Bioinformatics, 37(20), 2021, 3501–3508
doi: 10.1093/bioinformatics/btab245
Advance Access Publication Date: 11 May 2021
Original Paper

Gene expression

scConnect: a method for exploratory analysis of cell–cell communication based on single-cell RNA-sequencing data

Jon E.T. Jakobsson [1], Ola Spjuth [2] and Malin C. Lagerström [1,*]

[1]Department of Neuroscience, Uppsala University, 75124 Uppsala, Sweden and [2]Department of Pharmaceutical Biosciences, Uppsala University, 75124 Uppsala, Sweden

*To whom correspondence should be addressed.
Associate Editor: Inanc Birol

Received on July 10, 2020; revised on February 18, 2021; editorial decision on April 12, 2021; accepted on April 14, 2021

Abstract

Motivation: Cell to cell communication is critical for all multicellular organisms, and single-cell sequencing facilitates the construction of full connectivity graphs between cell types in tissues. Such complex data structures demand novel analysis methods and tools for exploratory analysis.

Results: We propose a method to predict the putative ligand–receptor interactions between cell types from single-cell RNA-sequencing data. This is achieved by inferring and incorporating interactions in a multi-directional graph, thereby enabling contextual exploratory analysis. We demonstrate that our approach can detect common and specific interactions between cell types in mouse brain and human tumors, and that these interactions fit with expected outcomes. These interactions also include predictions made with molecular ligands integrating information from several types of genes necessary for ligand production and transport. Our implementation is general and can be appended to any transcriptome analysis pipeline to provide unbiased hypothesis generation regarding ligand to receptor interactions between cell populations or for network analysis *in silico*.

Availability and implementation: scConnect is open source and available as a Python package at https://github.com/JonETJakobsson/scConnect. scConnect is directly compatible with Scanpy scRNA-sequencing pipelines.

Contact: malin.lagerstrom@neuro.uu.se

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Cell to cell communication is critical for all multicellular organisms. Most inter-tissue signaling, intra-tissue signaling, neuronal signaling, cell–cell adhesion and development rely on ligands and receptors (LRs) for their communication, and information about these LRs are inferable from transcriptomic datasets by creating gene to protein relationships (Efremova *et al.*, 2020). Integration of single-cell transcription data with known LR interactions has been used to detect cell–cell communication in tissues from mammalian lung (Raredon *et al.*, 2019), annelid brain (Williams *et al.*, 2017), tumors (Choi *et al.*, 2015; Kumar *et al.*, 2018; Zhou *et al.*, 2017), human liver buds (Camp *et al.*, 2017) and, mouse heart (Skelly *et al.*, 2018). However, tools specifically designed to study cell–cell communication between cell types in single-cell RNA-sequencing (scRNA-seq) datasets are just starting to emerge (Efremova *et al.*, 2020). Information about LRs, and interactions between them is collected in several databases, including IUPAR/BPS Guide to pharmacology (GTF) (Harding *et al.*, 2018), and the IMEx consortium (Orchard *et al.*, 2012), which integrate interactions from several other databases, such as database of interacting proteins (Salwinski *et al.*, 2004), interlogous interaction database (I2D) (Brown and Jurisica, 2005) and innateDB (Breuer *et al.*, 2013) among others. Moreover, cellPhoneDB (Efremova *et al.*, 2020) have specifically designed a database suitable for gene–protein linking and protein–protein interaction assessment. A current limitation in available methods includes the ability to link gene expression to molecular ligands, such as glutamate, GABA and acetylcholine, which expression is dependent on complex gene–ligand relationships. Information about these classical neurotransmitters is valuable for neuronal datasets, where much of the fast synaptic transmission is conducted using these ligands. Interaction between cell types can help to delineate the structure of complex neuronal circuits based on the results from *in vivo* tracing studies. Further, there is a lack of implementations compatible with the standard packages for scRNA-seq analysis, such as Scanpy (Wolf *et al.*, 2018) (Python) and Seurat (Butler *et al.*, 2018) (R).

Here, we introduce a method that estimates LR expression from gene expression profiles of a cell population. Expression of molecular ligands is calculated using an ensemble of enzyme and transporter genes. Using the GTF database (Harding *et al.*, 2018), we infer all putative interactions between populations in a scRNA-seq datasets and store this information in a multi-directional graph, which can be used for visualization and hypothesis generation. The method

and associated analytical tools have been implemented in a python package called scConnect, and we demonstrate it on two biological datasets to assess the ability to detect previously established connections and produce novel hypotheses for further studies.

2 Materials and methods

This section describes how single-cell transcriptomics data and the GTF database (Harding *et al.*, 2018) are integrated to facilitate inference of interactions between cell types. This includes a preparatory database construction step to link LRs with gene names for a specific species. For each subsequent analysis, the following steps are carried out (see also Fig. 1 for an overview): (i) gene calling on the scRNA-seq data to obtain group based gene expression; (ii) calculation of LR scores based on the gene expression; (iii) shuffling of group labels to estimate the random distribution of LR scores and calculation of z-scores, *P*-values and corrected *P*-values for each LR score; and (iv) interaction inference and graph construction. We refer to this method as scConnect. We further describe different methods to analyze the constructed graph to detect specific interactions, differential interactions between two populations, and for interactive exploratory analysis.

2.1 Database preparation

The database preparation procedure aims to link LRs with gene names for a specific species (S_x). From the GTF database, peptidergic LRs, such as G-protein-coupled receptors (GPCRs) and ligand-gated ion channels (LGICs) with gene annotations for human, rat or mouse (this includes the majority of annotations) are included (Harding *et al.*, 2018), and the orthologous genes in S_x are identified using g: Profiler (Raudvere *et al.*, 2019) and this gene–protein linkage is stored in two tables: LRs (Fig. 1a). However, molecular ligands, such as glutamate and acetylcholine are dependent on several sets of genes, and these have been manually annotated, guided by the approach of Zeisel *et al.* (2018) (Supplementary Table S1), and stored in the Ligand table (Fig. 1a). For further information regarding molecular ligands, see Section 2.3. Currently, 189 species are available in the gene orthology search (Raudvere *et al.*, 2019) and databases for these species can hence be constructed. However, the number of detectable orthologous genes decrease with evolutionary distance to the mammalian clade (Supplementary Table S2). With the assumption that ortholog LRs also interact (see Section 4 for why this is reasonable) information of interactions between LRs are stored on a protein level (Fig. 1a, Interactions DB). Only interactions shown to be endogenously occurring are included, but metadata related to receptor family, ligand species, receptor species and PubMed ID are also included.

2.2 Gene calling algorithms

Given a read or count matrix R, with *i* genes and *j* cells and a group annotation vector G with length *j* and a set of *n* groups, a gene call matrix GC can be constructed with *n* groups and *i* genes. We calculate mean expression over the groups. Hence, all downstream values relate to mean reads. Other methods are available but are not recommended by the authors. See online documentation introduced in Section 2.6 for available options.

2.3 LR calling

To facilitate the integration of the GTF interaction database, where LRs are proteins and molecular ligands, LR scores (L score and R score, respectively) are calculated from the gene call scores. All peptidergic ligands and most receptors have a one to one gene–protein relationship. Therefore, the LR scores are equal to the gene calls. As ortholog gene search can report more than one gene for one receptor, the maximum of those genes is used, as one of these might be a dysfunctional pseudogene. Synthesis of molecular ligands is dependent on one or more enzymes, and for these ligands to be accessible as signaling molecules vesicle transporters and reuptake transporters might be needed. Furthermore, certain enzymes, if present, can

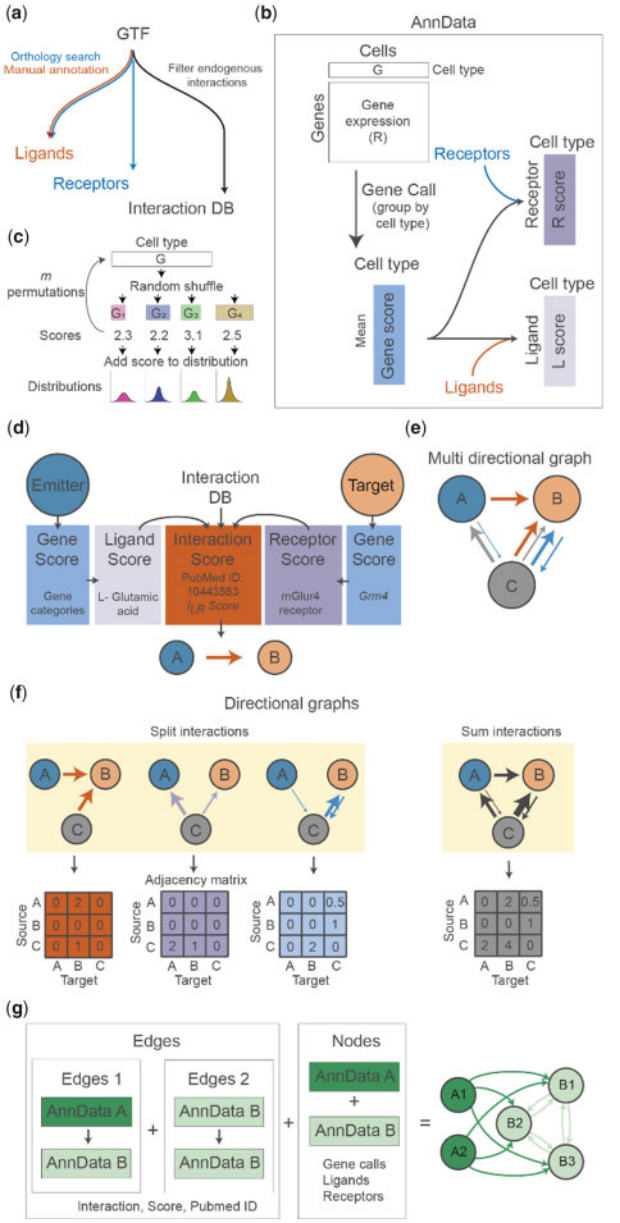

Fig. 1. Overview of scConnect. (**a**) Gene annotation for ligands includes ortholog gene search for peptidergic ligands and further manually curated gene categories for molecular ligands. Annotation for receptors include ortholog gene search. Filtering of interactions is performed including only endogenous interactions, and interactions involving LRs included in LR annotations. (**b**) Gene call is calculated as mean expression grouped by cell type. Ligand score is calculated based on the gene call scores. Molecular ligands incorporate several gene sets, whereas peptides only relate to one gene. Receptor score is calculated based on gene call score. (**c**) Permutation is used to estimate the random distribution of LR scores. Cell type annotations are randomly shuffled, and LR scores are calculated. Each LR score are stored to create a separate permutated distribution for each LR for each cell type. Number of scores in the distribution is equal to the number of permutations (*m*). (**d**) The emitter's ligand score and the target's receptor score are used to calculate an interaction score for a ligand–receptor interaction between the emitter and the target population. Valid ligand–receptor pairs are selected from the interaction DB, which only includes endogenous interactions. (**e**) After assessing the interaction score for all LR pairs between all populations, a multi-directional graph is constructed. Each interaction has a direction and a weight related to the interaction score. Multiple interactions can be found between two nodes. The color of the arrow represents a unique interaction type. (**f**) The multi-directional graph can be split into subgraphs only containing one type of interaction, or a sum of all interactions. These directional graphs can be represented as adjacency matrices, which can be plotted as heatmaps or used for downstream analysis. (**g**) A graph can be constructed from complex experimental setups. Here, interactions between populations in adata A to populations in adata B and interactions within populations in adata B is assessed by creating and appending individual queries. Metadata about the populations can be included by passing the two adata object to the graph construction

convert the ligand to another ligand, as with the dopamine → noradrenalin → adrenalin synthesis, and must not be present (see excluded genes, Supplementary Table S1). The L score for molecular ligands is calculated using the logic that one vesicle transporter and one reuptake transporter are sufficient for ligand transport, and hence the maximum gene expression in these gene categories is used $t = Max\{\text{Transporter}\}$, $r = \text{Max}\{\text{Reuptake}\}$. In contrast, all synthesis enzymes are needed to produce the ligand, and hence the geometric-mean of this gene category is used: $s = \sqrt[k]{\prod_{n=1}^{k} \text{Synthesis}_k}$. The geometric-mean of s, t and r represents the ligand promoting factor: $p = \sqrt[3]{t*r*s}$. The exclusion gene score (e) is subtracted from p, which nullify the term if the exclusion gene has higher expression than p. If $> p$: L score $= 0$. Hence:

$$L \text{ score} = p - e, \; L \text{ score} \geq 0 \tag{1}$$

LR score can be calculated from any gene score matrix, independent of gene score method (Fig. 1b).

To estimate how differentially expressed a ligand or receptor is in a cell type, m number of permutations with randomized cell type annotations are used to estimate a random permutation distribution of each LR for each cell type (Fig. 1c). An exact permutation P-value is calculated as $p = (b+1)/(m+1)$ where b is the number of permutated scores higher than the observed score within that cell type (Phipson and Smyth, 2010). P-values are corrected for multiple testing using the Benjamini/Hochberg method [statsmodels package (Seabold and Perktold, 2010)]. Furthermore, Z-scores for each LR were calculated as $(\text{score} - \mu)/s$ where μ is the mean and s is the SD of the permutation distribution.

2.4 Interaction inference and graph construction

Each cell type in the scRNA-seq dataset represents a node in a graph, and edges represent a directional LR pair interaction between an emitter node E and a target node T (Fig. 1d). Hence, the graph describes interactions between cell types. Valid LR pairs, where interactions have been proven to occur endogenously, are extracted from the GTP database (Harding *et al.*, 2018). The interaction score (I_{LR} score) is calculated as the geometric-mean of the emitter's L score and the target's R score.

$$I_{LR}\text{score}_E^T = \sqrt{E_{L \text{ score}} \times T_{R \text{ score}}} \tag{2}$$

The interaction specificity is based on the LR P-values. Both P-values must be low to produce a high specificity score. An LR P-value of 0.05 (corrected) produces an interaction specificity score of around 1.3. P-values of 1 produce a specificity score of 0.

$$\text{I}_{\text{LR}}\text{specificity}_{\text{E}}^{\text{T}} = -\text{log10}\left(\frac{\text{E}_{\text{L}_{\text{p-value}}} + \text{T}_{\text{R}_{\text{p-value}}}}{2}\right). \tag{3}$$

An estimate of interaction importance is calculated from the interaction score and interaction specificity.

$$I_{LR}\text{importance}_E^T = \text{log10}\left(I_{LR}\text{score}_E^T\right) * I_{LR}\text{specificity}_E^T. \tag{4}$$

By permutating though all $E \rightarrow T$ combinations for each LR pair, a multi-directional graph, containing the full putative connectome, can be inferred (Fig. 1e).

It is possible to convert the multi-directional graph to a directional graph by either splitting the graph based on interaction or by summarizing the interaction score for all interactions between two nodes (Fig. 1f). Directional graphs can be represented as adjacency matrices with source nodes as rows and target nodes as columns, thereby facilitating a simplified visualization of the graph. To facilitate contextual analysis, all metadata is stored in the graph structure. Gene calls, ligand scores and receptor scores are appended as metadata for the nodes. Interaction score, and other metadata related to an interaction is appended to each edge.

2.5 Analysis of the multi-directional graph

The data stored in a multi-directional graph is complex, and there are many ways to analyze this data. Initially, it might be desirable to enable dynamic and interactive visualizations. For this reason, we have developed an interactive web application that facilitates exploratory analysis of the data, further discussed in Section 2.5.2. Depending on the dataset examined, some features, such as treatment, gender or age can be annotated, and specific comparisons between these features might be of interest. This can be examined by comparing the interactions made between the two cell types, as described in Section 2.5.3.

2.5.1 Identifying specific and strong interactions

There are two main types of interactions that are of interest: strong interactions and specific interactions. Strong interactions are prominent features of the graph with high expression of the ligand and the receptor. They can be present between many or all cell types. Specific interactions connect emitters and targets where the LR is differentially expressed. These will only be present between a small number of cell types. Interactions displaying both of these properties are considered important interactions. We can identify these interactions using the interactive web application, filtering on score, specificity or importance. These interactions are also available as a table and can be filtered and plotted as desired.

2.5.2 Interactive web application

The information contained in a full connectome between different cell types can be vast, and to visualize this information in one graph would oversimplify the data. To this end, we have implemented a user-friendly web application that can be used to drill down into the dataset, and provide insight into strong and specific interactions as well as, LRs. The web application provides an overview of the full graph utilizing a summarized directional graph (Fig. 1f). Selecting one node in this graph plots all incoming and outgoing interactions to that node in a Sankey graph with the selected cell type in the middle. Furthermore, all LR scores are plotted against $-\log10(P \text{ value})$ for the selected cell type. Selecting an edge in the overview graph fills a plot with interaction $\log10(\text{score} + 1)$ against specificity and a table with all interactions between the emitter cell type and target cell type of the selected edge, including information about PubMed ID. This list can also be sorted based on score, significance or importance to identify strong and/or specific interactions.

2.5.3 Identifying differential interactions between two cell types

To detect interactions that differ between two populations (a and b) when making contact to other populations (x), a ratio $(a \rightarrow x)/(b \rightarrow x)$ is calculated, where $a \rightarrow x$ and $b \rightarrow x$ are the interaction score + 1. As the logaritmized interaction scores are represented as $\text{Log}_{10}(I_{LR} + 1)$, we define the ratio as:

$$\log\left(\frac{a \rightarrow x}{b \rightarrow x}\right) = \log(a \rightarrow x) - \log(b \rightarrow x). \tag{5}$$

Note that $a \rightarrow x$ and $b \rightarrow x$ are always ≥ 1, so if the interaction score is 0, we divide by 1. This underestimates the true ratio between interactions for low scores but does not break down when interaction scores approach 0. The ratio can be used to detect differences in interactions between two populations.

2.6 Implementation

The scConnect method has been implemented in a Python package called scConnect, which is compatible with Scanpy. scConnect utilizes AnnData objects (adata) to access scRNA-seq data and store gene calls and LR scores in the adata objects. Gene calls are grouped by any categorical observation variable in adata, such as a cell type annotation or Leiden group. We provide precomputed databases for species in Supplementary Table S2, but the user can also precompute

custom databases using the scConnect tool, and subsequently use these databases. scConnect infers interactions between cell types in a source adata and a target adata, only identifying interactions from the source to the target adata, providing the flexibility needed to create complex interaction setups (Fig. 1g). Inferred interactions are stored in an edge list, and metadata about the cell types are stored in a node list. These are used to construct the multi-directional graph using networkX package (Hagberg *et al.*, 2008), which features powerful graph analysis tools and export to other graph file formats. The interactive web application is implemented using dash (Plotly). The source code is available under MIT License at https://github.com/JonETJakobsson/scConnect together with documentation and a tutorial available as a Jupyter notebook. In this article, we used scConnect version 1.0.3.

3 Results

3.1 Postulating connections between brain regions

The brain is highly interconnected, with some brain regions sending projections to specific nuclei, and others perfusing most of the brain with their projections. To detect connectivity on tissue level, we investigated the mouse brain dataset from Saunders *et al.* (2018) and specifically used the meta cells, summarizing gene expression from thousands of cells of the same cell types. Meta cells originated from nine different brain regions, including thalamus (TH), striatum (STR), substantia nigra/ventral tegmental area (SN), posterior cortex (PC), hippocampus (HC), globus pallidus (GP), frontal cortex (FC), entopeduncular nucleus/subthalamic nucleus (ENT) and cerebellum (CB). We decided to evaluate the ability of scConnect to detect known connections between brain regions, focusing on the dopamine neurotransmitter system as these interactions are well documented.

After grouping meta cells based on tissue of origin, we performed gene calling using mean gene expression levels. To support most interactions in synapses, we included ligand-gated and voltage-gated ion channels and G-protein-coupled receptors in the database. The resulting graph contained 9 tissues connected by 19 879 interactions. Investigating specific interactions (specificity >1.3) we indeed found dopamine to D1 and D2 receptors from SN to STR among the most important interactions in the graph (Fig. 2a), which was expected and validates the ability of scConnect to detect interactions made using molecular ligands. Out of all interactions detected, 17.3% were made using molecular ligands (Supplementary Fig. S1a). SN had a dopamine ligand log(score) of 2.42 and the next highest was STR at 0.74, demonstrating that SN is the major producer of dopamine in the brain, displaying around 85× higher expression than other regions (Supplementary Fig. S1b). This is in concordance with previous studies, where substantia nigra pars compacta (SNc) and ventral tegmental area (VTA) host most dopaminergic neurons (Grimm *et al.*, 2004). The D1 receptor was primarily found in STR, GP and FC, but most brain tissues exhibited some expression (Supplementary Fig. S1c). The D2 receptor was expressed in STR, GP and SN (Supplementary Fig. S1d), and the D3 receptor was mostly found in STR (Supplementary Fig. S1e). The highest expression of the D5 receptor was found in HC (Supplementary Fig. S1f). Previous studies have shown that dopaminergic neurons in the SNc project to the STR and excite spiny projection neurons in the direct pathway (dSPN) via the Dopamine receptor D1 (G_s coupled), while they inhibit the spiny projection neurons in the indirect pathway (iSPN) via the D2 receptor (G_i coupled) (Bertran-Gonzalez *et al.*, 2008). Moreover, dopamine receptors are found in other basal ganglia nuclei and in cortex (Rommelfanger and Wichmann, 2010). Hence, scConnect could predict that dopamine originates from SN, and that the putative targets were STR and GP for both D1 and D2 receptors, and FC for D1 and ENT for D2, which is consistent with previous literature.

Next, we investigated which cell types in SN that could be predicted to contact cell types in STR. Similar cell types were grouped together based on annotated location and/or function, and only relevant cells types were included (Supplementary Table S3). The full connectome including all cell types was also built, and the overall

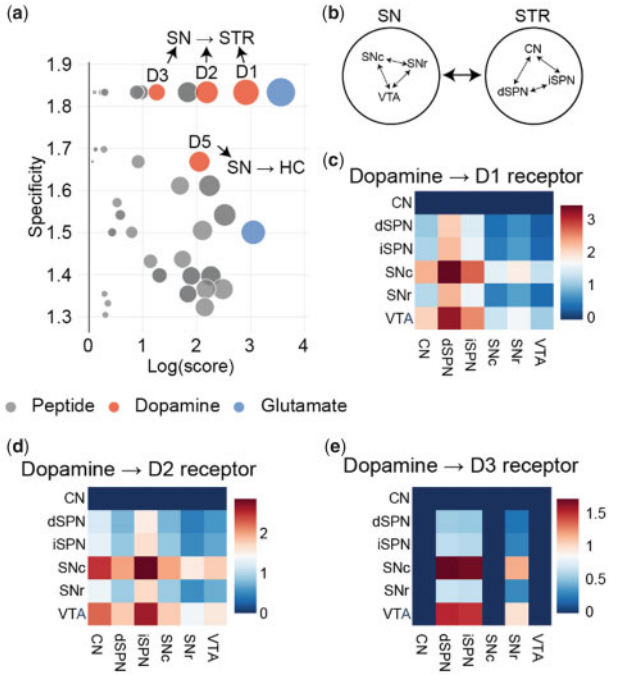

Fig. 2. Connectivity between tissues in the brain. (**a**) Interactions between brain tissues with specificity >1.3. Size of the dots represents the importance of the interaction, and the color represents the ligand type used. Dopamine to D1, D2 and D3 receptor were made from SN to STR, and dopamine to D5 receptor from SN to HC. (**b**) Overview of the simplified cell types in SN and STR tissues. (**c**) Adjacency matrix of dopamine to D1 receptor among simplified cell types. (**d**) Adjacency matrix of dopamine to D2 receptor among simplified cell types. (**e**) Adjacency matrix of dopamine to D3 receptor among simplified cell types. Color in (c–e) represents logarithmized interaction score. TH, thalamus; STR, striatum; SN, substantia nigra/ventral tegmental area; PC, posterior cortex; HC, hippocampus; GP, globus pallidus; FC, frontal cortex; ENT, entopeduncular nucleus/subthalamic nucleus; CB, cerebellum

results were similar (Supplementary Fig. S2a–c). The SN tissue contained the VTA, substantia nigra pars reticulata and SNc, whereas the STR tissue contained dSPN, iSPN and cholinergic neurons (CN) (Saunders *et al.*, 2018) (Fig. 2b). We used mean gene expression as gene calling, and the resulting graph contained 4905 interactions between the 6 cell types. Contrary to the tissues, the selected cells types segregated into either glutamatergic or GABAergic neurons, indicating homogeneity among the underlying cells (Supplementary Fig. S3a–b). SNc and VTA were the strongest dopamine-expressing cell types, with scores 3.1 and 2.7, respectively. dSPN displayed the strongest D1 expression (3.9) but iSPN and CN also showed some expression (2.48 and 1.64, respectively). CN and iSPN showed moderate expression of the D2 receptor (1.86 and 2.62, respectively). In contrast, dSPN displayed low expression of D2 (1.0). Consequently, scConnect predicted strong Dopamine → D1 interactions from VTA and SNc to dSPN (3.3 and 3.5) (Fig. 2c), whereas interactions detected for dopamine → D2 were from SNc and VTA to iSPN (2.9 and 2.7, respectively) (Fig. 2d). Note that dSPN received relatively low D2 input (Fig. 2d). This fits well with the findings of Bertran-Gonzalez *et al.* (2008). The dopamine →D3 receptor interaction was strongest from SNc and VTA to dSPN and iSPN (Fig. 2e).

Hitherto, we have probed the dataset for interactions that we expected to find, focusing on the dopaminergic system. We next aimed to investigate the interactions that scConnect could objectively identify as the most important interactions, and hence looked for the interactions with the highest importance score. The most important interaction to dSPN was dopamine to D1 and 5-hydroxytryptamine to D1 from SNc (7.9 and 5.8 importance score, respectively) and acetylcholine to M4 receptor from CN (5.6) (Supplementary Fig. S3c). Dopamine to D1 receptor was of course known, and scConnect also deemed it to be important. Moreover, acetylcholine has been shown to increase the excitability of dSPN neurons

following M4 receptor activation (Hernández-Flores *et al.*, 2015). The most important incoming interaction to iSPN was dopamine to D2 receptor (6.4) from SNc and acetylcholine to M3 receptor from CN (5.0) (Supplementary Fig. S3d). Not much is written about the M3 receptor in the striatal pathway, but as it is a specific interaction, it might be an interesting target for further research. CN received specific interactions from dSPN via substance P → neurokinin (NK) 1 receptor (8.2) (Supplementary Fig. S3e) and dopamine to D5 receptor from SNc (6.2) (Supplementary Fig. S3f). The expression of substance P is one of the defining features of dSPN (Reiner and Anderson, 1990), and previous studies have demonstrated the presence of NK1 receptors on striatal CN (Kaneko *et al.*, 1993). Presence of D5 receptor in CN in the STR is previously known (Bergson *et al.*, 1995; Berlanga *et al.*, 2005), and activation of the D5 receptor has been shown to activate the CN during L-DOPA treatment of Parkinson's (Castello *et al.*, 2020).

Here, we have demonstrated that scConnect could find expected interactions, such as that the dopamine system between SN and STR. Furthermore, when searching for the top important interactions scConnect detects the dopaminergic system but also the cholinergic system. Hence, scConnect could confirm previous knowledge and generate novel hypotheses. This displays the usefulness of scConnect in hypothesis generation prior to *in vivo* or *in situ* experiments.

3.2 Postulating interactions between cells in human tumors

To evaluate scConnect's ability to identify interactions between non-neuronal cells, we investigated a human melanoma scRNA-seq dataset (Tirosh *et al.*, 2016). Tirosh *et al.* sequenced over 4600 cells from 19 melanoma tumors and annotated non-malignant cells; T-cells, natural killer cells (NK), macrophages, endothelial cells, cancer associated fibroblasts (CAF) and B-cells and malignant cells for each tumor. They found that malignant cells were either associated with microphthalmia-associated transcription factor (MITF) expression or AXL (AXL receptor tyrosine kinase) expression, and that AXL-associated cells were comparatively more numerous after treatment with rapidly accelerated fibrosarcoma kinase and mitogen-activated protein kinase inhibitors (dabrafenib and trametinib) (Tirosh *et al.*, 2016). We used the increase in numerosity of the treatment resisting AXL population compared to the MITF population as a proxy for treatment effect and aimed to detect the difference in connectivity within the tumor following such treatment. The malignant cells were labeled as either MITF or AXL cells based on the gene set score for the MITF- or AXL-associated genes (Tirosh *et al.* Supplementary Tables S7 and S8). Most tumors had malignant cells of both AXL and MITF types, which is in concordance with Tirosh *et al.* findings (Fig. 3a). Using mean gene expression as gene call, scConnect generated a graph with 19 598 interactions between 8 cell types.

Both AXL and MITF received strong interactions mediated by the chemokine CXCL12 → CXCR4 from CAF (3.11 and 2.73, respectively), macrophages (2.75 and 2.37, respectively) and endothelial cells (2.67 and 2.29, respectively). AXL and MITF were also predicted to make contacts with B-cells (2.54, 2.06), T-cells (2.45, 1.98) and NK cells (2.44, 1.97) through CXCL12 → CXCR4 (Fig. 3b). Previous studies have demonstrated the importance of the CXCL12 → CXCR4 interaction in malignant melanoma growth and metastasis (André *et al.*, 2016; Balkwill, 2004; McConnell *et al.*, 2016; Murakami *et al.*, 2002).

After treatment of a tumor, interactions specific to the AXL cell type would be more prominent in the tissue, as they are treatment immune. To identify the interactions where AXL and MITF cell types differed the most, scConnect was used to calculate a ratio of the interaction scores that AXL and MITF made with other cell populations. We analyzed all interactions where either AXL or MITF cells displayed at least five times higher interaction score (Fig. 3c and d). Of the incoming interactions (Fig. 3c) CCL17 and CCL22 to CCR4 interactions were strongest to the AXL population and these ligands have been shown to increase metastasis formation of CCR4 expressing melanoma cells (Klein *et al.*, 2017). The CD30 ligand to CD30 interaction was also higher to the AXL population, which is interesting as CD30 has been shown to be up-regulated in T-cells following malignant melanoma treatment (Gill *et al.*, 2014). The neurotensin to neurotensin receptor 2 interaction is stronger to the MITF cells, and the major source of neurotensin is endothelial cells and AXL cells. Previous studies has demonstrated the role of neurotensin receptor 1 in malignant melanoma (Zhang *et al.*, 2014), but neurotensin receptor 2 has to our knowledge not been implicated. As AXL produce the ligand, this interaction might play a greater role in treated melanoma tumors. MITF also received stronger stem cell factor to KIT proto-oncogene receptor tyrosine kinase interactions. This receptor has been shown to be down-regulated with tumor progression (Montone *et al.*, 1997; Natali *et al.*, 1992). This result fits well with the notion that AXL are more malignant than MITF.

For the outgoing interactions from AXL and MITF, the chemokine CCL13 to CXCR3, CCR2, CCR1 and CCR5 interactions were higher from MITF than AXL (Fig. 3d). We did not find much evidence for this interaction to play a critical role in malignant melanoma pathology, but CCL13 was not expressed by AXL cells, which might warrant further investigation. AXL was shown to express elevated levels of chemokines CCL21 and CCL19, and both AXL and MITF expressed the CCR7 receptor (Fig. 3d). CCR7 is a receptor used by immune cells to detect lymph nodes, which produce ligands for this receptor (Willimann *et al.*, 1998). One such ligand, CCL21, has shown to attract CCR7-positive melanoma cells to lymph nodes (Takeuchi *et al.*, 2004), similarly CCL19 levels increase in blood plasma with melanoma progression, and results in increased levels of CCR7+ CD56+ NK cells in the blood (Cristiani *et al.*, 2019). AXL also produce the chemokine CCL17 which facilitate autocrine signaling via CCR4. This suggests that AXL population could promote motility in both AXL and MITF melanoma cells via CCL19 and CCL21 to CCR7 and CCL17 to CCR4. AXL had higher expression of neuropeptide Y (NPY) than MITF (Fig. 3d) and expression of NPY in melanoma has been associated with nodular melanomas with vertical growth, metastasis and decreased lymphocyte infiltration (Gilaberte *et al.*, 2012). Note however, that the same group later linked NPY expression to good prognosis outcome (Pérez Tato *et al.*, 2017), indicating a complex role of NPY in melanoma. AXL was also found to make interactions using the molecular ligand adrenaline to β2 and α1A adrenoreceptors (Fig. 3d). Studies have demonstrated a key role of the autonomic nervous system in activating certain cancers, such as hepatocellular carcinoma though adrenaline to β2 and α1A adrenoreceptors (Li *et al.*, 2014; Zhang *et al.*, 2017) and although this mechanism has not been established in melanoma, high expression of β2 adrenoreceptor in melanoma tumors is related to poor prognosis (Shimizu *et al.*, 2016). Here, the production of adrenaline by AXL, might bypass the need for innervation of the nervous system to utilize the same mechanism. AXL produced the granulocyte-macrophage colony-stimulating factor (GM-CSF) which is usually found in GM-CSF secreting tumors (Aliper *et al.*, 2014). Issues can arise as GM-CSF is used as an adjuvant in some melanoma treatments, and the GM-CSF secreting tumors can get an accelerated progression (Aliper *et al.*, 2014). Other outgoing interactions from AXL that might warrant further investigation were TL6 to glucocorticoid-induced TNF receptor and NGF to nerve growth factor receptor, which has both been the target of interest in melanoma and other cancer fields (Estrela *et al.*, 2019; Kasemeier-Kulesa *et al.*, 2018; Ramirez-Montagut *et al.*, 2006; Zhu *et al.*, 2020).

In conclusion, these results predict that the AXL cells were more involved in chemokine signaling than the MITF cells. MITF cells also displayed receptors for many of these chemokines, but the AXL cells could produce the ligands to these receptors and possibly induce motility in all melanoma cells. Moreover, no manual selection of interactions was performed here, and yet, almost all interactions detected had previously been shown to be relevant to the melanoma field. This indicates that the interactions detected by scConnect are biologically relevant, and that inferred interactions could be a sound basis to build a novel hypothesis on.

3.3 Comparison with similar tools

The most similar tool to scConnect is cellPhoneDB (Efremova *et al.*, 2020) and the main difference is that scConnect detects molecular

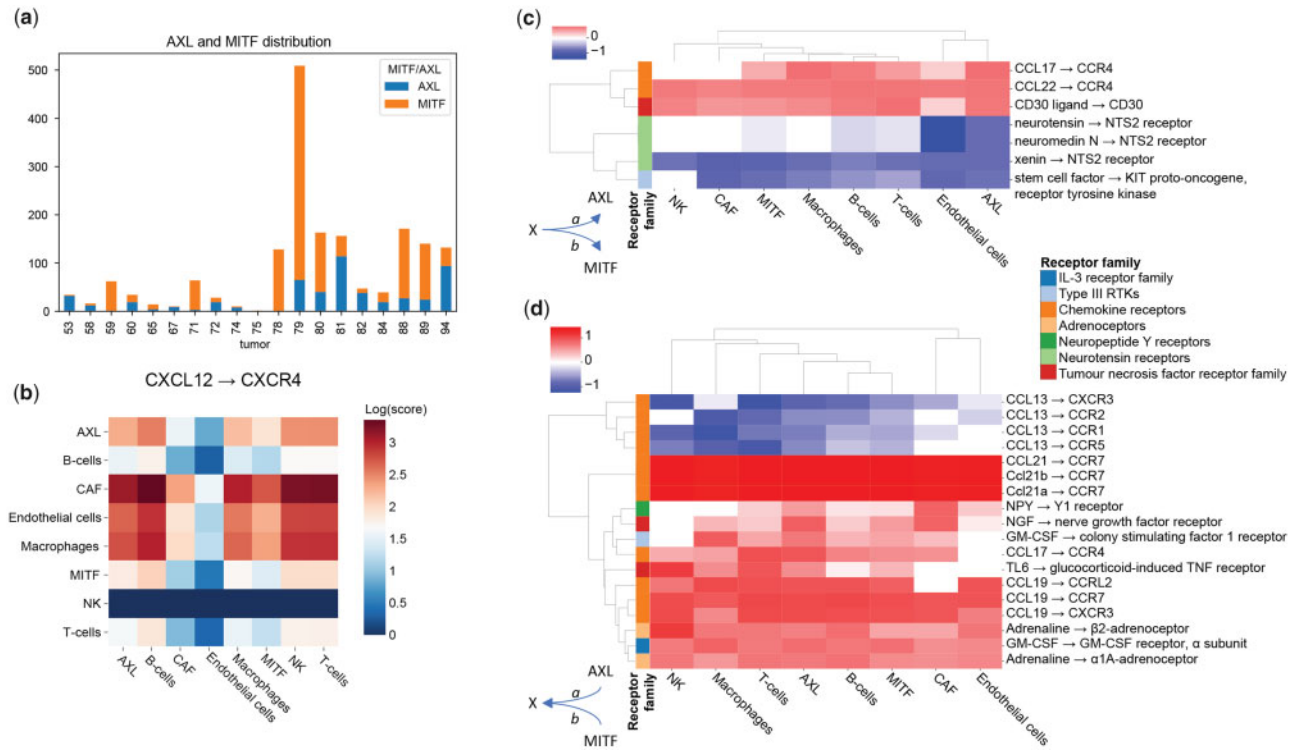

Fig. 3. Difference in interactions between AXL and MITF melanoma cell types. (**a**) Distribution of AXL and MITF cells in individual tumors. (**b**) Adjacency matrix of the strongest interaction (CXCL12 to CXCR4) in melanoma tumors. (**c**) Interaction ratio between AXL and MITF populations for incoming interactions. (**d**) Interaction ratio between AXL and MITF populations for outgoing interactions. Scale of heatmap in (c) and (d) is in log10(ratio)

ligands, which cellPhoneDB does not. ScConnect and cellPhoneDB detect interactions in a similar fashion. CellPhoneDB calculates interaction scores as the arithmetic-mean of LR expression if both LR expressions are >0, otherwise interaction score is 0. Here, scConnect use geometric-mean that naturally approaches 0 if any of the two factors approaches 0. CellPhoneDB calculates *P*-values for interaction scores compared to a random shuffle of cell type labels. Similarly, scConnect calculates *P*-values for LRs compared to a random shuffle of cell type labels and furthermore calculates a corrected *P*-value to deal with false discovery rates due to multiple testing. In scConnect, interaction specificity is instead inferred from the LR *P*-values. Using the melanoma dataset used previously (Tirosh *et al.*, 2016), but with all the malignant cells removed, scConnect and cellPhoneDB was used to detect specific interactions. Investigating the specific interactions that both cellPhoneDB and scConnect detected (*P*-value <0.05 for cellPhoneDB and specificity >1.3 for scConnect) (Supplementary Fig. S4a) showed that the same interactions had correlated specificity and interaction score, but scConnect had a greater dynamic range of specificity score. CellPhoneDB had on average higher interaction score than scConnect but there was a strong correlation between these scores (Supplementary Fig. S4b). For some interactions, cellPhoneDB gave an interaction score of 0, even though scConnect gave these interactions a score above 0. This is most likely due to the cell percentage filtering employed by cellPhoneDB where a set % of all cells in a cell type must express the receptor or ligand to be included (by default 10%). Furthermore, the higher interaction scores of cellPhoneDB were most likely explained by the difference in interaction score calculation (arithmetic-mean versus geometric-mean).

3.4 Preprocessing consideration

Preprocessing of scRNA-seq data can differ greatly between researchers and experiments. To test the robustness of the results, different preprocessing pipelines were used to produce an interaction graph of the melanoma dataset used previously (Supplementary Fig. S5). Six datasets were compared: original, normalized (cell size normalization), gene length corrected (RPKM), normalized and RPKM, subsample of 90% of the cells and finally subsample of 50% of the cells. The interaction strength was lower in the RPKM datasets, but this decrease was uniform over all interactions (Supplementary Fig. S5a). Specific interactions were not affected by normalization or gene length correction, although individual interaction specificity measures will differ slightly between runs as the permutation is random (Supplementary Fig. S5b). To evaluate how an interaction was affected, we compared the adjacency matrix of the CXC12β to CXCR4 interaction from the different datasets (Supplementary Fig. S6c–h). RPKM datasets showed the same pattern as the original dataset, but with lower interaction scores (Supplementary Fig. S5e and f). 90% subsampling provided results similar to the original (Supplementary Fig. S5g). About 50% subsampling failed to detect CXCL12β in B-cells, and hence lost this interaction totally but was otherwise similar to the original (Supplementary Fig. S5h). These results demonstrate that scConnect is robust to different preprocessing pipelines.

4 Discussion

In this article, we introduced the scConnect method for detection of interactions between cell types in scRNA-seq datasets and demonstrated that it can be used to investigate specific interactions, such as the dopaminergic system in the mouse brain, and to identify novel interactions in a human melanoma dataset. The method also facilitated hypothesis generation, which could be probed further with *in vitro* and *in vivo* studies.

Frameworks dedicated for the analysis of scRNA-seq data have been developed for both python [SCANPY (Wolf *et al.*, 2018)] and R [Seurat (Butler *et al.*, 2018)], and they provide functions dealing with data structure, preprocessing, clustering, trajectory inference and data integration. As scConnect utilizes the AnnData objects to store dataset-related information, such as LR expression it is fully compatible with Scanpy. This makes it easy to implement LR annotation in a Scanpy scRNA-seq pipeline and is especially useful to assess expression of molecular ligands as these are not a one to one map to gene expression. Furthermore, the connectome produced by

scConnect is stored in a multi-directional graph using networkX (Hagberg *et al.*, 2008), which can be exported in several graph formats to be analyzed in third party software, such as Cytoscape (Shannon *et al.*, 2003) and Gephi (Bastian *et al.*, 2009).

Single-cell transcription data is inherently noisy, in part due to technically induced dropouts (Kim *et al.*, 2015), and naturally occurring transcriptional bursting (Larsson *et al.*, 2019). Furthermore, inferring receptor and ligand expression on a protein/molecule level extrapolated from gene expression, introduces significant challenges, as the mRNA–protein correlation in single cells is weak (Darmanis *et al.*, 2016; Maier *et al.*, 2009). Noise reduction and improvement of gene–protein correlation can be achieved by analyzing gene expression in groups of cells, and that is why gene calling and downstream LR calculations are performed over cell types. In this implementation, we support gene calling using mean expression. Other gene calling algorithms, such as median expression, percent expression or betabinomial trinarization (Zeisel *et al.*, 2018) could be assessed and implemented in future versions.

As scConnect utilize the read/count matrix, it is important to first correct for technical biases, such as size factor, gene length and batches, as any bias in the input data will affect the constructed graph. For a review of the current analysis consensus, see Luecken and Theis (2019). For instance, failure to correct for gene length could systematically skew the interaction scores for interactions utilizing receptors or ligands with particularly long or short gene length.

There is no guarantee that an interaction involving proteins in one species holds true for the ortholog proteins in another species, however, it has been shown that ortholog proteins often retain their function and preserve critical amino acids used for interactions (Chen and Jeong, 2000; Zhou *et al.*, 2014). Genome duplication, resulting in the formation of paralog genes can make it difficult to identify ortholog genes between species, as some paralog genes will change function. This can make orthologous search between mammalians and fish particularly difficult as fish has undergone genome duplication since our last shared ancestor (Meyer and Van De Peer, 2005), thereby decreasing our confidence in the ortholog assignment. We can clearly see this in the databases that we have established, as we have comprehensive coverage of LRs for mammalians, but not for fish, and invertebrates (Supplementary Table S2). Another issue that arises from including interaction detected in several species is that some duplications can occur. For instance, Ccl21a → CCR7 and Ccl21b → CCR7 are interactions inferred from mouse, and CCL21 → CCR7 is inferred from humans. All these interactions were detected in the melanoma dataset as strong outgoing interactions from the AXL population (Fig. 3d).

CellPhoneDB utilizes both IMEx and GTF to retrieve interaction annotations and furthermore provide manually curated interactions between heteromeric complexes, making it possible to detect connections between heteromeric receptors and ligands (Efremova *et al.*, 2020). CellPhoneDB created a strong and robust database for interaction of protein–protein interactions, however, it lacks methods to link gene expression to expression of molecular ligands and does not include these in their resulting connectome. scConnect is hence specifically suited for analysis of neuronal datasets, where the classical neurotransmitters are molecular ligands.

Acknowledgements

We would like to thank Hannah Weman and Jörgen Jonsson for kindly reviewing the manuscript. We also want thank Anna Johansson, Markus Ringnér and Åsa Björklund from the National Bioinformatics Infrastructure Sweden at SciLifeLab for valuable feedback on the method.

Funding

This work was supported by grants from the Swedish Research Council [2016-00851]; Uppsala University; and The Brain foundation. M.C.L. is a Ragnar Söderberg Fellow in Medicine.

Conflict of Interest: none declared.

Data availability

Data underlying this work are available from Gene Expression Omnibus using GEO accession: GSE116470 and GSE72056. All scripts used for analysis are deposited at: https://github.com/JonETJakobsson/Connectivity-paper.

References

Aliper,A.M. *et al.* (2014) A role for G-CSF and GM-CSF in nonmyeloid cancers. *Cancer Med.*, **3**, 737–746.

André,N.D. *et al.* (2016) Knockdown of chemokine receptor CXCR4 gene by RNA interference: effects on the B16-F10 melanoma growth. *Oncol. Rep.*, **35**, 2419–2424.

Balkwill,F. (2004) Cancer and the chemokine network. *Nat. Rev. Cancer*, **4**, 540–550.

Bastian,M. *et al.* (2009) Gephi: an open source software for exploring and manipulating networks. In: *Third International AAAI Conference*. San Jose, CA (United States).

Bergson,C. *et al.* (1995) Regional, cellular, and subcellular variations in the distribution of D1 and D5 dopamine receptors in primate brain. *J. Neurosci.*, **15**, 7821–7836.

Berlanga,M.L. *et al.* (2005) Dopamine D5 receptor localization on cholinergic neurons of the rat forebrain and diencephalon: a potential neuroanatomical substrate involved in mediating dopaminergic influences on acetylcholine release. *J. Comp. Neurol.*, **492**, 34–49.

Bertran-Gonzalez,J. *et al.* (2008) Opposing patterns of signaling activation in dopamine D1 and D2 receptor-expressing striatal neurons in response to cocaine and haloperidol. *J. Neurosci.*, **28**, 5671–5685.

Breuer,K. *et al.* (2013) InnateDB: systems biology of innate immunity and beyond–recent updates and continuing curation. *Nucleic Acids Res.*, **41**, D1228–D1233.

Brown,K.R. and Jurisica,I. (2005) Online predicted human interaction database. *Bioinformatics*, **21**, 2076–2082.

Butler,A. *et al.* (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.*, **36**, 411–420.

Camp,J.G. *et al.* (2017) Multilineage communication regulates human liver bud development from pluripotency. *Nature*, **546**, 533–538.

Castello,J. *et al.* (2020) The Dopamine D5 receptor contributes to activation of cholinergic interneurons during L-DOPA induced dyskinesia. *Sci. Rep.*, **10**, 2542.

Chen,R. and Jeong,S.S. (2000) Functional prediction: identification of protein orthologs and paralogs. *Protein Sci.*, **9**, 2344–2353.

Choi,H. *et al.* (2015) Transcriptome analysis of individual stromal cell populations identifies stroma-tumor crosstalk in mouse lung cancer model. *Cell Rep.*, **10**, 1187–1201.

Cristiani,C.M. *et al.* (2019) Accumulation of circulating CCR7+ natural killer cells marks melanoma evolution and reveals a CCL19-dependent metastatic pathway. *Cancer Immunol. Res.*, **7**, 841–852.

Darmanis,S. *et al.* (2016) Simultaneous multiplexed measurement of RNA and proteins in single cells. *Cell Rep.*, **14**, 380–389.

Efremova,M. *et al.* (2020) CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.*, **15**, 1484–1506.

Estrela,J.M. *et al.* (2019) Glucocorticoid receptor antagonism overcomes resistance to BRAF inhibition in BRAFV600E-mutated metastatic melanoma. *Am. J. Cancer Res.*, **9**, 2580–2598.

Gilaberte,Y. *et al.* (2012) Neuropeptide Y expression in cutaneous melanoma. *J. Am. Acad. Dermatol.*, **66**, e201–e208.

Gill,K. *et al.* (2014) CD30-positive lymphoproliferative disorders arising after regional therapy for recurrent melanoma: a report of two cases and analysis of CD30 expression. *J. Surg. Oncol.*, **110**, 258–264.

Grimm,J. *et al.* (2004) Molecular basis for catecholaminergic neuron diversity. *Proc. Natl. Acad. Sci. USA*, **101**, 13891–13896.

Hagberg,A. *et al.* (2008) Exploring network structure, dynamics, and function using networkx (Conference). Los Alamos, NM (United States).

Harding,S.D. *et al.*; NC-IUPHAR. (2018) The IUPHAR/BPS Guide to PHARMACOLOGY in 2018: updates and expansion to encompass the new guide to IMMUNOPHARMACOLOGY. *Nucleic Acids Res.*, **46**, D1091–D1106.

Hernández-Flores,T. *et al.* (2015) Modulation of direct pathway striatal projection neurons by muscarinic M_4-type receptors. *Neuropharmacology*, **89**, 232–244.

Kaneko,T. *et al.* (1993) Substance P receptor-immunoreactive neurons in the rat neostriatum are segregated into somatostatinergic and cholinergic aspiny neurons. *Brain Res.*, **631**, 297–303.

Kasemeier-Kulesa,J.C. *et al.* (2018) NGF reprograms metastatic melanoma to a bipotent glial-melanocyte neural crest-like precursor. *Biol. Open*, **7**, bio030817.
Kim,J.K. *et al.* (2015) Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. *Nat. Commun.*, **6**, 8687.
Klein,A. *et al.* (2017) CCR4 is a determinant of melanoma brain metastasis. *Oncotarget*, **8**, 31079–31091.
Kumar,M.P. *et al.* (2018) Analysis of single-cell RNA-Seq identifies cell-cell communication associated with tumor characteristics. *Cell Rep.*, **25**, 1458–1468.e4.
Larsson,A.J.M. *et al.* (2019) Genomic encoding of transcriptional burst kinetics. *Nature*, **565**, 251–254.
Li,J. *et al.* (2014) Monoamine oxidase A suppresses hepatocellular carcinoma metastasis by inhibiting the adrenergic system and its transactivation of EGFR signaling. *J. Hepatol.*, **60**, 1225–1234.
Luecken,M.D. and Theis,F.J. (2019) Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.*, **15**, e8746.
Maier,T. *et al.* (2009) Correlation of mRNA and protein in complex biological samples. *FEBS Lett.*, **583**, 3966–3973.
McConnell,A.T. *et al.* (2016) The prognostic significance and impact of the CXCR4-CXCR7-CXCL12 axis in primary cutaneous melanoma. *Br. J. Dermatol.*, **175**, 1210–1220.
Meyer,A. and Van De Peer,Y. (2005) From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *Bioessays*, **27**, 937–945.
Montone,K.T. *et al.* (1997) Proto-oncogene c-kit expression in malignant melanoma: protein loss with tumor progression. *Mod. Pathol.*, **10**, 939–944.
Murakami,T. *et al.* (2002) Expression of CXC chemokine receptor-4 enhances the pulmonary metastatic potential of murine B16 melanoma cells. *Cancer Res.*, **62**, 7328–7334.
Natali,P.G. *et al.* (1992) Progression of human cutaneous melanoma is associated with loss of expression of c-kit proto-oncogene receptor. *Int. J. Cancer*, **52**, 197–201.
Orchard,S. *et al.* (2012) Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. *Nat. Methods*, **9**, 345–350.
Pérez Tato,B. *et al.* (2017) Neuropeptide Y expression in primary cutaneous melanoma. *J. Eur. Acad. Dermatol. Venereol.*, **31**, 443–449.
Phipson,B. and Smyth,G.K. (2010) Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn. *Stat. Appl. Genet. Mol. Biol.*, **9**, Article39.
Ramirez-Montagut,T. *et al.* (2006) Glucocorticoid-induced TNF receptor family related gene activation overcomes tolerance/ignorance to melanoma differentiation antigens and enhances antitumor immunity. *J. Immunol.*, **176**, 6434–6442.
Raredon,M.S.B. *et al.* (2019) Single-cell connectomic analysis of adult mammalian lungs. *Sci. Adv.*, **5**, eaaw3851.
Raudvere,U. *et al.* (2019) g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.*, **47**, W191–W198.
Reiner,A. and Anderson,K.D. (1990) The patterns of neurotransmitter and neuropeptide co-occurrence among striatal projection neurons: conclusions based on recent findings. *Brain Res. Brain Res. Rev.*, **15**, 251–265.
Rommelfanger,K.S. and Wichmann,T. (2010) Extrastriatal dopaminergic circuits of the Basal Ganglia. *Front. Neuroanat.*, **4**, 139.
Salwinski,L. *et al.* (2004) The database of interacting proteins: 2004 update. *Nucleic Acids Res.*, **32**, D449–D451.
Saunders,A. *et al.* (2018) Molecular diversity and specializations among the cells of the adult mouse brain. *Cell*, **174**, 1015–1030.e16.
Seabold,S. and Perktold,J. (2010) Statsmodels: econometric and statistical modeling with Python. In: *Proceedings of the 9th Python in Science Conference*. Austin, TX (United States). SciPy, pp. 92–96.
Shannon,P. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.
Shimizu,A. *et al.* (2016) Prognostic significance of β2-adrenergic receptor expression in malignant melanoma. *Tumour Biol.*, **37**, 5971–5978.
Skelly,D.A. *et al.* (2018) Single-Cell Transcriptional Profiling Reveals Cellular Diversity and Intercommunication in the Mouse Heart. *Cell Rep.*, **22**, 600–610.
Takeuchi,H. *et al.* (2004) CCL21 chemokine regulates chemokine receptor CCR7 bearing malignant melanoma cells. *Clin. Cancer Res.*, **10**, 2351–2358.
Tirosh,I. *et al.* (2016) Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*, **352**, 189–196.
Williams,E.A. *et al.* (2017) Synaptic and peptidergic connectome of a neurosecretory center in the annelid brain. *Elife*, **6**,
Willimann,K. *et al.* (1998) The chemokine SLC is expressed in T cell areas of lymph nodes and mucosal lymphoid tissues and attracts activated T cells via CCR7. *Eur. J. Immunol.*, **28**, 2025–2034.
Wolf,F.A. *et al.* (2018) SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.*, **19**, 15.
Zeisel,A. *et al.* (2018) Molecular architecture of the mouse nervous system. *Cell*, **174**, 999–1014.e22.
Zhang,L. *et al.* (2017) Sympathetic and parasympathetic innervation in hepatocellular carcinoma. *Neoplasma*, **64**, 840–846.
Zhang,Y. *et al.* (2014) Neurotensin receptor1 antagonist SR48692 reduces proliferation by inducing apoptosis and cell cycle arrest in melanoma cells. *Mol. Cell. Biochem.*, **389**, 1–8.
Zhou,J.X. *et al.* (2017) Extracting intercellular signaling network of cancer tissues using ligand-receptor expression patterns from whole-tumor and single-cell transcriptomes. *Sci. Rep.*, **7**, 8815.
Zhou,Q. *et al.* (2014) Stereoselectivity of chiral drug transport: a focus on enantiomer-transporter interaction. *Drug Metab. Rev.*, **46**, 283–290.
Zhu,M.M.T. *et al.* (2020) Evaluation of glucocorticoid-induced TNF receptor (GITR) expression in breast cancer and across multiple tumor types. *Mod. Pathol.*, **33**, 1753–1763.