



## ORIGINAL ARTICLE

# JAK-STAT core cancer pathway: An integrative cancer interactome analysis

Fettah Erdogan<sup>1,2</sup> | Tudor Bogdan Radu<sup>1,2</sup> | Anna Orlova<sup>3</sup> | Abdul Khawazak Qadree<sup>1,2</sup> | Elvin Dominic de Araujo<sup>1</sup>  | Johan Israelian<sup>1,2</sup> | Peter Valent<sup>4,5</sup> | Satu M. Mustjoki<sup>6,7,8</sup> | Marco Herling<sup>9</sup> | Richard Moriggl<sup>3</sup> | Patrick Thomas Gunning<sup>1,2</sup> 

<sup>1</sup>Department of Chemical and Physical Sciences, University of Toronto Mississauga, Mississauga, Ontario, Canada

<sup>2</sup>Department of Chemistry, University of Toronto, Toronto, Ontario, Canada

<sup>3</sup>Institute of Animal Breeding and Genetics, University of Veterinary Medicine, Vienna, Austria

<sup>4</sup>Division of Hematology and Hemostaseology, Department of Internal Medicine I, Medical University of Vienna, Vienna, Austria

<sup>5</sup>Ludwig Boltzmann Institute for Hematology and Oncology, Medical University of Vienna, Vienna, Austria

<sup>6</sup>Translational Immunology Research Program and Department of Clinical Chemistry and Hematology, University of Helsinki, Helsinki, Finland

<sup>7</sup>Hematology Research Unit, Helsinki University Hospital Comprehensive Cancer Center, Helsinki, Finland

<sup>8</sup>CAN Digital Precision Cancer Medicine Flagship, Helsinki, Finland

<sup>9</sup>Department of Hematology, Cellular Therapy, and Hemostaseology, University of Leipzig, Leipzig, Germany

## Correspondence

Patrick Thomas Gunning, Department of Chemical and Physical Sciences, University of Toronto Mississauga, 3359 Mississauga Rd N., Mississauga, Ontario L5L 1C6, Canada.

Email: patrick.gunning@utoronto.ca

Richard Moriggl, Institute of Animal Breeding and Genetics, University of Veterinary Medicine, A-1210 Vienna, Austria.

Email: richard.moriggl@vetmeduni.ac.at

## Funding information

NSERC, Grant/Award Number: RGPIN-2014-05767; CIHR, Grant/Award Number: MOP-130424 and MOP-137036; Canada Research Chair, Grant/Award Number: 950-232042; Canadian Cancer Society, Grant/Award Number: 703963; Canadian Breast Cancer Foundation, Grant/Award Number: 705456; CFI, Grant/Award Number: 33536; Ontario Research Fund, Grant/Award Number: 34876; EU consortia ERAPerMed 'JAKSTAT-TARGET'; ERANETPLL; European Research Council; Austrian Science Fund, Grant/Award Number: SFB-F04707 and SFB-F06105

## Abstract

Through a comprehensive review and in silico analysis of reported data on STAT-linked diseases, we analysed the communication pathways and interactome of the seven STATs in major cancer categories and proposed rational targeting approaches for therapeutic intervention to disrupt critical pathways and addictions to hyperactive JAK/STAT in neoplastic states. Although all STATs follow a similar molecular activation pathway, STAT1, STAT2, STAT4 and STAT6 exert specific biological profiles associated with a more restricted pattern of activation by cytokines. STAT3 and STAT5A as well as STAT5B have pleiotropic roles in the body and can act as critical oncogenes that promote many processes involved in cancer development. STAT1, STAT3 and STAT5 also possess tumour suppressive action in certain mutational and cancer type context. Here, we demonstrated member-specific STAT activity in major cancer types. Through systems biology approaches, we found surprising roles for EGFR family members, sex steroid hormone receptor ESR1 interplay with oncogenic STAT function and proposed new drug targeting approaches of oncogenic STAT pathway addiction.

## KEYWORDS

blood cancer, breast cancer, colorectal cancers, JAK/STAT pathway in cancers, liver cancers, lung cancer, protein-protein interactions, systems medicine

Fettah Erdogan and Tudor Bogdan Radu contributed equally.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Journal of Cellular and Molecular Medicine* published by Foundation for Cellular and Molecular Medicine and John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Transcription factors are the gatekeepers of cellular processes. They control the expression of genes encoding critical proteins of the entire proteome, from proteins involved in metabolism and cell communication to those that regulate the immune response and cell cycle. Signal transducer and activator of transcription (STAT) refers to a family of seven transcription factors that regulate the expression of genes controlling critical cellular processes in a mechanism distinct from secondary messengers.<sup>1,2</sup> Activation of STATs is typically initiated by ligand binding at cell surface receptors, followed by kinase-dependent phosphorylation of a conserved tyrosine (pY) residue.<sup>3</sup> This leads to Src 2 homology (SH2) domain mediated STAT-STAT parallel dimerization, rapid translocation into the nucleus, DNA binding and subsequent transcription of gene targets<sup>4-6</sup> (Figure 1).

All STAT members (Figure 1) display similar biochemical features controlling their subcellular localization and mode of action. Recent studies revealed that not only do they transduce signals to control transcription, they also regulate anabolism and catabolism at the mitochondria and are involved in nuclear compartmentalisation and genome integrity. Cytokine or growth factor signalling and their activation, however, is largely cell type dependent. Although cytokine or growth factor interactions with cellular components such as STATs are broadly deemed as protein-protein interactions, their effective mechanism is incompletely understood. Here, we investigate the STAT interactome in a gene product and cancer type-specific manner. Depending on the biological processes involved, the presence of cytokines, and underlying conditions and pathologies, they have a unique STAT interactome that illuminates the pleiotropic action of STAT family members. Numerous reports have shown that gain-of-function (GOF) mutations (i.e., somatic or acquired variations) in STAT proteins are a basis for oncogenesis in neoplastic cells.<sup>7-11</sup>

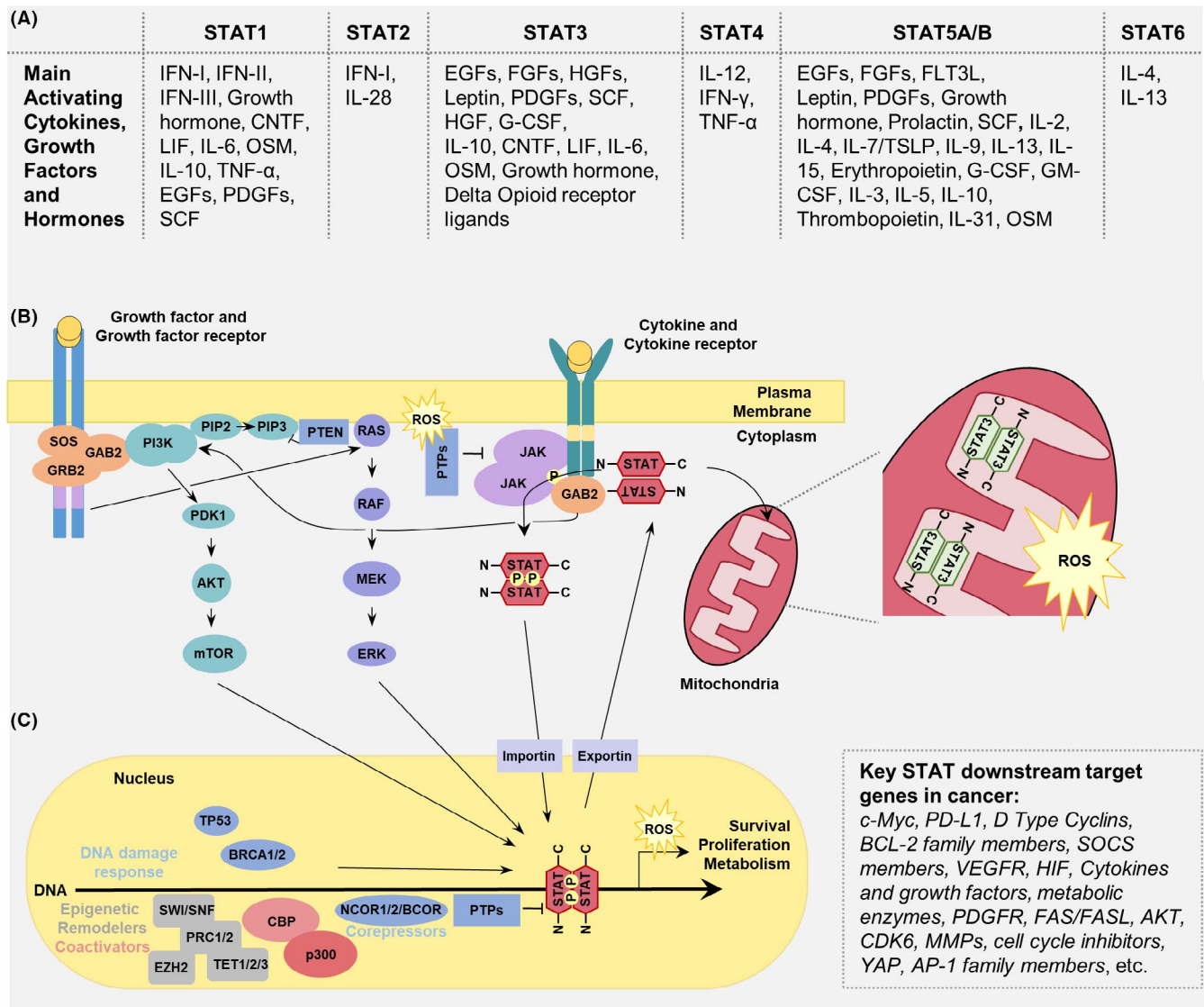
Our investigation focussed on the five major cancers ranked based on death rate, namely lung, breast, prostate, colorectal and liver cancers. Due to a high number of driver mutations in the STAT pathway, we further included blood cancers in our interactome analysis. We initially compiled all literature-based STAT interactions reported, analysed their member-specific role and generated a pathway-based figure depicting common STAT activation pathways. We then filtered STAT protein-protein interactions (PPIs) into those exclusively reported in association with a specific cancer type, either lung, breast, prostate, colorectal or liver cancer<sup>12</sup> and compared them with the global STAT interactome. We conducted a thorough analysis of the STAT interactome in blood cancers following these solid cancers. Overall, our analyses confirmed the central targets across multiple cancers such as the growth factor receptors (e.g., EGFR), while also revealing new viable targets such as the nuclear hormone receptor (e.g., ESR1) pathways in lung cancers, and supported the application of new single and combinatorial targeted approaches for the treatment of specific cancers.

## 2 | MATERIALS AND METHODS

The level of activity and protein interactions of each STAT member varies significantly. In the context of disease and specific cell type, some STAT members have a wider influence as demonstrated by a larger number of PPIs. In order to capture and visualize this influence and their differential activities, we summarized STAT interactions in a network map and segregated their PPIs into disease and tissue contexts. PPI data for most human proteins can be effectively used to identify pathogenic genes, drug targets and drug efficacy.<sup>13-18</sup> These data were collected using the integrated interactions database (IID) 2018 version (<http://iid.ophid.utoronto.ca/>)<sup>19</sup>. The IID gathers PPIs from nine curated databases and segregates these interactions into disease and tissue contexts. IID assigns context based on gene expression, if the mas5 normalized expression >200, with gene expression levels derived from 20 NCBI GEO gene expression datasets.<sup>19-22</sup> Given that PPIs are usually monitored in cells or in vitro, two proteins are considered to interact in a specific context if they have been shown to interact in two independent publications/databases/assays, and if both of those genes meet the above gene expression requirements in the specific context (e.g., breast cancer). The interactions presented here are curated to ensure minimum inclusion of false positive and negatives. To minimise the false positive rate, we only report interactions which are present in at least two publications or confirmed by at least two bioassays. To minimise the false negative rate, IID uses high confidence predictions to look for additional interactions.<sup>23,24</sup> To generate the interactome figures used here, we input the proteins we were interested in (STATs) into the Enter IDs field and retrieved all PPIs of the STATs in humans. We included experimental, predicted and protein orthology interactions, with thresholds for 2 or more studies, or two or more bioassays. We exported these data with source info, disease and subcellular localization context labelled. The data were then imported into NAViGaTOR 3, a network analysis software developed for protein-protein interactions (<http://navigator.ophid.utoronto.ca/navigatorwp/>)<sup>25</sup> which was used to generate the interactome maps.

Finally, it should be noted that while we can interpret these PPIs to achieve a deeper understanding of STAT communication, care should be taken when interpreting interactions that are not observed. For example, an interaction in the global interactome (Figure 2) can be absent in the context of a certain disease (Figure 3, Figure 4, and Figure 5) due to the rarity of the disease, lack of study analysis interest, lack of sufficient literature support or due to the complexity of given cancer subtypes as illustrated with our breast cancer insights.

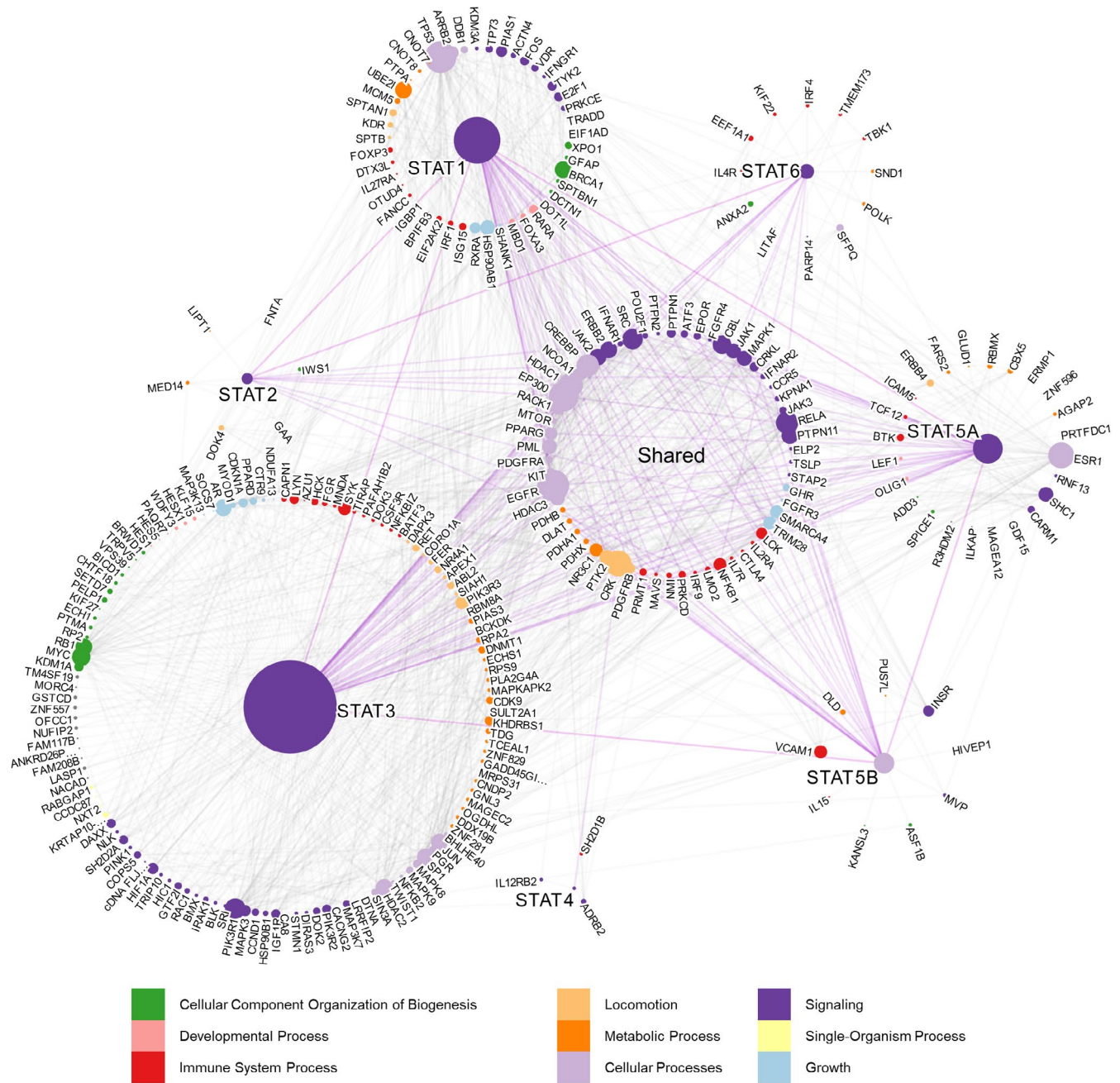
To analyse these PPI networks, we examine several factors including the 'degree' and 'centrality' of each protein with its interaction partners. 'Degree' correlates with how many interactions a protein makes, and 'degree' also reports on how influential a protein is in its immediate environment. The 'centrality'—measured through an undirected all pairs shortest paths (APSP) analysis—is indicative of how important a protein is for communication throughout the network. The higher the number of 'shortest paths



**FIGURE 1** Common STAT cellular activation-inactivation cycle pathways. (A) Variety of cytokines, hormones and growth factors can activate STAT family members. (B) Binding of the ligand to either the cytokine or growth factor (GF) receptor induces series of the activation cascades. Whereas JAK kinase activation is generally exclusive to cytokine receptor signaling, STAT action is triggered by both receptor types. Typically, auto-phosphorylated receptor-associated Janus kinases (JAKs) phospho-activate STATs which form STAT parallel dimer and are transported across nuclear membrane by Importin  $\alpha$ 's (KPNA1) and involve the action of Nucleophosmin I decamer (not shown). Binding of activated STATs to DNA is coupled to interactions with various transcriptional regulators discussed in the main text. Phosphatase-mediated inactivation of the DNA-bound STAT complex disengages the dimer from the DNA, breaks them into monomers and leads to their export via nuclear exportins (XPO1). Cytokine and growth factor receptors are both inducers of the RAS-RAF-MAPK and the PI3K-AKT-mTOR pathways. For simplicity, other signaling pathways such as activation of PKC or PLC are excluded. STAT3 is also capable of localizing into the mitochondria inner matrix and influencing ROS production. (C) STATs are also reported to be involved in the DNA damage response (e.g., TP53 and BRCA1/2) and interact with epigenetic modifiers (e.g., EZH2, TET1/2/3, SWI/SNF, PRC1/2, CBP-p300 and NCOA-1). CNTF, ciliary neurotrophic factor; EGF, Epidermal growth factor; FGF, fibroblast growth factor; FLT3L, FMS-like tyrosine kinase 3 ligand; G-CSF, granulocyte colony-stimulating factor; GM-CSF, granulocyte-macrophage colony-stimulating factor; HGF, hepatocyte growth factor; IFN, interferon; IL, interleukin; LIF, leukemia inhibitory factor; OSM, oncostatin M; PDGF, platelet-derived growth factor; SCF, stem cell factor; TGF, transforming growth factor; TNF, tumor necrosis factor; TSLP, thymic stromal lymphopoietin

pass' that pass through a protein, the more central it is, and thus, the more important it is for communication between proteins within the network. Together these parameters offer a strong indication of how important a node is to the network, and this can be extrapolated to determine how influential a protein is in signalling pathways, as well as identifying promising therapeutic

targets.<sup>13,14,16,17</sup> IID has been used recently as a part of CoVex to integrate protein-protein interactions into a system which helps identify already approved drugs to treat COVID-19.<sup>26,27</sup> A summary table of all interactions, interactor UniProt codes and the Pubmed ID(s) referencing each interaction can be found in Table S1.



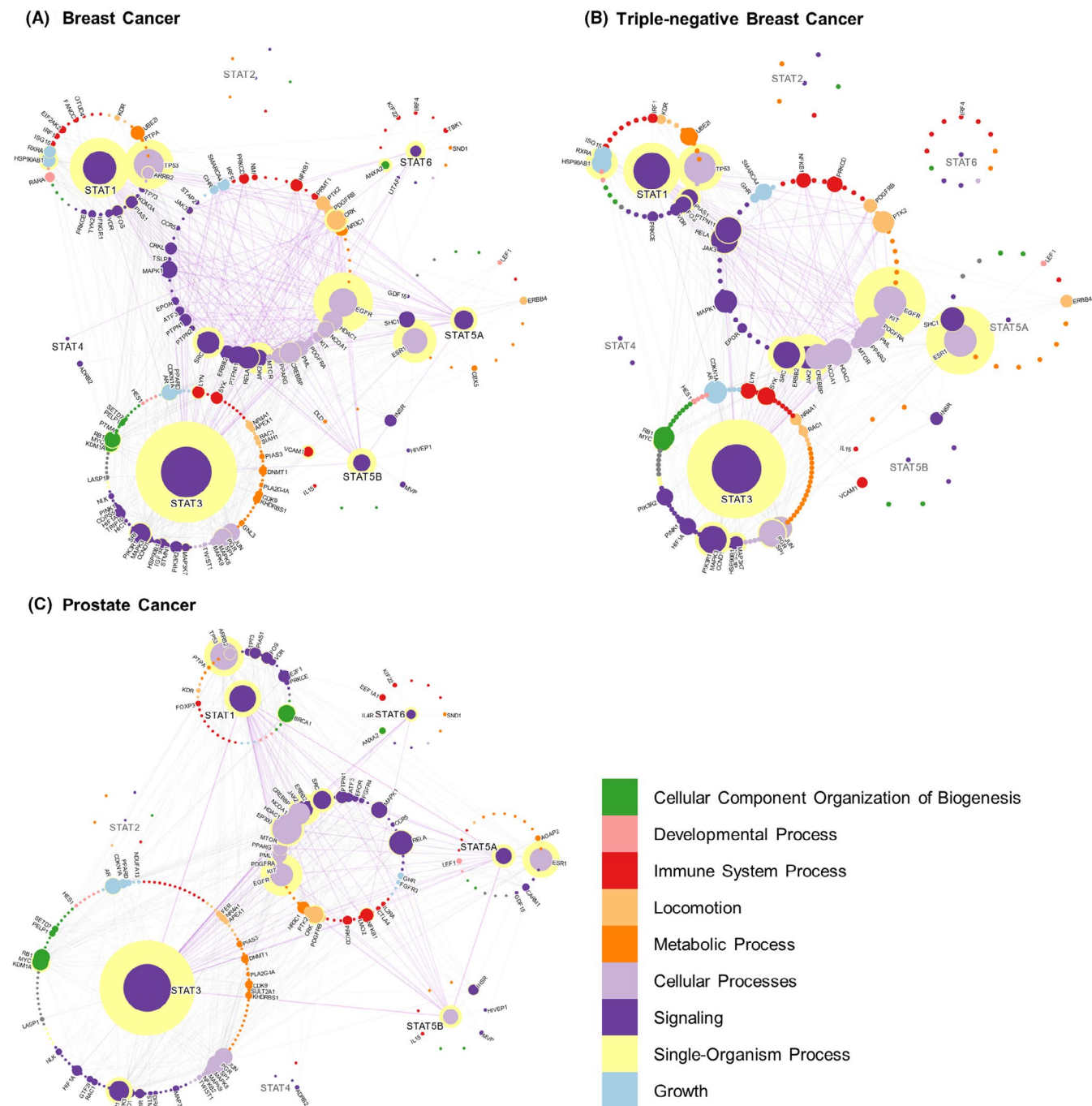
**FIGURE 2** STAT interactome. Global STAT interactome presenting all literature-based STAT PPIs. Each node (dot) is a protein, and each edge (line) is an interaction between two proteins. The size of the nodes is proportional to their degree, and the color of each node is representative of the biological process of the protein as described in the legend. The length of the edges connecting each node is arbitrary. Proteins which were found to interact with multiple STATs are grouped in the middle, and those exclusive to one STAT are grouped around that corresponding STAT member. A summary table of all interactions can be found in Table S1

### 3 | RESULTS AND DISCUSSION

#### 3.1 | Global STAT interactions

Although all STATs share a role in cell signalling and biogenesis, member-specific protein interactions reveal divergent activity of STAT members in various biological processes (Figure 2, Figure S1). Compilation of literature-based data on STAT PPIs shows that STAT3 interacts with the largest number/diversity of proteins (166

distinct proteins), followed closely by STAT1 (81 distinct proteins, Figure 2). The number of publications on a given STAT directly influences its appearance in the interactomes. For example, while STAT1 as a search term results in 10,272 publications, STAT3 results in ~28,419 hits in current PubMed database entry as of 14 July 2021. The search term STAT5 (for STAT5A/B) results only in 6616 hits. STAT3 possesses exclusive and extensive interaction evidence towards oncogenic partners such as c-Myc, Histone deacetylase 2 (HDAC2) and Mitogen-activated protein kinase 3 (MAPK3). Other

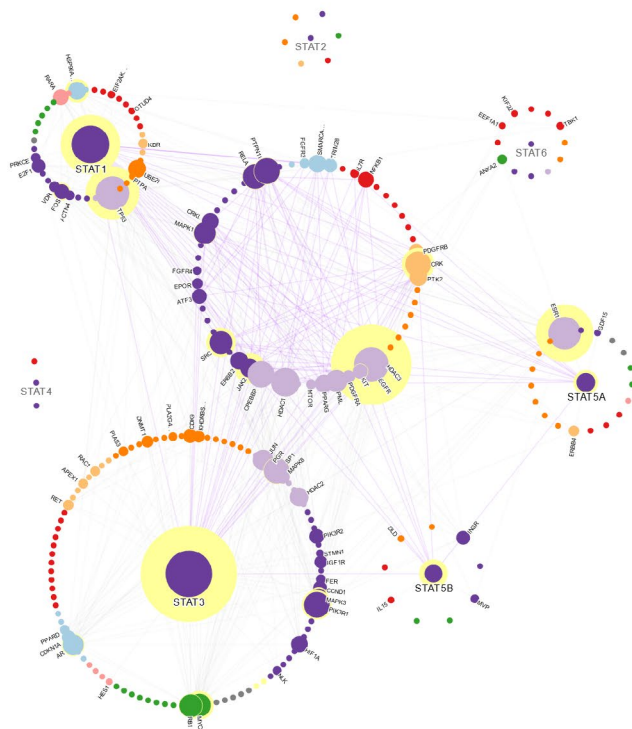


**FIGURE 3** STAT interactions in breast cancer, TNBC and prostate cancer. STAT interactions in two of the five most common death-causing cancers listed by the WHO; breast cancer and prostate cancer, as well as the breast cancer subtype TNBC. Any interaction shared between two or more STATs is grouped in the middle, and the edges connecting those interactions are highlighted. Proteins for which a label is shown are those which are reported to be over-expressed in that disease. Although not all proteins directly interact with each other, they may communicate indirectly through hub proteins. The proteins important for indirect communication are identified as 'central' to the networks and have a yellow halo in the network maps. Centrality measures how important the protein is for communication between the STAT interactome. It is measured using an undirected all-pairs shortest path algorithm which measures how many of the shortest paths between proteins pass through a node. Therefore, proteins which are more central act as hubs through which signals are efficiently transmitted from one end of the interactome to the other. The centrality of a protein is proportional to the size of the yellow halo

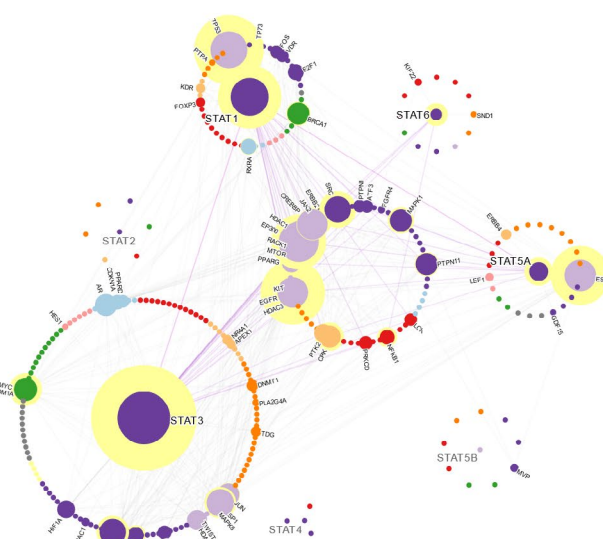
key STAT3-exclusive binding partners include the Androgen receptor (AR), Spleen tyrosine kinase (SYK) and the P110 $\alpha$ / $\beta$ / $\gamma$ / $\delta$ /P85 regulatory subunits of Phosphatidylinositol 3-kinase (PI3K). The constant P85 subunit of PI3K is bound by GAB1/2 scaffold proteins

that also bridge STAT5A/5B interaction in the cytoplasm, and STAT5 can transcribe an AKT isoform. These interactions, however, were filtered out by our interactome stringency rules. Mutations that lead to increased kinase activity of *PIK3CA* were found in up to 13% of

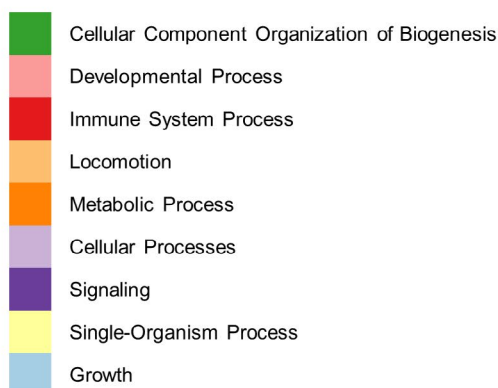
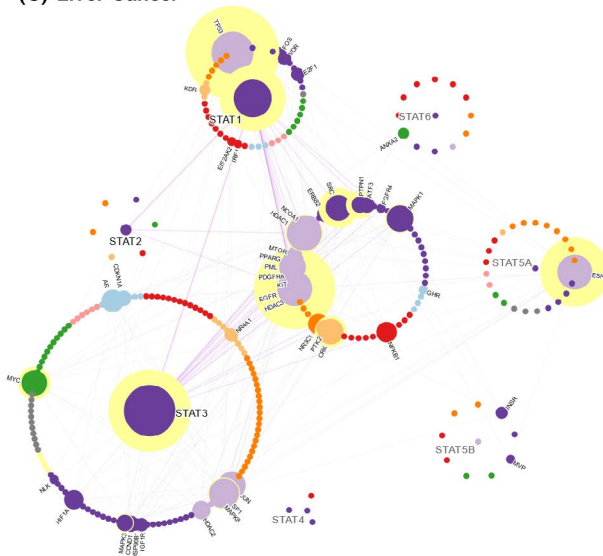
## (A) Lung Cancer



## (B) Colorectal Cancer



## (C) Liver Cancer



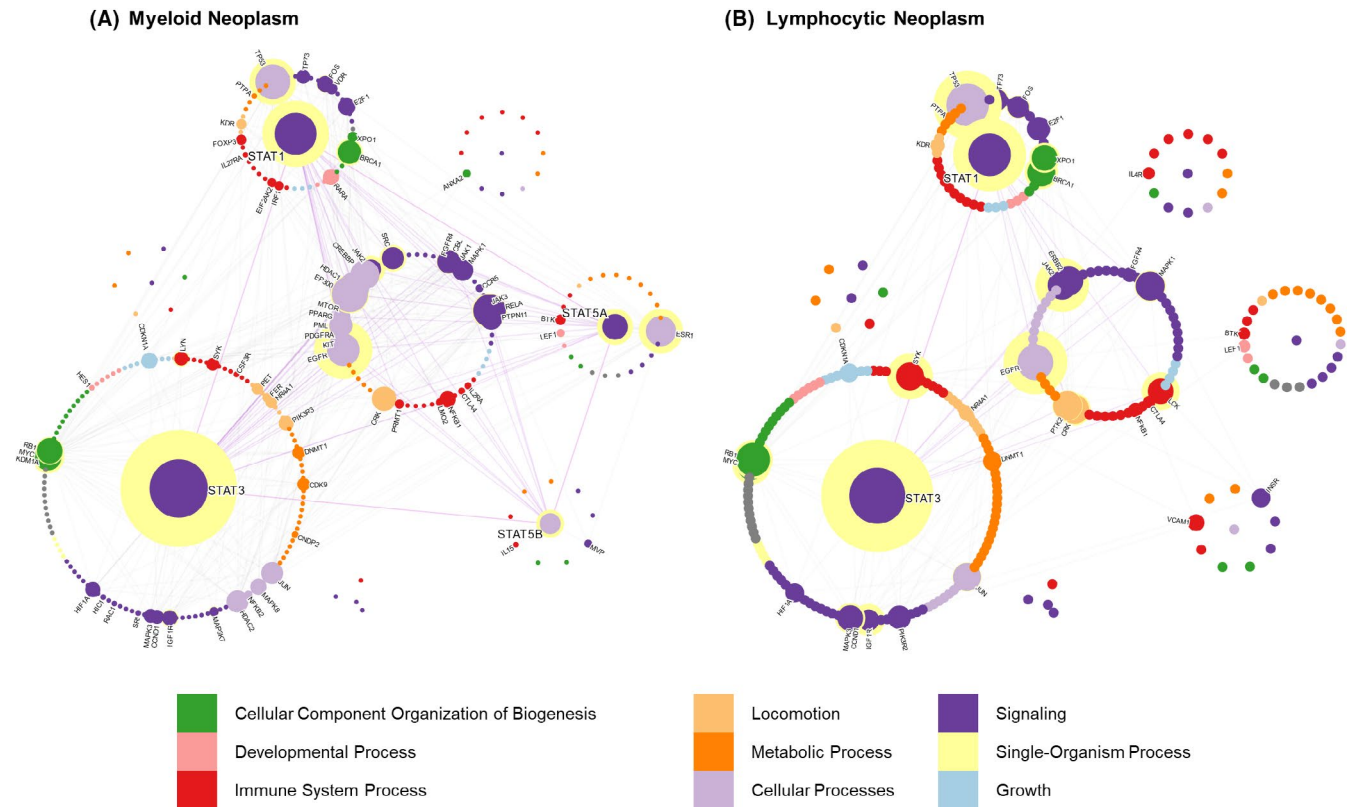
**FIGURE 4** STAT interactions in lung, colorectal and liver cancers. STAT interactions in the three of the five commonly death-causing cancers listed by the WHO lung cancer, colorectal cancer and liver cancer. For a descriptor of the interactome labels, see Figure 3

solid cancers including breast, colorectal, lung carcinomas, head and neck cancers among many others. As such, the involvement of STAT1/3/5 family members in the PI3K-AKT-mTOR-S6 pathway is key in these cancers.<sup>28</sup> All of these proteins also interact with other non-STAT proteins within the STAT interactome suggesting that while they are directly interacting with STAT3, they are indirectly communicating with other STATs. The fact that some of these interaction partners such as MAPK3 and c-Myc are important oncogenes indicates that STAT3 could be the vulnerable connector between these oncogenes, which also validates STAT3 as a promising cancer cell target.

Although STAT3 has the largest impact in the interactome, every other STAT also has exclusive interacting partners. STAT1 is largely an antagonist of oncogenic STAT3, also reflected through

its function as a heterodimerization partner of STAT3. The balancing role of STAT1 as a tumour suppressor protein over oncogenic STAT3 is well studied in various cancer types, including colorectal cancer.<sup>29</sup> STAT1 is the most prominent STAT protein interacting with the tumour suppressor TP53 by network analysis, although STAT3 and STAT5A/B have also been reported to interact with TP53 as well.<sup>30,31</sup>

Exclusive interaction of STAT5A and STAT6 with Bruton's tyrosine kinase (BTK) and Interleukin-4 receptor (IL-4R) proteins, respectively, highlights another point of member-specific STAT activity.<sup>32</sup> The IL-4R $\alpha$  chain is also a direct target gene of STAT5,<sup>33</sup> and both IL-4 and IL-13 can activate STAT5 and STAT6.<sup>34</sup> Surprisingly, STAT5A is the only STAT that has been shown to interact with the Estrogen receptor (ESR1) (Figure 2), and this is thought to occur via



**FIGURE 5** STAT interactions in hematopoietic cancers. STAT interactions in the categories of myeloid and lymphoid neoplasms. For a descriptor of the interactome labels, see Figure 4

the N-terminal domain of STAT5 in a fashion similar to many other nuclear hormone receptors that drive sex-specific cancers.<sup>35,36</sup> ESR1 has a large number of interactions in the global STAT interactome (Figure 2), and it has an important role in female reproductive tract cancers.<sup>37</sup> These representative key interactions predominantly differentiate the roles of STAT members and increase the diversity of STAT influence on biological processes.

Based on the shared proteins in the core of the interactome (58 proteins), the largest biological processes under which STATs converge are signal transduction or reprogramming of the nucleus, largely through the kinase interactome. Within the shared interactions, most of these proteins are tyrosine kinases such as JAK1/2/3, SRC, MAPK1, EGFR (human Epidermal growth factor receptor; HER1) (Figure S2), and phosphatases such as PTPN11, which is also part of the MAPK pathway, and is involved in signal transduction from the cytoplasm to the nucleus. These signalling proteins are important, as phosphorylation of STATs is a major contributor to their activation, subsequent dimerization, nuclear shuttling or transcriptional elongation. The second largest family of proteins in the core of the interactome are controllers of cellular processes, including coactivators or corepressors of gene expression, cell cycle regulators such as cell cycle inhibitors, cell cycle kinases or cyclins and physiological responses to circadian rhythm that is also controlled largely by GR stress hormone action. Of all the shared interacting partners mentioned above, EGFR and JAK2 are the most central between STAT proteins (Figure S1) and both modulate the STAT signalling cascade

through recruitment and phosphorylation of STATs. This is consistent with their reliance on phosphorylation for signalling. Another aspect of cellular processes in which STATs are involved appears to be histone acetylation-deacetylation. Histone acetylation is an epigenetic marker, which determines DNA transcription kinetics. CBP (CREB-binding protein), p300 and NCoA-1 (Nuclear receptor coactivator 1) display histone acetyl transferase activity while HDAC1 and HDAC3 deacetylate histones as well as some non-histone substrates. Contrastingly, PML (promyelocytic leukaemia) protein is a positive regulator of histone deacetylation and has been shown to interact with STATs in AML. This evidence of multiple positive and negative effectors of histone acetylation suggests that STATs have a large and unobserved role to play in epigenetic regulation.

The shared interactions between STATs further diversify based on their sequence and structure similarities. For example, STAT5A paired with STAT5B, and STAT1 paired with STAT3 have more PPIs in common than any other pair of STATs (Figure S1) and this is supported by their similarities in both the sequence alignment and structural alignment data.<sup>38</sup> Some of the shared key interaction partners of STAT3 and STAT1 include the mammalian Target of rapamycin (mTOR), HDAC3 and SRC kinase which phosphorylates both STAT1 and STAT3.<sup>39,40</sup> Importantly, STAT5A and STAT5B are direct phosphorylation targets of mTOR at a conserved S193 residue. Interestingly, STAT1 and STAT3 have more PPIs in common than the more structural homologous STAT5A and STAT5B though, this may again be biased by more literature content available for STAT1/3

than STAT5A/B. It should be noted that partial correlation of STATs having common PPIs does not mean that the two proteins are redundant in function. For example, despite over 90% structural identity, STAT5A has 24 exclusive interactions absent in STAT5B and, STAT5B has 9 exclusive interacting partners with insufficient evidence for STAT5A binding (Figure S1). This indicates that minor changes in a STAT protein could modulate its binding interface and interaction partners.

### 3.2 | Disease and tissue-specific STAT interactions

The plasticity of the five major cancer killers discussed here is mainly due to the biochemical and somatic heterogeneity that manifests in numerous cancer stem cell-derived subclones during cancer evolution and the tendency of these neoplasms to transdifferentiate and to escape the immune system in various subclones or metastases.<sup>41</sup> The tumour microenvironment and the status of the immune system play a decisive role in the success rates in cancer therapy.<sup>42</sup> Single cell sequencing and cancer pathway studies in primary cancer lesions, although conducted with the intention to define the core cancer pathway and simplify cancer for rational targeting efforts, have revealed a highly convoluted picture that awaits further simplification to fully comprehend cancer origins and progression.<sup>43,44</sup> New drivers, oncogenes, tumour suppressors and transcriptional regulatory elements are uncovered with the aim of adding to our understanding of cancer that has best to be illustrated in a simplified manner. Herein, our concept of disease filtered PPI allowed us to detect cancer driver pathways to highlight specific proteins as central hubs in these pathways. Simulating a systems approach concept, the presentation of cancer pathways could better guide research questions to focus on developing pathway blockers on highly convergent targets in the proteome. Systems approaches have emerged as a new way of thinking in cancer-targeting efforts, mainly in response to the accumulation of setbacks in the traditional approach of single versus combinatorial target therapy. This new approach is already being employed in blood cancer biology in association with chromatin remodelling by oncogenic STAT3/5 action.<sup>45</sup>

#### 3.2.1 | Lung cancer

Lung cancer is divided into two main histopathological types: small-cell lung cancer and non-small-cell lung cancer. Small-cell lung cancer arises from neuroendocrine cells and accounts for up to 15% of lung cancers. Most lung cancers are non-small-cell lung cancers (NSCLC), which are further divided into: adenocarcinoma, squamous cell carcinoma and large-cell carcinoma. Adenocarcinoma is stratified by different driver mutations such as activating mutations in *K-RAS* and *EGFR* and arises from alveoli-lining epithelial cells of the airways. The STAT lung cancer interactome (Figure 4A) focuses predominantly on adenocarcinomas, which accounts for up to 90% of lung cancers.<sup>46</sup> In the lung cancer interactome, we observe a larger number of central

proteins aside from *EGFR* and *STAT3*. All these key proteins are actively involved in lung cancer. *STAT3* and *EGFR* share a similar level of centrality in lung cancers, much more so than in the global interactome (Figure S1). This suggests a role for *EGFR* that is more important in the disease phenotype than in healthy lung tissue and is critical to lung cancer pathology. Furthermore, NSCLC pathway (Figure 4A) appears to be centralized around *STAT3* and its direct interaction partners, alongside *EGFR*. The majority of lung cancer studies described *STAT3* as an oncogene particularly with hyperactive *EGFR* signalling of autocrine IL-6 stimulation; however, in the context of *KRAS* mutations, *STAT3* possesses tumour suppressive activity. Such complications make targeting of *STAT3* in NSCLC a difficult decision.<sup>47</sup> While there are many proteins which act as partial communication hubs due to their centrality such as *STAT1*, *ESR1* or *TP53*, the concentration of the lung cancer pathway around *STAT3* and *EGFR* suggests that their inhibition would have the largest detrimental effect on the disease by preventing activation of drug resistance pathways upon combinatorial treatments. Targeting of oncogenic *RAS* or tumour suppressive/oncogenic *TP53* mutations is challenging and will require innovative therapy concepts; however, as suggested by the interactome, the nuclear hormone receptor (e.g. *ESR1*) pathways in lung cancer also appear to be central (represented by the yellow hollows, Figure 4A) and can serve as a new viable target in these cancers. Future studies could evaluate anti-estrogen drugs such as tamoxifen as part of combinatorial treatment approaches.

#### 3.2.2 | Breast cancer

There are many subtypes of breast cancers initiating in different regions of the breast, such as the ductal areas, lobules or in more rare cases, within the connecting cell types. Triple-negative breast cancer (TNBC) is the most aggressive and difficult to treat subtype, accounting for 15%–20% of breast cancers and is *de novo* resistant to estrogen therapy. Breast cancer biomarkers in TNBC include the Estrogen receptor (ER), Progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) with all three being absent in TNBC. The World Health Organization (WHO) distinguished at least 18 different histological breast cancer types.<sup>48</sup> Luminal breast cancers are the largest group, and they display the estrogen receptor (ER) as a biomarker of good prognosis if highly expressed. ER-positive carcinomas account for ~70% of all cases of breast cancers in Western societies.<sup>49,50</sup> Whereas 50% of ER positive breast cancers displays PR expression as a second biomarker, HER2 is the third biomarker in breast cancer accounting for 15%–25% of cases. HER2 amplification is associated with metastasis, and its persistent activation promotes *STAT3/5* oncoprotein activation; however, their role as being an oncogene or tumour suppressor protein in breast cancer is controversial due to complexity of the 18 different breast cancer subtypes or driver mutation context. Thus, biomarkers are highly relevant for the different breast cancer subtypes. Due to their heterogeneous driver mutation landscapes, treatment protocols involving personalized precision medicine have been described to be useful



and effective.<sup>51</sup> Because of tumour heterogeneity, different breast cancer subtypes are not accounted for, and subtype overlaps cannot be avoided in our network analysis of breast cancer. Nonetheless, analysis of simplified (or more homogenous) breast cancers (mainly luminal as opposed to the most aggressive TNBC) revealed that member-exclusive STAT activity becomes more apparent when STAT PPIs are filtered in relation to STAT-linked cancers.

In breast cancer, STAT3 has a large number of interactions and many of these interactions involve moderately central proteins in the STAT-breast cancer interactome (Figure 3A). STAT3 is the most central protein, followed closely by EGFR, revealing a surprising similarity to lung cancer. This may in part explain why breast cancer might favour lung cancer metastasis, an insight that could obtain attention in the different breast cancer subtypes with or without lung metastasis. Out of all STATs, STAT3 and STAT1 possess the largest number of interactions and a large centrality governing communication between other proteins, being the most important STATs in the breast cancer network. Although the non-STAT proteins such as TP53, ESR1, SRC and JAK2 are often linked to breast cancer, their dominance in the network is overshadowed by exclusion of non-STAT interacting proteins. A general comparison of PPIs in diseased and healthy mammalian tissue shows that these proteins become more central to communication in the disease context highlighting upregulated protein–protein communication in breast cancer.<sup>52,53</sup> Although inhibition of STAT3 is considered to severely hinder the viability of breast cancer cells,<sup>54</sup> this high degree of centrality and signalling redundancy between multiple central oncogenes again suggests that combination therapies targeting STAT3, and other key proteins including TP53, ESR1, EGFR or STAT1 may present with a higher degree of success in the clinic.<sup>55,56</sup>

Signal transducer and activator of transcription interactome analysis of TNBC was performed by investigating tumour-relevant pathways in this most aggressive highly metastasising subtype and by isolating the specific proteins. Compared with the overall breast cancer interactome (mainly comprising of the luminal subtype), STAT influence on the TNBC interactome is predominantly shifted towards STAT1 and STAT3 (Figure 3B). However, all of the central non-STAT proteins (TP53, EGFR and ESR1) in the breast cancer network are still interacting and they are central to the TNBC interactome. SRC and JAK2 kinases also display a more distinct influence in TNBC. The centrality of ESR1 in TNBC is again interesting as its epigenetic silencing has been observed in TNBC patients from India.<sup>57</sup> These findings suggest that while several STATs collectively govern breast cancer, STAT1 and STAT3 become more influential proteins to TNBC progression.

### 3.2.3 | Prostate cancer

Prostate cancers are androgen-driven tumours, and they are classified as luminal adenocarcinomas in the majority of cases.<sup>58</sup> Luminal adenocarcinomas are castration-sensitive (androgen-responsive) upon anti-androgen treatment. The other types of prostate cancers

are neuroendocrine prostate cancer or small-cell carcinoma, which account for less than 5%, and these are not covered by our analysis. In prostate luminal adenocarcinomas, we observe that STAT1/3/5A/5B are all interacting with a substantial number of proteins with moderate centrality (Figure 3C). STAT6 is also present although it has significantly fewer interactions with low centrality suggesting it is not a critical protein. STAT3 interacts with the largest number of proteins in a diverse array of gene ontologies and many of the shared interaction partners in the centre of the network. STAT3 is known to be constitutively activated in prostate tumours and related cell lines.<sup>59</sup> Again, similar to our observations in lung cancer or breast cancer, a key STAT3 interaction partner is an ERBB family member, namely EGFR, the second most central protein in the network (Figure 3C). The STAT interactome in prostate cancer supports that EGFR is a key signal transducer. Indeed, EGFR is already known for promoting the motility and growth of prostate cancer cells.<sup>60,61</sup> The interactions passing through the network are centralized on STAT3 and EGFR, with STAT1/5A/B, TP53 and ESR1 playing a supporting role.<sup>62,63</sup> Interestingly, both STAT3 and STAT5A interact with STAT1 and STAT5B but not with each other. EGFR, however, interacts with all STATs except STAT6. This appears to be the cause of the substantial centrality of EGFR as it interacts with all the major STATs present in prostate cancer. This also illuminates a divergence between STAT3 and STAT5A, and this is bridged by other STATs and EGFR interaction.

### 3.2.4 | Colorectal cancer

In contrast to prostate cancer, however, the non-STAT proteins appear to be more important in the communication between STATs in colorectal cancer (Figure 4B). Although STAT5A (15 interactions) and STAT6 (8 interactions) are interconnected, STAT1 (32 interactions) and STAT3 (45 interactions) have more interactions and show higher centrality (Figure 4B). Collectively, these findings suggest that among the STATs, STAT1 and STAT3 are by far the most important in colorectal cancer progression while other STATs such as STAT5A, STAT5B and STAT6 are vastly understudied.<sup>64,65</sup> EGFR, TP53 and ESR1 also have high centrality and degree in colorectal cancers. TP53 and ESR1 are both more central to the interactome than STAT1 and STAT5A, their respective exclusive interaction partners, both of which have been shown to be linked to colorectal cancer.<sup>66</sup> This situation arose because a larger number of TP53 and ESR1 interactions have been reported in association with colorectal cancer, while many of the physiological interactions of STAT1 and STAT5A in healthy cells were not identified in colorectal cancer. Notably, STAT1 is well known to interact with TP53 and both have been shown to collaborate in mediating the apoptosis response in colorectal cancer.<sup>67</sup>

### 3.2.5 | Liver cancer

Liver cancer shows one of the most surprising interactomes where the non-STAT proteins have a larger role than even STAT3

(Figure 4C). TP53 has by far the largest centrality of any protein shown here followed closely by EGFR. Connecting these two central liver cancer proteins are STAT1 and ESR1. These four proteins have the largest influence in the STAT interactome in liver cancer with STAT3 following closely behind. STAT1, EGFR, ESR1 and TP53 have all been shown to be important for the progression of liver cancer, along with the STAT3 signalling pathway.<sup>68–74</sup> Notably, two clinical trials, a phase III trial on tamoxifen (an ESR1 antagonist) and a phase II trial of tamoxifen in combination with cisplatin and doxorubicin hydrochloride, have already been performed for the treatment of liver cancers.<sup>75,76</sup> The postulated central role of the estrogen receptor should be further investigated to address whether both classes of estrogen receptors (ER $\alpha$  and ER $\beta$ ) are involved and relevant in hepatocellular carcinoma progression.

There is a dominant role for the IL-6-JAK-STAT component reported in many liver cancer studies, which clearly implicates the JAK1/JAK2-STAT3 axis as oncoproteins,<sup>77</sup> but expression of JAK kinases in general is much lower than the expression of STAT proteins, causing a bias in systems biology analysis. However, JAK kinases are frequently mutated, and if a gain or loss of function mutation would be found, then this analysis would strongly influence the systems biology insight. This can be best illustrated for the most abundant and frequent JAK mutation, which is found in JAK2V617F, driving myeloproliferative neoplasms (MPN). However, we did not include MPN in our blood cancer analysis since we wanted only to include acute leukaemia or lymphoma insights. The dominant role of JAK2 in MPN is described elsewhere, and we have not included these in our analysis of blood cancers that are discussed next due to the large diversity of hematopoietic cancers.<sup>78</sup>

### 3.2.6 | Blood cancers

Hematopoietic cancers are heterogeneous diseases of mesenchymal origin. The central role of perturbed JAK/STAT signalling in the initiation and progression of blood cancers is not surprising given its essential node in normal haematopoiesis—the development and function of blood cells of myeloid and lymphoid lineages. There, it is involved in the transmission of complex signalling pathways that are mediated by numerous cytokines and their receptors.<sup>79</sup>

Myeloid STAT signalling is distinct from lymphoid STAT networks. In the myeloid branch, binding of cytokines to their receptors initiates JAK activation, which in turn ignites particular STAT proteins, MAPK and PI3K-AKT-mTOR pathway in a receptor-specific fashion. In our analysis, EGFR and TP53 are the only central non-STATs in the two networks of hematopoietic cancers (Figure 5A, B) while ESR1 is central in only myeloid leukaemia. EGFR displays a comparable level of centrality to STATs, and as such, combination therapies targeting EGFR have been proven effective in AML cells. The centrality of TP53, a well-known interactor of STAT1/3/5 proteins, is the most important tumour suppressor transcription factor in human cancer with half the cases carrying a genetic mutation or loss. Quite interestingly, TP53 represses STAT5 transcription, but

mutated TP53 is shown to boost STAT5 transcription in AML cells.<sup>31</sup> On the contrary, lymphoid JAK-STAT signalling, cytokines that bind to common- $\gamma$ -chain containing receptors, such as IL-2/4/7/9/15/21, are central in physiological T-, NK- and B-cell development and function.<sup>80,81</sup> Hyperactivation of JAK1/JAK3 results in high, oncogenic STAT3 and STAT5 activity.<sup>82,83</sup>

Overall, STAT1/3/5A/5B are distinctly activated and/or overexpressed in hematologic neoplasms. STAT1 acts as a leukaemogenic oncogene in B-lymphoid leukaemia, and STAT3 mutations in the SH2 domain (affecting dimerization and activation) have been detected in numerous leukaemia samples in patients with LGL.<sup>84–86</sup> Whereas mutations in TP53 are associated with poor survival in AML patients, ESR1 methylation is reported to correlate with hypermethylation of several oncogenes linked to AML<sup>87,88</sup> and more work is needed to understand centrality of EGFR, TP53 and ESR1 in blood cancers. In conclusion, STAT1 and STAT3 are the only STATs interacting with the majority of proteins in the hematopoietic cancer interactome network, showing a high degree of centrality (Figure 5, Figure S4). Comparably, the other STATs (STAT2, STAT4 and STAT6) display a smaller number of interactions. This is unsurprising due to the specific expression pattern of STAT4 and STAT6, or a more restricted role of STAT2 in viral/bacterial defence.

### 3.3 | Subcellular localization of STAT proteins

Due to their common function of transferring signals from the cytoplasm to the nucleus, all STAT members appear in both the cytoplasm and the nucleus and they interact with many of the same proteins (Figure S3). Since the cytoplasm and nucleus is where the majority of the STAT pathway-associated proteins reside (see Figure 1), interactomes filtering these two subcellular compartments are very similar compared with the full STAT interactome. In the cytoplasmic STAT interactome, given their high centrality, STAT3, EGFR and STAT1 are responsible for the bulk of communication between proteins (Figure S3A). In the nucleus, however, TP53, STAT5A and ESR1 appear to be central communication hubs (Figure S3B). STAT3 and STAT5 play a regulatory role in the nucleus.<sup>89</sup> Taken together, these findings and associations suggest that the transmission of information is more delocalized in the nucleus than in the cytoplasm. Thus, for therapeutic strategies, targeting central cytoplasmic proteins would provide less room for the oncogenic signalling to be offloaded to another protein.

STAT3 delocalizes to all four subcellular compartments: the nucleus, cytoplasm, mitochondrion and plasma membrane, and it interacts with SRC, Receptor of activated protein C kinase 1 (RACK1), MAPK1/3, tyrosine-protein kinase Lyn (LYN), mitochondrial GRIM1 and Sorcin (SRI) across these compartments. Mitochondrial STAT3 action is required for RAS-RAF transformation and is triggered via actions such as the stress kinase-mediated serine phosphorylation of STAT3.<sup>90,91</sup> Apart from STAT3, both STAT1 and STAT5 are also found to translocate to mitochondria (Figure S3C). Unlike the nucleus and cytoplasm, one of the direct interaction partners of STAT3 in the mitochondria was described to be the proteins of complex I

and complex III of the respiratory chain and 2-Oxoglutarate dehydrogenase (OGDHL), which is observed in our interactomes and has a high degree of centrality. There are two other highly central proteins in the mitochondria, TP53 and Dihydropyridolipoyl dehydrogenase (DLD). DLD and OGDHL both function as oxidoreductases, and OGDHL is associated with cervical cancer.<sup>92</sup> The relatively large numbers of central proteins in this network suggests that STAT3 is not as important to mitochondrial signalling as cytosolic or nuclear localisation of STAT3. Recently, STAT3 has been shown to be located within a portion of the ER associated with the mitochondria instead of the mitochondria itself, putting its crucial role within the mitochondria into question.<sup>93</sup> Furthermore, the presence of multiple proteins in the above with similar functions (i.e., oxidoreductases) in the mitochondrion indicates the existence of a redundant molecular pathway. In the plasma membrane, STAT3, along with EGFR, dominates the network, with a similar number of interactions and degree of centrality (Figure S3D). Due to a high number of shared interactions in the plasma membrane interactome, both STAT3 and EGFR appear critical for signalling across the plasma membrane, followed consecutively by JAK1/2 and SRC. Due to its servitude as a reliable interaction partner and communication hub for STATs from the plasma membrane to the nucleus, EGFR disruption can clearly be foreseen and is shown to be beneficial in arresting or eradicating cancer cell growth.<sup>94</sup> Notably, STAT2 is the only other STAT delocalized to the plasma membrane although it does not appear to play a key role in the communication there. Overall, while the interactomes of the nucleus and cytoplasm appear remarkably similar to the whole STAT interactome, STAT3 appears to be the only influential STAT in the plasma membrane and displays minimal importance in the mitochondria.

#### 4 | FUTURE DIRECTIONS AND CONCLUSIONS

In all the carcinoma and hematopoietic cancer networks, STAT1 and STAT3 interact with a similar panel of proteins, including MTOR, SRC and EGFR (Figure 3, Figure 4, Figure 5 and Figure S4). The observed crosstalk between these STATs as well as between STAT5A and STAT5B suggests that the formation of a heterodimer is important in diseased states. Interestingly, although a STAT1-STAT5A interaction is reported in breast and lung cancers or leukaemia, there is no evidence for a STAT1-STAT5B or a STAT3-STAT5A interaction observed in any of the interactomes investigated. While all STATs interact with a larger number of oncogenes than tumour suppressors, the ratio of oncogenes to tumour suppressors is highest for STAT5A and STAT5B (Figure S5). For each disease investigated in this paper, a STAT1-STAT3 interaction has been consistently reported (Figure 3, Figure 4, Figure 5 and Figure S4). These interactions clearly indicate the importance of STAT crosstalk in diseased states, and hence, their ongoing exploration in drug discovery efforts will be important. Targeting of STATs is a long-standing strategy, and a variety of small-molecule drugs and degraders such as fludarabine, pimozone, sulforaphane,

pyrimethamine and SD-36 have been developed. Our analysis indicates that although inhibition of STATs is said to severely hinder the viability of certain cancers, combinatorial therapy of STATs alongside other key proteins including TP53, ESR1 and EGFR may result in a higher degree of success especially in solid cancers. Furthermore, based on our interactome, ESR1 appears to be central and it can serve as a new viable combinatorial target in lung cancers. In blood cancers, however, STAT1, STAT3 and STAT5 play a dominant role and single STAT-targeting therapies may suffice, also reflecting their less complex genetic driver mutation landscape. In most cases of STAT-linked diseases such as solid carcinomas, a combination therapy of STAT3/5 dual inhibitors or combination therapy with upstream TKI such as ERBB family member blockers or JAKinibs will likely be more successful.

#### ACKNOWLEDGEMENTS

P.T.G. is supported by research grants from NSERC (RGPIN-2014-05767), CIHR (MOP-130424, MOP-137036), Canada Research Chair (950-232042), Canadian Cancer Society (703963), Canadian Breast Cancer Foundation (705456) and infrastructure grants from CFI (33536) and the Ontario Research Fund (34876). This work was further supported by the EU consortia ERAPerMed 'JAKSTAT-TARGET' (S.M., M.H., and P.T.G.) and the 'ERANETPLL' (M.H. and R.M.). S.M. was funded by European Research Council (M-IMM project), Academy of Finland, Sigrid Juselius Foundation and the Cancer Foundation Finland. R.M. was supported by the Austrian Science Fund (FWF) (SFB-F04707 and SFB-F06105). P.V. was supported by the Austrian Science Fund (FWF) grant (F4704-B20).



#### CONFLICT OF INTEREST

The authors declare no competing interests.

#### AUTHOR CONTRIBUTIONS

**Fettah Erdogan:** Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – original draft (lead); Writing – review & editing (lead). **Tudor Bogdan Radu:** Data curation (lead); Formal analysis (equal); Methodology (supporting); Software (lead); Visualization (lead); Writing – original draft (supporting). **Anna Orlova:** Visualization (supporting); Writing – review & editing (supporting). **Abdul Khawazak Qadree:** Visualization (supporting); Writing – review & editing (supporting). **Elvin Dominic de Araujo:** Project administration (supporting); Resources (supporting); Writing – review & editing (supporting). **Johan Israelian:** Investigation (supporting). **Peter Valent:** Writing – review & editing (supporting). **Satu Mustjoki:** Writing – review & editing (supporting). **Marco Herling:** Writing – original draft (supporting); Writing – review & editing (supporting). **Richard Morigg:** Supervision (equal); Writing – review & editing (supporting). **Patrick Thomas Gunning:** Supervision (equal).

#### ORCID

Elvin Dominic de Araujo  <https://orcid.org/0000-0003-0716-2830>  
Patrick Thomas Gunning  <https://orcid.org/0000-0003-0654-735X>

## REFERENCES

- Yu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat Rev Cancer*. 2009;9:798-809.
- Wang Z, Zhu S, Shen M, et al. STAT3 is involved in esophageal carcinogenesis through regulation of Oct-1. *Carcinogenesis*. 2013;34:678-688.
- Schindler C, Darnell JE Jr. Transcriptional responses to polypeptide ligands: the JAK-STAT pathway. *Annu Rev Biochem*. 1995;64(1):621-652. doi:10.1146/annurev.bi.64.070195.003201
- Shuai K, Horvath CM, Huang LH, Qureshi SA, Cowburn D, Darnell JE Jr. Interferon activation of the transcription factor Stat91 involves dimerization through SH2-phosphotyrosyl peptide interactions. *Cell*. 1994;76(5):821-828. doi:10.1016/0092-8674(94)90357-3
- Darnell JE, Kerr LM, Stark GR. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*. 1994;264(5164):1415-1421. doi:10.1126/science.8197455
- Chen X, Vinkemeier U, Zhao Y, Jeruzalmi D, Darnell JE Jr, Kuriyan J. Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell*. 1998;93:827-839.
- Kücük C, Jiang B, Hu X, et al. Activating mutations of STAT5B and STAT3 in lymphomas derived from  $\gamma\delta$ -T or NK cells. *Nat Commun*. 2015;6:6025.
- Heppler LN, Frank DA. Rare mutations provide unique insight into oncogenic potential of STAT transcription factors. *J Clin Invest*. 2018;128:113-115.
- Rajala HLM, Eldfors S, Kuusanmäki H, et al. Discovery of somatic STAT5b mutations in large granular lymphocytic leukemia. *Blood*. 2013;121:4541-4550.
- de Araujo ED, Keserü GM, Gunning PT, Moriggl R. Targeting STAT3 and STAT5 in cancer. *Cancers (Basel)*. 2020;12:2002.
- de Araujo ED, Erdogan F, Neubauer HA, et al. Structural and functional consequences of the STAT5BN642H driver mutation. *Nat Commun*. 2019;10:2517.
- Ferlay J, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136:E359-E386.
- Navlakha S, Kingsford C. The power of protein interaction networks for associating genes with diseases. *Bioinformatics*. 2010;26:1057-1063.
- Lee I, Blom UM, Wang PI, Shim JE, Marcotte EM. Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Res*. 2011;21:1109-1121.
- Yeh S-H, Yeh H-Y, Soo V-W. A network flow approach to predict drug targets from microarray data, disease genes and interactome network - Case study on prostate cancer. *J Clin Bioinforma*. 2012;2:1.
- Isik Z, Baldow C, Cannistraci CV, Schroeder M. Drug target prioritization by perturbed gene expression and network information. *Sci Rep*. 2015;5:17417.
- Guney E, Menche J, Vidal M, Barabási A-L. Network-based in silico drug efficacy screening. *Nat Commun*. 2016;7:10331.
- Sang L, Miller JJ, Corbit KC, et al. Mapping the NPHP-JBTS-MKS protein network reveals ciliopathy disease genes and pathways. *Cell*. 2011;145:513-528.
- Kotlyar M, Pastrello C, Malik Z, Jurisica I. IID 2018 update: context-specific physical protein-protein interactions in human, model organisms and domesticated species. *Nucleic Acids Res*. 2019;47:D581-D589.
- Bossi A, Lehner B. Tissue specificity and the human protein interaction network *Mol Syst Biol*. 2009;5:260.
- Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 2013;41:D991-D995.
- Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix genechip data at the probe level. *Bioinformatics*. 2004;20:307-315.
- Elefsinioti A, Saraç ÖS, Hegele A, et al. Large-scale de novo prediction of physical protein-protein association. *Mol Cell Proteomics*. 2011;10:M111.010629.
- Kotlyar M, Pastrello C, Pivetta F, et al. In silico prediction of physical protein interactions and characterization of interactome orphans. *Nat Methods*. 2015;12:79-84.
- Brown KR, et al. NAViGaTOR: network analysis, visualization and graphing Toronto. *Bioinformatics*. 2009;25:3327-3329.
- Hufsky F, Lamkiewicz K, Almeida A, et al. Computational strategies to combat COVID-19: useful tools to accelerate SARS-CoV-2 and coronavirus research. *Brief Bioinform*. 2021;22:642-663.
- Sadegh S, Matschinske J, Blumenthal DB, et al. Exploring the SARS-CoV-2 virus-host-drug interactome for drug repurposing. *Nat Commun*. 2020;11:3518.
- Alqahtani A, Ayesh HSK, Halawani H. PIK3CA gene mutations in solid malignancies: association with clinicopathological parameters and prognosis. *Cancers (Basel)*. 2020;12(1):93.
- Nivarthi H, Gordziel C, Themanns M, et al. The ratio of STAT1 to STAT3 expression is a determinant of colorectal cancer growth. *Oncotarget*. 2016;7:51096-51106.
- Loh C-Y, Arya A, Naema AF, Wong WF, Sethi G, Looi CY. Signal Transducer and Activator of Transcription (STATs) proteins in cancer and inflammation: functions and therapeutic implication. *Front Oncol*. 2019;9:48.
- Girardot M, Pecquet C, Chachoua I, et al. Persistent STAT5 activation in myeloid neoplasms recruits p53 into gene regulation. *Oncogene*. 2015;34:1323-1332.
- Rahmati S, Abovsky M, Pastrello C, Jurisica I. pathDIP: an annotated resource for known and predicted human gene-pathway associations and pathway enrichment analysis. *Nucleic Acids Res*. 2017;45:D419-D426.
- Liao W, Schones DE, Oh J, et al. Priming for T helper type 2 differentiation by interleukin 2-mediated induction of interleukin 4 receptor alpha-chain expression. *Nat Immunol*. 2008;9:1288-1296.
- Rolling C, Treton D, Pellegrini S, Galanaud P, Richard Y. IL4 and IL13 receptors share the gamma c chain and activate STAT6, STAT3 and STAT5 proteins in normal human B cells. *FEBS Lett*. 1996;393:53-56.
- Kaltenecker D, Themanns M, Mueller KM, et al. Hepatic growth hormone - JAK2 - STAT5 signalling: metabolic function, non-alcoholic fatty liver disease and hepatocellular carcinoma progression. *Cytokine*. 2019;124:154569.
- Engblom D, Kornfeld JW, Schwake L, et al. Direct glucocorticoid receptor-Stat5 interaction in hepatocytes controls body size and maturation-related gene expression. *Genes Dev*. 2007;21:1157-1162.
- Bondesson M, Hao R, Lin C-Y, Williams C, Gustafsson J-Å. Estrogen receptor signaling during vertebrate development. *Biochim Biophys Acta*. 2015;1849:142-151.
- Erdogan F, Qadree AK, Radu TB, et al. Structural and mutational analysis of member-specific STAT functions. *Biochim Biophys Acta - Gen Subj*. 2021;1866(3):130058. doi:10.1016/j.bbagen.2021.130058
- Cirri P, Chiarugi P, Marra F, et al. c-Src Activates both STAT1 and STAT3 in PDGF-Stimulated NIH3T3 Cells. *Biochem Biophys Res Commun*. 1997;239:493-497.
- Goi T, Shipitsin M, Lu Z, Foster DA, Klinz SG, Feig LA. An EGF receptor/Ral-GTPase signaling cascade regulates c-Src activity and substrate specificity. *EMBO J*. 2000;19:623-630.
- Valent P, et al. Cancer stem cell definitions and terminology: the devil is in the details. *Nat Rev Cancer*. 2012;12:767-775.
- Valent P, Bonnet D, De Maria R, et al. Heterogeneity of neoplastic stem cells: theoretical, functional, and clinical implications. *Cancer Res*. 2013;73:1037-1045.

43. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med*. 2004;10:789-799.
44. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science*. 2013;339(6127):1546-1558.
45. Wingelhofer B, Neubauer HA, Valent P, et al. Implications of STAT3 and STAT5 signaling on gene regulation and chromatin remodeling in hematopoietic cancer. *Leukemia*. 2018;32:1713-1726.
46. Travis WD, Brambilla E, Burke AP, Marx A, Nicholson AG. Introduction to The 2015 World Health Organization classification of tumors of the lung, pleura, thymus, and heart. *J Thorac Oncol off Publ Int Assoc Study Lung Cancer*. 2015;10:1240-1242.
47. Mohrherr J, Uras IZ, Moll HP, Casanova E. STAT3: versatile functions in non-small cell lung cancer. *Cancers (Basel)*. 2020;12:E1107.
48. Hanby AM, Walker C, Tavassoli FA, Devilee P. *Pathology and Genetics: Tumours of the Breast and Female Genital Organs. WHO Classification of Tumours series - Volume IV. Lyon, France: IARC Press. Breast Cancer Research*. 6:133.
49. Howlader N, Altekruse SF, Li CI, et al. US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. *J Natl Cancer Inst*. 2014;106:dju055.
50. Łukasiewicz S, Czezelewski M, Forma A, Baj J, Sitarz R, Stanisławek A. Breast cancer-epidemiology, risk factors, classification, prognostic markers, and current treatment strategies-an updated review. *Cancers (Basel)*. 2021;13:4287.
51. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*. 2016;534:47-54.
52. Robinson DR, Wu YM, Vats P, et al. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat Genet*. 2013;45:1446-1451.
53. Marotta LL, Almendro V, Marusyk A, et al. The JAK2/STAT3 signaling pathway is required for growth of CD44<sup>+</sup>CD24<sup>-</sup> stem cell-like breast cancer cells in human tumors. *J Clin Invest*. 2011;121:2723-2735.
54. Zhang X, Yue P, Page BD, et al. Orally bioavailable small-molecule inhibitor of transcription factor Stat3 regresses human breast and lung cancer xenografts. *Proc Natl Acad Sci*. 2012;109:9623-9628.
55. Yang MY, Wang CJ, Chen NF, Ho WH, Lu FJ, Tseng TH. Luteolin enhances paclitaxel-induced apoptosis in human breast cancer MDA-MB-231 cells by blocking STAT3. *Chem Biol Interact*. 2014;213:60-68.
56. Liu CY, Chen KF, Chao TI, et al. Sequential combination of docetaxel with a SHP-1 agonist enhanced suppression of p-STAT3 signaling and apoptosis in triple negative breast cancer cells. *J Mol Med*. 2017;95:965-975.
57. Prabhu JS, Wahi K, Korlimarla A, et al. The epigenetic silencing of the Estrogen Receptor (ER) by hypermethylation of the ESR1 promoter is seen predominantly in triple-negative breast cancers in Indian women. *Tumor Biol*. 2012;33:315-323.
58. Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 WHO classification of tumours of the urinary system and male genital organs-part a: renal, penile, and testicular tumours. *Eur Urol*. 2016;70:93-105.
59. Mora LB, Buettner R, Seigne J, et al. Constitutive activation of stat3 in human prostate tumors and cell lines: direct inhibition of stat3 signaling induces apoptosis of prostate cancer cells. *Cancer Res*. 2002;62:6659-6666.
60. Singh RK, Lokeshwar BL. The IL-8-regulated chemokine receptor CXCR7 stimulates EGFR signaling to promote prostate cancer growth. *Cancer Res*. 2011;71:3268-3277.
61. Gan Y, Shi C, Inge L, Hibner M, Balducci J, Huang Y. Differential roles of ERK and Akt pathways in regulation of EGFR-mediated signaling and motility in prostate cancer cells. *Oncogene*. 2010;29:4947-4958.
62. Li LC, Shiina H, Deguchi M, et al. Age-dependent methylation of ESR1 gene in prostate cancer. *Biochem Biophys Res Commun*. 2004;321:455-461.
63. Fu C, Dong W-Q, Wang A, Qiu G. The influence of ESR1 rs9340799 and ESR2 rs1256049 polymorphisms on prostate cancer risk. *Tumor Biol*. 2014;35:8319-8328.
64. Yu L, Ye F, Li Y-Y, et al. Histone methyltransferase SETDB1 promotes colorectal cancer proliferation through the STAT1-CCND1/CDK6 axis. *Carcinogenesis*. 2020;41:678-688.
65. Rokavec M, Öner MG, Li H, et al. IL-6R/STAT3/miR-34a feedback loop promotes EMT-mediated colorectal cancer invasion and metastasis. *J Clin Invest*. 2014;124:1853-1867.
66. Gemignani F, Moreno V, Landi S, et al. A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. *Oncogene*. 2004;23:1954-1956.
67. McDermott U, Longley DB, Galligan L, Allen W, Wilson T, Johnston PG. Effect of p53 status and STAT1 on chemotherapy-induced, fas-mediated apoptosis in colorectal cancer. *Cancer Res*. 2005;65:8951-8960.
68. Zhao Y, Ye L, Yu Y. MicroRNA-126-5p suppresses cell proliferation, invasion and migration by targeting EGFR in liver cancer. *Clin Res Hepatol Gastroenterol*. 2020;44(6):865-873. doi:10.1016/j.clinre.2020.03.025
69. Zhou L, Xu Y, Wang D, et al. TP53 R249G/S mutation as an independent prognosis indicator in hepatocellular carcinoma. *J Clin Oncol*. 2020;38:e16622.
70. Zhou Y, Li Y, Xu S, et al. Eukaryotic elongation factor 2 kinase promotes angiogenesis in hepatocellular carcinoma via PI3K/Akt and STAT3. *Int J Cancer*. 2020;146:1383-1395.
71. Qin X, Song Y. Bioinformatics Analysis Identifies the Estrogen Receptor 1 (ESR1) Gene and hsa-miR-26a-5p as potential prognostic biomarkers in patients with intrahepatic cholangiocarcinoma. *Med Sci Monit*. 2020;26:e921815-1-e921815-15.
72. Kirchmeyer M, Servais F, Ginolhac A, et al. Systematic transcriptional profiling of responses to STAT1- and STAT3-activating cytokines in different cancer types. *J Mol Biol*. 2020;432:5902-5919.
73. Martí-Rodrigo A, Alegre F, Moragrega ÁB, et al. Rilpivirine attenuates liver fibrosis through selective STAT1-mediated apoptosis in hepatic stellate cells. *Gut*. 2020;69:920-932.
74. Liu Y, Li P-K, Li C, Lin J. Inhibition of STAT3 signaling blocks the anti-apoptotic activity of il-6 in human liver cancer cells\*. *J Biol Chem*. 2010;285:27429-27439.
75. Randomised trial of tamoxifen versus placebo for the treatment of inoperable hepatocellular carcinoma. 2013.
76. Cisplatin, doxorubicin and tamoxifen in the treatment of incurable soft tissue and endocrine malignancies. 2013.
77. Llovet JM, Kelley RK, Augusto V, et al. Hepatocellular carcinoma. *Nat Rev Dis Prim*. 2021;7:6.
78. Constantinescu SN, Vainchenker W, Levy G, Papadopoulos N. Functional consequences of mutations in myeloproliferative neoplasms. *HemaSphere*. 2021;5:e578.
79. Nosaka T, Kitamura T. Janus Kinases (JAKs) and Signal Transducers And Activators Of Transcription (STATs) in hematopoietic cells. *Int J Hematol*. 2000;71:309-319.
80. Rochman Y, Spolski R, Leonard WJ. New insights into the regulation of T cells by  $\gamma$ c family cytokines. *Nat Rev Immunol*. 2009;9:480.
81. Spolski R, Leonard WJ. Interleukin-21: basic biology and implications for cancer and autoimmunity. *Annu Rev Immunol*. 2008;26:57-79.
82. Hornakova T, Staerk J, Royer Y, et al. Acute lymphoblastic leukemia-associated JAK1 mutants activate the janus kinase/STAT pathway via Interleukin-9 receptor  $\alpha$  homodimers. *J Biol Chem*. 2009;284:6773-6781.
83. Demoulin J-B, Roost EV, Stevens M, Groner B, Renauld J-C. Distinct roles for STAT1, STAT3, and STAT5 in differentiation gene induction and apoptosis inhibition by interleukin-9 \*. *J Biol Chem*. 1999;274:25855-25861.
84. de Araujo ED, Orlova A, Neubauer HA, et al. Structural implications of stat3 and stat5 sh2 domain mutations. *Cancers (Basel)*. 2019;11(11):1757.

85. Koskela HL, Eldfors S, Ellonen P, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. *N Engl J Med*. 2012;366:1905-1913.
86. Kovacic B, Stoiber D, Moriggl R, et al. STAT1 acts as a tumor promoter for leukemia development. *Cancer Cell*. 2006;10:77-87.
87. Lainey E, Wolfroth A, Sukkurwala AQ, et al. EGFR inhibitors exacerbate differentiation and cell cycle arrest induced by retinoic acid and vitamin D3 in acute myeloid leukemia cells. *Cell Cycle*. 2013;12:2978-2991.
88. Hess CJ, Errami A, Berkhof J, et al. Concurrent methylation of promoters from tumor associated genes predicts outcome in acute myeloid leukemia. *Leuk Lymphoma*. 2008;49:1132-1141.
89. Ren Z, Aerts JL, Vandenplas H, et al. Phosphorylated STAT5 regulates p53 expression via BRCA1/BARD1-NPM1 and MDM2. *Cell Death Dis*. 2016;7:e2560.
90. Gough DJ, Corlett A, Schlessinger K, Wegrzyn J, Larner AC, Levy DE. Mitochondrial STAT3 supports Ras-dependent oncogenic transformation. *Science*. 2009;324:1713-1716.
91. Gough DJ, Marié IJ, Lobry C, Aifantis I, Levy DE. STAT3 supports experimental K-RasG12D-induced murine myeloproliferative neoplasms dependent on serine phosphorylation. *Blood*. 2014;124:2252-2261.
92. Sen T, Sen N, Noordhuis MG, et al. OGDHL Is a modifier of AKT-dependent signaling and NF- $\kappa$ B function. *PLoS One*. 2012;7:e48770.
93. Su Y, Huang X, Huang Z, Huang T, Xu Y, Yi C. STAT3 localizes in mitochondria-associated ER membranes instead of in mitochondria. *Front Cell Dev Biol*. 2020;8:274.
94. Darwin P, Joung YH, Kang DY, et al. Tannic acid inhibits EGFR/STAT1/3 and enhances p38/STAT1 signalling axis in breast cancer cells. *J Cell Mol Med*. 2017;21:720-734.

#### SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Erdogan F, Radu TB, Orlova A, et al. JAK-STAT core cancer pathway: An integrative cancer interactome analysis. *J Cell Mol Med*. 2022;26:2049-2062. doi:[10.1111/jcmm.17228](https://doi.org/10.1111/jcmm.17228)