Characterization of unknown adult stem cell samples by large scale data integration and artificial neural networks.

Ghislain Bidaut, C. J. Stoeckert

CHARACTERIZATION OF UNKNOWN ADULT STEM CELL SAMPLES BY LARGE SCALE DATA INTEGRATION AND ARTIFICIAL NEURAL NETWORKS

G.BIDAUT [1,2] AND C.J. STOECKERT JR.[1]

1: Center for Bioinformatics, Department of Genetics,
University of Pennsylvania School of Medicine
423 Guardian Drive, Philadelphia, PA 19104, USA
2: Centre de Recherche en Cancérologie de Marseille
INSERM U891 - Institut Paoli-Calmettes
Université de la Méditerranée
27 Boulevard Leï Roure, 13009 Marseille, France
E-mail: ghislain.bidaut@inserm.fr, stoeckrt@pcbi.upenn.edu

Stem cells represent not only a potential source of treatment for degenerative diseases but can also shed light on developmental biology and cancer. It is believed that stem cells differentiation and fate is triggered by a common genetic program that endows those cells with the ability to differentiate into specialized progenitors and fully differentiated cells. To extract the *stemness* signature of several cells types at the transcription level, we integrated heterogeneous datasets (microarray experiments) performed in different adult and embryonic tissues (liver, blood, bone, prostate and stomach in *Homo sapiens* and *Mus musculus*). Data were integrated by generalization of the hematopoietic stem cell hierarchy and by homology between mouse and human. The variation-filtered and integrated gene expression dataset was fed to a single-layered neural network to create a classifier to (i) extract the *stemness* signature and (ii) characterize unknown stem cell tissue samples by attribution of a stem cell differentiation stage. We were able to characterize mouse stomach progenitor and human prostate progenitor samples and isolate gene signatures playing a fundamental role for every level of the generalized stem cell hierarchy.

1. Introduction

A wide variety of stem cell types have been recently reported to exist in several adult organs and are suspected to be present in most tissues. Well known stem cells types include hematopoietic stem cells (HSCs), neural stem cells (NSCs), myogenic progenitors (muscle), and others having a more restricted potential (such as gut and skin[1]) The consensus among stem cell researchers is that stem cell fate decision and renewal are triggered by sev-

eral mechanisms that do not completely overlap among different stem cells type in the same organism[2], and in different species[3]. Two key questions are still unresolved: (i) Despite the knowledge of their existence, the location and differentiation capabilities of stem or progenitor cells are unknown for the vast majority adult organs ; methods to confirm the presence of stem/progenitor cells in adult tissues are needed. For instance, there is a large uncertainty on whether pancreatic beta-cell progenitor cells reside within the pancreatic ductal epithelium, the pancreatic small cells, or acinar tissue, or all of the previous[4]. (ii) The list of early markers that drive differentiation and self-renewal properties is still not agreed upon.

In this paper, we present a computational method to decipher common mechanisms of stem cells differentiation at the transcriptomic level using microarray data. We therefore integrated several datasets obtained from heterogeneous stem cells studies and trained an artificial neural network to extract a list of common gene markers triggering stem cell differentiation and fate decisions. Our hypothesis is that pathways triggering differentiation are at least partially conserved among adult stem cells/progenitors, forming a molecular signature reflecting embryonic and adult stem cells plasticity - the *stemness*. This hypothesis has been formulated several times[5], and is supported by measured transcriptome data[6], following a controversy about whether the *stemnesss* property could really be verified at the transcriptional level[7]. Discovering a shared signature should help characterize unknown stem cells populations, i.e. to determine whether a cell population contains stem cell/progenitors, and what is their differentiation potential/stage. In addition, the publication of a catalog of genes involved in differentiation of several cell types would be a valuable resource for stem cell researchers and developmentalists.

To extract molecular signatures for stem/progenitor cells developmental stages by gene expression profiling, we trained a multiclass single-layer linear ANN that subsequently allowed us to characterize unknown samples and position them in a hierarchy ranging from totipotent stem cells to fully differentiated cells. We tested the predictions made by our neural network using two tissues that were only partially characterized: Mouse stomach epithelium and human prostate (Figure 1). The ability of our system to generalize its classification to unknown stem cell types was assessed with a one-leave out cross-validation procedure on the training data.

ANNs represent a class of machine learning algorithm that were successfully applied to a large range of open ended problems. Their basic structure has been inspired from neurobiology and takes the form of a feed

forward network of neurons modeled by a transition function. In a typical setup, an initial network topology is chosen depending on the nature of the problem; neurons are then trained for several epochs, leading to an ANN model, which is in turn applied to the classification of unknown samples. In biology, they were applied early for multiclass tumor classification in cancer discovery of cancer molecular subtypes and identification of biomarkers[8].

To properly use a classification system on the selected dataset we consistently labeled the stem and differentiated cells in tissues according to their differentiation stage/properties using a controlled vocabulary. The hematopoietic stem cell hierarchy[5] was used as a model: in this differentiation system, the different cells identified and isolated thanks to functional assays, are hierarchically positioned according to their differentiation potentials. At top of the hierarchy is the HSC, which can give rise to all blood cells and has self-renewal capacity, and finally the various types of mature blood cells. We defined a more general model (see section *Data integration*) applicable to all stem cell types, including adult stem cells and embryonic stem cells. Five stages are recognized, as shown in Table 1

Table 1. Properties of the generalized stem cell hierarchy and the controlled vocabulary used for classification.

Code	Stem Cell Type	Properties
A	Totipotent Stem Cell	Capable of self-renewal and able to generate all cell types
B	Multipotent Stem Cell	Capable of self-renewal and able to generate most cell types
C	Progenitor Cell	Capable of generating several cell types
D	Lineage-Committed Progenitor (LCP) Cell	Capable of generating a single or a restricted number of cell types
E	Differentiated Cell	Cell displaying final phenotype

The whole scheme operates in three steps. First we generated a training dataset by integration of several gene expression datasets generated by the SCGAP (Stem Cell Genome Anatomy Projects) consortium in different tissues, and projected it in a predefined space of basis vectors using the technique of vector projection[9], to group genes having similar dynamics of differentiation. We then used this integrated dataset to train a single-layer artificial multi-class neural network to classify unknown tissues potentially containing stem/progenitors cells in one of the predefined categories. Cross validation over training dataset led to 31 independent ANN models, trained independently by iteratively pulling out a tissue from the training data. The optimal number of genes to be employed in the classification was found by performing the training while reducing the number of genes used for classification. At each iteration, the top genes were retained by sorting the ANN weights and were kept before next training iteration. Classification error rate was minimized for 63 genes, and this set of genes was retained

as a minimal core representing the *stemness* property shared by the set of stem/progenitor cells in the data. Finally, we tested the predicting power of the 31 models on two tissues not used for training that are potentially a mixture of adult and progenitor cells, namely mouse stomach progenitors[10] and human prostate progenitors[11].

The datasets analyzed measure gene expression in mouse and human, stem or progenitor cell in various tissues at different development stages. Gene coverage varies with the type of chip (from 6K-to 18K genes, see Table 2). Data were integrated at the gene level using the NCBI *Homologene* database[12] and at the experimental level using a controlled vocabulary of Table 1. The resulting data matrix contains 18720 genes profiled across 82 samples in total.

The 82 arrays are grouped by tissues. A unique tissue sample is characterized by a group of at least two arrays, up to five arrays covering previously stem/progenitor cells categories A-E - this is a necessary condition to be able to perform vector projection. This lead us to a final table of 40 tissues.

Table 2. Summary of tissues, platforms and cover. The two last tissues marked in bold are used for testing. Stem cell categories are marked with code (From A to E), corresponding to the previously defined stem cell stages in Table 1.

Author/Lab	Tissue	Platform	N genes	Categories
Darlington *et al.* (2007)[13]	Mouse Embryonic Liver	Affy. 430A	12798	C,E
Rowe *et al.* (unpublished)	Mouse Bone	Affy. U74 Av2, Bv2, Cv2	15127	C,D,E
Ivanova *et al.* (2002)[5]	Human Fetal Liver (HSCs)	Affy. U95 Av2,B,C,D,E	17024	B,D,E
Ivanova *et al.* (2002)[5]	Mouse Fetal Liver (HSC)	Affy. U74 Av2, Bv2, Cv2	15127	B,D,E
Ivanova *et al.* (2002)[5]	Mouse Adult Bone Marrow	Affy. U74 Av2, Bv2, Cv2	15127	B,C,D,E
Ivanova *et al.* (2002)[5]	Mouse Embryonic Stem Cells (ESCs)	Affy. U74 Av2, Bv2, Cv2	15127	A
Ivanova *et al.* (2002)[5]	Mouse Neural Stem Cells (NSCs)	Affy. U74 Av2, Bv2, Cv2	15127	B
Ivanova *et al.* (unpublished)	Human Coord Blood (HSCs)	Affy. U133 A,B	17275	B,D
Ivanova *et al.* (unpublished)	Human Adult Bone Marrow (HSCs)	Affy. U133 A,B	17275	B,D
Ivanova *et al.* (unpublished)	Mouse Adult Bone Marrow (HSCs)	Affy. 430 A,B	18626	B,C
Oudes *et al.* (2006)[11]	**Human Prostate Progenitors**	Affy. U133plus2.0	18806	X,E
Mills *et al.* (2002)[10]	**Mouse Stomach progenitors**	Affy. Mu11K A,B	6975	X,E
Total: 5 distinct	12 distinct	6 distinct	18720	5 distinct

The inherent architecture of the neural network-based single layer system allows for detailed exploration of two important parameters: (i) ranking weights for each of the five stages allow extraction of genes reported by the classifier to be a marker for this particular stem cell stage. (ii) ranking genes on a vector y resulting from the expression projection multiplied by the corresponding weight for a given cell population ($y_n = w_n.p_n$, w being the highest level neuron weight and p the expression projection) allows to delineate gene profiles critical for proper classification of this tissue. We additionally performed an analysis of highly ranked genes and statistically

enriched gene ontologies (GO) for each of the stem/progenitor stage (unpublished data). A hidden layer to the architecture was not added even though it might decrease classification error because of the concern of overtraining and losing the ability to interpret the gene weights associated with each stem/progenitor cell stage. Other classifier types have also been considered but ANNs have been retained to keep the ability to rank genes according to their weights.

2. Material and Methods

2.1. *Dataset*

The dataset is an integration of stem cell transcriptional data generated by the 7 members of the SCGAP (Stem Cell Genome Anatomy Projects) consortium to train and test the system. Every SCGAP members performed large scale gene expression analysis on AffyMetrix arrays in given cell types and in two organisms, *M. musculus* and *H. sapiens* after specific purification steps (see Table 2). Data are available for interactive search on the Consortium's web portal (http://www.scgap.org), for download from every SCGAP member's web site, and in an integrated form the supportive web site (http://scann.sourceforge.net).

2.2. *Data integration*

Data were integrated at two level to take into account the heterogeneous nature of the consortium platforms. Data were first normalized within single datasets using iterative *lowess* normalization to compensate for differences in hybridization among the different platforms. The *mas* summarization method implemented in the bioconductor (R) software package (library *affy*) provided us with the expression values. We labeled the training dataset with a controlled vocabulary describing the five stem cell/progenitor differentiation stages (See Figure 1(b) and Table 1).

2.3. *Vector Projection*

Vector projection is a technique previously used for feature extraction in time course pancreas development gene expression data[9]. This technique allows for quick gene identification from pre-defined expression profiles that model the expected behavior of the genes we wish to isolate as most representative. In our case, vector projection captures genes with expression profiles peaking in different stem/progenitor differentiation stages for a given

tissue (See Figure 1). The mathematical definition of vector projection is the inner product of a gene profile within a tissue to each model vector. We designed all model vector for the extraction of gene profile that were primarily expressed at one of the five progenitor/stem cell stage. These projection values were used to filter out genes without enough expression variation across tissues, and after filtering, projection values were presented to the ANN for training. Vector projection could be extended to extract gene profiles peaking in two or mode population - for instance to extract genes expressed during totipotency, turned off during self-renewal, and expressed again lated during lineage committed differentiation.

2.4. *Missing values and Data Organization*

A critical issue of the data we are analyzing are the missing values. This is particularly concerning here since projecting gene expression profiles characterized by missing value points can lead to results dramatically different from those expected. To cope with this issue, we performed vector projection using only the determined values, and re-normalize the model vector to a value of 1.0 using only the non-missing populations. After projection, the obtained dataset is organized both by tissue and projection values and reports the projection value for every stem/progenitor stem cell stage within every tissue. The projection values are then submitted to the ANN.

2.5. *System Architecture and Parameters*

The classifier used here is an extension of the well known single layer associative memory (Figure 1). It is characterized by a single neuronal layer holding the input weights to be learned from the data. Five neurons in total have been associated to the five levels of stem/differentiated cell hierarchy previously defined. The output for each neuron is the scalar product of the projection input by the weights: $y_i = \sum_{n=1}^{N} w_{i,n}.p_n$, with y being the neuron output, i the neuron number, n the gene index, N the total number of genes, $w_{i,n}$ the nth value of the ith neuron weight, and p_n the nth value of the gene expression projection to be classified. At every cross-validation step, one tissue is selected for testing and the others as a training set. A new neural network model is created (the weights are reinitialized) and the training set is presented to the network for 200 epochs. At each epoch, tissues from the training set are presented in a random order to the network, and weights are updated with a classic gradient descent-type training rule: $\Delta w_n = a(k)[y_n - yd_n]$, w_n being the current weight value, $a(k)$ the learning

rate (typically decreasing over time, we used $a(k) = 1/(k+1)$, k being the epoch number), y_n the output and yd_n the desired neuron output. Each ANN model trained during cross validation is kept and used at a later stage for final classification of unknown samples. The number of weights to be conserved is updated until a minimal classification error rate is reached during cross-validation.

We then performed the cross-validation and estimated the optimal number of features for tissue classification by iteratively reducing the number of weights (or input genes) used by the ANN during training. The sum of misclassified test samples on the 31 ANN models versus the number of genes is represented figure 2(a) and is minimized for 63 genes (16 genes on individual neurons).

2.6. *Classification algorithm*

For a given tissue that we wish to classify and attribute to one of the five cell types defined in Table 1, classification was performed as follows: The tissue tested must meet several conditions: At least two samples must be available to perform vector projection, a fully differentiated sample known *a priori*, and the second sample containing stem/progenitor cells to be characterized. The system then characterizes every sample by first assigning it to a category as described in Table 1 for each of the 31 ANN models generated during leave one out training. A consensus is given by majority voting on the 31 ANN models. To understand the biological functions involved in the various differentiation stages, we examined the genes used in the classification and the enriched ontologies with ClutrFree[14] (data not shown).

3. Results and Discussion

3.1. *classification results*

Figure 2(a) shows the global misclassification rate versus the number of genes. The quadratic error rate for the training set and tests sets are respectively represented Figures 2(b) and 2(c). The error rate of every individual ANN trained with 63 genes is shown Figures 2(d) for the training set and the corresponding held out tissue. Misclassified tissues are pointed with an arrow. The minimal misclassification rate was obtained for 63 genes (See arrow on Figure 2(a), yielding a set of genes representing the core of *stemness* genes necessary for cell differentiation from ESCs to fully differentiated cells in both *H. sapiens* and *M. musculus* data. Error progression

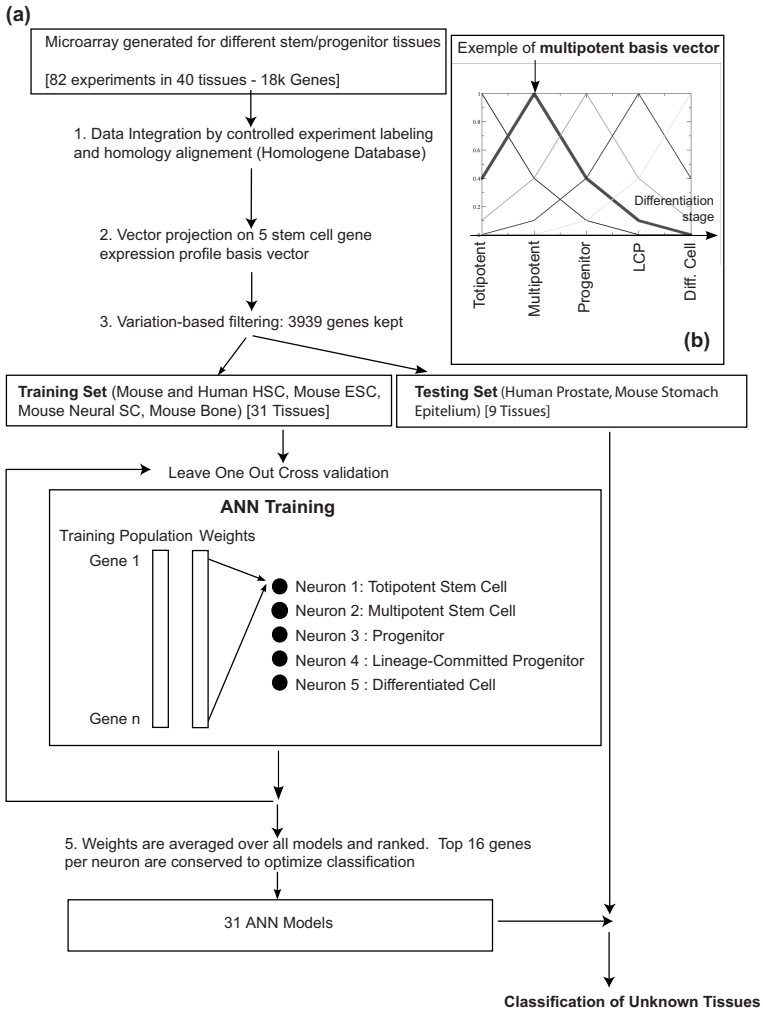

Figure 1. (a): ANN System Architecture used in the analysis: The data is initially projected onto the space defined by the 5 stem/progenitor cell basis vector. Splitting of the input dataset in Training/Testing dataset is shown. (b): Model vectors used for vector projection.

curves decreased rapidly and did not revealed overtraining in the neural network, even though we went up to 200 epochs, except for one neural stem cell tissues (Mouse NSC sample 2, not pictured).

We used the network to characterize two tissues potentially containing

progenitors, namely human prostate and mouse stomach. Those two tissues were classified as expected by the system and are thus likely a mixture of fully differentiated cells and stem cells (Table 3).

Table 3. Classification by majority vote on the 31 Neural Network models generated for 63 genes

Tissue to be tested	Majority Vote
Mouse Stomach Progenitor	Progenitor
Human Prostate Progenitor	Multipotent Stem Cells

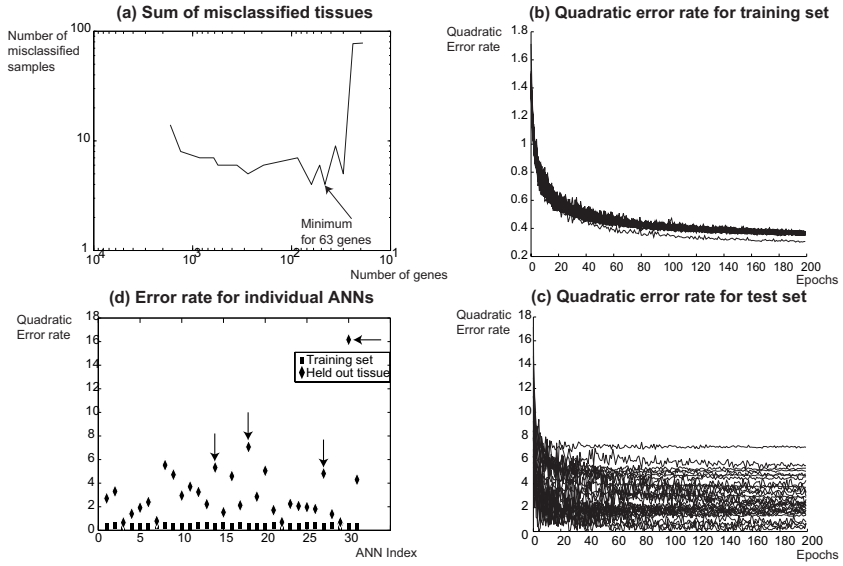

Figure 2. Classification results obtained for the ANN system.

The ANN models allowed us not only to characterize unknown tissues but also to extract genes that back the classification in those categories, by ranking the genes according to their weights - a higher weight meaning the gene plays an essential role in the classification. The training of 5 neurons allowed us to extract a total of 5 gene sets, each of them related to a stem cell developmental stage (totipotent, multipotent, progenitor, lineage-committed progenitors (LCP) and differentiated cell). The analysis of markers was done on the first model given by the leave-one out run for

a training on 63 genes. The top 10 genes lists ranked by weights for each population are presented Table 4.

Table 4. Top 10 genes sets for the 5 stem cell stages ranked by weights.

Totipotent	Multipotent	Progenitor	LCP	Diff. Cells
Gene List	Gene List	Gene List	Gene List	Gene List
Dbn1	Procr	Letmd1	Coq3	Aqp1
1110001A23Rik	Gprasp2	Lrp8	Mgst1	Rhced
BC053917	AI661017	Kpna3	PDZK8	Rhbdl4
5830405N20Rik	Ctso	Ptprm	Cybb	MGI:1933403
Iqwd1	Gkap1	Rbp4	Rbbp9	Wnt11
5730420B22Rik	Lrrc16	Ass1	Cst7	Eif2ak3
Rbp4	Nrbp2	Itgb2	4932441K18	Nfe2
Rpp40	Adam8	Med6	Fli1	Fech
Nfe2	Irak1bp1	Cd109	Anxa1	AI661017
Rbbp9	MGI:1916782	5830405N20Rik	Gpr124	Gzma

For the totipotent population, the top genes include *Rbbp9*, a gene that may play a role in cell proliferation and differentiation.

In the multipotent population, we found MGI:1916782 (*Hopx*) - expressed in embryonic myocardium and other mesoderm, but not in endocardium or great vessels. *HOP* is also highly expressed in the developing heart, dependent on the cardiac-restricted homeodomain protein Nkx2.5. Inactivation of Hop in mice by homologous recombination results in a partially penetrant embryonic lethal phenotype with severe developmental cardiac defects involving the myocardium.

In tissues containing progenitor cells, we found several interesting markers: *Letmd1* is an oncoprotein that has a role in the development of human breast cancer. The protein encoded by *ptprm* is a member of the protein tyrosine phosphatase (PTP) family which is known to be signaling molecules regulating a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. We also found the cell surface antigen *CD109*. This is a glycosylphosphatidylinositol (GPI)-linked glycoprotein found on a subset of hematopoietic stem and progenitor cells[15].

Other markers of interest were also found in the two other categories: In lineage-committed progenitors, *4932441K18* (*FIAT*) is a transcriptional regulator of osteoblastic functions. *Anxa3* encodes a protein member of the annexin family. Members of this family are involved in cellular growth, and this particular protein may play a role in anti-coagulation. *Sfrp4* is a Wnt pathway inhibitor and plays a central role in cell fate decisions. We also found *CD34*, known to be expressed in hematopoietic stem cells, and

a marker for stem cell purification (ranked 13th in the LCP category).

Interestingly, exploration of the gene list for 87 genes shows *Socs2* in the top 10 genes for the Multipotent category. *Socs2* is ubitiquously expressed in most tissues in *M. musculus* and was shown to play an important regulatory role in neural development, neural stem cell differentiation and neuronal growth[16]. It is also a growth-hormone inducible and novel inhibitor of intestinal epithelial cell proliferation and intestinal growth. It was also shown to play a role in mammary gland development and regulate the fate of mesenchymal precursor cells[17], which is the nature of the cells we are studying here. Phenotype of Socs2-/- mice are describe therein and includes enlargement of bone and skeletal muscles.

The full list of markers obtained with the training for 63 genes is available from the supporting web site (scann.sourceforge.net).

3.2. *Discussion*

We present a novel approach that involved large scale integration of heterogeneous microarray datasets and pattern recognition analysis based on a vector projection technique to create a neural network-based classifier for characterization the differentiation properties of unknown stem cell tissues and extraction a molecular signature of *stemness*. Analysis of genes obtained by weight ranking highlighted gene product involved in several steps of cell differentiation. Thus, differentiation recruits a large panel of different genes and pathways and results from subtle expression level changes in a large number of genes. However, there is obviously more work to precisely validate the markers found. We were able to perform correct classification of unknown tissues (mouse stomach progenitors and human prostate progenitors) with a signature of 63 genes representing a core of genes involved in the differentiation process. As a natural extension of this work, we plan to include other stem cell studies publicly available on repositories such as Gene Expression Omnibus and ArrayExpress. On the technical side, we plan to improve the generalization capabilities of the classifier by boosting techniques, and make it accessible through a publicly available web server to perform the classification of unknown stem cell type. Other classifiers types will also be considered and their performance on this dataset studied.

4. Acknowledgments

We thank the SCGAP Consortium members for sharing data and providing insights for the analysis, and Dr Françoise Birg, Dr Patrice Dubreuil and

Wahiba Gherraby for their valuable comments on the manuscript. Dr Ghislain Bidaut is funded by the Fondation pour la Recherche Médicale (FRM) and the Institut National du Cancer (INCa). This work was initially funded by grant U01 DK63481.

References

1. K. D. Bunting and R. G. Hawley, *Biol Cell* **95**, 563(Dec 2003).
2. A. V. Molofsky, R. Pardal and S. J. Morrison, *Curr Opin Cell Biol* **16**, 700(Dec 2004).
3. M. Rao, *Dev Biol* **275**, 269(Nov 2004).
4. K. L. Seeberger, J. M. Dufour, A. M. J. Shapiro, J. R. T. Lakey, R. V. Rajotte and G. S. Korbutt, *Lab Invest* **86**, 141(Feb 2006).
5. N. B. Ivanova, J. T. Dimos, C. Schaniel, J. A. Hackney, K. A. Moore and I. R. Lemischka, *Science* **298**, 601(Oct 2002).
6. A. C. Piscaglia, T. Shupe, A. Gasbarrini and B. E. Petersen, *Curr Pharm Biotechnol* **8**, 167(Jun 2007).
7. N. B. Ivanova, J. T. Dimos, C. Schaniel, J. A. Hackney, K. A. Moore, M. Ramalho-Santos, S. Yoon, Y. Matsuzaki, R. C. Mulligan, D. A. Melton and I. R. Lemischka, *Science* **302**, p. 393d(Oct 2003).
8. B. T. Greer and J. Khan, *Ann N Y Acad Sci* **1020**, 49(May 2004).
9. L. M. Scearce, J. E. Brestelli, S. K. McWeeney, C. S. Lee, J. Mazzarelli, D. F. Pinney, A. Pizarro, C. J. Stoeckert, S. W. Clifton, M. A. Permutt, J. Brown, D. A. Melton and K. H. Kaestner, *Diabetes* **51**, 1997(Jul 2002).
10. J. C. Mills, N. Andersson, C. V. Hong, T. S. Stappenbeck and J. I. Gordon, *Proc Natl Acad Sci U S A* **99**, 14819(Nov 2002).
11. A. J. Oudes, D. S. Campbell, C. M. Sorensen, L. S. Walashek, L. D. True and A. Y. Liu, *BMC Genomics* **7**, p. 92 (2006).
12. D. L. Wheeler, T. Barrett, D. A. Benson, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. Dicuccio, R. Edgar, S. Federhen, M. Feolo, L. Y. Geer, W. Helmberg, Y. Kapustin, O. Khovayko, D. Landsman, D. J. Lipman, T. L. Madden, D. R. Maglott, V. Miller, J. Ostell, K. D. Pruitt, G. D. Schuler, M. Shumway, E. Sequeira, S. T. Sherry, K. Sirotkin, A. Souvorov, G. Starchenko, R. L. Tatusov, T. A. Tatusova, L. Wagner and E. Yaschenko, *Nucleic Acids Res* **36**, D13(Jan 2008).
13. S. A. Ochsner, H. Strick-Marchand, Q. Qiu, S. Venable, A. Dean, M. Wilde, M. C. Weiss and G. J. Darlington, *Stem Cells* **25**, 2476(Oct 2007).
14. G. Bidaut and M. F. Ochs, *Bioinformatics* **20**, 2869(Nov 2004).
15. L. J. Murray, E. Bruno, N. Uchida, R. Hoffman, R. Nayar, E. L. Yeo, A. C. Schuh and D. R. Sutherland, *Exp Hematol* **27**, 1282(Aug 1999).
16. A. M. Turnley, *Trends Endocrinol Metab* **16**, 53(Mar 2005).
17. X. Ouyang, M. Fujimoto, R. Nakagawa, S. Serada, T. Tanaka, S. Nomura, I. Kawase, T. Kishimoto and T. Naka, *J Cell Physiol* **207**, 428(May 2006).