Stem Cell Research 17 (2016) 406–412

Contents lists available at ScienceDirect

Stem Cell Research

journal homepage: www.elsevier.com/locate/scr

Short report

A systems biology approach to identify niche determinants of cellular phenotypes☆

Srikanth Ravichandran, Satoshi Okawa, Susana Martínez Arbas, Antonio del Sol *

Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 7, Avenue des Hauts-Fourneaux, L-4362 Esch-sur-Alzette, Luxembourg

ARTICLE INFO

Article history:
Received 6 July 2016
Received in revised form 30 August 2016
Accepted 14 September 2016
Available online 15 September 2016

Keywords:
Stem cell niche
Signaling pathways
Transcriptional regulatory networks
Niche determinants

ABSTRACT

Recent reports indicate a dominant role for cellular microenvironment or niche for stably maintaining cellular phenotypic states. Identification of key niche mediated signaling that maintains stem cells in specific phenotypic states remains a challenge, mainly due to the complex and dynamic nature of stem cell-niche interactions. In order to overcome this, we consider that stem cells maintain their phenotypic state by experiencing a constant effect created by the niche by integrating its signals via signaling pathways. Such a constant niche effect should induce sustained activation/inhibition of specific stem cell signaling pathways that controls the gene regulatory program defining the cellular phenotypic state. Based on this view, we propose a computational approach to identify the most likely receptor mediated signaling responsible for transmitting niche signals to the transcriptional regulatory network that maintain cell-specific gene expression patterns, termed as niche determinants. We demonstrate the utility of our method in different stem cell systems by identifying several known and novel niche determinants. Given the key role of niche in several degenerative diseases, identification of niche determinants can aid in developing strategies for potential applications in regenerative medicine.

1. Introduction

Stem cells play a crucial role in maintaining tissue homeostasis. In this context, interplay between stem cells and their microenvironment, also known as the niche, is necessary for the maintenance of stem cell phenotypic state (Scadden, 2014). In general, distinct stem cell populations are influenced by specific niche elements with defined localization and composition of supportive stromal cells, extracellular matrix, gap junctions, and soluble factors (Scadden, 2014; Xin et al., 2016). The cellular niche acts as the intermediate that senses, integrates and translates information from the neighborhood of the stem cell by transfer of external cues into intracellular signaling events to determine its cellular state. The effect of niche on stem cell phenotype is shaped by physiological contexts determined by the tissue and organismal requirements (Xin et al., 2016). Further, the tight regulation of stem cell and progenitor cell turnover that contributes to tissue homeostasis characterizes healthy tissues. However, this fine balance is often found perturbed in cases of several degenerative diseases of liver, heart, lung and brain; where regenerative medicine hold immense promise (Lane et al., 2014). The key idea is targeted activation of endogenous stem cells to repair the damaged tissues. Currently, a major challenge in this area includes the limited functional integration (or engraftment) of transplanted stem cells into the target tissue. This could possibly be due to the negative regulatory effect of diseased niche on transplanted stem cells (Lane et al., 2014). In order to overcome this, an understanding of those signaling pathways or niche determinants that normally control stem cell functional state in response to the niche is essential. However, due to the inherent complexity and dynamic nature of stem cell-niche interactions it is often difficult to characterize them by experiments alone. Computational systems biology approaches to model stem cell-niche interactions could be very helpful in this regard. In fact, few computational methods to model population level dynamics of stem cell-niche interactions have been used to untangle their regulatory relationship at the level of cell-cell interactions (Lei et al., 2014; Szekely et al., 2014). Even though such models are useful, they encounter a number of limitations due to: incomplete characterization of the niche, fluctuations of the niche components, and a large number of non-linear interactions between the niche components and stem cells.

Here, we present a computational method to identify the niche determinants of stem cell phenotypic state. However, instead of attempting to model the stem cell-niche interactions (explicitly) in its entire complexity, we consider an alternative view by considering that each stem cell interacts with the niche via a constant effect created by all molecular and cellular signals coming from the niche. Such a constant effect of the niche on stem cell should induce sustained activation/inhibition of specific stem cell signaling pathways for maintaining

☆ Author contributions: A.d.S. conceived and initiated the study. S.R. and S.M.A. collected data and performed initial analysis. S.R. and S.O. implemented the method. All authors interpreted the results and contributed for the drafting of the manuscript.

* Corresponding author.
E-mail address: antonio.delsol@uni.lu (A. del Sol).

http://dx.doi.org/10.1016/j.scr.2016.09.006

their phenotypes. This view of stem cell-niche interactions shifts the focus of the problem towards the identification of signaling pathways constantly activated/inhibited due to the niche influence instead of accounting for niche composition and its interaction with the stem cell explicitly. The rationale behind our method is that the niche maintains stem cells in a stable state in the gene expression landscape by a sustained effect on their transcriptional regulatory network (TRN) despite niche dynamics and cellular heterogeneity. Such a sustained effect of the niche is transmitted through constantly activated/inhibited intracellular signaling pathways compatible with the phenotype specific TRN. Indeed, recent studies in different systems have demonstrated that stem cell state is determined by constant activation/inhibition of specific pathways by the constitutive influence of its niche (Codega et al., 2014; Kunisaki et al., 2013; Rompolas et al., 2013). Our method relies on gene expression differences between stem cells displaying different niche-dependent phenotypes, and aims to infer and rank signaling pathways that are sustained and differentially (in)active for stably maintaining their corresponding phenotypes.

In this regard, several existing computational pathway analysis methods can potentially identify deregulated signaling pathways from high-throughput datasets (Hung, 2013; Khatri et al., 2012). In particular, due to huge progress in microarray and sequencing techniques, computational methods tend to rely more on transcriptomic datasets to identify signaling pathways that differ between two conditions. However, these methods are generally not specific for identification of differentially active and sustained signaling that maintains the different phenotypes, but are more generic in nature to capture all possible deregulated signaling events that could also be transient in nature. On the contrary, our method is more specific, since it aims to identify only constantly activated/inhibited signaling pathways that are responsible for maintaining the phenotype-specific TRN state, and whose perturbations can destabilize this state. Indeed, sustained activation/inhibition of signaling pathways are shown to exhibit a clear influence on the expression of genes participating in such pathways, which is not always observed in cases of transient activation/inhibition (Whitmarsh, 2007). For example, gene expression signatures have been successfully employed to identify constantly active oncogenic addiction pathways (Bild et al., 2006). Furthermore, signaling pathways involved in differentiation and cellular growth are known to induce changes in expression of genes involved in the signal transduction (Codega et al., 2014; Zhu et al., 2006).

We applied our method to different stem cell systems, including neuronal, hair follicle, muscle and hematopoietic in order to demonstrate its utility in identifying niche determinants. Importantly, many of our top ranking predictions have been experimentally found to play a key role in niche mediated regulation of specific stem cell phenotypes. The proposed method is general in the sense that it can be applied to any stem cell system with existing transcriptome data corresponding to distinct niche mediated-phenotypes. Furthermore, our method overcomes the difficulty of explicitly characterizing niche components in order to describe their collective effect on maintenance of stem cell phenotypic states. The knowledge of niche determinants will be helpful for targeted manipulation of the niche effect on stem cells to control their phenotypic states, and therefore can aid in development of novel therapeutics based on regenerative medicine.

2. Results

2.1. Method framework

A schematic representation of our method is shown in Fig. 1. Our method considers that niche maintains stem cells in a stable state by a sustained effect on their TRN via constantly activated/inhibited intracellular signaling pathways compatible with the phenotype-specific TRN state. The fact that the cells exhibit differences in their phenotype due to differential effect of the niche suggests that the intracellular signaling events controlling the specific GRN for maintain the specific phenotype are differentially active. To identify key differences at the level of signaling intermediates due to differential effect of niche, we first compiled a database-derived raw signaling interactome originating from differentially expressed receptors (DERs) to differentially expressed transcription factors (DETFs) specific for the cell type under study. In order to recapitulate the control of the niche over the stem cell state, we consider that the upregulated receptors of a given stem cell phenotype are under the direct influence of the niche. Since the exact mechanisms of the niche effect are often not well known, we represent it by introducing a dummy niche node in the raw signaling network. This dummy node is then connected to all upregulated receptors for each phenotype under consideration (Fig. 1). Therefore, signal transduction from the niche to DETFs must be propagated through at least one of the upregulated receptors. Such a representation of unknown external influence by a dummy node has been applied earlier (Tuncbag et al., 2013). In our representation, we have two dummy niche nodes, each representing a specific niche condition (Fig. 1).

The edges in the signaling interactome were weighted using the gene expression data. This weighting scheme was implemented to maximize the compatibility between the expression data and interaction sign. By compatible, we mean consistency between the sign of the interactions (i.e. positive when activating and negative when inhibiting) and the effect (i.e. activation or inhibition) that the receptor has on its downstream target transcription factors (TFs). For example, sign of a signaling path from a receptor to a TF that is up regulated or overexpressed must be positive (activation), while it must be negative (inhibition) for down regulated or under-expressed TF. We calculated the differential edge weight such that for a given phenotype it reflected the probability of the target gene of the specific interaction to be relatively more active when compared to the other phenotype by considering the interaction sign and booleanized expression state of the interacting nodes.

Next, in this weighted raw signaling interactome we aimed to identify signaling paths that are regulated by the niche, and are potentially responsible for the observed expression pattern of the DETFs. For this, we employed Prize Collecting Steiner Tree (PCST) formalism to infer sub-networks with the dummy niche node as the root or origin and the DETFs as the terminal nodes. Steiner Tree formalism have been used earlier to reconstruct active signaling pathways (Bailly-Bechet et al., 2011). Since the dummy node is connected only to the upregulated receptors of a given cell type, the inferred sub-networks will encompass only those receptors that are both topologically favorable and compatible in the expression state to link the DETFs. Therefore, from several upregulated receptors, one could narrow down to the few linking the DETFs based on their unique network topological features. Such sub-networks (Steiner trees) are inferred for the two different phenotypes under consideration (Fig. 1). Further, the statistical significance of the receptors in the inferred sub-networks were assessed by randomizing (100 iterations) the edge weights of the raw signaling network.

In order to test if the signaling events initiated by the receptors shortlisted as significant are indeed sustained in their effect on the underlying transcriptional states, we performed in silico perturbation experiments of the specific receptors in a Boolean model of the TRN, and quantified the perturbation effect on the underlying attractor state. Further, this served as a ranking of the inferred signaling pathways (Methods).

In order to test the applicability of our method, we specifically chose systems where gene expression data for two distinct stem cell phenotypes from in vivo model systems were available. In particular, we focused on systems where the niche played an important role in maintaining the specific cellular phenotypic state. These include comparisons of inactive versus injury induced active stem cells, stem cells versus their differentiated progeny, quiescent versus active stem cells depending on their spatial location. Our computational predictions

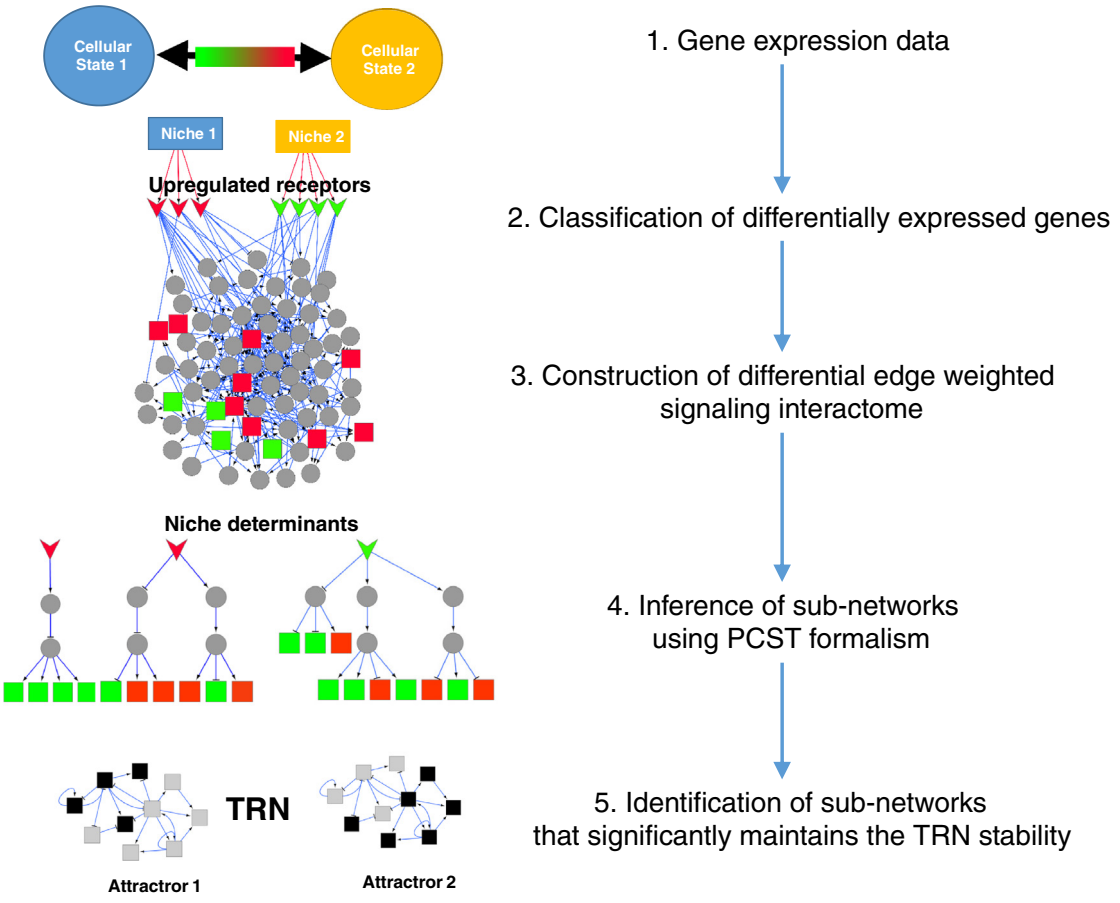

Fig. 1. Schematic of the computational method. Gene expression data of stem cells in two distinct phenotypic states is the input for the method. Then, differentially expressed genes (represented by red and green nodes) are identified, and are classified as receptors and transcription factors. In step 3, a signaling interactome is constructed by connecting the up regulated receptors (inverted triangles) to the DETFs (squares) from prior knowledge networks based on literature support. The edges of the signaling interactome are differentially weighted based on the gene expression data. Then, the receptors upregulated for each phenotype under consideration are linked to dummy niche nodes that represent the external influence of the niche. In step 4, sub-networks that link the upregulated receptors to the DETFs for each phenotype are extracted using a PCST formalism. Finally, in order to identify those sub-networks that are necessary for the maintenance of TRN stability, Boolean simulations of the combined signaling–TRN model are performed to assess the extent of change in the phenotype specific attractor state.

Table 1
Cellular phenotype comparisons and predicted niche determinants.

Stem cell system	Cellular phenotype comparison	Niche determinants		*p* value	
		Phenotype 1	*Phenotype 2*	*Phenotype 1*	*Phenotype 2*
Neuronal	Quiescent *vs* Active NSCs	*Quiescent*	*Active*	*Quiescent*	*Active*
		Cav1	**Egfr**	1.96E−25	4.24E−61
		Cd40	Trim27	2.01E−11	1.19E−16
		S1pr1		1.71E−02	
Neuronal oligodendrocyte precursors	Inactive *vs* Active OPCs	*Inactive*	*Active*	*Inactive*	*Active*
		-	**Cd44**	-	1.01E−03
Neuronal	NSCs *vs* Immature neurons	*NSCs*	*Immature neurons*	*NSCs*	*Immature neurons*
		Igf2	-	4.87E−02	-
Hair follicle	Long term *vs* short term HFSCs	*Long-term*	*Short-term*	*Long-term*	*Short-term*
		Tnfrsf18	**Bmpr1b**	1.37E−09	2.11E−08
		Itgb6	Notch3	1.18E−08	8.24E−06
		Ace	**Wnt**[a]	1.44E−04	4.34E−02
		Fgf2		6.75E−04	
		Osmr		2.70E−03	
Hematopoietic	Long term *vs* short term HSCs	*Long-term*	*Short-term*	*Long-term*	*Short-term*
		Tgfbr1	Il17rb	7.38E−06	3.81E−10
			Epha4		7.80E−07
			CD40		1.92E−02
Skeletal	Quiescent *vs* Active SCs	*Quiescent*	*Active*	*Quiescent*	*Active*
		Notch1	Il1rl1	2.30E−08	3.10E−10
		Fgfr1	**Cd44**	4.36E−03	1.20E−04

The table lists the cellular phenotype comparisons made in this study with the corresponding predictions for each phenotype. The last two columns show the *p*-values denoting the significance of the predictions. Those predictions with experimental validations are highlighted in bold.

[a] Represent three receptors perturbed simultaneously for the assessment of significance (see text for details).

rank the receptors most crucial for influencing the DETFs, and does not attempt to rank the inferred signaling intermediates (other downstream molecules like associated kinases and phosphatases) or the entire pathway as a whole. The list of computational predictions and associated literature support is listed in Table 1. In the following sections, we describe our computational predictions of niche determinants for each stem cell system in detail.

2.2. Quiescent versus active neural stem cells

Recent reports suggest that quiescent and active neural stem cells (NSCs) are localized in distinct compartments of the sub ventricular zone of the CNS (Codega et al., 2014). Further, the molecular make-up of these cells was found to be distinct. We found more than 3000 genes to be differentially expressed between quiescent and active NSCs, including 234 receptors and 245 TFs. In the case of quiescent stem cells, the receptor molecule, Cav1 was ranked the highest by our method (Table 1). Out of 245 differentially expressed TFs, Cav1 was inferred to regulate the expression state of 53 TFs. Importantly, our method could identify signaling from S1pr1 receptor as significant (Table 1, Fig. 2A), whose role in maintaining quiescence of NSCs has been experimentally validated (Codega et al., 2014). Cav1 (caveolin-1), a lipid raft associated protein, is a novel prediction of our method for regulating quiescent neural stem cell state that requires experimental validation. However, Cav1 is known to play certain crucial role in regulating neuronal stem cell differentiation (Baker and Tuan, 2013).

In the case of active neural stem cells, our method identified Egfr as the top candidate and it regulated 116 TFs (Table 1). Further, Egfr signaling is well known to regulate active NSCs and is in fact used as a marker to identify active NSC population (Codega et al., 2014).

2.3. Inactive versus active oligodendrocyte progenitor cells

Oligodendrocyte progenitor cells (OPCs) are the primary source of myelinating oligodendrocytes, and are responsible for restoration of myelin sheaths upon a demyelinating injury. Normally, this is a highly efficient process; however, in chronic disease conditions, such as multiple sclerosis, this process is not efficient. This is in part attributed to diseased niche causing a differentiation block of OPCs (Franklin and Ffrench-Constant, 2008). Therefore, in this example, we specifically focused on identifying niche determinants of OPCs activated in response to injury. Here, we have analyzed gene expression of OPCs before (inactive) and after (active) cuprizone induced demyelination in mice model to identify the niche determinants of active OPCs (Moyon et al., 2015). We identified about 800 genes to be differentially expressed, including 56 TFs and 33 receptors. For active OPCs, our computational analysis identified Cd44 as significant for controlling the active state of OPCs (Table 1, Fig. 2B). Signaling from Cd44 was found to control 41 TFs. Our method predicted a strong role of Cd44 in regulating cellular phenotype of active OPCs. In fact, Cd44 mediated signaling in response to demyelinating injury is experimentally observed to be necessary for migration and subsequent differentiation of the activated OPCs (Piao et al., 2013). Further, it was observed that knock down of this receptor leads to severe impairment of OPC migration to demyelinated lesions, implicating a strong role for Cd44 mediating signaling for active OPC phenotype (Piao et al., 2013).

2.4. Neural stem cells versus immature neurons

In the hippocampal dentate gyrus (DG), NSCs are responsible for generating neurons. Niche mediated signals play a crucial role in controlling the proliferation and differentiation of NSCs, and their progenitors in DG. In this case, we have analyzed the gene expression of NSCs

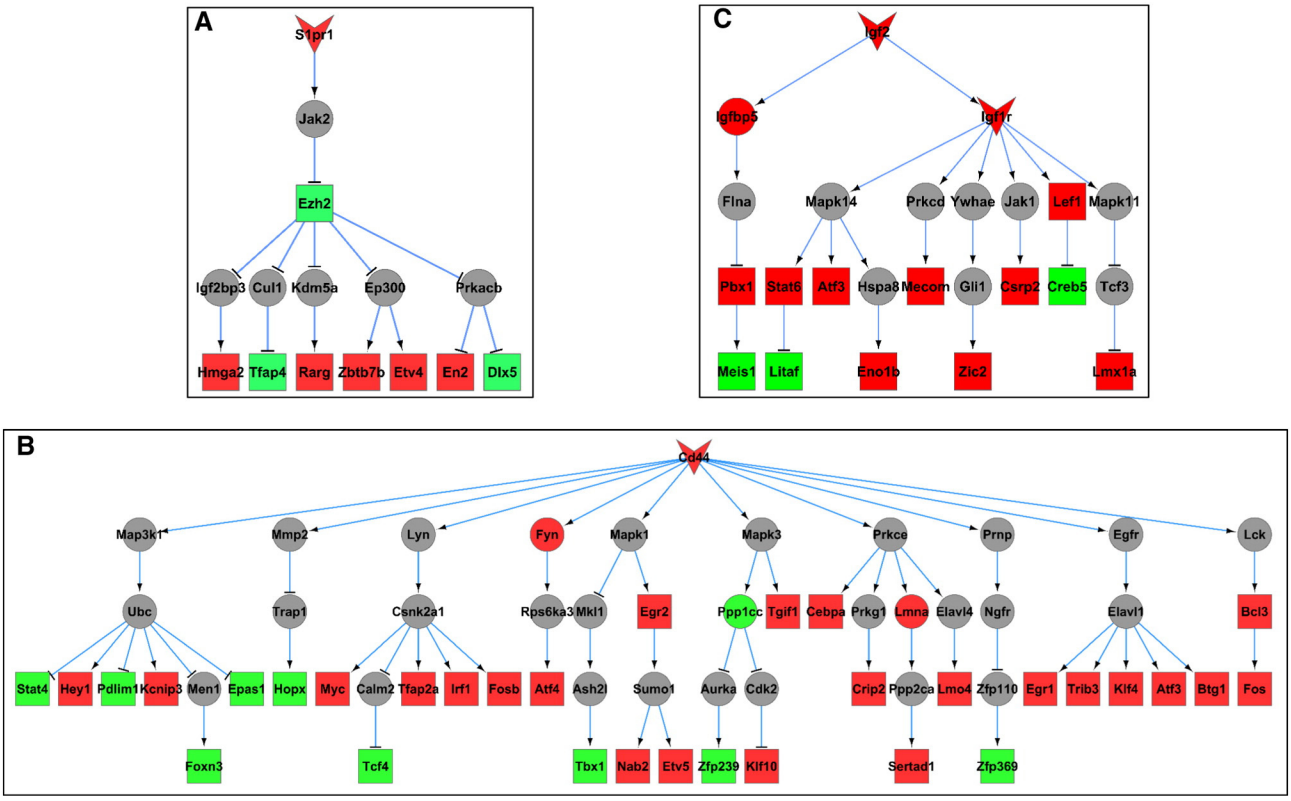

Fig. 2. Niche determinants of NSCs, OPCs and NSCs of DG. The figure shows the sub-networks of signaling pathway(s) identified by our method that are also experimentally validated. A) Quiescent vs Active NSCs. B) Inactive vs Active OPCs. C) NSCs vs immature neurons. Pointed arrows indicate activation and blunted arrows indicate inhibition. The node colors represent the expression status of the genes, red: upregulated and green: down regulated. The receptors molecules are represented as inverted triangle while the TFs are represented as squares.

and immature neurons (Bracko et al., 2012), and identified 1108 differentially expressed genes, including 66 receptors and 68 TFs. Igf2 was identified as the top ranking for NSC phenotype (Table 1, Fig. 2C). Igf2 mediated signaling was transmitted predominantly through Igf1r for regulating 12 TFs. The role of Igf2 in maintaining the NSC state in DG is well known and it is implicated in regulating the proliferation of these cells (Bracko et al., 2012). Further, shRNA mediated knock down of Igf2 lead to profound loss of proliferation of hippocampal NSCs demonstrating that Igf2 was a niche specific regulator of NSC (Bracko et al., 2012).

In the case of immature neurons, none of the receptors was found to significantly perturb the TRN attractor state.

2.5. Long term versus short term hair follicle stem cells

Recent report suggested that long term hair follicle stem cells (HFSCs) arise from embryonic progenitor cells depending on its spatial location (Xu et al., 2015). Embryonic progenitor cells residing in upper follicle gain expression of adult stem cell markers, and become long term HFSCs, whereas the cells remaining in the lower follicle region progressively lose their stem cell state and are defined as short term HFSCs. Analysis of the gene expression of these two cell types identified 1393 differentially expressed genes, including 134 receptors and 131 TFs. Bmpr1b and Notch3 were identified as top ranking receptors for regulating short term HFSCs and they regulated 26 and 23 DETFs respectively (Table 1). The role of Notch in short term HFSCs is not known, however, Bmp signaling is known for regulating hair follicle stem cell state (Kobielak et al., 2007). Further, our method identified Wnt pathway related receptors Fzd10 and Rspo1 in addition to Wnt10b (Fig. 3A). Since these three molecules were related to the same pathway, we performed Boolean simulations of perturbing them simultaneously and observed significant change in the attractor configuration (Table 1). In this context, Wnt mediated signaling has been shown to regulate the differentiation capabilities of short term HFSCs in vivo (Xu et al., 2015).

In the case of long term HFSCs, our method identified Tnfrsf18, Itgb6, Ace, Fgf2 and Osmr as significant (Table 1). Although the role of Itgb6 in hair follicle stem cell differentiation is known (Xie et al., 2012), the precise roles of these receptors for the phenotypic maintenance of long term HFSCs are not known.

2.6. Long term versus short term hematopoietic stem cells

The exquisite balance between self-renewal and differentiation of hematopoietic stem cells (HSCs) is required to maintain the homeostasis of lympho-hematopoietic system (Scadden, 2014). In this regard, regulation of quiescent or long-term HSC pool is important. Here, we have analyzed long term and short term HSCs in order to identify the niche determinants of these two distinct phenotypic states (Ficara et al., 2008). Gene expression analysis of these two cell types identified 1665 differentially expressed genes that include 98 receptors and 115 TFs. In the case of long term HSCs we found Tgfbr1 to be significant and regulated 22 TFs (Fig. 3B, Table 1). Importantly, the role of Tgfβ signaling in maintenance of long term quiescent HSCs is experimentally observed (Ficara et al., 2008) and Tgfbr1 is a key receptor involved in this signaling. Further, it has been suggested that cues from bone marrow niche are that activate this pathway is necessary for maintaining long-term quiescent HSCs (Ficara et al., 2008).

In the case of short term HSCs we found the receptors Il17rb, Epha4 and Cd40 to be significant, and they regulated 21, 17 and 8 TFs respectively (Table 1). The role of these receptors in regulating in short term HSC is not well established and are novel predictions of our method.

2.7. Quiescent versus active skeletal muscle stem cells

Skeletal muscle is a postmitotic tissue that shows very low turnover in the absence of disease or injury and its regenerative capacity is mediated by satellite cells (SC, skeletal muscle stem cells) (Liu et al., 2013). These cells remain generally in quiescent state and are activated upon disease or injury. Here, we have analyzed the gene expression of quiescent SCs versus activated (injury induced) SCs in order to identify the niche determinants of SC quiescence and active states (Liu et al., 2013). We identified 2961 differentially expressed genes, including

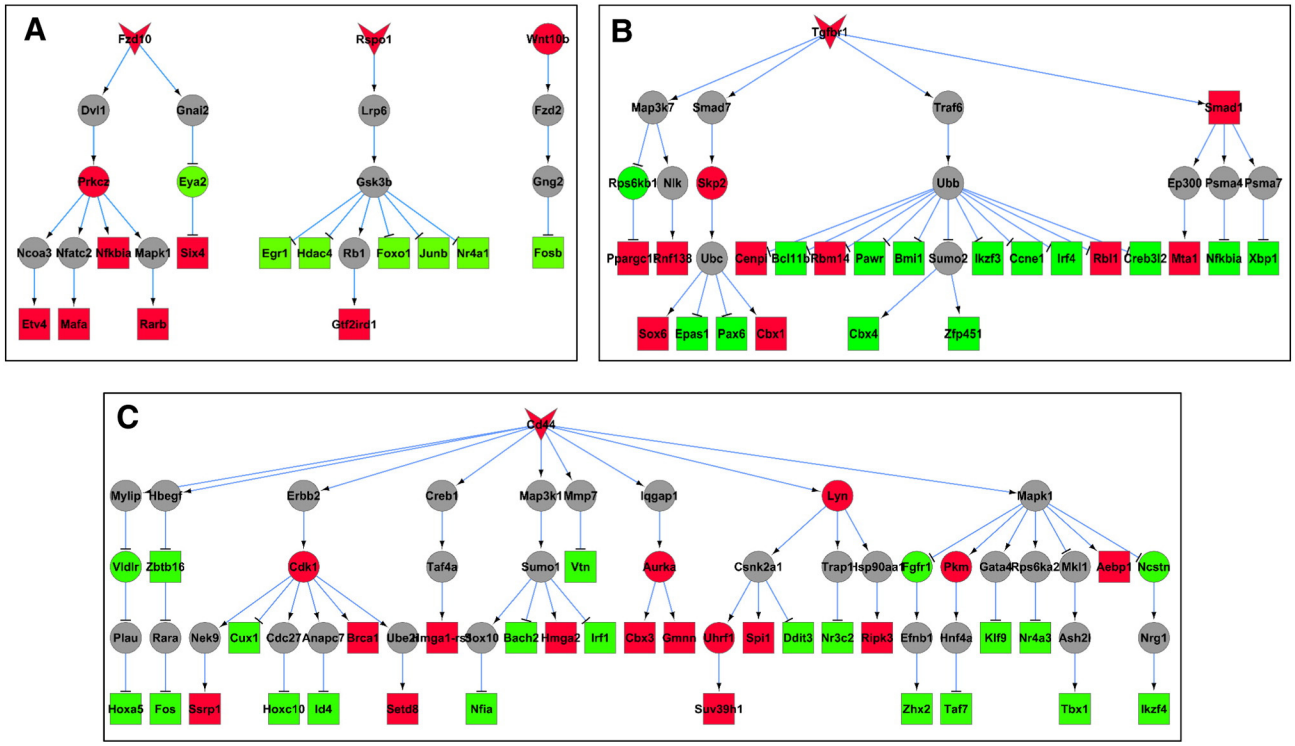

Fig. 3. Niche determinants of HFSCs, HSCs and STs. The figure is same as Fig. 2. A) Long term vs short term HFSCs. B) Long term vs short term HSCs and C) Quiescent vs Active SCs.

183 receptors and 176 TFs. In the case of quiescent SCs, our method identified Notch1 and Fgfr1 as significant (Table 1). Notch signaling was found to regulate 59 TFs whereas, Fgfr1 was found to regulate 28 TFs. Importantly, active Notch signaling is experimentally known to maintain quiescent SC state by preventing its activation (Bjornson et al., 2012). In fact, genetic ablation of Notch mediated signaling leads to concomitant aberrant activation of SCs eventually leading to their depletion (Bjornson et al., 2012). However, the relevance of Fgfr1 mediated signaling for quiescent SCs is not immediately clear.

In the case of injury induced active SCs, Ilrl1 and Cd44 were predicted to be significant (Table 1). Ilrl1 was found to regulate 46 TFs, whereas, Cd44 mediated signaling controlled 25 DETFs (Fig. 3C). In this regard, Cd44 is known to meditate SC migration and differentiation in response to muscular injury (Scimeca et al., 2015). Further, Cd44 knock down mice exhibited severely compromised migration and differentiation of SCs validating our predictions (Mylona et al., 2006). The role of Ilrl1 for regulating injury induced activation of SCs is not known.

3. Discussion

Characterizing the regulatory relationship between stem cells and their niches is fundamental for understanding tissue homeostasis and its implications in disease conditions. Rapid strides are made in this direction by explicit experimental characterization of niche elements and their respective role in maintaining specific stem cell phenotypes (Scadden, 2014). A few computational methods relying on this information have been proposed to describe stem cell-niche interactions. However, the immense complexity that arises due to the many-body nature of stem cell-niche interactions, as well as niche dynamics, urges the implementation of alternative computational approaches. In this direction, we have developed a computational method that attempts to overcome this complexity by capturing the net effect of the niche that is reflected at the level of constant activation/inhibition of stem cell signaling pathways responsible for the stable maintenance of stem cell phenotypes. Further, we showed that our method, though simple in its framework, is able to capture experimentally confirmed niche determinants for different stem cell systems from a large set of DERs. Notably, from more than 300 DERs, our method was able to capture S1pr1 and Egfr for maintenance of neural stem cell quiescent and active states, respectively (Codega et al., 2014). Prediction of Cd44 as a top ranking receptor for maintaining active OPCs state is experimentally known in the context of OPC migration upon injury-induced activation (Piao et al., 2013). Similarly, in the case of activated SCs in response to injury, our method predicted Cd44, whose role in SC migration was observed experimentally (Kobielak et al., 2007). Interestingly, Cd44 mediated signaling seems responsible for injury-induced activation of both OPCs and SCs, and it might represent a common mechanism of stem cell response to injury. Further, it points to the potential application of our method for understanding degenerative disease mechanisms in general and can possibly aid development of novel therapeutic strategies.

A precise quantitative assessment of overall sensitivity and specificity of our computational approach is not readily possible due to incomplete and ever increasing knowledge of niche mediated signaling pathways that regulate stem cell phenotypes. Particularly, the information about false positives and true negatives in stem cell systems are rarely available. However, in the case of false negatives, for some stem cell systems that were studied, we found that our approach could not identify certain known signaling pathways mediated by the niche. For instance, in the case of NSCs, prostaglandin signaling (mediated by Ptgdr) and Notch signaling are known to maintain the quiescent phenotype, and were not identified by our approach (Codega et al., 2014; Llorens-Bobadilla et al., 2015). Similarly, Wnt signaling which is known to play a role in long term maintenance of HSCs was not identified by our approach (Chotinantakul and Leeanansaksiri, 2012). Possible reasons for not identifying some of the known niche mediated signaling responsible for stem cell phenotype maintenance could be due to lack of information in the interactome databases, inherent noise in gene expression data, lack of good correlation between transcriptome and proteome/phosphoproteome levels and redundancy in signaling pathways. Despite these, our method, relying only on gene expression data could successfully identify several experimentally known candidates (true positives) of niche mediated signaling for maintaining specific stem cell phenotype.

In the current study, we used only gene expression data for the analysis; however, phosphoproteomics data could also be used to assign weights to the signaling interactome for the inference of niche determinants. An important limitation of the method is that it considers only DERs as the sources of niche induced signaling to regulate the expression of downstream TFs for a given phenotype. Nevertheless, even those receptors that are *not* differentially expressed can play a crucial role in regulating the target TFs for the stable maintenance of TRNs. Nevertheless, given only gene expression data with limited number of replicates, it is not feasible to quantitate absolute expression levels of genes without resorting to differential gene expression. However, with the advancements in single cell sequencing techniques, that can offer expression levels at single cell resolution, one can more reliably quantitate absolute gene expression levels for a given cell type. In such cases, our method does not need to rely on differential gene expression and can take advantage of absolute expression status of the receptors to infer the niche determinants. Further, this could allow considering multiple phenotypes simultaneously without the need for pair-wise comparison. Therefore, the proposed method is flexible in its application to different kinds of data including phosphoproteomics and single cell RNA sequencing.

In summary, we have proposed a computational method that simplifies the complexity of stem cell-niche interactions and enables identification niche determinants of cellular phenotypic states. Given the importance of the role played by deregulated niche components in several diseases, our computational method could aid the identification of novel therapeutic strategies for regenerative medicine by mimicking the niche effect on the target stem cells.

4. Materials and methods

A schematic representation of our method is shown in Fig. 1 and the detailed description of the methods is provided in the Supplementary Information.

4.1. Phenotype specific signaling pathway inference

A raw signaling interactome originating from DERs to DETFs in each cell type is retrieved from MetaCore (GeneGo Inc. (Nikolsky et al., 2005)) using the Trace Path algorithm. Then, we employed a PCST formalism to infer a minimal sub-network (Steiner tree) in the raw interactome that connects the dummy niche node with the DETFs via specific receptors. For this, we used a message-passing algorithm, based on belief propagation for the inference of Steiner trees (Bailly-Bechet et al., 2011). Formally, the PCST problem is defined as, given a graph $G = (V,E)$, representing the raw signaling interactome, with defined edge costs (weights), c_e and node prizes b_v find a connected subgraph $T = (V',E')$, $V' \subseteq V$, $E' \subseteq E$, that minimizes the following function:

$$T = \min_{(E',V')\,connected} \left(\sum_{e\in E'} c_e - \lambda \sum_{v\in V'} b_v\right) \quad (1)$$

The costs of the edge c_e reflect the probability of the target node to be differentially active for the given interaction (described in Supporting Information). The node prizes are computed by $b_v = |\log \text{fold change } (V)|$ from the gene expression data. The constant λ determines the tradeoff of adding new proteins to the inferred network by balancing the cost of new edges and the prize gained by adding a new protein. We chose $\lambda = 0.01$ for our simulations. Essentially, minimizing this function implies, collecting the largest set of high prize nodes while minimizing the set of high cost edges in a tradeoff tuned by λ that

results in a connected subgraph. We employed a heuristic method based on a message-passing algorithm to infer the PCSTs (Bailly-Bechet et al., 2011).

Since our signaling interactome is differentially weighted to account for the two phenotypes, the algorithm was employed twice for inferring sub-networks specific to each phenotype. This will result in two different sub-networks from the upregulated receptors in each phenotype to the entire set of DETFs (Fig. 1). The significance of the inferred receptors in the sub-network were assessed by randomizing the edge weights iteratively, and z-scores for statistical significance of the sub-network was calculated based on the number of TFs controlled by the receptor under randomized edge weights. The z-score was calculated by $(x - x')/\sigma$, where x is the number of TFs regulated by a receptor in the original weighted network and x' is the average and σ is the standard deviation of the number of TFs regulated by the same receptor with 100 randomized edge weights. Only those receptors with z-score greater than 2.5 were shortlisted for downstream analysis.

4.2. Integration of signaling with transcriptional network

Each receptor-to-target DETF path inferred from PCSTs is collapsed into a single network edge, whose sign is defined as the product of the signs in the original path. In parallel, the interaction network among DETFs is retrieved from MetaCore (GeneGo Inc. (Nikolsky et al., 2005)), which is then combined with the collapsed receptor-target DETF signaling edges. To model the TRN stabilized by the niche effect through intracellular signaling, this combined raw network was contextualized (pruned) against the stable gene expression pattern using the method proposed by (Crespo et al., 2013). Briefly, this algorithm assumes that each cellular phenotype is a stable steady state attractor of a given network, and removes edges that are inconsistent with the Booleanized mRNA expression data. Receptor-target DETF edges are treated as fixed interactions. The Boolean simulation is carried out using the synchronous updating scheme. The node weights are set to all 1. The inhibitor dominant logic rule is applied (i.e., the presence of at least one inhibitory interaction is sufficient to suppress the target gene expression).

4.3. Identification of sub-networks necessary for TRN stability

To identify sub-networks that play key roles in the maintenance of TRN stability and therefore whose perturbation could trigger a significant change in the TRN state, target DETFs in each sub PCST are perturbed in the combined signaling–TRN. For each inferred sub-network, the significance p-value is computed with the t-test on the number of genes whose Boolean states have changed after the perturbation. The null distribution is formed by perturbing the median number of target DETFs over all sub-networks in each cell type. This random perturbation is repeated 100 times and sub-networks with the p-value below 0.05 are considered significant (i.e., necessary for TRN stability).

Acknowledgements

S.O. is supported by an FNR AFR Postdoctoral Grant (7682104/PDR).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.scr.2016.09.006.

References

Bailly-Bechet, M., Borgs, C., Braunstein, A., Chayes, J., Dagkessamanskaia, A., Francois, J.M., Zecchina, R., 2011. Finding undetected protein associations in cell signaling by belief propagation. Proc. Natl. Acad. Sci. U. S. A. 108, 882–887.

Baker, N., Tuan, R.S., 2013. The less-often-traveled surface of stem cells: caveolin-1 and caveolae in stem cells, tissue repair and regeneration. Stem cell research & therapy 4, 90.

Bild, A.H., Yao, G., Chang, J.T., Wang, Q., Potti, A., Chasse, D., Joshi, M.B., Harpole, D., Lancaster, J.M., Berchuck, A., et al., 2006. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 439, 353–357.

Bjornson, C.R., Cheung, T.H., Liu, L., Tripathi, P.V., Steeper, K.M., Rando, T.A., 2012. Notch signaling is necessary to maintain quiescence in adult muscle stem cells. Stem Cells 30, 232–242.

Bracko, O., Singer, T., Aigner, S., Knobloch, M., Winner, B., Ray, J., Clemenson Jr., G.D., Suh, H., Couillard-Despres, S., Aigner, L., et al., 2012. Gene expression profiling of neural stem cells and their neuronal progeny reveals IGF2 as a regulator of adult hippocampal neurogenesis. J. Neurosci. 32, 3376–3387.

Chotinantakul, K., Leeanansaksiri, W., 2012. Hematopoietic stem cell development, niches, and signaling pathways. Bone marrow research 2012, 270425.

Codega, P., Silva-Vargas, V., Paul, A., Maldonado-Soto, A.R., Deleo, A.M., Pastrana, E., Doetsch, F., 2014. Prospective identification and purification of quiescent adult neural stem cells from their in vivo niche. Neuron 82, 545–559.

Crespo, I., Krishna, A., Le Bechec, A., del Sol, A., 2013. Predicting missing expression values in gene regulatory networks using a discrete logic modeling optimization guided by network stable states. Nucleic Acids Res. 41, e8.

Ficara, F., Murphy, M.J., Lin, M., Cleary, M.L., 2008. Pbx1 regulates self-renewal of long-term hematopoietic stem cells by maintaining their quiescence. Cell Stem Cell 2, 484–496.

Franklin, R.J., Ffrench-Constant, C., 2008. Remyelination in the CNS: from biology to therapy. Nat. Rev. Neurosci. 9, 839–855.

Hung, J.H., 2013. Gene set/pathway enrichment analysis. Methods Mol. Biol. 939, 201–213.

Khatri, P., Sirota, M., Butte, A.J., 2012. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput. Biol. 8, e1002375.

Kobielak, K., Stokes, N., de la Cruz, J., Polak, L., Fuchs, E., 2007. Loss of a quiescent niche but not follicle stem cells in the absence of bone morphogenetic protein signaling. Proc. Natl. Acad. Sci. U. S. A. 104, 10063–10068.

Kunisaki, Y., Bruns, I., Scheiermann, C., Ahmed, J., Pinho, S., Zhang, D., Mizoguchi, T., Wei, Q., Lucas, D., Ito, K., et al., 2013. Arteriolar niches maintain haematopoietic stem cell quiescence. Nature 502, 637–643.

Lane, S.W., Williams, D.A., Watt, F.M., 2014. Modulating the stem cell niche for tissue regeneration. Nat. Biotechnol. 32, 795–803.

Lei, J., Levin, S.A., Nie, Q., 2014. Mathematical model of adult stem cell regeneration with cross-talk between genetic and epigenetic regulation. Proc. Natl. Acad. Sci. U. S. A. 111, E880–E887.

Liu, L., Cheung, T.H., Charville, G.W., Hurgo, B.M., Leavitt, T., Shih, J., Brunet, A., Rando, T.A., 2013. Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. Cell Rep. 4, 189–204.

Llorens-Bobadilla, E., Zhao, S., Baser, A., Saiz-Castro, G., Zwadlo, K., Martin-Villalba, A., 2015. Single-cell transcriptomics reveals a population of dormant neural stem cells that become activated upon brain injury. Cell Stem Cell 17, 329–340.

Moyon, S., Dubessy, A.L., Aigrot, M.S., Trotter, M., Huang, J.K., Dauphinot, L., Potier, M.C., Kerninon, C., Melik Parsadaniantz, S., Franklin, R.J., et al., 2015. Demyelination causes adult CNS progenitors to revert to an immature state and express immune cues that support their migration. J. Neurosci. 35, 4–20.

Mylona, E., Jones, K.A., Mills, S.T., Pavlath, G.K., 2006. CD44 regulates myoblast migration and differentiation. J. Cell. Physiol. 209, 314–321.

Nikolsky, Y., Ekins, S., Nikolskaya, T., Bugrim, A., 2005. A novel method for generation of signature networks as biomarkers from complex high throughput data. Toxicol. Lett. 158, 20–29.

Piao, J.H., Wang, Y., Duncan, I.D., 2013. CD44 is required for the migration of transplanted oligodendrocyte progenitor cells to focal inflammatory demyelinating lesions in the spinal cord. Glia 61, 361–367.

Rompolas, P., Mesa, K.R., Greco, V., 2013. Spatial organization within a niche as a determinant of stem-cell fate. Nature 502, 513–518.

Scadden, D.T., 2014. Nice neighborhood: emerging concepts of the stem cell niche. Cell 157, 41–50.

Scimeca, M., Bonanno, E., Piccirilli, E., Baldi, J., Mauriello, A., Orlandi, A., Tancredi, V., Gasbarra, E., Tarantino, U., 2015. Satellite cells CD44 positive drive muscle regeneration in osteoarthritis patients. Stem Cells Int. 2015, 469459.

Szekely Jr., T., Burrage, K., Mangel, M., Bonsall, M.B., 2014. Stochastic dynamics of interacting haematopoietic stem cell niche lineages. PLoS Comput. Biol. 10, e1003794.

Tuncbag, N., Braunstein, A., Pagnani, A., Huang, S.S., Chayes, J., Borgs, C., Zecchina, R., Fraenkel, E., 2013. Simultaneous reconstruction of multiple signaling pathways via the prize-collecting Steiner forest problem. J. Comput. Biol. 20, 124–136.

Whitmarsh, A.J., 2007. Regulation of gene transcription by mitogen-activated protein kinase signaling pathways. Biochim. Biophys. Acta 1773, 1285–1298.

Xie, Y., McElwee, K.J., Owen, G.R., Hakkinen, L., Larjava, H.S., 2012. Integrin beta6-deficient mice show enhanced keratinocyte proliferation and retarded hair follicle regression after depilation. Journal of Investigative Dermatology 132, 547–555.

Xin, T., Greco, V., Myung, P., 2016. Hardwiring stem cell communication through tissue structure. Cell 164, 1212–1225.

Xu, Z., Wang, W., Jiang, K., Yu, Z., Huang, H., Wang, F., Zhou, B., Chen, T., 2015. Embryonic attenuated Wnt/beta-catenin signaling defines niche location and long-term stem cell fate in hair follicle. eLife 4.

Zhu, X., Zhang, J., Tollkuhn, J., Ohsawa, R., Bresnick, E.H., Guillemot, F., Kageyama, R., Rosenfeld, M.G., 2006. Sustained Notch signaling in progenitors is required for sequential emergence of distinct cell lineages during organogenesis. Genes Dev. 20, 2739–2753.