Identification of differentiation regulators based on transcriptional heterogeneity

転写の不均質性に基づく分化制御因子の特定

Tatsuya Ando

安藤 達哉

Department of Biotechnology, Graduate School of Engineering,

Nagoya University, Japan

名古屋大学大学院 工学研究科 化学・生物工学専攻

生物機能工学分野

Contents

Chapter 1

General Introduction

1.1. Preface

Heterogeneity within cell populations has been difficult to be experimentally proven and neglected because it was often treated as technical noise rather than biological variability in biological experiments [1]. However, heterogeneity in cell populations is not a new concept. It has been studied as clonal populations of bacteria [2]. They described as non-genetic individuality that is variability in cell populations with the same genetic background. Cellular heterogeneity became partially measurable as differences in cellular phenotypes and intermediate phenotypes among

cell populations [3]–[5]. Some studies tried to uncover the sources of heterogeneity within cell populations [6],[7]. One of the simple explanations of the source is that phenotype heterogeneity comes from a mixture of a limited number of distinct subpopulations (Fig. 1-1A) [8]. The cellular states in the subpopulations are not static, which reflects biological complexity. The cellular states are changed or fluctuated dynamically across time due to cyclic biological processes such as cell cycle, circadian rhythm and ultradian rhythm (Fig. 1-1B) [9],[10]. The cyclic states vary across cell populations at any time points, which is another source of the heterogeneity. Although pioneering studies have explored heterogeneity in cell populations using a limited number of preselected genes [11]–[13], the complexity of cellular heterogeneity is not understood well. Understanding of the cellular phenotype heterogeneity and the underlying mechanism is one of the next frontiers of biological science.

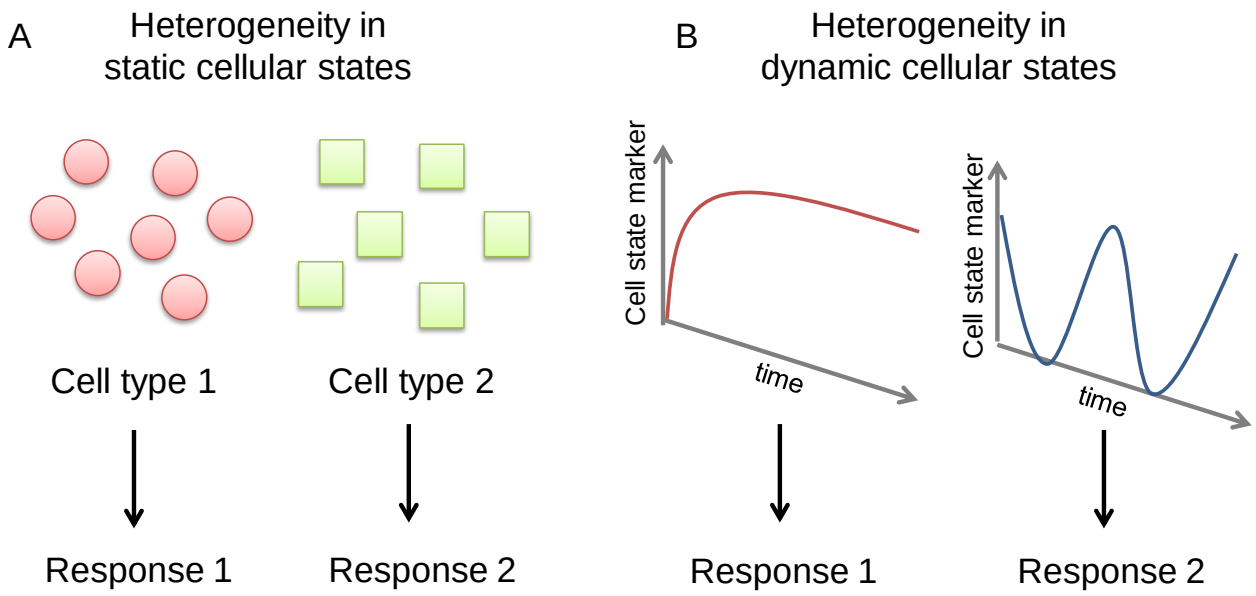


Figure 1-1. The source of phenotype heterogeneity in cell populations. (A) Difference of cellular response in the distinct cell types. **(B)** Difference of cellular response by dynamic change of temporal cellular states

1.2. The current challenges for heterogeneity researches

More recently, phenotype heterogeneity in cell populations has been observed in diverse physiological processes, pathophysiological conditions, and responses to therapeutics [14]–[17], which poses challenges to develop treatment of diseases (Fig. 1-1). Several research steps are considered to be required to achieve homogeneous

response against treatment at a cell population level. The first step is to understand the molecular mechanism underlying cellular heterogeneity. Then, we can identify regulators of cellular phenotype heterogeneity based on the regulatory mechanism. After that, modulating the regulators enables us to regulate cellular phenotypes and have effects on tissues in animals and humans.

One of the application fields is cancer drug development. Drug responses in cancer cells are heterogeneous [16],[17]. Drug-resistant cancer cells prevent effective treatments for cancer patients. Therefore, a deeper understanding of homogeneous drug responses in cancer is beneficial for better drug development for cancers. Another application field is regenerative medicine. It is difficult to supply cells suitable for the personalized cell therapy on demands because cellular differentiation is a time-consuming process, and differentiated cell populations can consist of heterogenic population of functionally undefinable cells which can lead to inefficient therapy [18]. Rapid and homogeneous differentiation systems could improve the efficacy and cost of cell therapy.

Although the molecular mechanism underlying cellular heterogeneity is largely

unknown, multiple clues have discovered in the cell differentiation and development research area. One of the potential mechanisms of cellular heterogeneity is considered to the diversity of cell cycle states. The differences between cell-cycle phases of each cell in a cell population, which considered as "heterogeneity of cell cycles", can inhibit synchronous differentiation of stem cells [19]. The other mechanism is the oscillation of the differentiation regulators in the stem cells [9]. The oscillation of such regulator genes is commonly not synchronized even in individual cells, which makes the cellular heterogeneity more complex [20]. The important implication from these studies is that oscillation phenomenon itself is required for multi-potency and maintenance of stem cells [21]. Although these cell differentiation studies shed light on heterogeneity mechanism from the limited aspects, systematic and unbiased researches on cellular heterogeneity could not be performed until measurement and analytical methodologies have recently developed.

1.3. The current opportunities for heterogeneity analysis

The recent development of both measurement technologies and analytical theory has been accelerating the science of cellular heterogeneity. The first wave of advance in measurement technology was the development and spread of Omics technologies [22]. The genome sequences of humans and mice were reported and genes were annotated on the genomes with massive transcript sequences [23]–[25]. It enabled us to design the synthetic oligonucleotides which hybridize to the target transcripts and measure transcripts in genome-wide manner (transcriptome) [22]. The explosive development of sequencing technology offered an opportunity for quantitative measurement of transcripts by sequencing a huge number of RNAs and mapping to the genomes [26]. By applying these sequence technologies, epigenomics studies have been increased. The recent advances in mass spectrometry technology expanded the measurable molecules at once in proteomics and metabolomics experiments [27],[28]. Among these omics technologies, the transcriptomics technology was advancing with its quantitative determination performance, reproducibility and low error rate [29].

The second wave was an invention of single cell RNA-seq, which enabled quantitative genome-wide measurement of transcripts at the individual cellular level

compared to the previous population average level [30]. Transcriptional variability in cell populations became measurable with higher precision, which contributed to establishing an accurate evaluation of cellular heterogeneity [31],[32]. Single cell technology is now applying to epigenomics [33]. These applications will evaluate cellular heterogeneity from multiple layers of intermediates phenotypes such as transcriptomics and epigenomics.

In parallel with the development of measurement technology, analytical theories have been developed. Dynamical systems theory treats heterogeneity and oscillation as one of the patterns of fluctuations [34],[35]. The theory provides a uniform framework to describe an abrupt shift of a complex dynamic system from one state to another [35]. This “abrupt shift”, designated as “critical transition”, can be seen in various types of systems around us, not only in the cellular research field but also in other scientific fields: asthma attacks or epileptic seizures in our body system; systemic market crashes in finance in the economic system; changes in ocean circulation or climate on earth as a global system [36]–[38].

Historically, dynamical systems theory has suggested indicators that the

system state starts to transit to another, especially just before the critical transition. Such indicators are known as 'critical slowing down' [39]. It was also suggested that the slowing down should lead to some pattern of certain characteristic changes of fluctuations in a system. There are two possible characteristics suggested being strengthened in such change of fluctuations accompanied by critical slowing down (Fig. 1-2). One of the possible characteristic change is the increase of variance in the pattern of fluctuation [40], and the other is the increase of correlation between neighboring components [41]. A highly connected system shows an abrupt change by a strong perturbation because all components shift their state in synchrony. Such possible characteristics in the fluctuation can be mathematically shown [41]. In cellular biology, the components of a system are the genes/transcripts in the cell population. Therefore, as a hypothesis, the expression pattern of genes can show high variability (variance) and/or correlation. As a practical example, the Notch signaling genes (Hes1 and Dll1) are known to oscillate, and co-expressed before neuronal differentiation in the stem cells [20]. The oscillation is not synchronized among individual cells and could show high variability in the cells at a certain time point before differentiation. he analogy

between dynamical systems theory and the biological observation in the neuronal differentiation motivated us to further examine the dynamic state transition in the cellular system from the view point of cellura heterogeneity. The recent accumulation of transcriptome data in the cell differentiation process also provide us an opportunity for unbiased analysis of cell state transition with cellular heterogeneity.

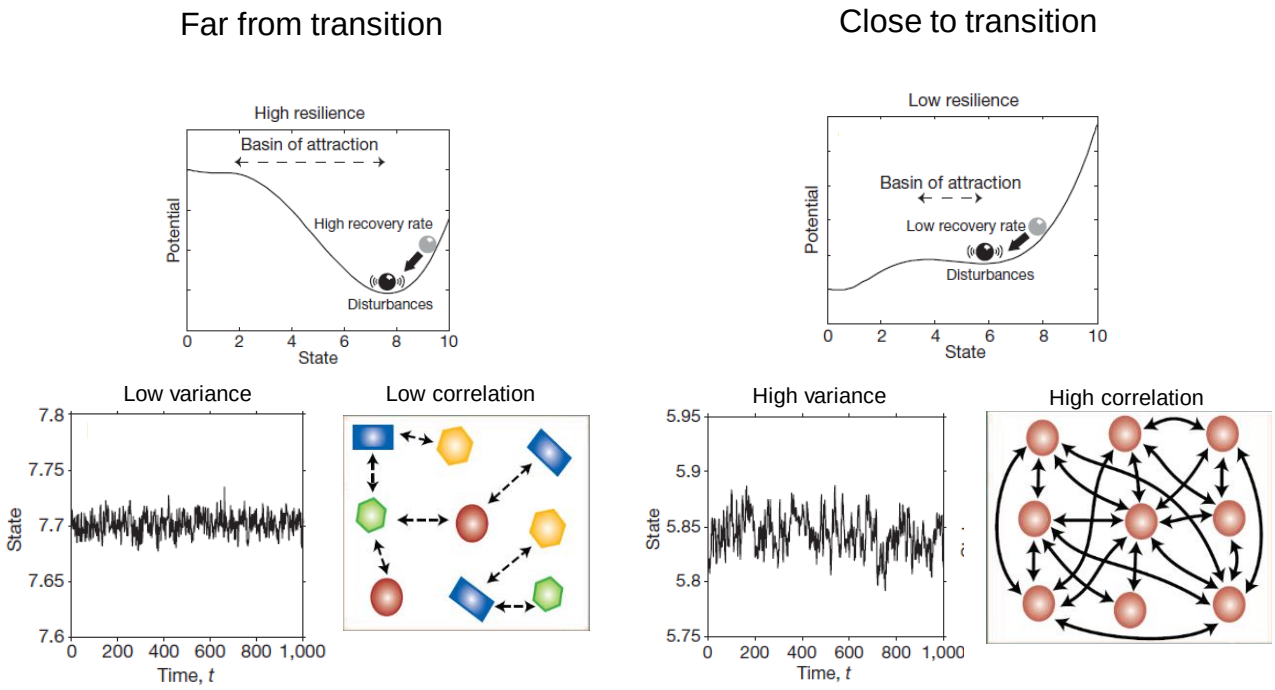


Figure 1-2. Characteristic changes as a system approaches a critical transition. Far from the transition, the system is characterized by low variance of fluctuations and low correlation between the system components. When the system is close to the transition, the system shows high variance of fluctuations and high correlation between the

system components as a consequence of the slowing down. The figure was modified from Scheffer et al. [34], [35].

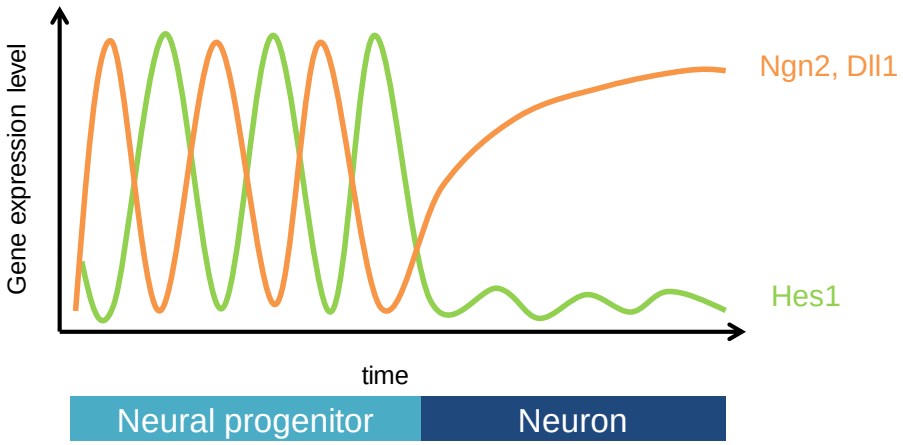


Figure 1-3. The expression profile of Notch signaling genes during neuron differentiation in a single cell. Expression of Hes1, Ngn2 and Dll1 oscillates in dividing neural progenitors. In neurons, Hes1 is downregulated, whereas Ngn2 and Dll1 are upregulated. The figure was modified from Shimojo et al. [20].

1.4. Aim of this research

In this work, I aimed to identify the regulators in a cellular system, especially in the neuronal and glial differentiation, by investigating transcriptional variability

and co-regulation pattern of global gene expressions. The reason why I selected neuronal and glial differentiation is the accumulated biological knowledge from the pioneer studies and its feasibility of unbiased analysis with accumulated transcriptome data [21]. For the biologically validated target gene in this field, Notch signaling genes including Hes1 was found to oscillate during the neuronal differentiation from the prior works. Multiple transcriptome data sets in the neuronal and glial differentiation are also required to replicates the results from the unbiased analysis [42]–[45], which was another factor made me focus on this differentiation process.

I examined fundamental characteristics of transcriptional variability during neuronal and glial differentiation aiming to identify the novel regulators in the neuronal and glial differentiation process as a primary focus of this thesis. In Chapter 2, I examined the cellular heterogeneity in neuronal differentiation. The primary purpose of this first study is to examine the existence of events according to the dynamical systems theory, by the variability analysis of transcriptome data reflecting the “averaged expressions of cell population” measured by microarray. After such evaluation, I also tried to identify the key neuronal differentiation regulators based on

the regulatory characteristics of highly variable genes. The preservation of the characteristics of identified regulators between humans and mice was also evaluated. In Chapter 3, the transcriptional heterogeneity in the astrocyte differentiation was examined. In this chapter, I expanded my formerly examined analysis concept, the evaluation of transcriptional variability analysis, to identify the differentiation regulators from the single cell RNA-seq the data, which can describe the heterogeneity of cell population in more detail. Through this work, I investigated the similarities of transcriptional variability changes between neuronal and astrocyte differentiations. The potential regulators in astrocyte differentiation could also be identified similarly to the analytical methods established in the Chapter 2. In this work, not only the data analysis but also the experimental validation was conducted to prove the potency of the identified regulator from the analysis.

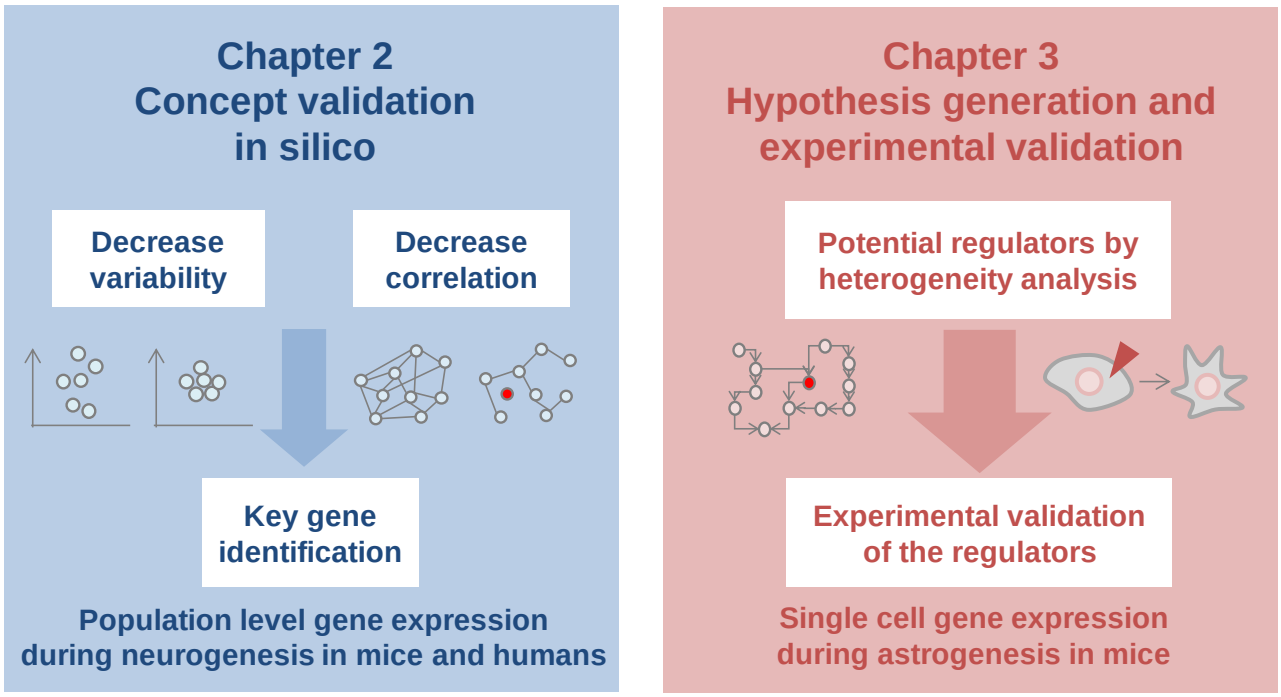


Figure 1-4. Overview of this thesis. In Chapter 2, I identified key genes of neuron differentiation based on transcriptional variability and correlation using population level gene expression data during neurogenesis in mice and humans. In Chapter 3, I identified potential regulators of astrocyte differentiation by the method established in Chapter 2 using single cell gene expression data during mouse astrogenesis and validate them experimentally.

I believe that the thesis opens new era of cellular heterogeneity research in cell differentiation and help us understand the molecular mechanism underlying cellular heterogeneity. I also believe that it will also contribute to improving the therapeutic

effects in the disease treatment and cell therapy process by controlling the cellular heterogeneity based on the findings from this work.

1.5. References

1. Huang S. Non-genetic heterogeneity of cells in development: more than just noise. Development 2009; 136: 3853–62.

2. Spudich JL KDJ. Non-genetic individuality: chance in the single cell. Nature 1976; 262: 467–471.

3. Levsky JM, Shenoy SM, Pezo RC, Singer RH. Single-cell gene expression profiling. Science 2002; 297: 836–40.

4. Takasuka N, White MR, Wood CD, Robertson WR, Davis JR. Dynamic changes in prolactin promoter activation in individual living lactotrophic cells. Endocrinology 1998; 139: 1361–8.

5. Warren L, Bryder D, Weissman IL, Quake SR. Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR. Proc Natl Acad Sci U S A 2006; 103: 17807–12.

6. Raj A, van Oudenaarden A. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. Cell 2008; 135: 216–226.

7 . Kaern M, Elston TC, Blake WJ, Collins JJ. Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet 2005; 6: 451–64.

8 . Slack MD, Martinez ED, Wu LF, Altschuler SJ. Characterizing heterogeneous cellular responses to perturbations. Proc Natl Acad Sci U S A 2008; 105: 19306–11.

9 . Hirata H, Yoshiura S, Ohtsuka T, Bessho Y, Harada T, Yoshikawa K, Kageyama R. Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop. Science 2002; 298: 840–3.

10 . Purvis JE, Lahav G. Encoding and decoding cellular information through signaling dynamics. Cell 2013; 152: 945–56.

11 . Tay S, Hughey JJ, Lee TK, Lipniacki T, Quake SR, Covert MW. Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. Nature 2010; 466: 267–71.

12 . Spencer SL, Gaudet S, Albeck JG, Burke JM, Sorger PK. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. Nature 2009; 459: 428–32.

13 . Feinerman O, Jentsch G, Tkach KE, Coward JW, Hathorn MM, Sneddon MW,

Emonet T, Smith KA, Altan-Bonnet G. Single-cell quantification of IL-2 response by effector and regulatory T cells reveals critical plasticity in immune response. Mol Syst Biol 2010; 6: 437.

14 . Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature 2008; 453: 544–7.

15 . Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, Pesce CG, Brent R. Regulated cell-to-cell variation in a cell-fate decision system. Nature 2005; 437: 699–706.

16 . Campbell LL, Polyak K. Breast tumor heterogeneity: cancer stem cells or clonal evolution? Cell Cycle 2007; 6: 2332–8.

17 . Gascoigne KE, Taylor SS. Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs. Cancer Cell 2008; 14: 111–22.

18 . Heathman TRJ, Nienow AW, McCall MJ, Coopman K, Kara B, Hewitt CJ. The translation of cell-based therapies: clinical landscape and manufacturing

challenges. Regen Med 2015; 10: 49–64.

19 . Dalton S. Linking the Cell Cycle to Cell Fate Decisions. Trends Cell Biol 2015; 25: 592–600.

20 . Shimojo H, Ohtsuka T, Kageyama R. Oscillations in notch signaling regulate maintenance of neural progenitors. Neuron 2008; 58: 52–64.

21 . Imayoshi I, Isomura A, Harima Y, Kawaguchi K, Kori H, Miyachi H, Fujiwara T, Ishidate F, Kageyama R. Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. Science 2013; 342: 1203–8.

22 . Nelson NJ. Microarrays have arrived: gene expression tool matures. J Natl Cancer Inst 2001; 93: 492–4.

23 . Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M,

Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, et al. The sequence of the human genome. Science 2001; 291: 1304–51.

24 . Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409: 860–921.

25 . Mouse Genome Sequencing Consortium, Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG,

Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, et al. Initial sequencing and comparative analysis of the mouse genome. Nature 2002; 420: 520–62.

26. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009; 10: 57–63.

27. Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics 2008; 92: 255–64.

28. Griffiths WJ, Wang Y. Mass spectrometry: from proteomics to metabolomics and lipidomics. Chem Soc Rev 2009; 38: 1882–96.

29. SEQC/MAQC-III Consortium. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. Nat Biotechnol 2014; 32: 903–14.

30. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 2013; 14: 618–30.

31 . Shalek AK, Satija R, Shuga J, Trombetta JJ, Gennert D, Lu D, Chen P, Gertner RS, Gaublomme JT, Yosef N, Schwartz S, Fowler B, Weaver S, Wang J, Wang X, Ding R, Raychowdhury R, Friedman N, Hacohen N, Park H, May AP, Regev A. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature 2014; 510: 363–9.

32 . Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, Schwartz S, Yosef N, Malboeuf C, Lu D, Trombetta JJ, Gennert D, Gnirke A, Goren A, Hacohen N, Levin JZ, Park H, Regev A. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 2013; 498: 236–40.

33 . Schwartzman O, Tanay A. Single-cell epigenomics: techniques and emerging applications. Nat Rev Genet 2015; 16: 716–26.

34 . Scheffer M, Carpenter SR, Lenton TM, Bascompte J, Brock W, Dakos V, van de Koppel J, van de Leemput I a, Levin S a, van Nes EH, Pascual M, Vandermeer J. Anticipating critical transitions. Science 2012; 338: 344–8.

35 . Scheffer M, Bascompte J, Brock W a, Brovkin V, Carpenter SR, Dakos V, Held H,

van Nes EH, Rietkerk M, Sugihara G. Early-warning signals for critical transitions. Nature 2009; 461: 53–9.

36 . Venegas JG, Winkler T, Musch G, Vidal Melo MF, Layfield D, Tgavalekos N, Fischman AJ, Callahan RJ, Bellani G, Harris RS. Self-organized patchiness in asthma as a prelude to catastrophic shifts. Nature 2005; 434: 777–82.

37 . Litt B, Esteller R, Echauz J, D'Alessandro M, Shor R, Henry T, Pennell P, Epstein C, Bakay R, Dichter M, Vachtsevanos G. Epileptic seizures may begin hours in advance of clinical onset: a report of five patients. Neuron 2001; 30: 51–64.

38 . Lenton TM, Held H, Kriegler E, Hall JW, Lucht W, Rahmstorf S, Schellnhuber HJ. Tipping elements in the Earth's climate system. Proc Natl Acad Sci U S A 2008; 105: 1786–93.

39 . Wissel C. A universal law of the characteristic return time near thresholds. Oecologia 1984; 65: 101–107.

40 . Carpenter SR, Brock WA. Rising variance: a leading indicator of ecological transition. Ecol Lett 2006; 9: 311–8.

41 . Ives AR. Measuring Resilience in Stochastic Systems. Ecol Monogr 1995; 65: 217–233.

42 . Theunissen PT, Pennings JL a, Robinson JF, Claessen SMH, Kleinjans JCS, Piersma AH. Time-response evaluation by transcriptomics of methylmercury effects on neural differentiation of murine embryonic stem cells. Toxicol Sci 2011; 122: 437–47.

43 . Paşca SP, Portmann T, Voineagu I, Yazawa M, Shcheglovitov A, Paşca AM, Cord B, Palmer TD, Chikahisa S, Nishino S, Bernstein J a, Hallmayer J, Geschwind DH, Dolmetsch RE. Using iPSC-derived neurons to uncover cellular phenotypes associated with Timothy syndrome. Nat Med 2011; 17: 1657–62.

44 . Llorens-Bobadilla E, Zhao S, Baser A, Saiz-Castro G, Zwadlo K, Martin-Villalba A. Single-Cell Transcriptomics Reveals a Population of Dormant Neural Stem Cells that Become Activated upon Brain Injury. Cell Stem Cell 2015; 17: 329–340.

45 . Dulken BW, Leeman DS, Boutet SC, Hebestreit K, Brunet A. Single-Cell Transcriptomic Analysis Defines Heterogeneity and Transcriptional Dynamics in

the Adult Neural Stem Cell Lineage. Cell Rep 2017; 18: 777–790.

Chapter 2

Differential variability and correlation of gene expression identifies key genes involved in neuronal differentiation

2.1. Introduction

Cell differentiation is a complex process requiring precise dynamic regulation of cellular components. The spatio-temporal heterogeneity of ES cells and iPS cells makes it hard to determine the molecular mechanisms of cell differentiation and establish efficient differentiation protocols [1,2]. From the standpoint of dynamic systems theory, differentiation processes, like societies and ecological or biological networks, are

systems that shifts abruptly from one state to another, often in response to external stimulation; such shifts are referred to as “critical transitions” [3-5]. ES and iPS cells are in a balanced stable state that can shift to multiple other states representing differentiated cell types [5,6]. Although it may exhibit little change beforehand, a system close to a critical transition usually shows signs of fragility. For example, high variance and correlation of system components are empirical indicators of upcoming transitions [4]. Thus, the mRNA and protein components of a gene regulatory network may exhibit highly variable expression and correlated expression patterns prior to ES and iPS cell differentiation. Such variance is thought to be related to spatio-temporal fluctuation (“noise”) in gene expression [7]. Controlled temporal fluctuation or oscillation is required for maintenance of stem cell self-renewal [8]. Some gene sets related to self-renewal not only oscillate, but are also co-expressed during neural differentiation [9]. The co-expression network provides a comprehensive picture of the correlation relationships between gene products and reveals the functional organization of the transcriptome [10,11]. The structure of the transcriptional regulatory network may be altered during ES cell differentiation. Fluctuations in the

levels of important differentiation regulators may affect network structure, thereby controlling cell fate decisions and population heterogeneity [12]. Recently developed systematic approaches can identify such changes in network structure[13]. These approaches identify regulators or marker genes of disease pathophysiology by comparing gene network structures and variances of the genes in the network between healthy and disease progression states [14]–[20]. Application of these methods to ES and iPS cell may reveal alterations in network structures and variance in gene expression levels during ES cell differentiation process, ultimately leading to identification of important regulators of differentiation. In support of this idea, several known regulators exhibit co-regulated fluctuations during differentiation [9]. Understanding of these gene regulatory networks could help dissect the complex molecular mechanisms underlying stem cell biology. Although pioneering work has revealed the behaviors of dynamic gene fluctuations, especially in developmental biology, genome-wide discovery of genes exhibiting dynamic fluctuation during differentiation has not been comprehensively performed. In this study, we developed an analytical framework for investigating the dynamics of transcriptional networks and

applied it to the differentiation processes of ES and iPS cells. Examination of the gene expression profile during mouse neural differentiation revealed that the variability of a group of genes that were co-expressed in the undifferentiated state decreased after neural differentiation. We then ranked the individual genes using an integrative scoring method (Fig. 1) that simultaneously assessed the changes in gene expression variances and co-expression relationships between the undifferentiated and differentiated states. This analysis identified 671 highly ranked genes, including Hes1, previously shown to oscillate prior to neural differentiation. The common biological functions among these genes are related to neural differentiation, and act downstream of pluripotency-related transcription factors. This group was also enriched in genes that cause phenotypic alternations of developmental processes in KO/Tg mice. We demonstrated that these genes significantly overlapped with the set of genes that exhibited differential variance and correlation during neural differentiation of human iPS cells. This study suggests that analysis of network dynamics can be used to identify genes important for the differentiation process, as well as yield insights into dynamic molecular mechanisms in both mice and humans.

2.2. Materials and Methods

2.2.1 Transcriptional data

Transcriptional data of neural differentiation in mouse ES cells were obtained from the ArrayExpress database (E-TABM-1108) [21]. Transcriptional data of neural differentiation in human iPS cells were obtained from the Gene Expression Omnibus (GSE25542) [22]. The raw transcriptional data were $\log_2$ transformed and subjected to quantile normalization. Probes corresponding to genes that were expressed (presence call > 50%) and exhibited variance (SD > 0) across replicates in each time point or differentiation stage were used in the analysis. Differential co-expression and variation analyses were performed based on 30,035 and 12,364 probes from mouse ES cell and human iPS cell data, respectively. The gene expression signals were standardized to the Z-score (average = 0, SD =1) for each gene across replicates of each time point or differentiation stage.

2.2.2. Differential variance and co-expression analysis

The differential co-expression and variation analysis was conducted using the Bioconductor package in the R language [23]. To identify modules, hierarchical clustering was applied to the standardized expression values from mouse ES cells at day 0 or from human iPS cells. The hierarchical clustering was performed based on the Pearson correlation coefficient (PCC) and average linkage method. Modules were detected using a dynamic tree-cutting algorithm (hybrid mode, minimal module size of 100). In each module, the average PCC of each gene with other genes in the module was calculated. Differential correlation was calculated as the absolute value of the difference between PCC after day 0 and PCC at day 0 in the module as defined on day 0. The average SD of each gene among replicates in each time point or cellular state was calculated. The differential variance was calculated as the absolute value of the difference between SD after day 0 and SD at day 0. The system transition score, based on a previously described composite index, was used to rank the genes and identify those with high differential correlation and variance during neural differentiation [15]. This score was calculated using in the following formula:

$$\max\left\{\log\left(\frac{PCC * \frac{SD}{OPCC} \quad at \quad day \quad 0}{PCC * \frac{SD}{OPCC} \quad at \quad day \quad 3,4,5,6,7}\right)\right\}$$

where OPCC is the average PCC of each gene with the genes outside the module. Highly ranked genes were defined those whose scores were 2 SDs higher than the average score over all genes. Highly ranked genes in the same module displayed in network representation using Cytoscape 3.1.2 [24]. Correlation coefficients above 0.95 are shown as connections in the network visualization figure.

2.2.3. Enrichment analysis

Functional enrichment and upstream regulator analysis was performed using Ingenuity Pathways Analysis (IPA®, Qiagen, www.qiagen.com/ingenuity) software. Genes associated with differentiation-related phenotypes in knockout and transgenic mice were identified based on the Mouse Genome Informatics database [25]. Enrichment of genes associated with differentiation-related phenotypes among DVC genes was assessed by cumulative hypergeometric probability using the *phyper* function in R.

2.2.4. Statistical analysis

Welch's t-test was applied to transcriptional data to identify genes differentially expressed between day 0 and subsequent days during neural differentiation process. F-test was carried out to evaluate differential variances of genes between day 0 and subsequent days. P-values were adjusted by the Benjamini–Hochberg method [26]. Differentially expressed genes were defined as those whose fold changes were more than 2 SDs higher than the mean of all genes, with adjusted p-value < 0.05, as in the definition of DVC genes. To compare the mouse ES cell and human iPS cell data, mouse orthologs of the DVC genes from the human iPS cell study were identified based on information in the HUGO Comparison of Orthology Predictions database [27]. Overlap analysis of DVC genes in mice and human was performed by hypergeometric test, as in the enrichment analysis.

2.3. Results

2.3.1. Network dynamics detection in the differentiation process

We analyzed microarray data collected at six time points (days 0, 3, 4, 5, 6, and 7) from mouse ES cells undergoing neural differentiation; each time point was measured in eight replicates (E-TABM-1108) [21]. Two indicators were calculated to identify genes whose expression patterns predicted the transition from the undifferentiated state to the neural lineage (Fig. 1). One indicator was differential variance, representing the difference in gene expression variance between day 0 and each time point after day. The other was differential correlation, representing the difference between the average value of correlations within co-expressed gene sets at day 0 and those on subsequent days. These co-expressed gene sets, so-called 'modules', were defined at day 0 using hierarchal clustering. A system transition score was assigned to each gene by combining the differential variance and differential correlation (Fig. 2-1 and Methods). Hierarchical clustering analysis of gene expression data from ES cells at day 0 revealed 76 modules (Fig. 2-2). Comparisons of expression

variances between day 0 and subsequent days identified 315 differentially variable genes. Genes for which the variances decreased after day 0 were the most significantly enriched in the skyblue module (colored sky blue in Figure 2), which contained the most highly correlated genes at day 0 ($p = 3.61e\text{-}50$). This result indicates that genes specifically expressed in ES cells in the undifferentiated state were more variable than those expressed in differentiating cells, even though it is likely that multiple types of cells are present in the population during neural differentiation. On the other hand, the genes that were differentially expressed between day 0 and subsequent days were not enriched in the skyblue module (Fig. 2) as much as genes with differential variance. As noted above, high differential variance and correlation are observed in fragile systems before a critical transition [4], and these features may represent an early warning signal of imminent differentiation.

2.3.2. Genes with differential variance and correlation are involved in neural differentiation

To identify fluctuating genes that contribute to dynamic changes in the

transcriptional network, we ranked all genes based on their system transition scores. We referred to the highly ranked genes as DVC (differential variance and correlation) genes. We detected 671 such genes. One of the top-ranked DVC genes, Hes1, exhibited high variance at day 0 (Fig. 2-3C). Co-expression relationships with Hes1 were diminished at day 4 relative to day 0 (Fig. 2-3D). Functional analysis of the DVC genes revealed that genes involved in body axis development, neuron movement, and transcription were enriched among the DVC genes (Fig. 2-3A). There were 822 differentially expressed genes that satisfied (1) $p < 0.05$ by t-test, and (2) fold change > mean + 2SD (standard deviation), when their mean expression levels were tested between differentiating states (Days 3, 4, 5, 6, and 7) and Day 0. Compared to the functions enriched in the DVC genes, the enriched functions of the differentially expressed genes included cancer-related functions (malignant solid tumor, proliferation of cells, digestive tumor, cell death of tumor). Moreover, the DVC and differentially expressed genes did not significantly overlap ($p = 0.793$), but the DVC genes did significantly overlap with the set of genes (Fig. 2-3B) regulated by the Yamanaka factors (Myc, Sox2, and Pou5f1(Oct4)) expressed at the highest levels in the

undifferentiated state (Day 0). Hoxb3, Fgf4, and Pax6 are downstream of both Pou5f1 (Oct4) and Sox2. These results suggest that the DVC genes not only represent early warning signals for neural differentiation, but are also functionally involved in the differentiation process.

2.3.3 Gene expression variance and correlation are altered at a neuroectodermal stage

Next, we assessed the expression patterns of differentiation marker genes in order to understand when during the neuronal differentiation process their individual gene expression variance and co-expression relationship with associating genes changed. The cell populations initially expressed markers of the undifferentiated state, such as Pou5f1(Oct4) and Nanog, and began to express the primitive ectoderm marker Fgf5 on day 3 (Fig. 2-4). The early neuroectodermal marker Sox1 was expressed after day 4, and the neural markers Ascl1 (Mash1) and Tubb3 (Tuj1) were elevated after day 5 under neural differentiation conditions. Over half (55.7%) of the DVC genes exhibited their largest changes in variability and correlation at Day 4. For instance, the expression variance of Hes1 decreased after day 4, when Sox1 was dramatically

up-regulated. The co-expression networks of six time points show that the correlations between the genes in the skyblue module decreased the most at day 4. Because the early neuroectodermal marker Sox1 was highly expressed at Day 4, we believe that this time point represents a neuroectodermal stage. Based on these results, we infer that the transition to the neuroectodermal stage may involve an abrupt genetic system shift in differentiating cells.

2.3.4 DVC genes are genetically associated with the differentiation process

We next performed enrichment analyses to determine whether the DVC genes play important roles in differentiation processes. To this end, we compared the DVC with the genes associated with differentiation-related phenotypes in knockout and transgenic mice. This analysis revealed that genes involved in embryogenesis, embryonic lethality, neural differentiation and neural progenitor cell differentiation were significantly more enriched among the DVC genes than among the differentially expressed genes (Fig. 2-5). Thus, the DVC genes could play important roles in neural differentiation. In other words, the genetic factors underlying neural differentiation

could control differentiation by affecting the transcriptional network and its fluctuations, which are measured as dynamic changes in correlations and variances.

2.3.5. DVC genes conserved in mouse ES and human iPS cells

Next, we assessed the conservation between humans and mice of differential variance and correlation during neural differentiation. To this end, the analysis used to detect network dynamics was applied to gene expression data collected at three differentiating states during neural differentiation of human iPS cells: iPS cells, neural precursor cells, and neurons [22]. We identified 284 DVC genes in neural differentiation of human iPS cells, which overlapped significantly with those in mice ($p = 0.0204$, Fig. 2-6A); for example, Hes1 and Ccng2 in the mouse co-expression network were also identified as DVC genes in neural differentiation of human iPS cells (Fig. 2-3D). Proliferation- and morphology-related genes were over-represented both in mouse and human (Fig. 2-6B). Furthermore, in both species, DVC genes were commonly regulated by MYC and SOX1 (Fig. 2-6C). For example, SOX1 is a transcriptional factor for HES1 [28], and the abrupt up-regulation of SOX1 may drive

the dynamic changes of HES1 expression and co-expression relationships at day 4. Both genes play important roles in stem cell maintenance [29].

2.4. Discussion

We carried out the first genome-wide analysis aimed at detecting dynamical changes in gene-expression variance and co-expression relationships during neural differentiation of mouse ES and human iPS cells. Our results demonstrate that genes that were highly correlated in ES cells exhibited significant changes in expression variance. Functional analysis revealed that genes exhibiting both differential variances and differential correlations may encode the regulators of neural differentiation. Although differentially expressed genes are normally used to identify genes that play important roles in differentiation [24-26], analysis of network dynamics allows us to identify potential key regulators that cannot be detected by differential expression analysis.

The DVC genes tended to be downstream of the Yamanaka factors (Fig 2-3), and could

therefore be related to self-renewal of stem cells and maintenance of pluripotency. Hoxb3, Fgf4, and Pax6 are downstream of both Pou5f1 (Oct4) and Sox2. Hoxb3 plays a role in maintaining self-renewal [33], and Fgf4 is involved in pluripotency [34]. On the other hand, Pax6 is a master regulator of neuronal differentiation [35]. Sox1, one of the common transcription factors upstream of the DCV genes in both the mouse and human studies (Fig 2-6), promotes neurogenesis[28]. Hes1 and Pitx2 are downstream of Sox1. Hes1-null mice exhibit premature neurogenesis and neural tube defects during embryogenesis [36]. The Hes1 protein is a transcriptional repressor that inhibits differentiation of ES cells into the neural lineage and delays mesoderm and endoderm differentiation [37]. Pitx2 is important for mesodermal and neuroectodermal development *in vivo* [38]. These results suggest that the DVC genes include not only pluripotency genes but also the genes specifically related to neuronal differentiation.

SOX1 also acts upstream of DVC genes in human iPS cells. HES1, FRZB, and WLS are common DVC genes downstream of SOX1 in both humans and mice. Reduction in FRZB expression is required for neural progenitor proliferation and the acceleration of neuron development [39]. FRZB can bind extracellular Wnt and inhibits Wnt signaling.

WLS is important for Wnt-mediated neuronal development [40]. These genes might have regulatory links and be involved in a neuronal differentiation.

HES1 expression is dynamically regulated during neuronal differentiation. In particular, Hes1 expression oscillates in ES cells and neural progenitor cells, and is transiently down-regulated during the transition to epiblast stem cells (epiSCs) in neural and non-neural lineages [9,28,29]. The heterogeneity of Hes1 expression was greater in the ES cell population than in epiSCs. This heterogeneity may be due to oscillatory expression, observed as expression variances in ES cell populations. Our analysis was able to detect the change in expression variances during the transition from ES cells to the neural lineage.

Inactivation of HES1 in ES cells promotes rapid and homogeneous differentiation into neural progenitors [41]. CCNG2 was also identified as a DVC gene in both mouse ES and human iPS cells. Both Hes1 and Ccng2 prolong G1 phase to reduce cell proliferation [30,31]. Smad2 is a transcriptional regulator of Hes1 and Ccng2, and the inhibition of Smad2 promotes immediate differentiation into functional neurons [44]. These findings suggest that modulation of the DVC genes might facilitate the

development of experimental protocols for rapid and homogeneous differentiation of neurons. Because the heterogeneities of ES and iPS cells preclude the use of these materials as a stable supply of rapidly differentiated neurons, such an approach could contribute greatly to the understanding of brain functions and the development of regenerative medicine. These heterogeneities could arise in part from the dynamical behaviors of cellular components such as protein expression and localization. Recently, a growing number of attempts have been made to control dynamical patterns by targeted perturbations using chemical compounds and other interventions [45]. Perturbation of the genes identified by the methods in this study could contribute further to understanding of the molecular basis of stem cell differentiation. Moreover, dynamical regulations of the genes by such perturbations might control neural differentiation. Specific dynamical patterns are associated with various cellular responses such as apoptosis and immune response; therefore, the application of this method to other biological responses could identify important regulators of specific cellular responses.

However, it remains challenging to identify dynamic changes in molecular networks.

The approach used in this study might not detect all the early warning signals of upcoming transitions. For example, genes previously shown to exhibit oscillatory expression in stem cells, such as Ascl1 and Dll1, were not identified in this study. One possible reason for this is that we measured mean values of gene expression in cell populations, potentially resulting in underestimation of gene expression variance among individual cells. A single-cell analysis of gene-expression profiles with a large number of replicates would help us to observe the transcriptional distribution of each gene across individual cells. Characterizing the transcriptional distribution in this manner could provide more accurate estimation of a gene expression variance between cellular states.

There are, of course, genetic and epigenetic differences between humans and mice. For example, the molecular machineries that maintain the stemness of ES and iPS cells are not completely identical [26,33]. In addition, various human iPS cell lines come from different genetic backgrounds [22]. Despite such differences, a significant number of common DVC genes were identified during neural differentiation in mouse ES and human iPS cells. This finding supports the idea that the machineries responsible for

dynamic changes in gene expression variances and correlations during neural differentiation are conserved between humans and mice.

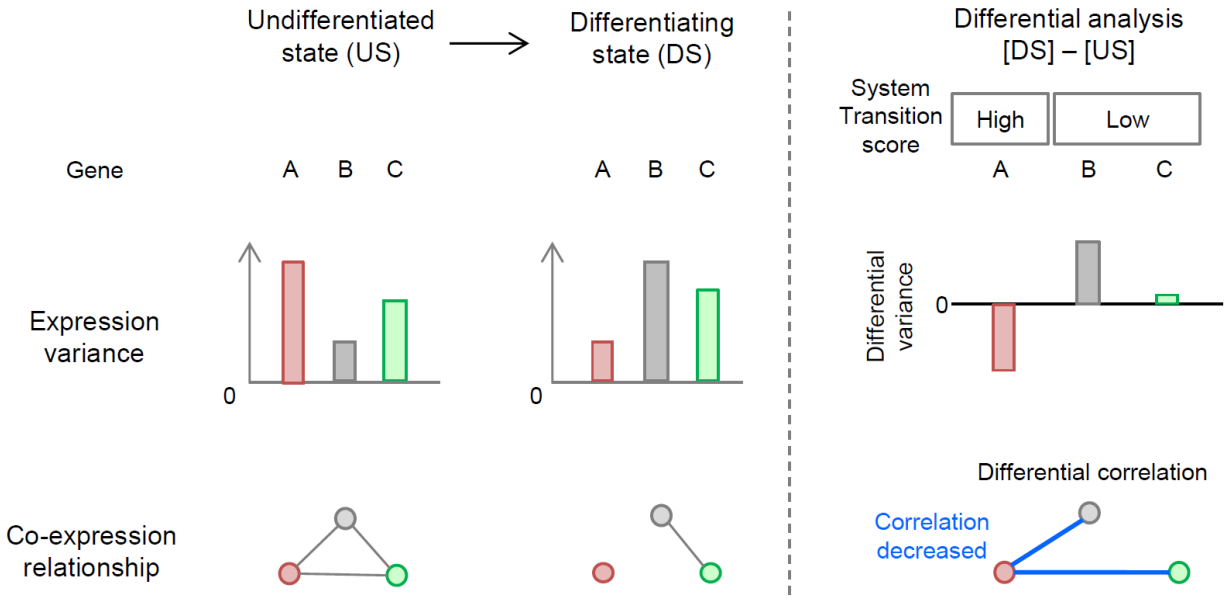


Figure 2-1. Methodological overview of system transition scoring based on network dynamics. Two indexes were used for system transition scoring: gene expression variance within replicates in each cellular state, and the co-expression relationship between genes in each cellular state. After calculating these indexes for each state, a differential analysis was performed to compare the indexes of the undifferentiated and differentiated states. A conceptual gene expression variance and co-expression network of three genes (genes A, B, and C) is shown. In the left panel, gene A (in red) exhibits

the largest change in expression variance within replicates at the undifferentiated state (US), however the variance greatly decreases following the shift to the differentiating state (DS). This change in variance is defined as differential variance. Co-expression relationship (grey lines connecting three genes) is defined as the correlation between genes within replicates; therefore, the relationship between gene A and gene B/C diminishes in the DS. Such a correlation difference between states is defined as differential correlation (blue line indicates “decrease” of correlation” in the right panel). When both the differential variance and differential correlation are large, the system transition score is high. Gene B (in grey) and gene C (in green) are member genes that co-express with gene A in the undifferentiated state. Although gene B exhibits differential variance, the differential correlation of gene B is smaller than that of gene A. The differential variance of gene C is much smaller than that of gene A. The system transition scores of genes B and C are lower than that of gene A.

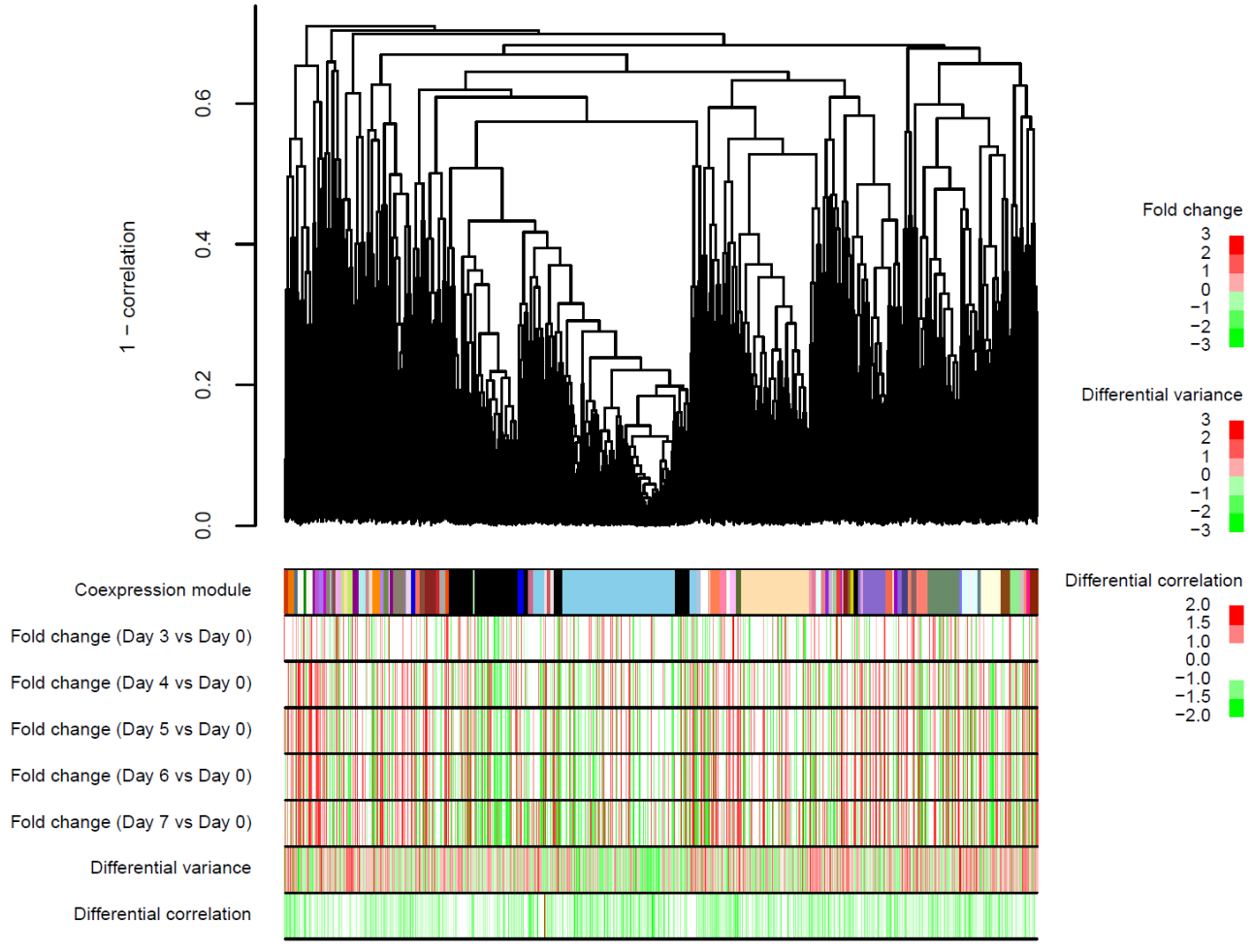
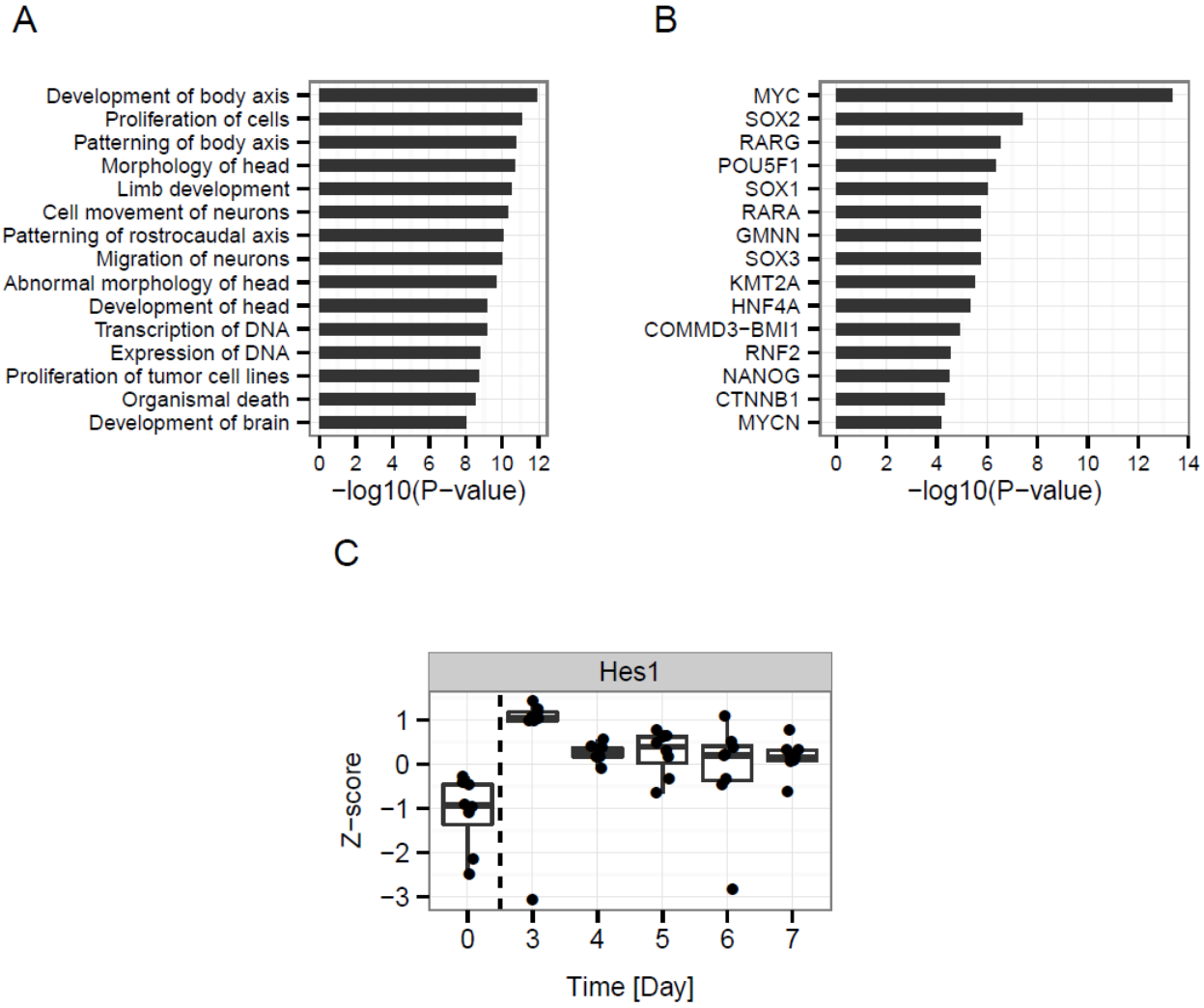


Figure 2-2. Co-expression modules identified in the undifferentiated state (day 0) in mouse ES cells. Dendrogram shows modules of co-expressed genes identified by

hierarchical clustering of gene expression profiles on day 0. Individual colors represent single modules. The heatmap indicates fold change, differential variance, and differential correlation between day 3/4/5/6/7 and day 0. Differential variance indicates the absolute value of the difference between SD after day 0 and SD on day 0. Differential correlation indicates the absolute value of the difference between PCC after day 0 and PCC on day 0.



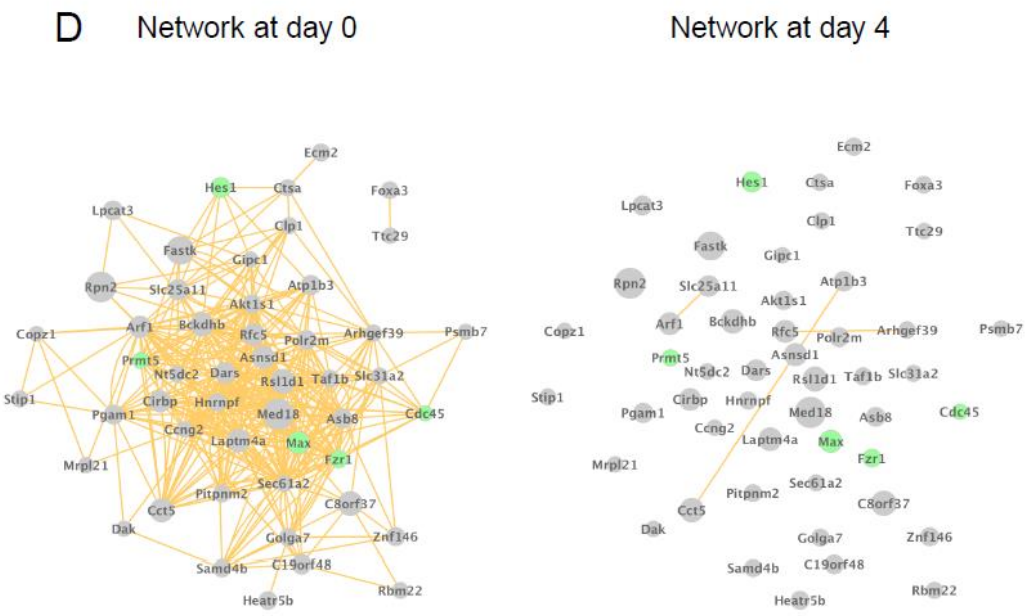


Figure 2-3. Functional analysis of DVC genes. **(A)** Functional categories enriched in the DVC genes during neural differentiation. **(B)** Upstream regulator analysis revealed transcriptional factors that could regulate the DVC genes. **(C)** Expression profile of a representative DVC gene, HES1. **(D)** Co-expression network of the module exhibiting the highest differential correlation and variance. Links between nodes represent strong correlation relationships (correlation coefficient ≥ 0.95). Green nodes indicate genes associated with differentiation-related phenotype in knockout or transgenic mice. Node size indicates system transition score.

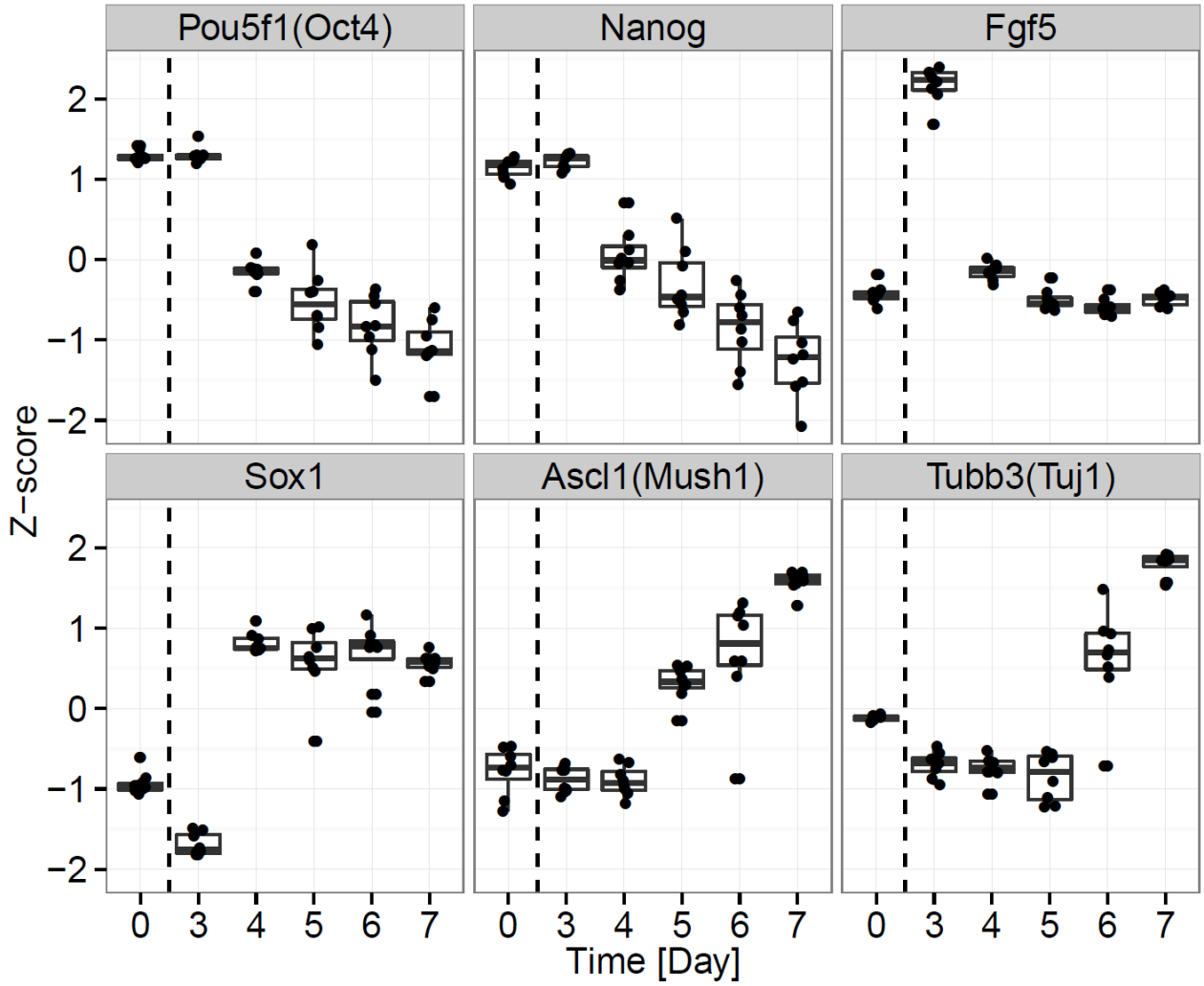


Figure 2-4. Expression profiles of neural differentiation markers in mouse ES cells.

The Z-score indicates relative expression differences in each marker gene throughout the neural differentiation period (day 0 to day 7). Pou5f1 (Oct4) and Nanog are undifferentiated pluripotency markers. Fgf5 is a primitive ectoderm marker. Sox1 is an early neuroectodermal marker. Ascl1 (Mash1) and Tubb3 (Tuj1) are neural differentiation markers. Mouse ES cells pass through the epiblast-like stage at day 3

and convert to neurons at day 6.

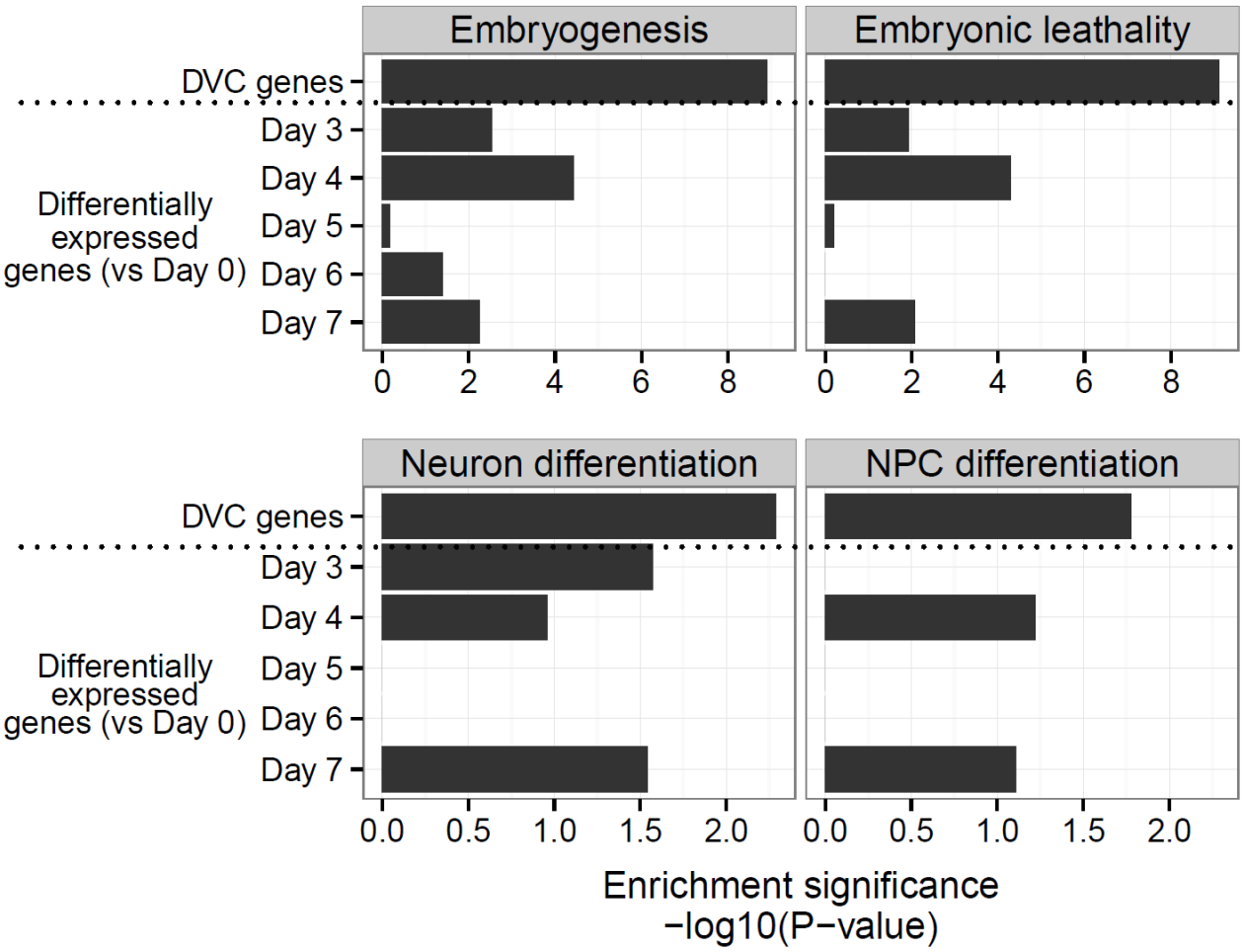


Figure 2-5. Genetic association of DVC genes with the differentiation process. Enrichment of gene sets associated with differentiation-related phenotypes in KO/Tg mice among the DVC and differentially expressed genes.

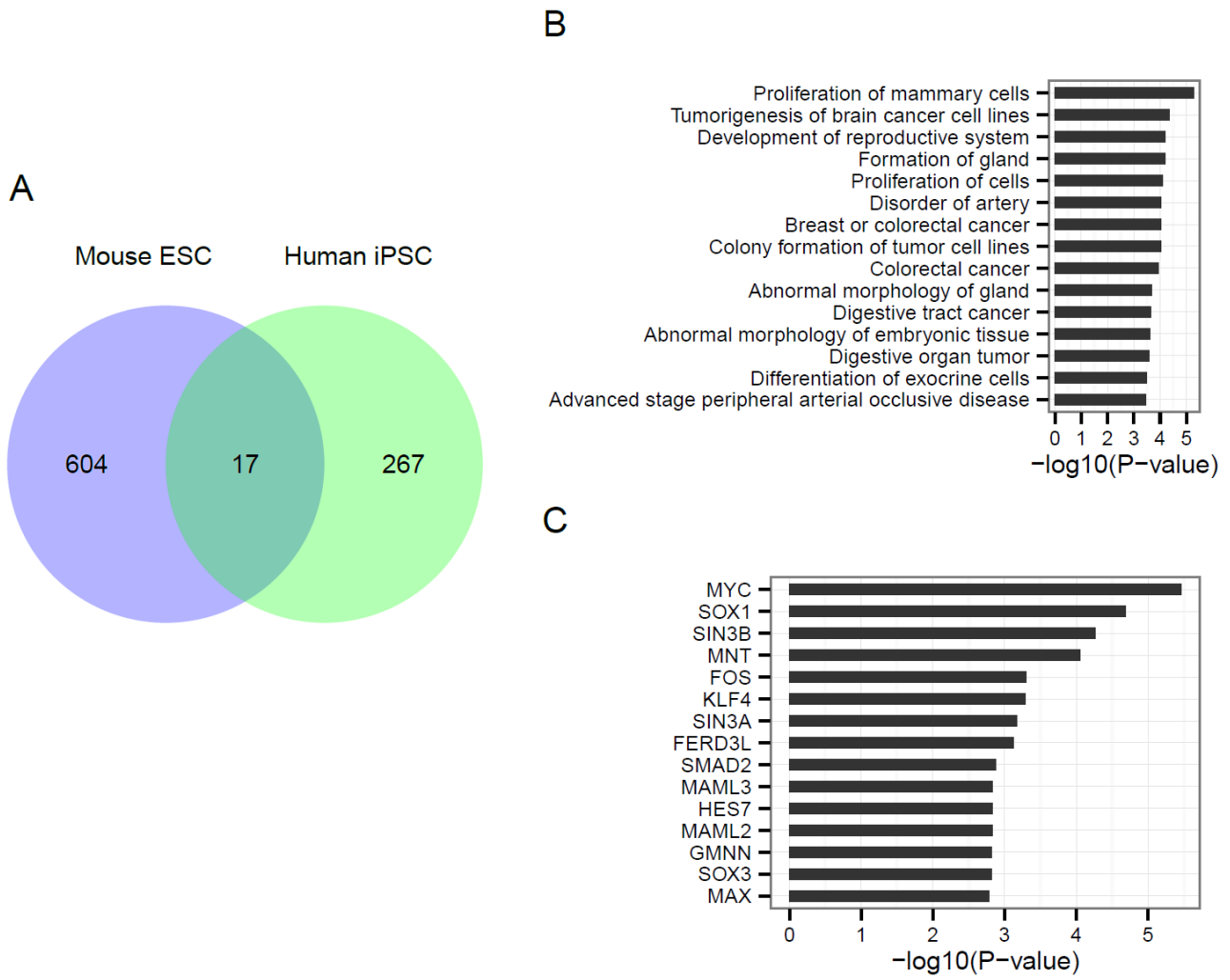


Figure 2-6. Common transcriptional regulators and functions of DVC genes in mouse ES and human iPS cells. **(A)** Venn diagram indicating the overlap between DVC genes from analyses of mouse ES and human iPS cells. DVC genes overlapped significantly between mouse ES and human iPS cells ($p = 0.0204$). **(B)** Functional categories enriched in overlapped DVC genes during neural development. **(C)** Upstream regulator analysis revealed transcriptional factors that could regulate the overlapping genes.

2.5. Summary

Understanding the dynamics of stem cell differentiation processes at the molecular level is a central challenge in developmental biology and regenerative medicine. Although the dynamic behaviors of differentiation regulators have been partially characterized, the architecture regulating the underlying molecular systems remains unclear. System-level analysis of transcriptional data was performed to characterize the dynamics of molecular networks in neural differentiation of stem cells. Expression of a network module of genes tightly co-expressed in mouse embryonic stem (ES) cells fluctuated greatly among cell populations before differentiation, but became stable following neural differentiation. During the neural differentiation process, genes exhibiting both differential variance and differential correlation between undifferentiated and differentiating states were related to developmental functions such as body axis development, neuronal movement, and transcriptional regulation. Furthermore, these genes were genetically associated with neuronal differentiation, providing support for the idea they are not only differentiation markers but could also

play important roles in neural differentiation. Comparisons with transcriptional data from human induced pluripotent stem (iPS) cells revealed that the system of genes dynamically regulated during neural differentiation is conserved between mouse and human. The results of this study provide a systematic analytical framework for identifying key genes involved in neural differentiation by detecting their dynamical behaviors, as well as a basis for understanding the dynamic molecular mechanisms underlying the processes of neural differentiation.

2.6. References

1. Singer ZS, Yong J, Tischler J, Hackett JA, Altinok A, Surani MA, Cai L, Elowitz MB. Dynamic Heterogeneity and DNA Methylation in Embryonic Stem Cells. Mol Cell 2014; 55: 319–331.

2. Kumar RM, Cahan P, Shalek AK, Satija R, Jay DaleyKeyser A, Li H, Zhang J, Pardee K, Gennert D, Trombetta JJ, Ferrante TC, Regev A, Daley GQ, Collins JJ. Deconstructing transcriptional heterogeneity in pluripotent stem cells. Nature 2014; 516: 56–61.

3. Scheffer M, Bascompte J, Brock W a, Brovkin V, Carpenter SR, Dakos V, Held H, van Nes EH, Rietkerk M, Sugihara G. Early-warning signals for critical transitions. Nature 2009; 461: 53–9.

4. Scheffer M, Carpenter SR, Lenton TM, Bascompte J, Brock W, Dakos V, van de Koppel J, van de Leemput I a, Levin S a, van Nes EH, Pascual M, Vandermeer J. Anticipating critical transitions. Science 2012; 338: 344–8.

5. Macarthur BD, Ma'ayan A, Lemischka IR. Systems biology of stem cell fate and

cellular reprogramming. Nat Rev Mol Cell Biol 2009; 10: 672–81.

6 . Huang S. Non-genetic heterogeneity of cells in development: more than just noise. Development 2009; 136: 3853–62.

7 . Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. Nature 2010; 467: 167–73.

8 . Imayoshi I, Isomura A, Harima Y, Kawaguchi K, Kori H, Miyachi H, Fujiwara T, Ishidate F, Kageyama R. Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. Science 2013; 342: 1203–8.

9 . Shimojo H, Ohtsuka T, Kageyama R. Oscillations in notch signaling regulate maintenance of neural progenitors. Neuron 2008; 58: 52–64.

10 . Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 2005; 4: Article17.

11 . Geschwind DH, Konopka G. Neuroscience in the era of functional genomics and systems biology. Nature 2009; 461: 908–15.

12 . MacArthur BD, Sevilla A, Lenz M, Müller F-J, Schuldt BM, Schuppert A a, Ridden SJ, Stumpf PS, Fidalgo M, Ma'ayan A, Wang J, Lemischka IR.

Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity. Nat Cell Biol 2012; 14: 1139–47.

13 . Ideker T, Krogan NJ. Differential network biology. Mol Syst Biol 2012; 8: 565.

14 . Zhang B, Gaiteri C, Bodea L-G, Wang Z, McElwee J, Podtelezhnikov A a, Zhang C, Xie T, Tran L, Dobrin R, Fluder E, Clurman B, Melquist S, Narayanan M, Suver C, Shah H, Mahajan M, Gillis T, Mysore J, Macdonald ME, Lamb JR, Bennett D a, Molony C, Stone DJ, Gudnason V, Myers AJ, Schadt EE, Neumann H, Zhu J, Emilsson V. Integrated Systems Approach Identifies Genetic Nodes and Networks in Late-Onset Alzheimer's Disease. Cell 2013; 153: 707–720.

15 . Chen L, Liu R, Liu Z-P, Li M, Aihara K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. Sci Rep 2012; 2: 342.

16 . Liu R, Li M, Liu Z-P, Wu J, Chen L, Aihara K. Identifying critical transitions and their leading biomolecular networks in complex diseases. Sci Rep 2012; 2: 813.

17 . Liu R, Wang X, Aihara K, Chen L. Early Diagnosis of Complex Diseases by Molecular Biomarkers , Network Biomarkers , and Dynamical Network

Biomarkers. 2013; : 1–24.

18 . Liu R, Yu X, Liu X, Xu D, Aihara K, Chen L. Identifying critical transitions of complex diseases based on a single sample. Bioinformatics 2014; 30: 1579–86.

19 . Li Y, Jin S, Lei L, Pan Z, Zou X. Deciphering deterioration mechanisms of complex diseases based on the construction of dynamic networks and systems analysis. Sci Rep 2015; 5: 9283.

20 . Yu X, Zeng T, Wang X, Li G, Chen L. Unravelling personalized dysfunctional gene network of complex diseases based on differential network model. J Transl Med 2015; 13: 189.

21 . Theunissen PT, Pennings JL a, Robinson JF, Claessen SMH, Kleinjans JCS, Piersma AH. Time-response evaluation by transcriptomics of methylmercury effects on neural differentiation of murine embryonic stem cells. Toxicol Sci 2011; 122: 437–47.

22 . Paşca SP, Portmann T, Voineagu I, Yazawa M, Shcheglovitov A, Paşca AM, Cord B, Palmer TD, Chikahisa S, Nishino S, Bernstein J a, Hallmayer J, Geschwind DH, Dolmetsch RE. Using iPSC-derived neurons to uncover cellular phenotypes

associated with Timothy syndrome. Nat Med 2011; 17: 1657–62.

23 . Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004; 5: R80.

24 . Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498–504.

25 . Blake J a, Bult CJ, Eppig JT, Kadin J a, Richardson JE. The Mouse Genome Database: integration of and access to knowledge about the laboratory mouse. Nucleic Acids Res 2014; 42: D810-7.

26 . Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med 1990; 9: 811–8.

27 . Wright MW, Eyre T a, Lush MJ, Povey S, Bruford E a. HCOP: the HGNC comparison of orthology predictions search tool. Mamm Genome 2005; 16: 827–8.

28 . Kan L, Israsena N, Zhang Z, Hu M, Zhao L-R, Jalali A, Sahni V, Kessler J a. Sox1 acts through multiple independent pathways to promote neurogenesis. Dev Biol 2004; 269: 580–94.

29 . Genethliou N, Panayiotou E, Panayi H, Orford M, Mean R, Lapathitis G, Gill H, Raoof S, De Gasperi R, Elder G, Kessaris N, Richardson WD, Malas S. SOX1 links the function of neural patterning and Notch signalling in the ventral spinal cord during the neuron-glial fate switch. Biochem Biophys Res Commun 2009; 390: 1114–20.

30 . Evseenko D, Zhu Y, Schenke-Layland K, Kuo J, Latour B, Ge S, Scholes J, Dravid G, Li X, MacLellan WR, Crooks GM. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. Proc Natl Acad Sci U S A 2010; 107: 13742–7.

31 . Yung S, Ledran M, Moreno-Gimeno I, Conesa A, Montaner D, Dopazo J, Dimmick I, Slater NJ, Marenah L, Real PJ, Paraskevopoulou I, Bisbal V, Burks D, Santibanez-Koref M, Moreno R, Mountford J, Menendez P, Armstrong L, Lako M. Large-scale transcriptional profiling and functional assays reveal important roles

for Rho-GTPase signalling and SCL during haematopoietic differentiation of human embryonic stem cells. Hum Mol Genet 2011; 20: 4932–46.

32 . Kim JJ, Khalid O, Namazi A, Tu TG, Elie O, Lee C, Kim Y. Discovery of consensus gene signature and intermodular connectivity defining self-renewal of human embryonic stem cells. Stem Cells 2014; 32: 1468–79.

33 . Pazianos G, Uqoezwa M, Reya T. The elements of stem cell self-renewal: a genetic perspective. Biotechniques 2003; 35: 1240–7.

34 . Yuan H, Corbi N, Basilico C, Dailey L. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes Dev 1995; 9: 2635–2645.

35 . Manuel MN, Mi D, Mason JO, Price DJ. Regulation of cerebral cortical neurogenesis by the Pax6 transcription factor. Front Cell Neurosci 2015; 9: 70.

36 . Kageyama R, Ohtsuka T, Kobayashi T. The Hes gene family: repressors and oscillators that orchestrate embryogenesis. Development 2007; 134: 1243–51.

37 . Zhou X, Smith AJH, Waterhouse A, Blin G, Malaguti M, Lin C-Y, Osorno R, Chambers I, Lowell S. Hes1 desynchronizes differentiation of pluripotent cells by

modulating STAT3 activity. Stem Cells 2013; 31: 1511–22.

38 . Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, Schafer X, Lun Y, Lemischka IR. Dissecting self-renewal in stem cells with RNA interference. Nature 2006; 442: 533–8.

39 . Jang MH, Bonaguidi M a., Kitabatake Y, Sun J, Song J, Kang E, Jun H, Zhong C, Su Y, Guo JU, Wang MX, Sailor K a., Kim JY, Gao Y, Christian KM, Ming GL, Song H. Secreted frizzled-related protein 3 regulates activity-dependent adult hippocampal neurogenesis. Cell Stem Cell 2013; 12: 215–223.

40 . Fu J, Ivy Yu HM, Maruyama T, Mirando AJ, Hsu W. Gpr177/mouse Wntless is essential for Wnt-mediated craniofacial and brain development. Dev Dyn 2011; 240: 365–371.

41 . Kobayashi T, Mizuno H, Imayoshi I, Furusawa C, Shirahige K, Kageyama R. The cyclic gene Hes1 contributes to diverse differentiation responses of embryonic stem cells. Genes Dev 2009; 23: 1870–5.

42 . Baek JH, Hatakeyama J, Sakamoto S, Ohtsuka T, Kageyama R. Persistent and high levels of Hes1 expression regulate boundary formation in the developing

central nervous system. Development 2006; 133: 2467–76.

43. Bennin D a, Don ASA, Brake T, McKenzie JL, Rosenbaum H, Ortiz L, DePaoli-Roach A a, Horne MC. Cyclin G2 associates with protein phosphatase 2A catalytic and regulatory B' subunits in active complexes and induces nuclear aberrations and a G1/S phase cell cycle arrest. J Biol Chem 2002; 277: 27449–67.

44. Zhang M, Schöler HR, Greber B. Rapid and efficient generation of neurons from human pluripotent stem cells in a multititre plate format. J Vis Exp 2013; : e4335.

45. Purvis JE, Lahav G. Encoding and decoding cellular information through signaling dynamics. Cell 2013; 152: 945–56.

46. Wang A, Huang K, Shen Y, Xue Z, Cai C, Horvath S, Fan G. Functional modules distinguish human induced pluripotent stem cells from embryonic stem cells. Stem Cells Dev 2011; 20: 1937–50.

Chapter 3

Identification of an early cell fate regulator by detecting dynamics in transcriptional heterogeneity and co-regulation during astrocyte differentiation

3.1. Introduction

The variability in gene expression among individual cells is known to increase the complexity of cellular phenotypes due to the population effect [1]–[6]. Such an effect of cellular heterogeneity has been observed in various complex biological processes including disease progression, drug responses, and cell differentiation. To understand the complexity of biological processes, it is important to investigate the mechanism of variability in gene expression and the co-expression relationships.

Transcriptional variability is caused by factors that greatly affected by the complexity of the cell population, such as individual differences in the cell state, the cell cycle, and in their biological profiles. One of the features which best describes such heterogenic states in a cell population is the mis-synchronization of oscillations in gene expression [7]. During neuronal and glial differentiation, genes in the notch signaling cascade have been reported to oscillate, however there is both spatio- and temporal-heterogeneity between cells [8]. In neuronal stem cells (NSCs), oscillations in gene expression are also less synchronized, therefore gene expression levels among individual cells show a large degree of diversity, resulting in a high degree of variability in gene expression in the cell population.

With such oscillation in gene expression, synchronization among different genes, even within a single cell, is another important feature that must be considered. During neuronal differentiation for example, it is known there are pairs of genes (e.g. *Dll1* and *Ngn2*) that are co-expressed in order to synchronize their gene expression oscillations in a single cell [8],[9]. Such transcriptional co-regulation changes as the cell progresses from the progenitor cellular state toward the differentiation state is known to play an important role in the decision of cell fate [9].

From the standpoint of dynamical system theory, an increase in both variability and a correlation of system components, such as gene expression, is a sign of an upcoming transition of the cellular state [10]–[12]. The dynamic change in the cellular differentiation process in response to external stimulation is thought to be determined by some critical cell state transitions, which rapidly shifts the cellular state from one state to another. We hypothesized that both high variability in gene expression and a high correlation of expression between genes are commonly observed in such a transition state during the differentiation process. By analyzing genome-wide transcriptional data from the differentiation process in embryonic stem (ES) cells and induced pluripotent stem (iPS) cells, we previously reported that gene sets co-expressed in the undifferentiated state showed a large difference of variability in expression levels during the transition from the undifferentiated state to the differentiated state. Detecting these dynamic changes in the variability of gene expressions was therefore proposed be an early signal capable of predicting the cellular transition in neural differentiation[13].

In order to evaluate the controlled heterogeneity in gene expression, we previously proposed the analytical concept of evaluating “differential variability and

correlation (DVC)" [13]. This analysis focuses on the differential transition of two index between the undifferentiated state and the differentiated state: (1) the variability of gene expression within the same state, which reflects the potential variation in the expression of a particular gene measured for a number of times using a single cell or bulk of cells, and (2) the correlation, which reflects a similar pattern of expression for multiple genes (co-expression) in the same state. The concept of DVC analysis is illustrated in Fig. 3-1. Briefly, each gene expression is evaluated by two indexes ("change of expression variability" and "change of co-expression pattern") calculated from cells which change from one state to another (in our case undifferentiated sate to differentiated state). This DVC analysis was designed to identify specific biological responses in a group of cells according to the hypothesis that the responses of a group of cells are not homogeneous at the start of event but can harmonize as the stages of the event change. Changes in biological responses are typically expressed as the average of gene expression rates for specific genes. In these analyses, bulk group of cells are lysed and measured to obtain their average gene expression, and such expression data has been known to reflect their biological status. However, it is difficult to use this concept to explore genes expressed at low average levels but greatly impact

the biological event as such as a trigger. When gene expression rates are averaged within a group of cells, minor genes expressed only in few sub-populations of cells or over short time in individual cells, such gene effect detection is difficult. Considering that such genes have either large variance among groups of cells or exhibit the sudden appearance/disappearance among time series, our DVC analysis is designed to provide new aspects for analyzing transcriptomic data. Our previous work indicated that even using the average expression data from bulk cells, the gene sets identified by DVC analysis (DVC genes) were found to play important roles in neuronal differentiation [13] (illustrated in Fig. 3-1).

Recently, single cell transcriptomics has been found to have sufficient sensitivity to interpret the heterogeneity of gene expression in cellular differentiation [14]–[18]. However, the importance of variability in the expression developmental regulator genes remains unclear. In this study, in order to extend our previous work on finding key functional genes in the neuronal differentiation process through a DVC analysis, we analyzed single cell RNA sequencing data obtained from mouse neural stem cells (NSCs), transit amplifying progenitors (TAPs), and astrocytes[19]. The astrocyte developmental process is still poorly defined due to a lack of lineage stage

specific markers[20]. We hypothesized that the controlled heterogeneity of gene expression profiles can be one of the key events to explain astrocyte differentiation. DVC genes, potentially including oscillatory genes, can be potential early marker genes for predicting upcoming astrocyte differentiation. Using our DVC analysis to compare three different cellular states, we found that the DVC gene signatures are a potential predictive biomarker to indicate an upcoming critical cell state transition. A functional analysis of the DVC gene signature suggested that these genes have an impact on the astrocyte differentiation process. In addition, from the study in which we used an antagonist to block Ntsr2, one of our DVC gene candidates, we found that this gene is involved in controlling the direction of astrocyte differentiation. We propose a framework for DVC analysis that can identify regulator genes that have an impact on upcoming cellular transition events, and have gained insights into the molecular mechanism underlying biological state transitions.

3.2 Materials and Methods

3.2.1. Single cell RNA-seq data and its preprocessing

The single cell RNA-seq data from three types of cells namely 92 NSCs, 27 TAPs, and 22 astrocytes were obtained from the Sequence Read Archive (SRP057125). According to archived data, $Glast^{+}Prom^{+}$ cells (designated as NSCs) and $Glast^{-}Prom^{-}Egfr^{+}$ cells (designated as TAPs) were isolated by FACS from the su-ventricular zone (SVZ). $Glast^{hi}$ (designated as astrocytes) were isolated from the striatum and somatosensory cortex according to archived data. NSCs can produce neural progenitor cells (TAPs or type C cells), which are a proliferative cell population expressing markers of early neuronal differentiation. Some NSCs can generate both neurons and astrocytes. TAPs are known to give rise to neuroblasts (type A cells) that differentiate into primarily interneurons. Although the complete lineages of neuronal and glial cells in the mammalian brain remains unclear, according to such lineage information, we hypothetically ordered the three types of cells (NSCs, TAPs, and astrocytes) from the early lineage to the differentiated state. Another dataset for validating our findings by DVC analysis included cell populations isolated by FACS

from GFAP-GFP transgenic mice: GFAP-GFP^{+}PROM1^{+}EGFR^{+} (astrocytes), GFAP-GFP^{+}PROM1^{+} (NSCs) [21]. To further confirm our DVC analysis concept, we evaluated recent data (defined as the second dataset) from Llorens-Bobadilla *et al.* [19] which showed similar cell population data when representative biomarkers (Glast (Slc1a3) and Cd9) were compared. This data included 53 NSCs and 13 astrocytes after data filtering. The second dataset was preprocessed and DVC analysis was carried out in the same manner as for the first dataset. The reads were mapped to the mouse genome (ENSEMBL Release 83) using STAR [22]. FeatureCounts was used to count the mapped reads for genes [23]. TMM (trimmed mean of the M value) normalization and CPM (counts per million) transformations were performed using EdgeR to compare the expression levels across the samples [24]. A principal component analysis was applied to the CPM data to remove oligodendrocyte like cells from the NPCs. Filtering out the genes with low expression levels (a read count < 2) for each cell type resulted in the detection of 12,147 commonly analyzable genes within the three cell states, and was used for the following analyses. The CPM data were $\log_2$ transformed and standardized to the Z-score for each gene across individual cells for each cell type.

3.2.2 Co-expression analysis

Hierarchical clustering was applied to identify co-expressed genes as cluster modules. The gene expression values were standardized within individual cells. For the clustering, only the single cell data in the NSC state, the first state in our analysis, were used to assemble the clustering tree. For the clustering, both the Pearson correlation coefficient (PCC) and Wald linkage method were used. For module detection, a dynamic tree-cutting algorithm (hybrid mode, minimal module size of 100) was used. All calculations were coded by R.

3.2.3. Differential variability and correlation (DVC) analysis

The detailed procedure for the differential variability and correlation (DVC) analysis has been described in our previous study [13]. The analysis was conducted using the Bioconductor package in the R language. The system transition score (STS) was used to rank the genes and identify those with high differential variability and correlation between two different cellular statuses (NSCs vs. TAPs, or NSCs vs. astrocytes). The detailed description of STS calculation is shown in the supplementary note. The co-expression network of the DVC genes included in the same module was

displayed using Cytoscape 3.1.2 [25]. PCCs above 0.7 were shown as the connection between genes in the network figure.

3.2.4. Functional enrichment analysis of candidate gene signature

The astrocyte differentiation related gene sets (GO:0048708) for the functional enrichment analysis were obtained from the Gene Ontology database [26],[27]. The ChIP-seq data showing the transcription factor binding sites and DNase-seq data showing open chromatin sites in NSCs were obtained from a previously published study [28]. The genomic regions overlapping between ChIP-seq and DNase-seq data were discovered using bedtools. The genes within 1M bases from the overlapping regions were used in the transcription factor enrichment analysis to incorporate potential enhancer regions into the analysis. The enrichment significance was assessed using the cumulative hypergeometric probability with the phyper function in R. The enrichment test is one-sided. Ingenuity Pathways analysis (IPA®, Qiagen, http://www.ingenuity.com) was used to examine the upstream regulators of the DVC genes. The reference data set was set as the 12,147 genes that represented all the genes used in the functional enrichment analyses. As a comparison, we applied a

functional enrichment analysis on the conventional “differently expressed genes (DEGs).” The DEG definition is described in the supplementary note.

3.2.5. Cell culture and Ntsr2 inhibition assay

Fetal-derived mouse NSCs (mNSCs, Cell Application Inc., San Diego, CA, USA) at passage two were seeded at a density of 1.0×10^5 cells/cm^2 in T25 flasks coated with poly-L-ornithine hydrobromide (Sigma–Aldrich, St. Louis, MO, USA) and natural mouse laminin (Thermo Fisher Scientific, Waltham, MA, USA) for maintenance and differentiation. The maintenance culture and differentiation culture (astrocyte differentiation) was performed according to the manufactures’ protocol, with some modifications as described in our previous work [29]. mNSC’s were seeded in triplicate into a coated 6-well plate for the real-time PCR experiment, and into a 12-well plate for immunohistochemical staining. For the real-time PCR, the following primers were applied: ß-actin (forward: GGCCCAGAGCAAGAGAG GTATCC, reverse: ACGCACGATTTCCCTCTCAGC), GFAP (forward: GCCA CCAGTAACATGCAAGA, reverse: CGGCGATAGTCGTTAGCTTC), and SOX2 (forward: GGCGGCAACCAGAAGAACAG, reverse: GCTTGGCCTGCGTCGATG AAC).

For RT-PCR, total RNA was extracted using RNeasy kit (QIAGEN, Germantown, MD, USA), and cDNA was generated using Superscript II (Invitrogen, Carlsbad, CA, USA). PCR was performed over 30 cycles for all genes except ß-actin (25 cycles). For Gfap immunohistochemistry, an anti-glial fibrillary acidic protein (Gfap) antibody (GR15465010, ab53554; Abcam, Cambridge, MA, USA), and the secondary antibody anti-goat DAG-IgG-Alexa Fluor 488 (GR2460881, ab150129; Abcam) were used. The protocol for immunohistochemical staining is described in our previous work[29]. The Ntsr2 antagonist (JMV449), a pan neurotensin receptor antagonist, was purchased from TOCRIS (Avonmouth, Bristol, UK), and added to the mNSCs at final concentrations of 0.15 pM, 1.5 pM, 15 pM, 0.15 nM, and 1.5 nM.

3.2.6. Statistical analysis

The Leven-test was applied to the RNA-seq data to evaluate the variability among individual cells in each cell type. The Voom-limma method was used to identify the DEGs between cell types[30]. P-values were adjusted using the Benjamini-Hochberg method. These procedures were conducted in R. The statistical tests are two-sided.

3.3. Results

3.3.1. Comparison of transcriptional variability and correlation in three cell states (NSCs, TAPs, and astrocytes)

To understand the differential dynamics of gene expression variability, we first classified three types of cells as being representative of three states: NSCs representing the most undifferentiated state, TAPs representing a potentially intermediate differentiation state, and astrocytes representing the most differentiated state. By comparing transcriptional variability and a correlation of NSCs with two different states (TAPs and astrocytes), we evaluated alterations in the co-expressed gene sets in NSCs.

First, to obtain the co-expressed genes in the NSC state, which we defined as the most undifferentiated state, a hierarchical clustering analysis was performed on their single cell gene expression data. Using a correlation coefficient, which describes the similarity of gene expression profiles in each individual cell, the co-expressed gene modules in the NSC state were identified. Within the 12,147 genes, 17 modules of

co-expressed genes were identified (Fig 3-2A).

By comparing TAPs/astrocytes vs. NSCs, we evaluated the differences in gene expression variability using the single cell gene expression data (Fig. 3-1A). Between the two cell states, TAPs and NSCs, 3423 genes (28.2%) were found to show differential variability (Levene test $q < 0.001$). Between NSCs and astrocytes, 13.5% of all genes also showed differential variability (Levene test $q < 0.001$). These data show that 1637 genes changed their expression profile, either in a harmonized or in a heterogeneous pattern, as the cell state changed from undifferentiated to differentiated. By examining the overlap between genes with differential variability in expression and each co-expressed gene set, some modules were found to be enriched in genes that changed their variability as a result of the cell state transition. When all the modules were examined for the direction of differential variability (Fig. 3-2B), it was found that there were only a few modules that contained genes that increased their expression variability compared to the NSC state (Fig. 3-1B upper two rows). In contrast, there were several modules that showed a large member of genes that had a decrease in their expression variability compared to the NSC state. More than half of the genes in the black and magenta modules decreased their expression variability in TAPs

(enrichment significance was $p = 4.35e^{-79}$) and astrocytes ($p = 3.69e^{-70}$), respectively. The red and purple modules showed that more than 1/3 of the module member genes decreased their expression variability in both TAPs ($p = 4.79e^{-39}$, $p = 6.03e^{-73}$) and astrocytes ($p = 8.24e^{-45}$, $p = 1.30e^{-28}$).

By comparing TAPs/astrocytes vs. NSCs, differential correlations were also evaluated (Fig. 3-2A). When we focused on the modules which showed decreased expression of variability (black, magenta, purple, and red module), we found that most of their member genes also showed a decrease in their correlation. These data indicate that both the transcriptional variability and the co-expressed gene relationships in the NSC state decreased as astrocyte differentiation progressed. With the differential correlation heatmap, the profile of "TAP vs. NSC" and "Astrocyte vs. NSC" was found to be very similar. Therefore, we checked the difference among their detailed gene correlation networks. When detailed correlation networks were confirmed, we found that most of the correlations between different pairs of genes were different. Therefore, the similar "change profile" illustrated by heatmap is just showing the brief total tendency of numerous correlation scores per each gene, and their individual correlation networks are more complex.

3.3.2. DVC genes in the cell state transition of astrocyte differentiation

For further the analysis of differential gene expression profiles between cell states, we carried out our DVC analysis to identify candidate genes that could be predictive of an up-coming drastic cell state transition. We measured the change in two parameters with our measure score, the system transition score (STS), which combines both an evaluation of the differential variability and the differential correlation of gene expression. We defined genes with a high STS as "DVC genes," and selected candidate genes from a comparison of pairs of transition states, TAPs vs. NSCs, and astrocytes vs. NSCs. From the comparison of TAPs vs. NSCs, 474 DVC genes were found, and from astrocytes vs. NSCs comparison, 504 DVC genes were found. This result suggests that there are more genes related to the response to the state transition from the NSC state to the differentiated astrocyte state.

In the DVC genes, the cell cycle genes which is important for early proliferative phase was overrepresented (enrichment p-value = 1.79e-32). For example, Cdk6, which contributes mainly in G1 phase and proliferation, was found as the top DVC gene. This result suggest that the DVC analysis reflects the commonly known functional genes

that is predominant differentiation pathway from progenitor to committed cell.

To examine whether candidate DVC genes are functionally involved in astrocyte differentiation, a functional enrichment analysis was performed. As a result, astrocyte differentiation related genes were highly enriched in the DVC genes identified between the astrocyte vs NSC states ($p = 6.15e^{-6}$) compared with DVC genes identified between the TAP vs. NSC states (Fig. 3-3A).

In comparison, using a conventional analysis, which compares the average (averaged value of single cell data) expression rates of genes among different cell states, to identify differentially expressed genes (DEGs) (definition described in the Material and Methods), we rarely found astrocyte differentiation-related genes, even from both state comparisons (TAPs vs. NSC or astrocyte vs. NSC). This result suggests that the genes which play important roles in astrocyte differentiation vary in expression variability and correlation and not in average expression level.

To further investigate the robustness of our DVC analysis applied to cells which changed their state from stem cells to astrocytes, we evaluated the second dataset (53 NSCs and 13 astrocytes) obtained from an independent study [22]. Our DVC analysis identified 97 highly ranked DVC genes from the second dataset. There were 15

overlapping genes among the top-ranking DVC genes between the first and second datasets (enrichment significance: p-value = 3.06e-11).

3.3.3. Cell-fate marker genes found as DVC genes

To further understand the function of the DVC genes, we examined if known biomarkers and regulators involved in neuron and astrocyte differentiations were also DVC genes. In this regard, *Ascl1*, which promotes neuronal fate determination, was identified as being a DVC gene between the astrocyte vs. NSC states [31]. As previously reported [9], *Ascl1* expression oscillates in NSCs, although it becomes stably suppressed following astrocyte differentiation. Similar changes in *Ascl1* gene expression were observed in this study. *Ascl1* showed a large variability in expression among individual cells in the NSC state, although this variability decreased in the TAP state. In the astrocyte state, it was expressed at a low level (Fig. 3-3B).

The notch signaling gene, *Dll1* was also found to be a DVC gene between the astrocyte vs. NSCs states. By plotting its expression levels, this DVC gene also showed a large variability in expression in the NSC state cell population, and a low level of expression in the astrocyte state. Such variability in gene expression is NSC is

consistent with a previous report indicating that *Dll1* shows oscillatory expression in the NSC state [32]. When *Dll1* was used as a representative gene, previously reported marker genes were also identified by searching the co-expression gene module (Fig. 3-1B). For example, in the red module, both *Dll3* and *Sox9* were identified as co-expression members for *Dll1*. *Dll3* is another notch-signaling gene, which also shows a large variability in expression among individual cells in the NSC state, and loses this variability in expression in the astrocyte state. Such an involvement of notch signaling, is consistent with previous work which has reported the oscillation and co-expression of notch signaling genes in a single NSC [8]. *Sox9*, which is known to be a glial fate determination marker, was also found in the red module [33]. From its expression profile, it was also noted that Sox9 showed a large variability in expression levels in the NSC state, although this variability was lost in astrocytes.

3.3.4. DVC genes are potentially regulated by Sox9, Ascl1 and Max

We performed an upstream regulator analysis to obtain regulatory insights into the DVC genes. First, a data-driven interpretation approach was applied based on the broad collection of transcriptional regulatory relationships from the published

literature using the Ingenuity Pathways Analysis software. From this analysis, Ascl1, one of the DVC genes, was identified as the most significant upstream regulator of other DVC genes (Fig. 3-4A). Second, a more focused approach was performed to confirm whether Ascl1 and other transcription factors important for neuronal and glial differentiation could potentially regulate the expression of DVC genes. The enrichment analysis of the DVC genes compared with the target gene candidates of 11 transcription factors (Ascl1, Ctcf, Fox3, Max, Nfi, Olig2, Smac1a, Sox2, Sox9, Sox21 and Tcf3) by ChIP-seq in the NSC state [28]. A transcription factor enrichment analysis then showed that Sox9, Ascl1, and Max could be candidates that regulate the transcription of other DVC genes. These consistent data for Ascl1, obtained from two separate and distinct approaches, indicate that Ascl1 is the most likely regulator of the DVC genes found in our analysis, and that it is involved in the cell state transition from NSCs to astrocytes.

3.3.5. Experimental validation of the role of the DVC candidate gene (*Ntsr2*) in determining cell fate

To validate the functional importance of our DVC gene candidates, we searched

for DVC genes which could potentially be involved in the *Ascl1* regulation gene network. From this gene network analysis, *Ntsr2* became a focus because it was one of the DVC genes that were co-expressed with *Ascl1* (Fig. 3-5A). The co-regulated relationship in expression between *Ntsr2* and *Ascl1* become weak in the astrocyte state compared to the NSC state (Fig. 3-5A). *Ntsr2* also showed a greater degree of variability in single cell expression in the NSC state than in the astrocyte state (Fig. 3-4B). However, the expression level of *Ntsr2* was increased in the astrocyte state compared with that in the NSC state. Therefore, we assumed that inhibition of Ntsr2 function would have a significant effect on Ascl1-related signaling in the NSC state, whereas Ntsr2 inhibition in the astrocyte state would have little effect on Ascl1-related signaling. In the NSC state, the addition of the Ntsr antagonist JMV449 clearly inhibited expression of the undifferentiation and astrocyte marker *Gfap* to levels lower than the control, without any sign of cytotoxicity (Fig. 3-6A). This result indicates that the antagonist disrupted the essential variability in the NSC state for astrocyte differentiation potential. When the expression level of the early neuronal fate marker, *Sox2*, was measured, the effect of the Nstr2 antagonist was very weak (Fig. 3-6B). This indicates that the antagonist showed a greater effect on the *Ntsr2* gene network, which

suggests that gene network of DVC genes plays a critical role in the state maintenance in NSCs for upcoming astrocyte differentiation. However, when JMV449 were added to the astrocyte differentiation medium, we did not find any significant effect (data not shown). These data also suggest that a disturbance of DVC genes is effective when their variability is large and their gene network is tight.

3.4. Discussion

Cell fate decision in the differentiation processes is proposed to be a system that transits abruptly from one state to another in response to external stimulation based on dynamic systems theory. Such cell states transitions are referred to as critical transitions [10],[11]. The 'fragility' of various biological profiles, is a new concept adopted to help understand complex biological phenotypes that are found during such critical transitions. It is now thought that both 'variability' in gene expression levels and 'co-expression' among heterogeneous populations of cells are empirical indicators which are involved in any upcoming biological transition. Based on this theory, our DVC analysis offers objective measurement criterion which correlate with the critical

cell state transition. In this study, in order to investigate the use of our DVC analysis in understanding astrocyte differentiation, we analyzed single cell transcriptome data to measure both the variability and correlation between cell states more accurately than cell population transcriptional data. Moreover, the robustness of DVC analysis was further confirmed in the independent second dataset.

By focusing on the evaluation of STS, a score incorporating both variability and co-expression, our analysis identified several candidate DVC genes, as being central regulating genes important in the transition from the NSC state to the astrocyte state. The functional enrichment of DVC genes important in astrocyte differentiation was more significant than that for conventional DEGs, indicating the importance of evaluating heterogeneity of gene expression data. Moreover, using a gene network analysis followed by pharmacological inhibition of a single DVC gene we demonstrated that the DVC analysis could identify key players in the transition from the NSC state. Our data also suggest that signatures that are involved in state transition are not easily identified using the conventional comparison of "expression averages." This study therefore has an impact by improving our understanding of transcriptional regulation in differentiation processes.

Through our DVC analysis, the importance of expression 'fragility' was clarified especially in the most undifferentiated NSC state. When we compared three states, represented by NSCs, TAPs, and astrocytes, the differences between NSCs and astrocytes were clear. However, the differences between TAPs and NSCs provided very few informative genes. Therefore, as proposed by Molofsky et al.[34], we consider that TAPs do not lie on the direct line of lineage from NSCs to astrocytes. In such considerations of lineage type differences, our STS score in DVC analysis, which rank the genes and identify the high differential variability and correlation between cellular states, can be informative.

Ascl1, the central gene identified here from the DVC analysis, is a well-defined transcription factor. For example, the notch signaling genes, *Dll1* and *Dll3*, are known to be targets of Ascl1 [35]. Notch signaling also up-regulates the expression of *Sox9*, and induces differentiation into astrocytes [36]. Moreover, Sox9 is known to bind to the genomic regions close to *Ascl1*, *Dll1*, and *Dll3* from a Chip-seq study [28]. Taken together, these data suggest that Sox9 may be both upstream and downstream of genes involved in notch signaling suggesting that a transcriptional loop could be formed. The change in the co-expression network between cell states might imply there is a change

in the network regulatory loop during the cell state transition.

Ntsr2 is a G-protein coupled receptor that binds neurotensin [37], and is expressed in NSCs and astrocytes[38]. However, its functional role in the cell state transition has not been previously described. An Ntsr antagonist suppressed *Gfap* expression and in addition had a small effect on *Sox2* expression in NSCs. Type 1 NSCs are characterized by presence of both Gfap and Sox2 expression in the undifferentiated state [39]. This type 1 state triggers the cellular state transit to type 2 in NSCs that express Sox2, but not Gfap [39], and is thought to be the state of self-renewal. It has been found that NSCs have the potential to differentiate into both neurons and astrocytes in their type 2 state[40]. Use of the Ntsr antagonist might guide type 1 NSCs to become type 2 NSCs, which have a high capacity to give rise to neural lineages. This study suggests that Ntsr2 could be involved in the cell state transition in the early cell fate decision making process. In the second data analysis, Ntsr2 was ranked at 304th and was not lost in the independent data. The sequencing depth of the second data set was lower than that of the first data set. This may have resulted in the loss of co-expression structure. Therefore, considering the difference in sequencing depth between the first and second data sets, we consider that our method provided

reproducible results. However, since the Ntsr2 inhibition potentially affects Stat3 signaling that may lead Gfap expression change, we should evaluate Stat3 signaling more in detail to gain insight on Ntsr2 biology during astrocyte differentiation. To further analyze the Ntsr2-related cascade, we believe further development of inhibitor libraries are required. First, we could not find an appropriate small molecule inhibitor for directly inhibiting Ascl1. Second, although JMB449 are known to inhibit both types of Ntsr receptors, Ntsr1 and Ntsr2, there were no molecule inhibits only Ntsr2. However, the expression read counts were not detectable for Ntsr1 in the data of all cell types; therefore, our experimental design was appropriate for studying Ntsr2 using JMB449.

Although the DVC analysis provided several possible clues, this study could not definitively explain why the variability in DVC gene expression is high in the undifferentiated state. A study of variability in the hematopoietic differentiation system demonstrated that both cell cycle and variations in cell size could, to a small extent, explain this variability in gene expression[15]. This report also suggested that variability in gene expression could be caused by other mechanisms. In this study, we identified oscillatory genes as DVC genes in the process of astrocyte differentiation.

Our previous study showed that *Hes1* is one DVC gene that is important during neuronal differentiation. *Hes1* was also found to have oscillation in its expression levels before neuronal differentiation. Combining the data from our neuronal and astrocyte differentiation studies enhances our hypothesis that oscillations in gene expression in the undifferentiated state causes a high transcriptional variability before differentiation occurs. These oscillations may be one reason why the transcriptional system becomes 'fragile' before an upcoming cell state transition such as cell fate decision. In the future, we intend to add more time-course data representing other cellular states in the differentiation process in order to extend DVC analysis to investigate the cause of high transcriptional variability in the differentiation process.

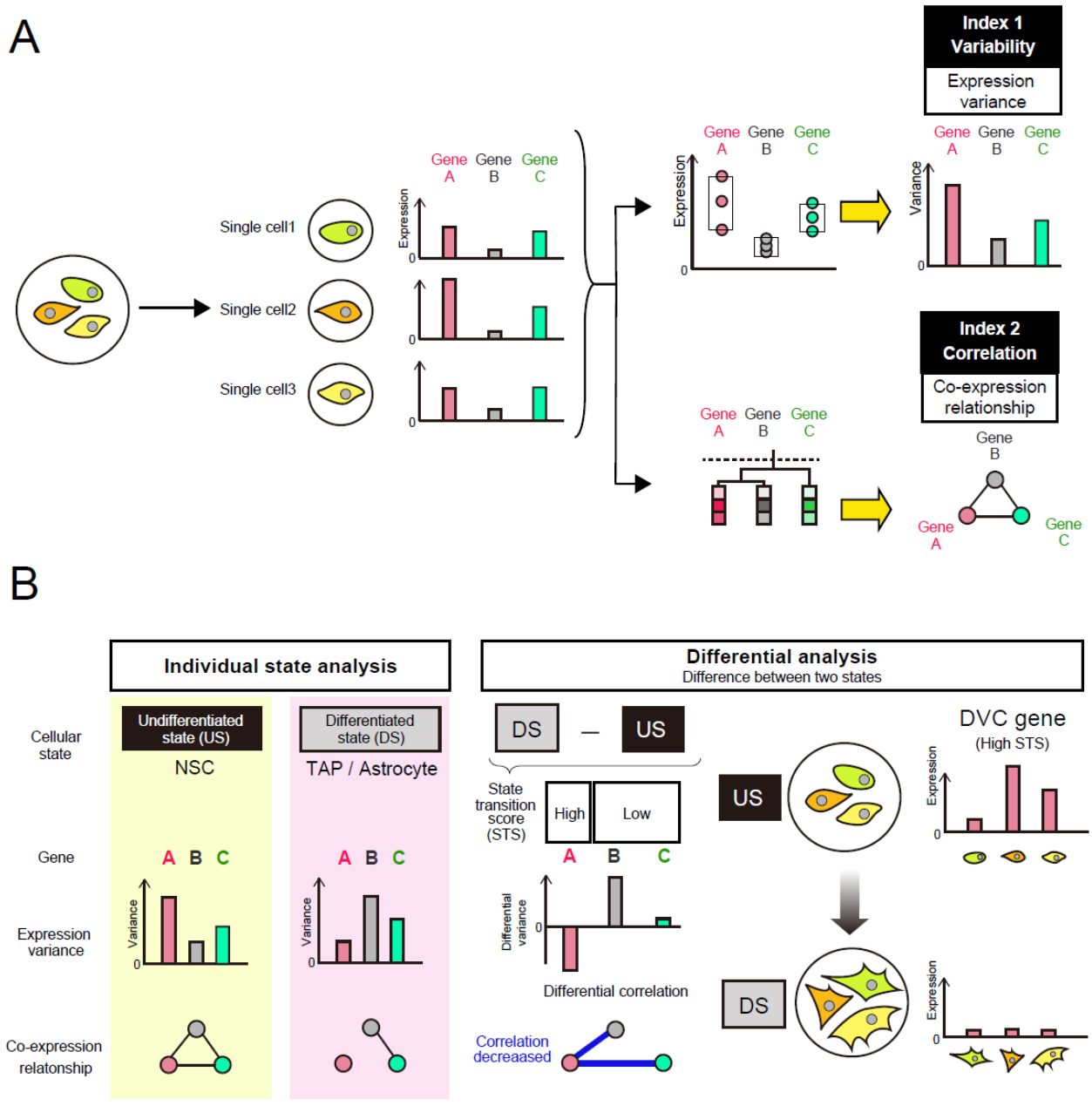


Figure 3-1. Conceptual illustration of the DVC analysis based on gene expression variability and correlation. Calculation concept of DVC analysis. **(A)** First step of calculating two indexes. In each state of cells, two types of criteria, variability and correlation, is calculated. For index 1 (variability), the standard deviation (SD) of each

gene expression between cells (3 single cells in this example) are calculated. For index 2 (correlation), the Pearson correlation coefficient of pair of expression patterns from group of cells (3 single cells in this example) is calculated between each pair of genes. By clustering, the correlated genes are grouped as co-expressed modules. **(B)** Second step of calculating state transition score (STS). The change rate between two state of cells is calculated using both indexes to obtain STS. DVC gene is the gene with high STS. An example image of DVC gene is also illustrated, which shows oscillation-like variability in the early state of cells, although settle/harmonize after the transition to the next state.

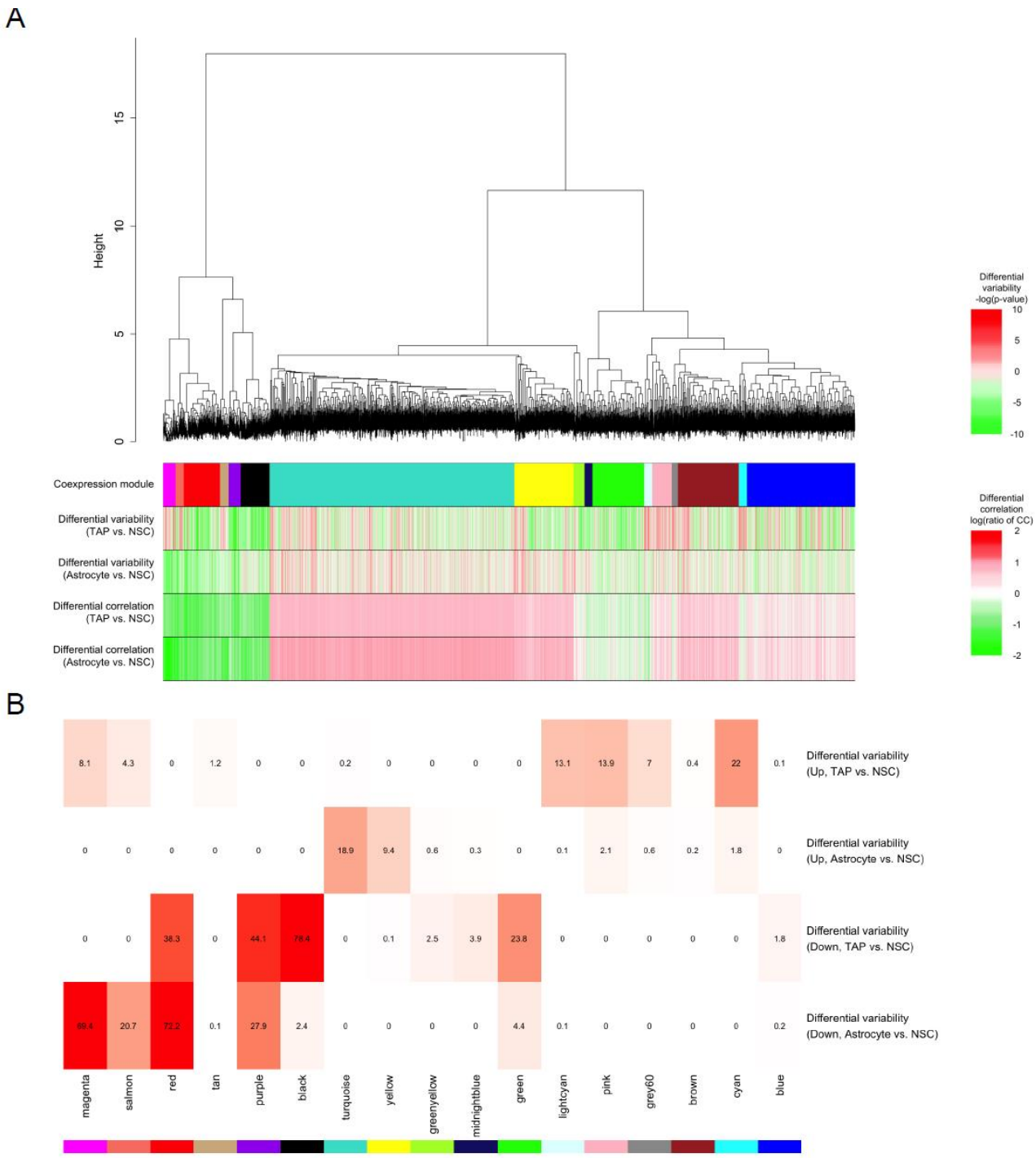


Figure 3-2. Identification of co-expression modules in the NSC state with their differential profile. (A) Clustered modules of co-expressed genes in NSC cells, and their differential profile compared to two differentiated cell states (TAPs/astrocytes). (Top tree) Hierarchical clustering tree shows the clustered genes based on co-expression

patterns among individual single cell transcriptomes (92 cells) in the NSC state. (Heatmaps) Co-expression modules: the divided clusters of co-expressed genes obtained from the above hierarchical clustering are represented by the colored classifiers. Differential variabilities: the average differential variability in single cell transcriptomes between two cell states, TAPs vs. NSCs or astrocytes vs. NSCs. Differential correlations: the average differential correlation of single cell transcriptomes between two cell states, TAPs vs. NSCs or astrocytes vs. NSCs. The color chart represents the significance of variability (differential variability chart), and the ratio of correlation coefficient (differential correlation chart) between cell states. The green color indicates lower values, and the red color indicates higher values. **(B)** Enrichment rate indication of differential variability genes in combination with the color of clustered modules in Fig. 3-2A. The clustered modules are aligned from the left to the right in the same order as shown as co-expression modules in Fig. 3-2A. Up regulated or down regulated genes are counted separately. The colored matrix reflects the enrichment rate along with statistical significance. The negative logarithm of the p-values are shown in the matrix.

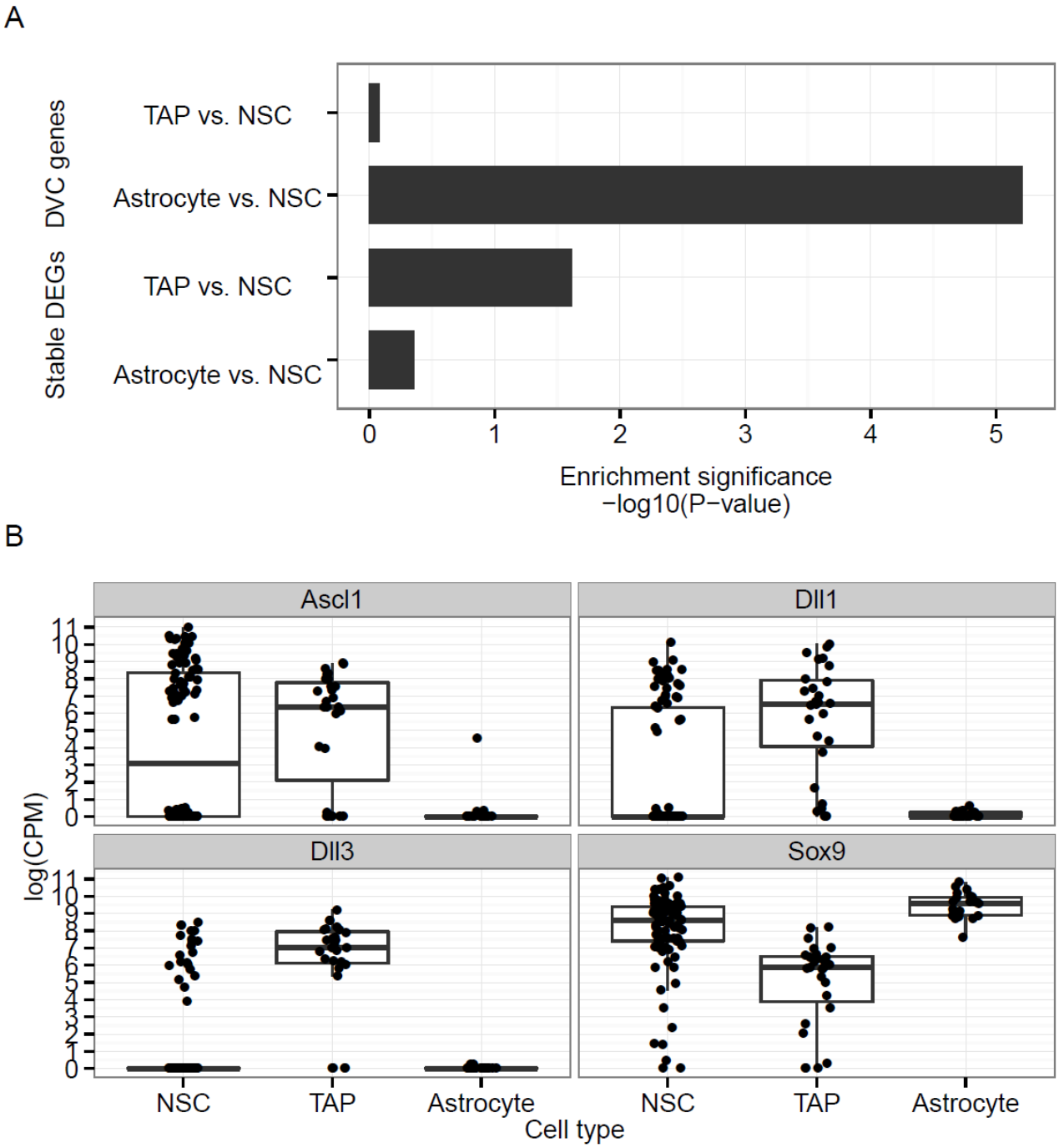


Figure 3-3. DVC genes between two cell states in astrocyte differentiation. (A) Bar plot showing the enrichment significance of astrocyte differentiation related genes compared between two cell states; TAPs/astrocytes vs. NSCs. DVC genes, and conventional DEGs are compared to one another. **(B)** Gene expression profiles of

representative DVC genes identified from the DVC analysis.

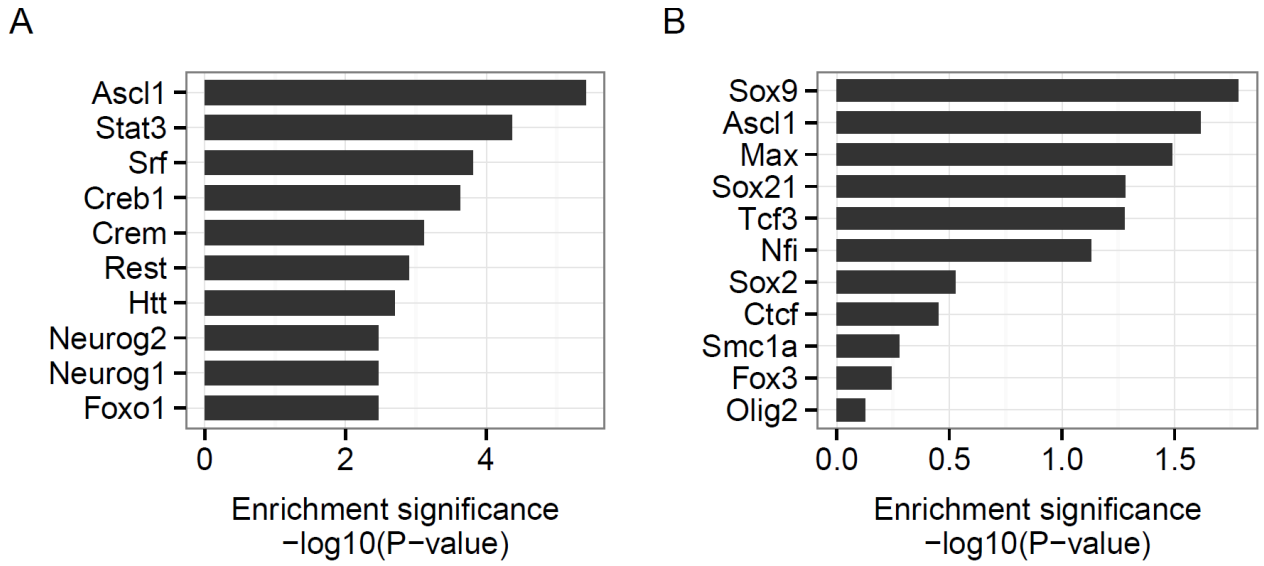


Figure 3-4. Upstream regulators of DVC genes between astrocyte and NSC states. (A) Enrichment significance of transcription factor candidates that can regulate DVC genes obtained using an unbiased upstream regulator analysis. **(B)** Enrichment significance of transcription factor candidates that can regulate DVC genes obtained from the ChIP-seq data.

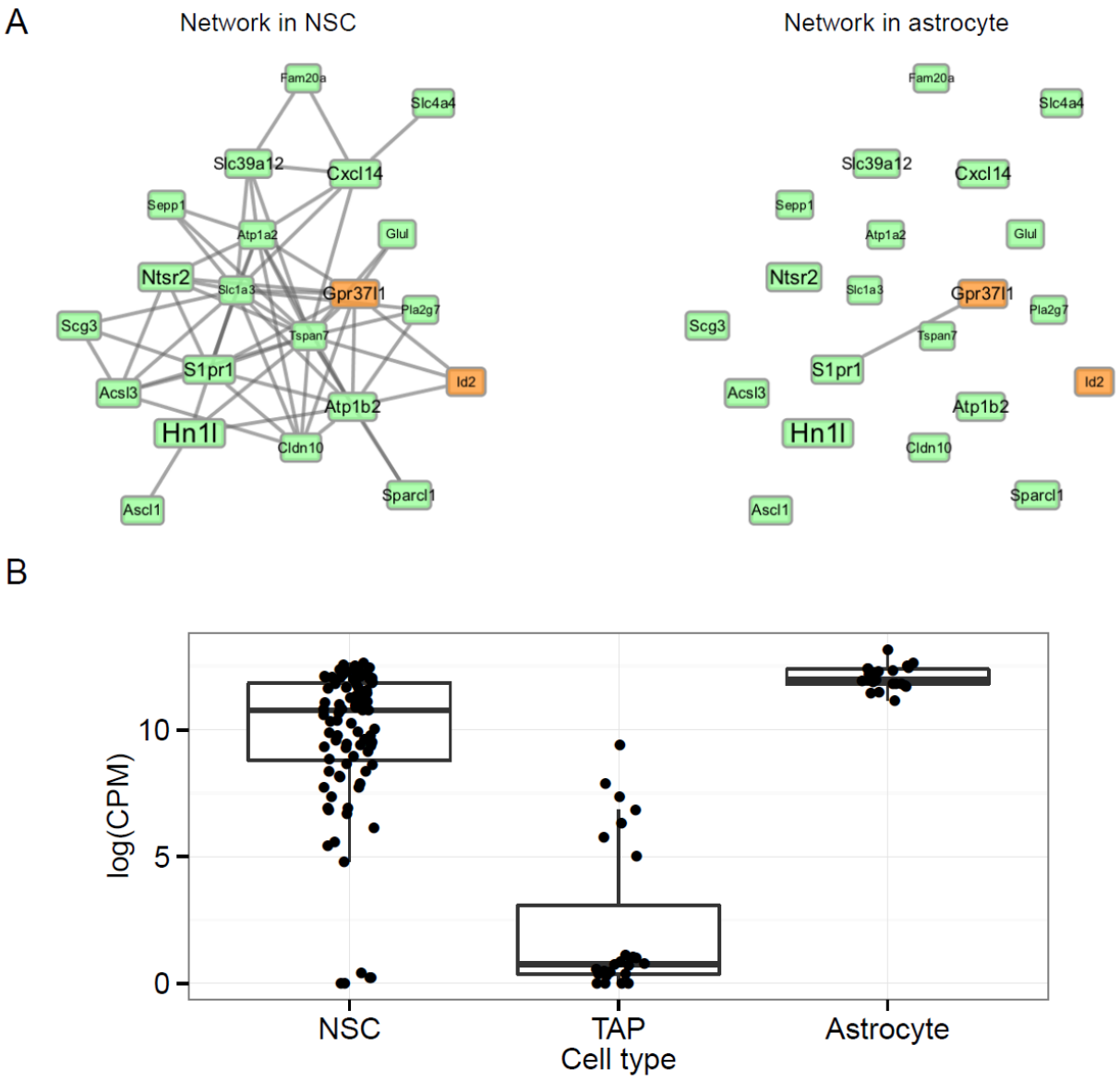


Figure 3-5. Gene network analysis of *Ascl1* and expression profile of *Ntsr2*. **(A)** Co-expression network predicted to be involved with *Ascl1*. The links between nodes represent a strong correlation (correlation coefficient ≥ 0.7). The orange nodes indicate astrocyte differentiation related genes based on Gene Ontology. The node size indicates the system transition score. **(B)** Single cell gene expression profile of *Ntsr2* in the three cell states (NSC, TAP and astrocyte).

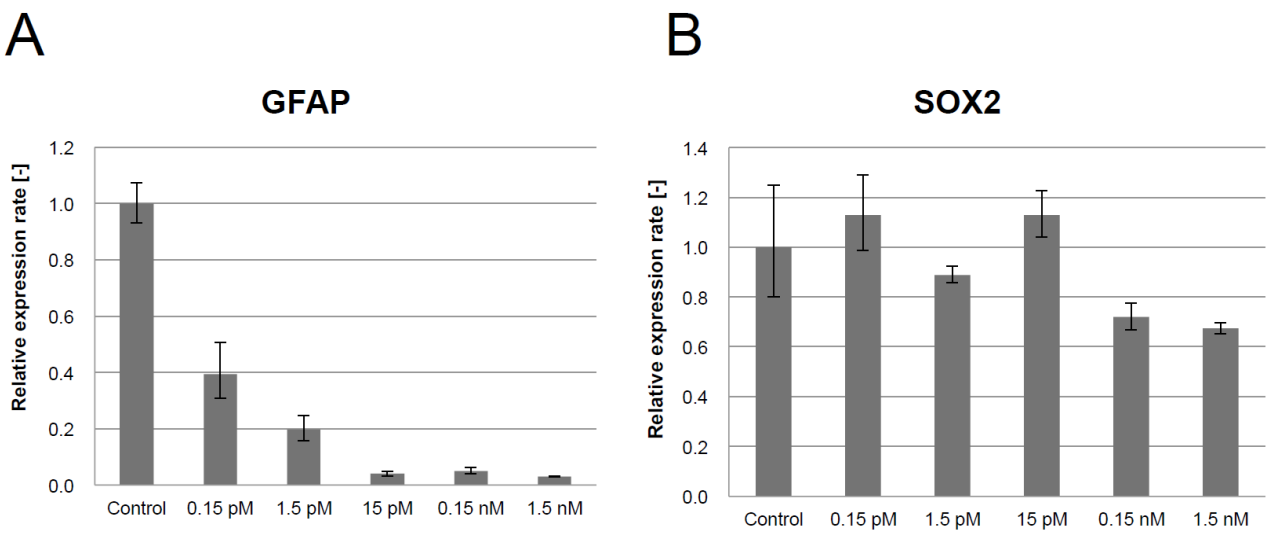


Figure 3-6. Effect of the Ntsr2 inhibitor on the NSC state. (A) Relative expression rate of *Gfap* mRNA in the presence and absence of the Ntsr2 inhibitor (JMV449). **(B)** Relative expression rate of *Sox2* mRNA in the presence and absence of the Ntsr2 inhibitor (JMV449).

3.5. Summary

There are an increasing number of reports that characterize the temporal behavior of gene expression at the single cell level during cell differentiation. Despite accumulation of data describing the heterogeneity of biological responses, the dynamics of gene expression heterogeneity and its regulation during the differentiation process have not been studied systematically. To understand transcriptional heterogeneity during astrocyte differentiation, we analyzed single cell transcriptional data from cells representing the different stages of astrocyte differentiation. When we compared the transcriptional variability of co-expressed genes between the undifferentiated and differentiated states, we found that there was significant increase in transcriptional variability in the undifferentiated state. The genes showing large changes in both "variability" and "correlation" between neural stem cells (NSCs) and astrocytes were found to be functionally involved in astrocyte differentiation. We determined that these genes are potentially regulated by *Ascl1*, a previously known oscillatory gene in NSCs. Pharmacological blockade of *Ntsr2*, which is transcriptionally co-regulated with *Ascl1*, showed that *Ntsr2* may play an important role in the differentiation from NSCs to

astrocytes. This study shows the importance of characterizing transcriptional heterogeneity and rearrangement of the co-regulation network between different cell states. It also highlights the potential for identifying novel regulators of cell differentiation that will further increase our understanding of the molecular mechanisms underlying the differentiation process.

3.6. References

1. Slack MD, Martinez ED, Wu LF, Altschuler SJ. Characterizing heterogeneous cellular responses to perturbations. Proc Natl Acad Sci U S A 2008; 105: 19306–11.

2. Raj A, Rifkin SA, Andersen E, van Oudenaarden A. Variability in gene expression underlies incomplete penetrance. Nature 2010; 463: 913–8.

3. Singh DK, Ku C-J, Wichaidit C, Steininger RJ, Wu LF, Altschuler SJ. Patterns of basal signaling heterogeneity can distinguish cellular populations with different drug sensitivities. Mol Syst Biol 2010; 6: 369.

4. Laslo P, Spooner CJ, Warmflash A, Lancki DW, Lee H-J, Sciammas R, et al. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. Cell 2006; 126: 755–66.

5. Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, Kuperwasser C, et al. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell 2011; 146: 633–44.

6. Tyson DR, Garbett SP, Frick PL, Quaranta V. Fractional proliferation: a method

to deconvolve cell population dynamics from single-cell data. Nat Methods 2012; 9: 923–8.

7 . Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, et al. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat Biotechnol 2015; 33: 155–60.

8 . Shimojo H, Ohtsuka T, Kageyama R. Oscillations in notch signaling regulate maintenance of neural progenitors. Neuron 2008; 58: 52–64.

9 . Imayoshi I, Isomura A, Harima Y, Kawaguchi K, Kori H, Miyachi H, et al. Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. Science 2013; 342: 1203–8.

10 . Scheffer M, Bascompte J, Brock W a, Brovkin V, Carpenter SR, Dakos V, et al. Early-warning signals for critical transitions. Nature 2009; 461: 53–9.

11 . Scheffer M, Carpenter SR, Lenton TM, Bascompte J, Brock W, Dakos V, et al. Anticipating critical transitions. Science 2012; 338: 344–8.

12 . Liu R, Li M, Liu Z-P, Wu J, Chen L, Aihara K. Identifying critical transitions and their leading biomolecular networks in complex diseases. Sci Rep 2012; 2: 813.

13. Ando T, Kato R, Honda H. Differential variability and correlation of gene expression identifies key genes involved in neuronal differentiation. BMC Syst Biol 2015; 9: 82.

14. Hasegawa Y, Taylor D, Ovchinnikov D a., Wolvetang EJ, de Torrenté L, Mar JC. Variability of Gene Expression Identifies Transcriptional Regulators of Early Human Embryonic Development. PLOS Genet 2015; 11: e1005428.

15. Richard A, Boullu L, Herbach U, Bonnafoux A, Morin V, Vallin E, et al. Single-Cell-Based Analysis Highlights a Surge in Cell-to-Cell Molecular Variability Preceding Irreversible Commitment in a Differentiation Process. PLoS Biol 2016; 14: e1002585.

16. Stumpf PS, Smith RCG, Lenz M, Schuppert A, Müller FJ, Babtie A, et al. Stem Cell Differentiation as a Non-Markov Stochastic Process. Cell Syst 2017; 5: 268-282.e7.

17. Mojtahedi M, Skupin A, Zhou J, Castaño IG, Leong-Quong RYY, Chang H, et al. Cell Fate Decision as High-Dimensional Critical State Transition. PLoS Biol 2016; 14: 1–28.

18. Semrau S, Goldmann JE, Soumillon M, Mikkelsen TS, Jaenisch R, van

Oudenaarden A. Dynamics of lineage commitment revealed by single-cell transcriptomics of differentiating embryonic stem cells. Nat Commun 2017; 8: 1096.

19 . Llorens-Bobadilla E, Zhao S, Baser A, Saiz-Castro G, Zwadlo K, Martin-Villalba A. Single-Cell Transcriptomics Reveals a Population of Dormant Neural Stem Cells that Become Activated upon Brain Injury. Cell Stem Cell 2015; 17: 329–340.

20 . Gallo V, Deneen B. Glial development: The crossroads of regeneration and repair in the CNS. Neuron 2014; 83: 283–308.

21 . Dulken BW, Leeman DS, Boutet SC, Hebestreit K, Brunet A. Single-Cell Transcriptomic Analysis Defines Heterogeneity and Transcriptional Dynamics in the Adult Neural Stem Cell Lineage. Cell Rep 2017; 18: 777–790.

22 . Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013; 29: 15–21.

23 . Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014; 30: 923–30.

24 . Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010; 26: 139–40.

25 . Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498–504.

26 . Blake J a, Bult CJ, Eppig JT, Kadin J a, Richardson JE. The Mouse Genome Database: integration of and access to knowledge about the laboratory mouse. Nucleic Acids Res 2014; 42: D810-7.

27 . Gene Ontology Consortium: going forward. Nucleic Acids Res 2015; 43: D1049–D1056.

28 . Mateo JL, van den Berg DLC, Haeussler M, Drechsel D, Gaber ZB, Castro DS, et al. Characterization of the neural stem cell gene regulatory network identifies OLIG2 as a multifunctional regulator of self-renewal. Genome Res 2015; 25: 41–56.

29 . Fujitani M, Huddin NS, Kawai S, Kanie K, Kiyota Y, Shimizu K, et al. Morphology-based non-invasive quantitative prediction of the differentiation

status of neural stem cells. J Biosci Bioeng 2017; 124: 351–358.

30 . Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 2014; 15: R29.

31 . Kageyama R, Ohtsuka T, Kobayashi T. The Hes gene family: repressors and oscillators that orchestrate embryogenesis. Development 2007; 134: 1243–51.

32 . Shimojo H, Isomura A, Ohtsuka T, Kori H, Miyachi H, Kageyama R. Oscillatory control of Delta-like1 in cell interactions regulates dynamic gene expression and tissue morphogenesis. Genes Dev 2016; 30: 102–116.

33 . Rowitch DH, Kriegstein AR. Developmental genetics of vertebrate glial–cell specification. Nature 2010; 468: 214–222.

34 . Molofsky AV, Deneen B. Astrocyte development: A Guide for the Perplexed. Glia 2015; 63: 1320–9.

35 . Castro DS, Martynoga B, Parras C, Ramesh V, Pacary E, Johnston C, et al. A novel function of the proneural factor Ascl1 in progenitor proliferation identified by genome-wide characterization of its targets. Genes Dev 2011; 25: 930–45.

36 . Martini S, Bernoth K, Main H, Ortega GDC, Lendahl U, Just U, et al. A critical role for Sox9 in notch-induced astrogliogenesis and stem cell maintenance. Stem

Cells 2013; 31: 741–51.

37 . Vincent JP, Mazella J, Kitabgi P. Neurotensin and neurotensin receptors. Trends Pharmacol Sci 1999; 20: 302–9.

38 . Dulken BW, Leeman DS, Boutet SC, Hebestreit K, Brunet A. Single-Cell Transcriptomic Analysis Defines Heterogeneity and Transcriptional Dynamics in the Adult Neural Stem Cell Lineage. Cell Rep 2017; 18: 777–790.

39 . Mu Y, Lee SW, Gage FH. Signaling in adult neurogenesis. Curr Opin Neurobiol 2010; 20: 416–423.

40 . Suh H, Consiglio A, Ray J, Sawai T, D'Amour KA, Gage FH. In Vivo Fate Analysis Reveals the Multipotent and Self-Renewal Capacities of Sox2+ Neural Stem Cells in the Adult Hippocampus. Cell Stem Cell 2007; 1: 515–528.

Chapter 4

Concluding remarks

Heterogeneity in cell populations has been known to potentially affect various biological processes from the recent works. However, it is still difficult to analyze such heterogeneity to further interpret the underlying molecular mechanism without prior knowledge of the genes involved in phenotype heterogeneity. Inspired by the dynamical systems theory, my goal of this thesis was to establish a novel unbiased and systematic analysis method of cellular heterogeneity during cellular differentiation. I generated the data driven hypothesis to identify the potential regulator genes showing heterogeneity in cell populations and validated it experimentally. In this study, I focused to establish the framework from data driven hypothesis generation to experimental validation on cellular heterogeneity research with the model data of

neuronal and glial differentiation.

In **Chapter 1**, a general introduction offered current challenges and opportunities of cellular heterogeneity researches to be discussed. The objective of this thesis was described based on them

In **Chapter 2**, the transcriptional variability and co-expression analysis were applied to the gene expression data of neuronal differentiation in mouse ES and human iPS cells to investigate dynamic alternation of cellular heterogeneity in the neuronal differentiation process. Although the gene expression data was the average data from the bulk cellular population, the analysis was able to detect transcriptional variability change and revealed that most of the genes co-expressed in the neuronal stem cell decrease their transcriptional variability. The DVC genes are not only predictive biomarker for neuronal differentiation but also play an important role in neuronal differentiation. The dynamic behavior of the genes is preserved between mice and human in neuronal differentiation.

In **Chapter 3**, the analytical concept established in Chapter 2 was extended to the single-cell transcriptional data in astrocyte differentiation. The key findings that the genes co-expressed in the stem cells decreased their transcriptional variability were also observed in the astrocyte differentiation. The genes with such dynamic behavior included the know oscillation genes including Ascl1 which regulated neuronal and glial cell fate. The modulation of Ntsr2 transcriptionally co-regulated with Ascl1 showed that Ntsr2 was involved in an early cell fate decision.

Throughout this work, the dynamic behavior of cellular heterogeneity during differentiation process was examined and used as the core concept to identify the regulators in the critical transition of cellular states. The unbiased and systematic analysis of transcriptional data gained novel mechanistic insight into cellular heterogeneity in neuronal differentiation (Chapter 2) and astrocyte differentiation (Chapter 3). I conclude that the analytical framework proposed in this thesis made a significant step forward into cellular heterogeneity science, and will accelerate the development of effective therapeutics and regenerative medicine.

List of publications for dissertation

Research Articles

1. Tatsuya Ando, Ryuji Kato and Hiroyuki Honda. Differential variability and correlation of gene expression identifies key genes involved in neuronal differentiation. BMC Syst Biol. 2015; 19;9:82.
2. Tatsuya Ando, Ryuji Kato and Hiroyuki Honda. Identification of an early cell fate regulator by detecting dynamics in transcriptional heterogeneity and co-regulation during astrocyte differentiation. NPJ Syst Biol Appl. 2019; 5:18.

Acknowledgments

This study has been carried out in Laboratory of Bioprocess Engineering, Department of Biotechnology, Graduate School of Engineering, Nagoya University, Japan.

First of all, the author would like to express the sincerest gratitude to Professor Hiroyuki Honda (Innovative Research Center for Preventive Medical Engineering (PME)/Department of Biomolecular Engineering, Graduate School of Engineering, Nagoya University) for his supreme leading, support, and assistance for this research. The research could never be completed without Professor Honda's precise thesis selection and his advice leading to many opportunities to grow as a researcher.

The author would like to express deep gratefulness to Professor Shigeki Kiyonaka (Department of Department of Biomolecular Engineering, Graduate School of Engineering, Nagoya University), Associate Professor Kazunori Shimizu (Department of Department of Biomolecular Engineering, Graduate School of Engineering, Nagoya University), and Associate Professor Masahiro Nakatochi (Department of Nursing, Nagoya University Graduate School of Medicine), for their

thoughtful advice to the research.

The author would like to express deep appreciation to Associate Professor Ryuji Kato (Department of Basic Medicinal Sciences, Graduate School of Pharmaceutical Sciences, Nagoya University) for his precise leading, thoughtful advises, and many encouragements. This research could never be completed without his helpful suggestions.

Finally, the author is expressing the deepest appreciation to his wife and his daughter and all of the family members who have supported his life for devoting everything to the doctor course.

May, 2019, Tatsuya Ando