Front Cell Dev Biol. 2020;8:708. PMID: 32850828. doi: 10.3389/fcell.2020.00708.
140. Niu X, Ma F, Li F, Wei C, Zhang J, Gao Z, et al. Integration of bioinformatics and cellular experiments unveils the role of SYT12 in gastric cancer. BMC cancer. 2024 Oct 29;24(1):1331. PMID: 39472897. doi: 10.1186/s12885-024-13077-w.
141. Wu J, Li J, Chen K, Liu G, Zhou Y, Chen W, et al. Atf7ip and Setdb1 interaction orchestrates the hematopoietic stem and progenitor cell state with diverse lineage differentiation. Proc Natl Acad Sci U S A. 2023 Jan 3;120(1):e2209062120. PMID: 36577070. doi: 10.1073/pnas.2209062120.
142. Littleton SH, Trang KB, Volpe CM, Cook K, DeBruyne N, Maguire JA, et al. Variant-to-function analysis of the childhood obesity chr12q13 locus implicates rs7132908 as a causal variant within the 3' UTR of FAIM2. Cell Genom. 2024 May 8;4(5):100556. PMID: 38697123. doi: 10.1016/j.xgen.2024.100556.
143. Cai J, Ye Z, Hu Y, Wang Y, Ye L, Gao L, et al. FAIM2 is a potential pan-cancer biomarker for prognosis and immune infiltration. Front Oncol. 2022;12:998336. PMID: 36185230. doi: 10.3389/fonc.2022.998336.
144. Dobish KK, Wittorf KJ, Swenson SA, Bean DC, Gavile CM, Woods NT, et al. FBXO21 mediated degradation of p85alpha regulates proliferation and survival of acute myeloid leukemia. Leukemia. 2023 Nov;37(11):2197–208. PMID: 37689825. doi: 10.1038/s41375-023-02020-w.
145. Yang W, Jing T, Wu C, Lu M, Yao X, Xia D, et al. F-box protein FBXO21 overexpression inhibits the proliferation and metastasis of clear cell renal cell carcinoma and is closely related to

the CREB pathway and tumor immune cell infiltration. J Transl Med. 2025 Mar 15;23(1):335. PMID: 40089779. doi: 10.1186/s12967-025-06356-y.
146. Hajjari M, Salavaty A. HOTAIR: an oncogenic long non-coding RNA in different cancers. Cancer Biol Med. 2015 Mar;12(1):1–9. PMID: 25859406. doi: 10.7497/j.issn.2095-3941.2015.0006.
147. Bhan A, Mandal SS. LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. Biochim Biophys Acta. 2015 Aug;1856(1):151–64. PMID: 26208723. doi: 10.1016/j.bbcan.2015.07.001.
148. Hakami MA, Hazazi A, Abdulaziz O, Almasoudi HH, Alhazmi AYM, Alkhalil SS, et al. HOTAIR: A key regulator of the Wnt/beta-catenin signaling cascade in cancer progression and treatment. Pathol Res Pract. 2024 Jan;253:154957. PMID: 38000201. doi: 10.1016/j.prp.2023.154957.
149. Bhatlekar S, Fields JZ, Boman BM. Role of HOX Genes in Stem Cell Differentiation and Cancer. Stem Cells Int. 2018;2018:3569493. PMID: 30154863. doi: 10.1155/2018/3569493.
150. Corgiat EB, List SM, Rounds JC, Corbett AH, Moberg KH. The RNA-binding protein Nab2 regulates the proteome of the developing Drosophila brain. J Biol Chem. 2021 Jul;297(1):100877. PMID: 34139237. doi: 10.1016/j.jbc.2021.100877.
151. Huang YH, Wang WL, Wang PH, Lee HT, Chang WW. EXOSC5 maintains cancer stem cell activity in endometrial cancer by regulating the NTN4/integrin beta1 signalling axis. Int J Biol Sci. 2024;20(1):265–79. PMID: 38164180. doi: 10.7150/ijbs.86275.
152. Suthon S, Perkins RS, Bryja V, Miranda-Carboni GA, Krum SA. WNT5B in Physiology and Disease. Front Cell Dev Biol. 2021;9:667581. PMID: 34017835. doi: 10.3389/fcell.2021.667581.
153. Wu Y, Liu W, Chen J, Liu S, Wang M, Yang L, et al. Nuclear Exosome Targeting Complex Core Factor Zcchc8 Regulates the Degradation of LINE1 RNA in Early Embryos and Embryonic Stem Cells. Cell Rep. 2019 Nov 19;29(8):2461–72 e6. PMID: 31747613. doi: 10.1016/j.celrep.2019.10.055.
154. Kim JA. Peroxisome Metabolism in Cancer. Cells. 2020 Jul 14;9(7). PMID: 32674458. doi: 10.3390/cells9071692.
155. Hu X, Zhang L, Li Y, Ma X, Dai W, Gao X, et al. Organoid modelling identifies that DACH1 functions as a tumour promoter in colorectal cancer by modulating BMP signalling. EBioMedicine. 2020 Jun;56:102800. PMID: 32512510. doi: 10.1016/j.ebiom.2020.102800.
156. Ma Y, Wei J, Song J, Hu Z, Zhang R, Li Z, et al. The DACH1 Gene Transcriptional Activation and Protein Degradation Mediated by Transactivator Tas of Prototype Foamy Virus. Viruses. 2023 Sep 8;15(9). PMID: 37766305. doi: 10.3390/v15091899.
157. Gilliland DG, Griffin JD. The roles of FLT3 in hematopoiesis and leukemia. Blood. 2002 Sep 1;100(5):1532–42. PMID: 12176867. doi: 10.1182/blood-2002-02-0492.
158. Ye B, Dai Z, Liu B, Wang R, Li C, Huang G, et al. Pcid2 inactivates developmental genes in human and mouse embryonic stem cells to sustain their pluripotency by modulation of EID1 stability. Stem Cells. 2014 Mar;32(3):623–35. PMID: 24167073. doi: 10.1002/stem.1580.
159. Kretzschmar K, Boonekamp KE, Bleijs M, Asra P, Koomen M, Chuva de Sousa Lopes SM, et al. Troy/Tnfrsf19 marks epidermal cells that govern interfollicular epidermal renewal and cornification. Stem Cell Reports. 2021 Sep 14;16(9):2379–94. PMID: 34358453. doi: 10.1016/j.stemcr.2021.07.007.
160. Luo Z, Gao X, Lin C, Smith ER, Marshall SA, Swanson SK, et al. Zic2 is an enhancer-binding factor required for embryonic stem cell specification. Molecular cell. 2015 Feb 19;57(4):685–94. PMID: 25699711. doi: 10.1016/j.molcel.2015.01.007.
161. Gabel AM, Crosse EI, Belleville AE, Hogg SJ, McKellar SA, Abdel-Wahab O, et al. Muscleblind-like proteins are novel modulators of the tumor-immune microenvironment. PLoS One. 2025;20(4):e0321148. PMID: 40305509. doi: 10.1371/journal.pone.0321148.

162. Setiawan AM, Kamarudin TA, Abd Ghafar N. The role of BMP4 in adipose-derived stem cell differentiation: A minireview. Front Cell Dev Biol. 2022;10:1045103. PMID: 36340030. doi: 10.3389/fcell.2022.1045103.
163. Fan S, Jiao Y, Khan R, Jiang X, Javed AR, Ali A, et al. Homozygous mutations in C14orf39/SIX6OS1 cause non-obstructive azoospermia and premature ovarian insufficiency in humans. Am J Hum Genet. 2022 Jul 7;109(7):1343. PMID: 35803236. doi: 10.1016/j.ajhg.2022.06.006.
164. Paul EN, Carpenter TJ, Fitch S, Sheridan R, Lau KH, Arora R, et al. Cysteine-rich intestinal protein 1 is a novel surface marker for human myometrial stem/progenitor cells. Commun Biol. 2023 Jul 3;6(1):686. PMID: 37400623. doi: 10.1038/s42003-023-05061-0.
165. Sun X, Zhang R, Lin X, Xu X. Wnt3a regulates the development of cardiac neural crest cells by modulating expression of cysteine-rich intestinal protein 2 in rhombomere 6. Circ Res. 2008 Apr 11;102(7):831–9. PMID: 18292601. doi: 10.1161/CIRCRESAHA.107.166488.
166. Sato Y, Shibata N, Hashimoto C, Agata K. Migratory regulation by MTA homologous genes is essential for the uniform distribution of planarian adult pluripotent stem cells. Dev Growth Differ. 2022 Apr;64(3):150–62. PMID: 35124813. doi: 10.1111/dgd.12773.
167. Zhang J, Wang Y, Zhang J, Wang X, Liu J, Huo M, et al. The feedback loop between MTA1 and MTA3/TRIM21 modulates stemness of breast cancer in response to estrogen. Cell Death Dis. 2024 Aug 17;15(8):597. PMID: 39154024. doi: 10.1038/s41419-024-06942-w.
168. Panda SK, Kim DH, Desai P, Rodrigues PF, Sudan R, Gilfillan S, et al. SLC7A8 is a key amino acids supplier for the metabolic programs that sustain homeostasis and activation of type 2 innate lymphoid cells. Proc Natl Acad Sci U S A. 2022 Nov 16;119(46):e2215528119. PMID: 36343258. doi: 10.1073/pnas.2215528119.
169. Nagpal I, Wei LN. All-trans Retinoic Acid as a Versatile Cytosolic Signal Modulator Mediated by CRABP1. Int J Mol Sci. 2019 Jul 24;20(15). PMID: 31344789. doi: 10.3390/ijms20153610.
170. Chen PC, Kuo YC, Chuong CM, Huang YH. Niche Modulation of IGF-1R Signaling: Its Role in Stem Cell Pluripotency, Cancer Reprogramming, and Therapeutic Applications. Front Cell Dev Biol. 2020;8:625943. PMID: 33511137. doi: 10.3389/fcell.2020.625943.
171. Ma L, Huang M, Liao X, Cai X, Wu Q. NR2F2 Regulates Cell Proliferation and Immunomodulation in Whartons' Jelly Stem Cells. Genes (Basel). 2022 Aug 16;13(8). PMID: 36011369. doi: 10.3390/genes13081458.
172. Zhang J, Li X, Hu J, Cao P, Yan Q, Zhang S, et al. Long noncoding RNAs involvement in Epstein-Barr virus infection and tumorigenesis. Virology journal. 2020 Apr 9;17(1):51. PMID: 32272952. doi: 10.1186/s12985-020-01308-y.
173. Constam DB. Regulation of TGFbeta and related signals by precursor processing. Semin Cell Dev Biol. 2014 Aug;32:85–97. PMID: 24508081. doi: 10.1016/j.semcdb.2014.01.008.
174. Ouyang X, Shi X, Huang N, Yang Y, Zhao W, Guo W, et al. WDR72 Enhances the Stemness of Lung Cancer Cells by Activating the AKT/HIF-1alpha Signaling Pathway. J Oncol. 2022;2022:5059588. PMID: 36385964. doi: 10.1155/2022/5059588.
175. LeBlanc ME, Wang W, Caberoy NB, Chen X, Guo F, Alvarado G, et al. Hepatoma-derived growth factor-related protein-3 is a novel angiogenic factor. PLoS One. 2015;10(5):e0127904. PMID: 25996149. doi: 10.1371/journal.pone.0127904.
176. Richter K, Wirta V, Dahl L, Bruce S, Lundeberg J, Carlsson L, et al. Global gene expression analyses of hematopoietic stem cell-like cell lines with inducible Lhx2 expression. BMC Genomics. 2006 Apr 6;7:75. PMID: 16600034. doi: 10.1186/1471-2164-7-75.
177. Qu L, Tang Y, Wu J, Yun X, Lo HH, Song L, et al. FBXL16: a new regulator of neuroinflammation and cognition in Alzheimer's disease through the ubiquitination-dependent

degradation of amyloid precursor protein. Biomark Res. 2024 Nov 21;12(1):144. PMID: 39568047. doi: 10.1186/s40364-024-00691-w.
178. Dou Z, Son JE, Hui CC. Irx3 and Irx5 - Novel Regulatory Factors of Postnatal Hypothalamic Neurogenesis. Front Neurosci. 2021;15:763856. PMID: 34795556. doi: 10.3389/fnins.2021.763856.
179. Fedorova MS, Snezhkina AV, Lipatova AV, Pavlov VS, Kobelyatskaya AA, Guvatova ZG, et al. NETO2 Is Deregulated in Breast, Prostate, and Colorectal Cancer and Participates in Cellular Signaling. Front Genet. 2020;11:594933. PMID: 33362854. doi: 10.3389/fgene.2020.594933.
180. Sahu SK, Fritz A, Tiwari N, Kovacs Z, Pouya A, Wullner V, et al. TOX3 regulates neural progenitor identity. Biochim Biophys Acta. 2016 Jul;1859(7):833–40. PMID: 27080130. doi: 10.1016/j.bbagrm.2016.04.005.
181. Quek LS, Grasset N, Jasmen JB, Robinson KS, Bellanger S. Dual Role of the Anaphase Promoting Complex/Cyclosome in Regulating Stemness and Differentiation in Human Primary Keratinocytes. J Invest Dermatol. 2018 Aug;138(8):1851–61. PMID: 29526760. doi: 10.1016/j.jid.2018.02.033.
182. Rausch P, Steck N, Suwandi A, Seidel JA, Kunzel S, Bhullar K, et al. Expression of the Blood-Group-Related Gene B4galnt2 Alters Susceptibility to Salmonella Infection. PLoS Pathog. 2015 Jul;11(7):e1005008. PMID: 26133982. doi: 10.1371/journal.ppat.1005008.
183. Zhao M, Nakada Y, Wei Y, Bian W, Chu Y, Borovjagin AV, et al. Cyclin D2 Overexpression Enhances the Efficacy of Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes for Myocardial Repair in a Swine Model of Myocardial Infarction. Circulation. 2021 Jul 20;144(3):210–28. PMID: 33951921. doi: 10.1161/CIRCULATIONAHA.120.049497.
184. Xiao Y, Yang Y, Xiong H, Dong G. The implications of FASN in immune cell biology and related diseases. Cell death & disease. 2024 Jan 25;15(1):88. PMID: 38272906. doi: 10.1038/s41419-024-06463-6.
185. Knobloch M, Braun SM, Zurkirchen L, von Schoultz C, Zamboni N, Arauzo-Bravo MJ, et al. Metabolic control of adult neural stem cell activity by Fasn-dependent lipogenesis. Nature. 2013 Jan 10;493(7431):226–30. PMID: 23201681. doi: 10.1038/nature11689.
186. Zhang YW, Mess J, Aizarani N, Mishra P, Johnson C, Romero-Mulero MC, et al. Hyaluronic acid-GPRC5C signalling promotes dormancy in haematopoietic stem cells. Nat Cell Biol. 2022 Jul;24(7):1038–48. PMID: 35725769. doi: 10.1038/s41556-022-00931-x.
187. Lawson H, Sepulveda C, van de Lagemaat LN, Durko J, Barile M, Tavosanis A, et al. JMJD6 promotes self-renewal and regenerative capacity of hematopoietic stem cells. Blood Adv. 2021 Feb 9;5(3):889–99. PMID: 33560400. doi: 10.1182/bloodadvances.2020002702.
188. Saha SK, Choi HY, Kim BW, Dayem AA, Yang GM, Kim KS, et al. KRT19 directly interacts with beta-catenin/RAC1 complex to regulate NUMB-dependent NOTCH signaling pathway and breast cancer properties. Oncogene. 2017 Jan 19;36(3):332–49. PMID: 27345400. doi: 10.1038/onc.2016.221.
189. Saha SK, Yin Y, Chae HS, Cho SG. Opposing Regulation of Cancer Properties via KRT19-Mediated Differential Modulation of Wnt/beta-Catenin/Notch Signaling in Breast and Colon Cancers. Cancers (Basel). 2019 Jan 15;11(1). PMID: 30650643. doi: 10.3390/cancers11010099.
190. Kyrousi C, O'Neill AC, Brazovskaja A, He Z, Kielkowski P, Coquand L, et al. Extracellular LGALS3BP regulates neural progenitor position and relates to human cortical complexity. Nat Commun. 2021 Nov 2;12(1):6298. PMID: 34728600. doi: 10.1038/s41467-021-26447-w.
191. Stampolidis P, Ullrich A, Iacobelli S. LGALS3BP, lectin galactoside-binding soluble 3 binding protein, promotes oncogenic cellular events impeded by antibody intervention. Oncogene. 2015 Jan 2;34(1):39–52. PMID: 24362527. doi: 10.1038/onc.2013.548.

192. Cai W, Chen G, Luo Q, Liu J, Guo X, Zhang T, et al. PMP22 Regulates Self-Renewal and Chemoresistance of Gastric Cancer Cells. Mol Cancer Ther. 2017 Jun;16(6):1187–98. PMID: 28336807. doi: 10.1158/1535-7163.MCT-16-0750.
193. Lopez-Anido C, Poitelon Y, Gopinath C, Moran JJ, Ma KH, Law WD, et al. Tead1 regulates the expression of Peripheral Myelin Protein 22 during Schwann cell development. Hum Mol Genet. 2016 Jul 15;25(14):3055–69. PMID: 27288457. doi: 10.1093/hmg/ddw158.
194. Panelo LC, Machado MS, Rubio MF, Jaworski F, Alvarado CV, Paz LA, et al. High RAC3 expression levels are required for induction and maintaining of cancer cell stemness. Oncotarget. 2018 Jan 19;9(5):5848–60. PMID: 29464039. doi: 10.18632/oncotarget.23635.
195. Aguilar-Medina M, Avendano-Felix M, Lizarraga-Verdugo E, Bermudez M, Romero-Quintana JG, Ramos-Payan R, et al. SOX9 Stem-Cell Factor: Clinical and Functional Relevance in Cancer. J Oncol. 2019;2019:6754040. PMID: 31057614. doi: 10.1155/2019/6754040.
196. Lu Y, Zhang Z, Wang S, Qi Y, Chen F, Xu Y, et al. Srebf1c preserves hematopoietic stem cell function and survival as a switch of mitochondrial metabolism. Stem Cell Reports. 2022 Mar 8;17(3):599–615. PMID: 35148846. doi: 10.1016/j.stemcr.2022.01.011.
197. Zhang Y, Gao Q, Wu Y, Peng Y, Zhuang J, Yang Y, et al. Hypermethylation and Downregulation of UTP6 Are Associated With Stemness Properties, Chemoradiotherapy Resistance, and Prognosis in Rectal Cancer: A Co-expression Network Analysis. Front Cell Dev Biol. 2021;9:607782. PMID: 34485268. doi: 10.3389/fcell.2021.607782.
198. Jin Y, Zhang M, Tong Y, Qiu L, Ye Y, Zhao B. DCXR promotes cell proliferation by promoting the activity of aerobic glycolysis in breast cancer. Mol Med Rep. 2023 Feb;27(2). PMID: 36562355. doi: 10.3892/mmr.2022.12918.
199. Choi HY, Saha SK, Kim K, Kim S, Yang GM, Kim B, et al. G protein-coupled receptors in stem cell maintenance and somatic reprogramming to pluripotent or cancer stem cells. BMB Rep. 2015 Feb;48(2):68–80. PMID: 25413305. doi: 10.5483/bmbrep.2015.48.2.250.
200. Cheng CJ, Nizar JM, Dai DF, Huang CL. Transport activity regulates mitochondrial bioenergetics and biogenesis in renal tubules. FASEB J. 2024 May 31;38(10):e23703. PMID: 38805156. doi: 10.1096/fj.202400358RR.
201. Murillo-de-Ozores AR, Rodriguez-Gama A, Bazua-Valenti S, Leyva-Rios K, Vazquez N, Pacheco-Alvarez D, et al. C-terminally truncated, kidney-specific variants of the WNK4 kinase lack several sites that regulate its activity. J Biol Chem. 2018 Aug 3;293(31):12209–21. PMID: 29921588. doi: 10.1074/jbc.RA118.003037.
202. Mao H, Brown HE, Silver DL. Mouse models of Casc3 reveal developmental functions distinct from other components of the exon junction complex. RNA. 2017 Jan;23(1):23–31. PMID: 27780844. doi: 10.1261/rna.058826.116.
203. Ribeiro V, Martins SG, Lopes AS, Thorsteinsdottir S, Zilhao R, Carlos AR. NFIXing Cancer: The Role of NFIX in Oxidative Stress Response and Cell Fate. International journal of molecular sciences. 2023 Feb 21;24(5). PMID: 36901722. doi: 10.3390/ijms24054293.
204. Hu R, Zhao J, Lai KC, Wang S, Zheng J, Stoddard C, et al. CHD7 regulates definitive endodermal and mesodermal development from human embryonic stem cells. Stem Cell Res Ther. 2025 Jun 17;16(1):311. PMID: 40528267. doi: 10.1186/s13287-025-04437-9.
205. Tang CT, Liang Q, Yang L, Lin XL, Wu S, Chen Y, et al. RAB31 Targeted by MiR-30c-2-3p Regulates the GLI1 Signaling Pathway, Affecting Gastric Cancer Cell Proliferation and Apoptosis. Front Oncol. 2018;8:554. PMID: 30534536. doi: 10.3389/fonc.2018.00554.
206. Lodhi IJ, Chiang SH, Chang L, Vollenweider D, Watson RT, Inoue M, et al. Gapex-5, a Rab31 guanine nucleotide exchange factor that regulates Glut4 trafficking in adipocytes. Cell Metab. 2007 Jan;5(1):59–72. PMID: 17189207. doi: 10.1016/j.cmet.2006.12.006.

207. Dong K, Geng C, Zhan X, Sun Z, Pu Q, Li P, et al. GREB1L overexpression is associated with good clinical outcomes in breast cancer. Eur J Med Res. 2023 Nov 14;28(1):510. PMID: 37964281. doi: 10.1186/s40001-023-01483-y.
208. Yu Y, Wang Z, Zheng Q, Li J. GREB1L overexpression correlates with prognosis and immune cell infiltration in lung adenocarcinoma. Sci Rep. 2021 Jun 24;11(1):13281. PMID: 34168239. doi: 10.1038/s41598-021-92695-x.
209. Chen X, Jiang Z, Wang Z, Jiang Z. The prognostic and immunological effects of ZBTB7C across cancers: friend or foe? Aging (Albany NY). 2021 May 4;13(9):12849–64. PMID: 33946045. doi: 10.18632/aging.202955.
210. Seiler C, Gebhart N, Zhang Y, Shinton SA, Li YS, Ross NL, et al. Mutagenesis Screen Identifies agtpbp1 and eps15L1 as Essential for T lymphocyte Development in Zebrafish. PLoS One. 2015;10(7):e0131908. PMID: 26161877. doi: 10.1371/journal.pone.0131908.
211. Walker MR, Goel HL, Mukhopadhyay D, Chhoy P, Karner ER, Clark JL, et al. O-linked alpha2,3 sialylation defines stem cell populations in breast cancer. Sci Adv. 2022 Jan 7;8(1):eabj9513. PMID: 34995107. doi: 10.1126/sciadv.abj9513.
212. Seo L, Kim YI, Kim H, Hyun K, Kim J, Lee JE. Discovery of Klf2 interactors in mouse embryonic stem cells by immunoprecipitation-mass spectrometry utilizing exogenously expressed bait. Biochim Biophys Acta Proteins Proteom. 2021 Sep;1869(9):140672. PMID: 34000451. doi: 10.1016/j.bbapap.2021.140672.
213. Harris L, Zalucki O, Gobius I, McDonald H, Osinki J, Harvey TJ, et al. Transcriptional regulation of intermediate progenitor cell generation during hippocampal development. Development. 2016 Dec 15;143(24):4620–30. PMID: 27965439. doi: 10.1242/dev.140681.
214. Woelk J, Narasimhan H, Pfeifhofer-Obermair C, Schraml BU, Hermann-Kleiter N. NR2F6 regulates stem cell hematopoiesis and myelopoiesis in mice. Front Immunol. 2024;15:1404805. PMID: 39840064. doi: 10.3389/fimmu.2024.1404805.
215. Besson V, Kyryachenko S, Janich P, Benitah SA, Marazzi G, Sassoon D. Expression Analysis of the Stem Cell Marker Pw1/Peg3 Reveals a CD34 Negative Progenitor Population in the Hair Follicle. Stem Cells. 2017 Apr;35(4):1015–27. PMID: 27862634. doi: 10.1002/stem.2540.
216. Arno G, Agrawal SA, Eblimit A, Bellingham J, Xu M, Wang F, et al. Mutations in REEP6 Cause Autosomal-Recessive Retinitis Pigmentosa. Am J Hum Genet. 2016 Dec 1;99(6):1305–15. PMID: 27889058. doi: 10.1016/j.ajhg.2016.10.008.
217. Ishiuchi T, Ohishi H, Sato T, Kamimura S, Yorino M, Abe S, et al. Zfp281 Shapes the Transcriptome of Trophoblast Stem Cells and Is Essential for Placental Development. Cell Rep. 2019 May 7;27(6):1742–54 e6. PMID: 31067460. doi: 10.1016/j.celrep.2019.04.028.
218. Addison JB, Koontz C, Fugett JH, Creighton CJ, Chen D, Farrugia MK, et al. KAP1 promotes proliferation and metastatic progression of breast cancer cells. Cancer Res. 2015 Jan 15;75(2):344–55. PMID: 25421577. doi: 10.1158/0008-5472.CAN-14-1561.
219. Zhu W, Feng D, Shi X, Li D, Wei Q, Yang L. A pan-cancer analysis of the oncogenic role of zinc finger protein 419 in human cancer. Front Oncol. 2022;12:1042118. PMID: 36578929. doi: 10.3389/fonc.2022.1042118.
220. Barenz F, Kschonsak YT, Meyer A, Jafarpour A, Lorenz H, Hoffmann I. Ccdc61 controls centrosomal localization of Cep170 and is required for spindle assembly and symmetry. Mol Biol Cell. 2018 Dec 15;29(26):3105–18. PMID: 30354798. doi: 10.1091/mbc.E18-02-0115.
221. Vicioso-Mantis M, Fueyo R, Navarro C, Cruz-Molina S, van Ijcken WFJ, Rebollo E, et al. JMJD3 intrinsically disordered region links the 3D-genome structure to TGFbeta-dependent transcription activation. Nat Commun. 2022 Jun 7;13(1):3263. PMID: 35672304. doi: 10.1038/s41467-022-30614-y.

222. Chou WC, Chen WT, Kuo CT, Chang YM, Lu YS, Li CW, et al. Genetic insights into carbohydrate sulfotransferase 8 and its impact on the immunotherapy efficacy of cancer. Cell Rep. 2024 Jan 23;43(1):113641. PMID: 38165805. doi: 10.1016/j.celrep.2023.113641.
223. Park JR, Kim RJ, Lee YK, Kim SR, Roh KJ, Oh SH, et al. Dysadherin can enhance tumorigenesis by conferring properties of stem-like cells to hepatocellular carcinoma cells. J Hepatol. 2011 Jan;54(1):122–31. PMID: 20952084. doi: 10.1016/j.jhep.2010.06.026.
224. Urtatiz O, Van Raamsdonk CD. Gnaq and Gna11 in the Endothelin Signaling Pathway and Melanoma. Front Genet. 2016;7:59. PMID: 27148356. doi: 10.3389/fgene.2016.00059.
225. Zhou W, Yan K, Xi Q. BMP signaling in cancer stemness and differentiation. Cell Regen. 2023 Dec 5;12(1):37. PMID: 38049682. doi: 10.1186/s13619-023-00181-8.
226. Lala-Tabbert N, AlSudais H, Marchildon F, Fu D, Wiper-Bergeron N. CCAAT/enhancer-binding protein beta promotes muscle stem cell quiescence through regulation of quiescence-associated genes. Stem Cells. 2021 Mar;39(3):345–57. PMID: 33326659. doi: 10.1002/stem.3319.
227. Maharjan BD, Ono R, Nosaka T. Eya2 is critical for the E2A-HLF-mediated immortalization of mouse hematopoietic stem/progenitor cells. Int J Oncol. 2019 Mar;54(3):981–90. PMID: 30628662. doi: 10.3892/ijo.2019.4673.
228. Quarmyne M, Doan PL, Himburg HA, Yan X, Nakamura M, Zhao L, et al. Protein tyrosine phosphatase-sigma regulates hematopoietic stem cell-repopulating capacity. J Clin Invest. 2015 Jan;125(1):177–82. PMID: 25415437. doi: 10.1172/JCI77866.
229. Nakaoka HJ, Hara T, Yoshino S, Kanamori A, Matsui Y, Shimamura T, et al. NECAB3 Promotes Activation of Hypoxia-inducible factor-1 during Normoxia and Enhances Tumourigenicity of Cancer Cells. Sci Rep. 2016 Mar 7;6:22784. PMID: 26948053. doi: 10.1038/srep22784.
230. Schulten HJ, Hussein D. Array expression meta-analysis of cancer stem cell genes identifies upregulation of PODXL especially in DCC low expression meningiomas. PLoS One. 2019;14(5):e0215452. PMID: 31083655. doi: 10.1371/journal.pone.0215452.
231. Fortugno P, Monetta R, Belli M, Botti E, Angelucci F, Palmerini MG, et al. RIPK4 regulates cell-cell adhesion in epidermal development and homeostasis. Hum Mol Genet. 2022 Aug 17;31(15):2535–47. PMID: 35220430. doi: 10.1093/hmg/ddac046.
232. Cao X, Lu Z, Fang Y, Huang F, Tang M, Wu B, et al. RIPK4 promotes epidermal differentiation through phase separation and activation of LATS1/2. Dev Cell. 2025 Jun 19. PMID: 40570855. doi: 10.1016/j.devcel.2025.05.017.
233. Horton AJ, Brooker J, Streitfeld WS, Flessa ME, Pillai B, Simpson R, et al. Nkx2-5 Second Heart Field Target Gene Ccdc117 Regulates DNA Metabolism and Proliferation. Sci Rep. 2019 Feb 11;9(1):1738. PMID: 30742009. doi: 10.1038/s41598-019-39078-5.
234. Wang Y, Zhou T, Wan J, Yang Y, Chen X, Wang J, et al. Comparative transcriptome analysis reveals a regulatory network of microRNA-29b during mouse early embryonic development. Oncotarget. 2016 Aug 16;7(33):53772–82. PMID: 27449102. doi: 10.18632/oncotarget.10741.
235. Funato N, Yanagisawa H. TBX1 targets the miR-200-ZEB2 axis to induce epithelial differentiation and inhibit stem cell properties. Sci Rep. 2022 Nov 23;12(1):20188. PMID: 36418889. doi: 10.1038/s41598-022-24604-9.
236. Simonin P, Guerrero GL, Bardin S, Gannavarapu RV, Krndija D, Boyd J, et al. The GTPase RAB6 is required for stem cell maintenance and cell migration in the gut epithelium. Development. 2024 Nov 1;151(21). PMID: 39431301. doi: 10.1242/dev.203038.

Supplementary Table S2 Functions of gene control sites methylated in non-malignant oral keratinocytes that are not methylated in the breast cancer cohort

Chromosome	Gene	Class	Encoded Function
2	*CYP1B1*	Microenvironment	Detoxification
2	*CYP27A1*	Other	Cholesterol metabolism
2	*EMX1*	Stem cell	Stem cell self renewal and proliferation
2	*GPR39*	Microenvironment	Zinc homeostasis
2	*MCEE*	Microenvironment	Metabolic enzyme essential for energy production in mitochondria
4	*CHRNA9*	Microenvironment	Ion channel receptor for acetyl choline expressed in stem cells
5	*EGFLAM*	Microenvironment	Extracellular matrix organization
5	*STC2*	Microenvironment	Calcium phosphate homeostasis
5	*ZNF354A*	Other	Transcription regulator cancer biomarker
6	*MB21D1*	Other	Antiviral defense, DNA damage sensor
7	*KCNH2*	Microenvironment	Voltage gated potassium channel
8	*HEY1*	Stem cell	Downstream target of Notch signaling pathway, regulates self renewal and differentiation
8	*NAT2*	Other	Drug metabolism and detoxification
8	*ZNF1*	Other	Transcription regulation
9	*GLDC*	Microenvironment	Glycine catabolism, redox balance
9	*SLC1A1*	Microenvironment	Glutamate transporter essential for neural stem cell self-renewal [1]
9	*TPRN*	Other	Inner ear function, cell structure and function
10	*GOT1*	Microenvironment	Glutamate level regulation, redox balance
10	*RNLS*	Microenvironment	Regulates cardiac function, blood pressure and immune microenvironment.
10	*UTF1*	Stem cell	Maintains pluripotency in embryonic stem cells [2]
11	*DAGLA*	Microenvironment	Affects hormone and immune microenvironment in the CNS
11	*MAP4K2*	Stem cell	Member of a pathway that affects cell fate of pluripotent stem cells

11	*MUC6*	Microenvironment	Forms gelatinous barrier to protect epithelial cells, alters cell signaling, modifies ECM and immunity
14	*GPX2*	Microenvironment	Redox balance and detoxification
15	*ANPEP*	Microenvironment	Endocytosis, amino acid and peptide transport. Used as a stem cell marker
16	*MT3*	Microenvironment	Detoxification, cell response to stress, zinc homeostasis. May be involved in deciding cell fate.
16	*POLR3K*	Other	Transcription, Subunit in RNA polymerase III.
17	*CLDN7*	Microenvironment	Mitochondrial tight junctions
17	*RAC3*	Stem cell	Required to maintain pluripotency of normal stem cells [3]
17	*RTN4RL1*	Microenvironment	Signaling cascades that impact actin cytoskeleton. May have stem cell like properties in nervous system but too little information available
19	*EPOR*	Stem cell	Stem cell differentiation especially in erythrocyte and neural progenitor cells [4].
19	*GAMT*	Other	Creatine synthesis, muscle function, energy metabolism
19	*ZNF576*	Other	Transcription regulation
22	*A4GALT*	Microenvironment	Production of a glycolipid called globotriaosylceramide Membrane composition and receptor availability for cell interactions.
22	*MAPK8IP2*	Stem cell	Affects cell fate decisions in pluripotent neural progenitor cells, probably affects JNK signal pathway [5]
22	*RASD2*	Stem cell	Stem cell differentiation pathways
22	*SOX10*	Stem cell	Stem cell regulator and marker, mammary stem cells [6]

References

1. Rieskamp JD, Rosado-Burgos I, Christofi JE, Ansar E, Einstein D, Walters AE, et al. Excitatory amino acid transporter 1 supports adult hippocampal neural stem cell self-

renewal. iScience. 2023 Jul 21;26(7):107068. PMID: 37534178. doi: 10.1016/j.isci.2023.107068.
2. Raina K, Dey C, Thool M, Sudhagar S, Thummer RP. An Insight into the Role of UTF1 in Development, Stem Cells, and Cancer. Stem Cell Rev Rep. 2021 Aug;17(4):1280–93. PMID: 33517544. doi: 10.1007/s12015-021-10127-9.
3. Panelo LC, Machado MS, Rubio MF, Jaworski F, Alvarado CV, Paz LA, et al. High RAC3 expression levels are required for induction and maintaining of cancer cell stemness. Oncotarget. 2018 Jan 19;9(5):5848–60. PMID: 29464039. doi: 10.18632/oncotarget.23635.
4. Niemann T, Greiner JFW, Kaltschmidt C, Kaltschmidt B. EPO regulates neuronal differentiation of adult human neural-crest derived stem cells in a sex-specific manner. BMC Neurosci. 2023 Mar 6;24(1):19. PMID: 36879191. doi: 10.1186/s12868-023-00789-1.
5. Roessler R, Goldmann J, Shivalila C, Jaenisch R. JIP2 haploinsufficiency contributes to neurodevelopmental abnormalities in human pluripotent stem cell-derived neural progenitors and cortical neurons. Life Sci Alliance. 2018 Aug;1(4):e201800094. PMID: 30456368. doi: 10.26508/lsa.201800094.
6. Dravis C, Spike BT, Harrell JC, Johns C, Trejo CL, Southard-Smith EM, et al. Sox10 Regulates Stem/Progenitor and Mesenchymal Cell States in Mammary Epithelial Cells. Cell reports. 2015 Sep 29;12(12):2035–48. PMID: 26365194. doi: 10.1016/j.celrep.2015.08.040.