Self-Renewal Pattern-Associated Genes and Their Role in Adult Stem Cell Function

by

Minsoo Noh

B.S. Manufacturing Pharmacy, 1993; M.S. Pharmacology, 1995
Seoul National University, Seoul, Republic of Korea

Submitted to the Division of Biological Engineering in
Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Molecular and Systems Toxicology and Pharmacology
at the
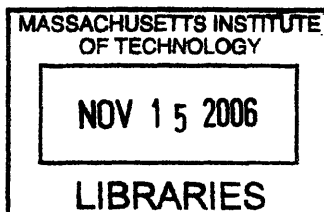Massachusetts Institute of Technology

June 2006

Signatue of Author:

Minsoo Noh
Division of Biological Engineering
May 16, 2006

Certified by:

James L. Sherley, M.D. Ph.D.
Associate Professor of Biological Engineering
Thesis Supervisor

Accepted by:

Alan J. Grodzinsky, Ph.D.
Professor of Electrical, Mechanical, and Biological Engineering

Self-Renewal Pattern-Associated Genes and Their Role in Adult Stem Cell Function

by

Minsoo Noh

B.S. Manufacturing Pharmacy, 1993; M.S. Pharmacology, 1995
Seoul National University, Seoul, Republic of Korea

Submitted to the Division of Biological Engineering in
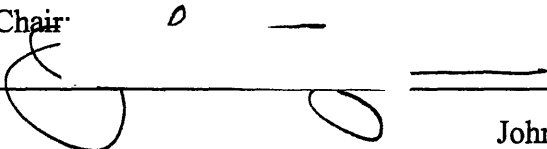Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Molecular and Systems Toxicology and Pharmacology
at the
Massachusetts Institute of Technology
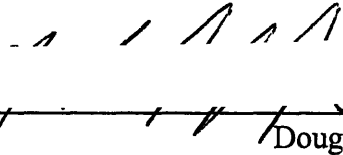
June 2006

Signature of Committee Chair·

John M. Essigmann, Ph.D.
Professor of Chemistry, Toxicology, and Biological Engineering

Signatures of Committee Members:

Douglas A. Lauffenburger, Ph.D.
Whitaker Professor of Biological Engineering, Chemical Engineering, and Biology

David B. Schauer, Ph.D.
Professor of Biological Engineering and Comparative Medicine

Table of Contents

Abstract

Molecular markers for adult stem cells (ASCs) are highly demanded for research and clinical applications. The development of specific molecular markers for ASCs has been difficult mainly due to the technical barriers in the identification and isolation of rare ASCs. Previously, reported transcriptional profiling studies for defining molecular features of ASCs were compromised by the use of impure ASC preparations. Thesis for this research was that the study of asymmetric self-renewal, a defining property of ASCs, might provide key clues to understanding ASC function and lead to discovery of novel molecular markers for ASCs. Fifty two self-renewal pattern associated (SRPA) genes were identified by cDNA microarray analysis with cell culture models whose self-renewal pattern could be reversibly regulated, instead of using heterogeneous ASC-enriched populations. From evaluation of whole genome transcript levels to expand the SRPA gene pool, 543 SRPA genes were discovered. Both microarray studies showed that asymmetric self-renewal associated (ASRA) genes were highly represented in ASC-enriched populations but not in embryonic stem cells. The SRPA gene expression signature successfully distinguished isolated ASC-enriched populations from non-stem cell populations by principal component analysis (PCA). The SRPA gene signature clustered and classified putative epidermal stem cell-enriched populations better than reported stemness gene signatures in PCA. Therefore, gene microarray analyses for studying self-renewal pattern *per se* confirmed for the first time that asymmetric self-renewal is an essential molecular feature of ASCs *in vivo*. Chromosome mapping of the SRPA genes identified two SRPA chromosome gene cluster regions. One chromosome cluster contained primarily ASRA genes, whereas the other contained primarily symmetric self-renewal associated (SSRA) genes. These two SRPA chromosome cluster regions are frequently rearranged or deleted in particular human cancers. Functional and expression analysis of several selected ASRA and SSRA gene-encoded proteins implicated them in control of asymmetric self-renewal and non-random chromosome co-segregation, respectively. Moreover, one plasma membrane bound ASRA protein, CXCR6, had properties of one of the most specific molecular markers for ASCs described to date. In conclusion, this research strongly supported the precept that asymmetric self-renewal is a unique molecular feature for understanding ASCs, their relation to cancer, their unique function, and for their eventual exclusive identification.

Introduction: asymmetric self-renewal of adult stem cells

Stem cells are essential for the life-time maintenance of normal adult tissues. Functionally, stem cells have been defined by self-renewal and ability to produce differentiated progeny cells in adult tissues (Reya et al., 2001; Sherley, 2002). Because of great potential for cell replacement therapy, gene therapy, and tissue engineering, stem cells have been drawing more attention for recent years. However, there is no current experimental tool to distinguish a single stem cell from any given cells in an adult mammalian tissue (Loeffler and Potten, 1997; Sherley, 2002; Zipori, 2004). In stem cell biology and its engineering, many highly sought after questions still remain unanswered: What molecular features define a stem cell? What determines a stem cell fate decision to self-renewal and to differentiation into a progeny? How does one identify and isolate stem cells from adult tissues? The studies to address these issues have been difficult because of the inability to purify homogenous stem cell populations from adult tissues or expand them to sufficient quantities for analyses in culture (Reya et al., 2001; Sherley, 2002).

Stem cell self-renewal models and current problems in the application of stem cell definitions to self-renewing cell populations

The two stem cell defining properties, self-renewal and ability to produce differentiated progeny cells, were originated from the study of unique cells responsible for active tissue renewal in hematopoietic system and other adult tissues (Loeffler and Potten, 1997; Weissman et al., 2001; Zipori, 2004). The stem cell definition in adult tissues has

been adopted by embryonic stem cells (ESCs) and cancer stem cells (Fuchs and Segre, 2000; Reya et al., 2001; Weissman et al., 2001; Pardal et al., 2003). Before ESCs emerged (Evans and Kaufmann, 1981), the concept of stem cell self-renewal had been evaluated based on functions in adult tissues (Lajtha, 1979; Loeffler and Potten, 1997; Sherley, 2002). In order to explain how stem cells play a role in adult tissue maintenance, scientists have suggested several models for tissue renewal. Although there are minor differences in specific details, adult tissue renewal can be classified into three models with a stem cell concept: clonal succession, stochastic self-renewal, and deterministic self-renewal (Loeffler and Potten, 1997).

The clonal succession model postulates a pool of dormant stem cells from which individual stem cells can be triggered to generate progeny cells when needed (Figure 1a). Stem cells in this model are not self-renewing and can be considered as original parent cells of differentiated progeny cells. This model has been developed from the studies of somatic mutation of a specific marker in hematopoietic systems (Abkowitz et al., 1993). However, the compartment which may support the clonal succession model has not been found in animal tissues (Loeffler and Potten, 1997).
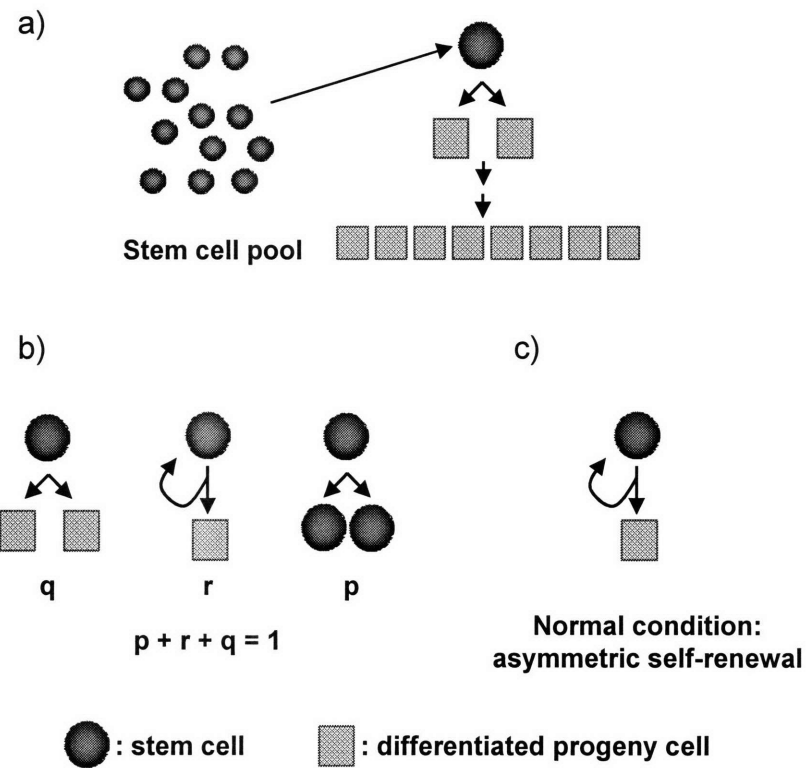
Figure 1. **Tissue stem cell renewal models**. a) clonal succession model. b) stochastic stem cell self-renewal model. The fate decision for stem cell extinction, q, asymmetric stem cell self-renewal, r, and symmetric stem cell expansion, p, is modeled as a random process. c) deterministic stem cell self-renewal model. In each stem cell division, two qualitatively different daughter cells are generated. One is a stem cell and another is a differentiated progeny cell.

The stochastic model for stem cell self-renewal was proposed to explain the spleen colony forming assay of hematopoietic cells (Figure 1b) (Vogel et al., 1969; Matioli et al., 1970). In contrast to the clonal succession model, a self-renewing stem cell exists in the stochastic model (Loeffler and Potten, 1997). A stem cell can perform basically three cell division processes, the symmetric renewal of two stem cells, the asymmetric generation of a stem cell and a differentiated daughter cell, and the symmetric production of two differentiated daughter cells. In the basic assumption of this model, a stem cell can divide with each of all three possible division processes by random choice, and each division cannot be affected by the previous division patterns. In the condition that zero, one, or two stem cells are generated with probabilities of p, r, or q, respectively, the model stipulates the sum of p, r, and q as 1. Because p is not zero ($p \neq 0$) in this model, there is a chance that stem cell pools become extinct by at some frequency generating only differentiated daughter cells. The experimental observation on intestinal crypt generation and loss for 3 weeks indicated that asymmetric self-renewing probability could not be excluded in any stochastic model (Bjerknes, 1994). Most interestingly, from the model simulation with mouse intestinal crypts, the asymmetric self-renewing probability, r, was over 0.95, which suggested that the most crypt stem cell divisions are asymmetric (Loeffler et al., 1993 and 1997).

In the deterministic model, a stem cell in adult tissues gives rise to two functionally distinct cells by each cell division, supporting asymmetric self-renewal in normal physiological condition (Loeffler and Potten, 1997). One cell retains stem cell properties and stays in the stem cell pool whereas another cell becomes a differentiated

cell and leaves the stem cell pool (Figure 1c). In this model, self-renewal is not a random process and overall cell population size is maintained at a specific level under normal conditions. In contrast, in the stochastic model, in which stem cell fate to take one of three division patterns is randomly selected, there may be fluctuations in the cell population size in adult tissues (Loeffler and Potten, 1997). Currently, there is no direct experimental evidence of cell population variation in any normal organ.

Under normal physiological conditions, experimental observations supported the hypothesis that stem cells generally undergo asymmetric self-renewal to maintain specific tissue mass whether by a stochastic model or a deterministic model. This unique self-renewal was also demonstrated in ex vivo human hematopoietic stem cell system by time-lapse microscopic studies (Punzel and Ho 2001; Punzel et al., 2003). In addition, evidence for the adult stem cell self-renewal was shown in human esophageal and intestinal epithelium (Seery and Watt, 2000; Yatabe et al., 2001).

Historically, the self-renewal models of adult stem cells have been discussed under the context of in vivo physiological tissue function in adult tissues. Recently, many scientists have applied the concept of self-renewal to cell states like embryo-derived cells or cancer-derived cells (Fuchs and Segre, 2000; Reya et al., 2001; Pardal et al., 2003; Huntley and Gilliland, 2005). In contrast to adult stem cells, the self-renewal of ESCs is based entirely on in vitro studies. ESCs are maintained by symmetric self-renewal in cell culture. Although stem-like cells exist in several human cancers, it is evident that the regulation of cancer stem cell self-renewal differs from that of adult stem cell self-renewal (Pardal et al., 2003; Radtke and Clevers, 2005). ASC self-renewal maintains mass balance

8

in adult tissues whereas cancer stem cell self-renewal breaks it. Therefore, knowledge of the differences in the control mechanisms of self-renewal between adult stem cells and cancer stem cells may reveal new mechanisms of carcinogenesis.

Self-renewal symmetry of adult stem cells

Adult stem cells (ASCs) self-renew asymmetrically *in vivo*. In deterministic models, one daughter cell retains the self-renewal potential of its stem cell parent, whereas the second daughter cell produces a differentiated non-dividing cell or a transit amplifying cell. The transit amplifying cell may undergo a limited number of cell divisions before becoming terminally differentiated cells (Loeffler and Potten, 1997; Merok and Sherley, 2001; Sherley, 2002). Adult stem cells also undergo symmetric self-renewal in specific conditions like wound healing and early active growth period in life. In addition, ASCs may stay dormant tissues and begin to self-renew during replenishment of tissue loss or tissue regeneration (Figure 2). The fate of ASCs into three possible self-renewal symmetry states may be decided by highly regulated mechanisms. Disruption of the self-renewal symmetry regulation in ASCs may cause many hyperproliferative conditions like cancer (Radtke and Clevers, 2005). This is especially true in continually renewing tissues, if stem cells acquire sufficient gene mutations for eventual progression to neoplastic cell growth (Carins, 1975 and 2002). However, the in vivo self-renewal symmetry control of ASCs is poorly understood.

9
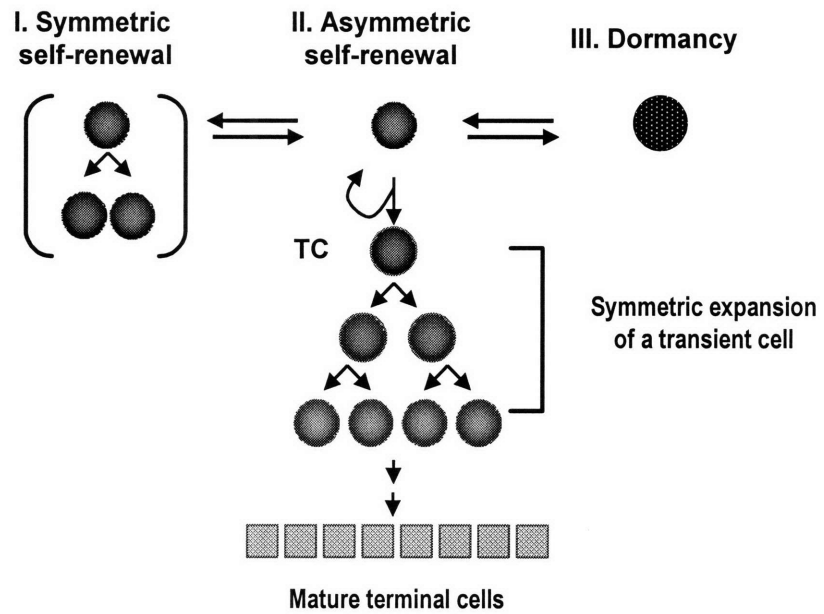
Figure 2. Types of self-renewal in ASCs in a tissue turnover unit. In vivo, ASCs can undergo three self-renewal programs: I) limited symmetric self-renewal, e.g. tissue repair, early embryonic development, etc.; II) asymmetric self-renewal; and III) dormancy. Depending on tissue types, the progeny cell of transient amplifying cell (TC) divisions commits to a different fate.

In spite of its importance, the asymmetric self-renewal of ASCs is rarely recognized in stem cell research (Sherley 2002). Most research efforts have been directed to find methods to transform ESCs into differentiated adult cells or to transdifferentiate ASCs in vivo (Fuchs and Segre, 2000; Wagers and Weissman, 2004; Hoffman and Carpenter, 2005). However, even in the therapeutic application of ESCs, the asymmetric self-renewal of engineered adult cells is important in reconstituted adult tissues (Sherley, 2002). When ESCs are engineered into specific differentiated lineage cells, the engineered cells should divide with asymmetric self-renewal to prevent disastrous tumorigenesis by the uncontrollable symmetric expansion and to provide long-term cell replacement. However, this viewpoint of asymmetric self-renewal has been over-looked in stem cell research, partly due to difficulties in identification of ASCs from adult tissues in vivo (Sherley, 2002).

Asymmetric chromosome co-segregation of adult stem cells

In addition to the fate determination during each cell division, another level of asymmetry exists when ASCs divide. During mitosis, ASCs continuously co-segregate a complement of chromosomes with the oldest parental DNA strands specifically to their stem cell daughters. These unique parental DNA molecules are referred to as immortal DNA strands, hypothesized by Carins (Cairns, 1975). This unique co-segregation of the oldest parental DNA strands in ASCs has been demonstrated in several mammalian cell systems (Merok et al., 2002; Potten et al., 2002; Rambhatla et al., 2005; Karpowicz et al.,

11

2005). In contrast, in transient amplifying tissue cells and in cancer cells, random chromosome segregation occurs during mitosis (Potten et al., 2002, Merok et al., 2002).

By adopting non-random chromosome co-segregation, stem cells may protect themselves from genetic mutations derived from DNA replication errors (Cairns, 1975 and 2002; Merok et al., 2002; Rambhatla et al., 2005). Many proteins regulating cytokinesis and proper chromosome segregation during mitosis have been implicated in various human proliferative diseases. Aurora-B kinase and survivin are examples of such proteins (Altieri, 2003; Carmena and Earnshaw, 2003). However, this aspect of non-random chromosome co-segregation in ASCs and its significance in human diseases has not been thoroughly studied. Recently, the association between asymmetric stem cell division with centrosome function during mitosis was uncovered in dividing Drosophila male germline stem cells (Yamashita et al., 2003). This result suggests that chromosome segregation in ASCs is much more organized and controlled than that of transit amplifying cells or other symmetrically expanding cells. Therefore, genes associated with asymmetric self-renewal in ASCs may provide many key clues to understand non-random chromosome co-segregation and its relation to stem cell function.

Major challenges for stem cell biology and gene expression signature studies for stem cells

The isolation of ASCs and their initial identification in vivo have been major challenges in stem cell research and biomedical applications, mainly due to their rarity in tissues and lack of specific molecular markers for ASCs (Sherely, 2002; Zipori, 2004;

Eckfeldt et al., 2005). Currently, stem cells can only be defined and characterized retrospectively by limited tissue repopulation assays (Loeffler and Potten, 1997; Zipori, 2004). Stem cell-associated markers in adult tissues have been conventionally identified from time consuming and labor intensive studies like *in vivo* reconstitution assay with cells purified with CD (cluster of differentiation)-based immuno-selection. For example, CD34, Sca-1, and Lin-("the absence of differentiation specific markers") for hematopoietic stem cells (HSCs) are representative of markers discovered through this approach (Weissman et al., 2001). In spite of significance in the current characterization for stem cells, conventional approaches to characterize molecular entities associated with stem cells have been time consuming, low-throughput, and somewhat dependent on serendipity. Even though many molecular candidates have been studied as HSC markers, no single molecular entity has guaranteed the initial identification of HSCs so far. These markers have been widely used for enriching HSCs to certain level, but they have not allowed purification to homogeneity.

Recently, transcriptional profiling studies with stem cell-enriched populations have been performed to find a set of commonly expressed genes (so-called 'stemness' genes) in any stem cell population, adult or embryonic (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003). Theoretically, discovery of a specific common gene expression signature for stem cells may be expected through these genome-wide transcript assessments with stem cell-enriched populations. Although experimental designs and procedures were virtually identical among these independent microarray studies, poor gene overlaps among these studies have brought lots of controversy regarding the

reliability of the data and also the general soundness of the theory (Figure 3a) (Burns and Zon, 2002; Zipori, 2004; Cai et al., 2004; Eckfeldt et al., 2005).

The initial identification of stem cells in adult tissues has been impossible and, as a result, the isolation of pure adult stem cells from mammalian tissues has not been accomplished. This accounts for confounded data in transcriptional profiling studies (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003; Ivanova et al., 2003). The more reliable data for describing a stem cell-specific gene expression signature is highly dependent on the reliability and purity of stem cell preparations. However, these problems issued from identification and purification of stem cells cannot be resolved with conventional stem cell-enriched preparations (Burns and Zon, 2002; Zipori, 2004; Cai et al., 2004). Even authors who reported the transcriptional profiling studies acknowledged limitations due to the use of heterogeneous cell preparations for neural stem cells and the use of in vitro adapted embryonic stem cells (Ivanova et al., 2003). In addition, the lack of overlap among these stemness studies may be due to the arbitrary choice of non-stem cell controls to select stem cell-specific genes. For instance, the Ramalho-Santos et al. study compared the gene expression of stem cell-enriched preparations with that of differentiated cells from lateral ventricles of the brain, whereas the Ivanova et al. study used adult mature blood cells (Ivanova et al., 2002; Ramalho-Santos et al., 2002).

Beyond these technical issues, there is more fundamental limitation in terms of the self-renewal symmetry of an individual stem cell-enriched population, which has previously gone without recognition. Stem cells in HSC-enriched and other isolated stem

cell-enriched populations from mammalian tissues are under the control of asymmetric self-renewal programs in vivo (Punzel et al., 2002 and 2003), whereas ESCs and neurospheres are cultured to self-renew symmetrically in vitro (Evans and Kaufmann, 1981; Wiles and Johansson, 1999; Schuldiner et al., 2000; Carpenter et al., 1999; Gottlieb 2002) (Figure 3b). Therefore, previous studies that sought to identify genes commonly expressed in the selected stem cell population may have inevitably excluded genes that are ASC-specific, but related to the pattern of self-renewal type (asymmetric vs. symmetric).

Asymmetric self-renewal as a theme for ASC research

In adult tissues, transient amplifying cells multiply their number by limited symmetric expansions and also have the ability to produce daughter cells of multiple lineage (Loeffler and Potten, 1997; Sherley 2002; Zipori, 2004). If the self-renewal and differentiation potential are considered in stem cell analyses, it is extremely difficult to discriminate the stem cell from transient cells without mentioning asymmetric self-renewal. Properties of stem cell-enriched populations which self-renew symmetrically in culture may be more close to characteristics of transit amplifying cells. Therefore, the asymmetric self-renewal of ASCs should be a major consideration when new transcriptional profiling studies are performed to find common molecular and cellular signatures that identify diverse cell populations specifically.

15
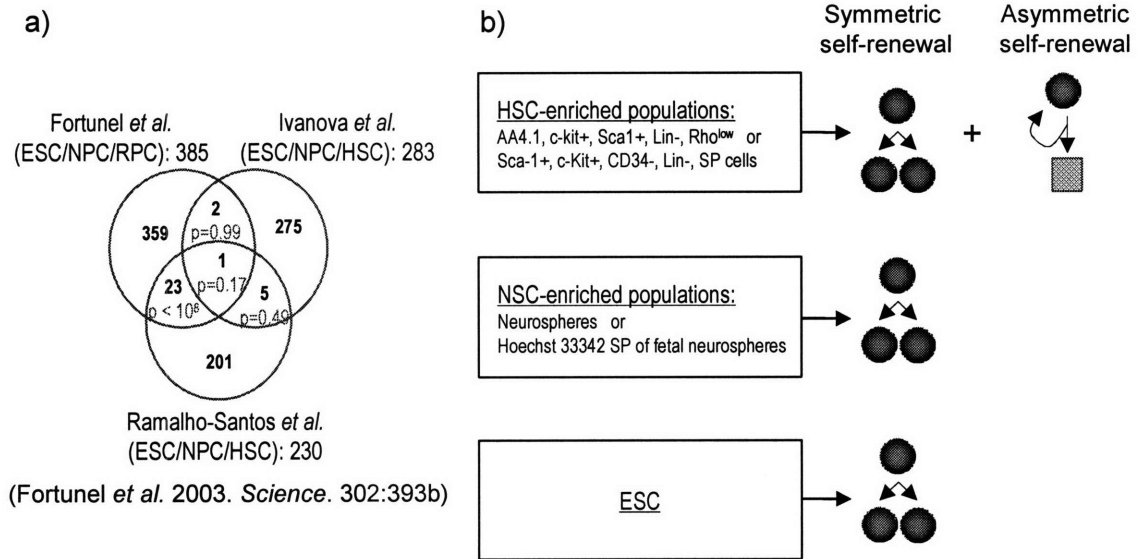
Figure 3. Gene overlap between stemness studies (a) Overlapping genes among the three stemness transcriptional profiling studies (Fortunel et al., 2003) (b) Predicted self-renewal symmetry pattern of stem cell-enriched populations and embryonic stem cells on the "stemness" studies.

16

Several genes have been described to have an association with the regulation of asymmetric self-renewal (Sherley, 2002). To date, at least two cellular factors have been demonstrated to regulate the asymmetric self-renewal in cell culture. The tumor suppressor p53 is one factor (Sherely, 1991; Sherley et al., 1995a ; Rambhatla et al., 2001; Sherley 2002; Merok et al., 2002). The p53 gene is well known for its role in normal cell growth regulation in diverse mammalian tissues. However, the normal cellular function of the p53 protein on self-renewal decision is not fully understood (Sharpless and DePinho, 2002). Several cultured murine cell lines that exhibit reversible asymmetric self-renewal in response to controlled expression of the wild-type murine p53 have been reported (Sherely, 1991; Sherley et al., 1995a; Rambhatla et al., 2001).

The second factor, inosine-monophosphate dehydrogenase (IMPDH), is the rate-limiting enzyme for guanine nucleotide biosynthesis in normal cells (Liu et al., 1998a and 1998b; Rambhatla et al., 2005). IMPDH is highly expressed in many human cancers, and down-regulation of IMPDH with specific inhibitors suppresses cell growth in culture. Under physiological conditions, p53 regulates the activity of the enzyme and controls the guanine ribonucleotide pool in dividing cells. Xanthine nucleosides, salvage precursors for cellular IMPDH products, convert cells dividing with asymmetric self-renewal to symmetric self-renewal in culture (Sherley 2002; Lee et al., 2003). Importantly, adult hepatic stem cell lines were successfully isolated and established by the modulation of p53/IMPDH pathways (Lee et al., 2003). This result indicated a role for p53 and IMPDH in regulating the asymmetric self-renewal of ASCs in vivo, too.

Strategy to find novel molecular markers for adult stem cells

Molecular markers which define ASCs in situ or in vivo are extremely valuable targets in stem cell research and engineering. It has been practically impossible to isolate ASCs up to 100% purity, making it difficult to investigate molecular mechanisms of the self-renewal symmetry regulation of ASCs and their related functions (Sherley, 2002). At present, no unique marker for identifying and purifying ASCs is available. Only adult stem cell enrichment markers like CD133 and CD34 have been widely used for HSC research (Weissman et al., 2001; Hess et al., 2006). These markers have limited use for general application to all ASCs. As seen from the stemness studies, many attempts to find genetic signature for stem cells have been complicated by using heterogeneous stem cell-enriched populations like neurospheres and HSC-enriched cell preparations (Burns and Zon, 2002; Cai et al., 2004).

My thesis research is based on the premise that the study of asymmetric self-renewal, the defining property of adult stem cells, may provide key clues to understand adult stem cell function and lead to discovery of novel molecular markers for adult stem cells. The purpose of this dissertation was to identify genes with potential for involvement in self-renewal symmetry mechanisms by microarray analyses, to study the role of self-renewal symmetry associated genes in the regulation of ASC functions, and to evaluate self-renewal symmetry-associated genes for their potential to be novel molecular markers that identify ASCs exclusively.

As a strategy to discover self-renewal symmetry associated genes, in vitro cell culture models whose self-renewal pattern mimics the asymmetric self-renewal of ASCs

were employed, instead of using heterogeneous stem cell-enriched preparations. Through this approach, a new set of genes associated with deterministic asymmetric self-renewal was identified. Members of this gene set provide novel insights to adult stem cell function, are implicated to play a role in human carcinogenesis, and act as a new molecular signature for adult stem cells derived from diverse tissues in diverse mammalian species.

Chapter 1. Development of a Population Doubling Cycle Ratio Method as a Quality Control Parameter for Self-Renewal Patterns of Engineered Cell Models Used in Microarray Analyses

<u>Rationale</u>:

Cell impurities in ASC-enriched populations have been considered as a main culprit of technical limitations in several transcriptional profiling studies for stem cells (Ivanove et al., 2003; Cai et al., 2004). In this study, the gene expression signature of self-renewal symmetry regulation was postulated to provide a key feature of adult stem cells that might serve as a basis for their identification. For transcriptional profiling studies to find gene signatures which may be associated with ASCs, cultured model cells whose self-renewal symmetry could be reversibly regulated between asymmetric and symmetric self-renewal. The advantage of using these experimental model systems is that the self-renewal patterns of these cell populations are nearly pure, in contrast to heterogeneous stem cell-enriched populations used by previous investigators.

For replicate microarray analyses, it was essential to insure precision and accuracy in the cell kinetics states of examined cell populations. In terms of cell division kinetics, ideal asymmetric self-renewal results in linear population growth kinetics in culture, whereas symmetric self-renewal manifests as exponential kinetics. Common analyses for mammalian cell proliferation in cell culture have been based on the exponential growth equation (EGE), $N_t = N_0 e^{kt}$. In symmetric self-renewal conditions, the EGE can be used to calculate parameters like the culture's population doubling time (DT). However, for asymmetric self-renewal, both dividing daughter cells and non-dividing daughters can be

produced (Sherley et al., 1995b; Rambhatla et al., 2001). The EGE is misleading when applied to the cell division kinetics of asymmetrically self-renewing cell populations (Sherley et al., 1995b). Based on a cell growth progression series that incorporated the idea of the generation of both dividing and non-dividing daughter cells in a asymmetrically self-renewing population, the following equation was derived (Sherley et al., 1995b, Equation [1])

$$N_t = N_0 \left[0.5 + \frac{1 - (2 F_d)^{(t/GT + 1)}}{2(1 - 2 F_d)} \right] \qquad [1]$$

$$N_t = (N_0/2GT) \, t + N_0, \quad (F_d = 0.5)$$

In the equation [1], N_0 and N_t were cell number at the time 0 and at a given time t. Generation time (GT) was the time to require to complete each cell cycle in continuously dividing cells. F_d was defined as dividing cell fraction in the cell population growth. Population doubling cycle (PDC) has been used to calculate how many cell cycles are completed by an initial dividing cell population. In other words, PDC removes the confounding effect of non-dividing cell fractions on the calculation of cell division kinetic parameters (Sherley et al., 1995b). The PDC of symmetric self-renewal was calculated with the equation [2] derived from the EGE. The PDC for asymmetric self-renewal was calculated with the equation [3]

$$PDC_{sym} = \frac{Ln(N_t / N_0)}{Ln2} \qquad [2]$$

$$PDC_{asym} = \frac{N_t - N_0}{F_d \times N_0} \qquad [3]$$

PDC is the measure of the number of times that an initial dividing cell population completes a cell cycle. PDC quantifies cycling rate when F_d is known independent of non-cycling cells unlike DT. However, PDC is not useful for specifying whether a cell culture is dividing by asymmetric vs symmetric cell kinetics. In other words, in order for the calculation of the correct PDC for a cell population at any given time, the dividing fraction (F_d) in cell culture must be known or assumed. For the purpose of the quality control in the self-renewal patterns of engineered model cells in microarray analyses, a new cell kinetics parameter, the PDC_{sym}/PDC_{asym} ratio, was developed for estimating the dividing fraction or the non-dividing fraction without knowing or assuming the states of self-renewal pattern.

To validate the new quality control parameter, cell kinetic studies were performed by using the engineered mouse embryonic fibroblasts whose self-renewal pattern was regulated by conditional p53 gene expression (Liu et al., 1998b; Rambhatla et al., 2001).

Materials and Methods:

Cell Culture

For cell division kinetics analyses, engineered mouse embryonic fibroblasts (MEFs) that conditionally express p53 gene (Ind-8 cells) and control-vector transfected MEFs (Con-3) cells were used. As a result of the stable transfection of p53 gene into the p53 null-MEFs under the control of a zinc (Zn) responsive promoter (Rambhatla et al.

22

2001, Merok et al. 2002), the dividing fraction of MEFs in culture can be controlled by varying zinc chloride concentration. Ind-8 cells and Con-3 cells were maintained as described (Rambhatla et al., 2001). For the analyses of cell division kinetics, cells were grown in Zn-free media over a 3-day period to about 50 % confluency, trypsinized, and replated in zinc-free medium (DMEM, 10% dialyzed fetal bovine serum (DFBS), 5 μg/ml puromycin) at a cell:plating area:medium volume ratio of 1×10^5:75 cm^2:20 ml. This ratio was held constant for all experiments, unless specified otherwise. Sixteen to 24 hours later, the culture medium was replaced with the same volume of zinc-free medium or medium containing the specified concentration of $ZnCl_2$. This time was designated as 0 hour in the analyses. Cells were harvested by trypsinization and counted with a Model ZM Coulter Counter to quantify cell kinetics.

Analysis of Cell Kinetics and Estimation of Dividing Fraction in Cell Culture

Ideal cell growth curves were generated for exponential cell growth equation, $N_t = N_0 e^{kt}$, and linear cell growth equation at various F_d (Equation [1]). From the theoretical cell numbers at each generation time, PDC_{sym} and PDC_{asym} were calculated by equations [2] and [3], respectively (Table 1 and Figure 4). The cell kinetics parameters of Ind-8 and Con-3 cells were calculated for the estimation of the dividing fraction in cell cultures with the theoretical PDC_{sym}/PDC_{asym} ratio curves.

GT \ F₀	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
1	0.50	0.51	0.53	0.55	0.57	0.58	0.61	0.63	0.66	0.69
2	0.33	0.36	0.39	0.42	0.46	0.50	0.54	0.59	0.64	0.68
3	0.21	0.25	0.29	0.33	0.38	0.44	0.50	0.57	0.63	0.68
4	0.13	0.17	0.21	0.26	0.32	0.40	0.48	0.56	0.62	0.68
5	0.08	0.11	0.15	0.20	0.27	0.36	0.46	0.55	0.62	0.68
6	0.05	0.07	0.11	0.16	0.23	0.33	0.44	0.54	0.62	0.68
7	0.03	0.05	0.07	0.12	0.20	0.31	0.43	0.54	0.62	0.68
8	0.02	0.03	0.05	0.10	0.17	0.29	0.42	0.54	0.62	0.68
9	0.01	0.02	0.04	0.07	0.15	0.27	0.42	0.54	0.62	0.68

Table 1. Theoretical PDC_{sym}/PDC_{asym} ratio values. Cell numbers at the corresponding generation time were derived based on the EGE or the linear cell growth equation varying dividing fraction (F_d), where the default F_d is 0.5 in equations [2] and [3]. F_d: dividing cell fraction in actual cell culture, GT: generation time
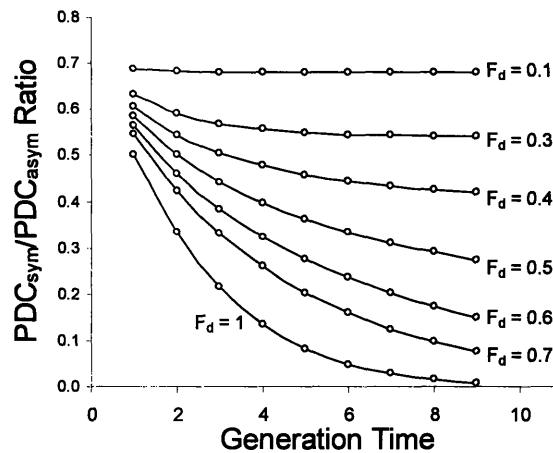

Figure 4. Theoretical PDC_{sym}/PDC_{asym} ratio curves. Cell numbers at the corresponding generation time were derived based on the EGE or the linear cell growth equation varying dividing fraction, F_d. After calculating the PDC_{sym}/PDC_{asym} ratio from ideal cell growth values, PDC_{sym}/PDC_{asym} ratio values were plotted as a function of generation time.

24

Results:

Cell division kinetics of zinc-responsive p53-inducible mouse embryonic fibroblasts

It was important to define exact cell culture conditions for reproducible microarray studies for replicate analyses with low variance. The well defined model cell systems whose self-renewal pattern was controllable in culture were used for studying cell kinetics parameters (Rambhatla et al., 2001). Ind-8 mouse embryonic fibroblasts (MEFs) stably transfected with conditional p53 expression vectors divided with asymmetric self-renewal in the presence of the inducer, $ZnCl_2$, and empty vector control-MEFs Con-3 cells symmetrically self-renew in all conditions. From the EGE, $N_t = N_0e^{kt}$, the average doubling time (DT) of Con-3 and Ind-8 cells without induction condition were 18 hours (SD $\pm$ 0.9, n = 5) and 21 hours (SD $\pm$ 0.7, n=5), respectively. In the presence of 65 μM $ZnCl_2$, the DT of Con-3 cells was 17 hours (SD $\pm$ 1.4, n=5) as almost the same as the condition without $ZnCl_2$. However, the average DT of Ind-8 cells in the presence of 65 μM $ZnCl_2$, was 40 hours (SD $\pm$ 11.0, n=5). Fig. 5a shows examples of each cell growth curve for the zinc-responsive p53-inducible and non-responsive p53 null MEFs. Time-lapse cell division pedigree studies showed that these MEFs that Ind-8 cells divided according to asymmetric self-renewal in the presence of zinc and that the cell cycle interval of dividing Ind-8 cells was almost same as their symmetric counterpart Con-3 cells (Rambhatla et al., 2001). These studies confirmed that cell culture that generated both dividing and non-dividing populations in each cell cycle and that the DT calculated from the EGE was not an appropriate parameter for the quality control of cell cultures for

microarray comparisons. In other words, the dividing or non-dividing fractions could not be measured directly from the DT.

Population doubling cycle (PDC) can remove the confounding effect of non-dividing cell populations in the calculation of critical cell kinetic parameters. Thus, PDC was used to estimate the number of cell cycles in cell cultures where both dividing and non-dividing cell populations produced simultaneously. As mentioned earlier, the PDC approach also has limitations for application to cell culture quality control for microarray studies. For accurate PDC calculation, the dividing fraction in evaluated cell cultures should be known (Equations [2] and [3]). However, the exact value of F_d in culture was inconvenient to determine in real time for every microarray analyses. To overcome these limitations, we developed PDC_{asym}/PDC_{sym} ratio curve analyses (Table 1 and Figure 4). The PDC ratio approach was a mathematical strategy to address this technical limitation.

PDC_{sym}/PDC_{asym} ratio curves to estimate dividing fractions in cell culture

The PCD ratio analysis required the GT of each evaluated cell population. The GT can be measured by time lapse microscopy studies or estimated from the average DT of cells dividing with exponential cell kinetics. The F_d of cell cultures can be estimated by simply fitting experimental PDC_{sym}/PDC_{asym} ratio curves to the theoretical PDC_{sym}/PDC_{asym} ratio curves.
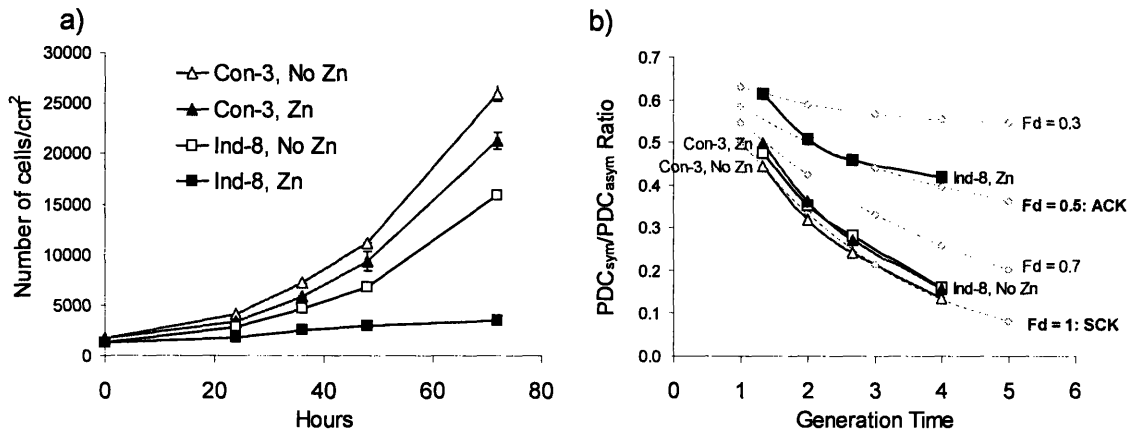
Figure 5. Example of PDC$_{sym}$/PDC$_{asym}$ ratio analysis. a) Cell growth curves for the zinc (Zn) responsive p53-inducible fetal mouse fibroblasts (Ind-8) dividing with asymmetric self-renewal and their isogenic p53-null control lines (Con-3). b) PDC ratio curve analysis. Cell growth curves in a) were transformed (solid lines) and compared to the ideal PDC ratio curves (dashed lines). After calculating PDC$_{sym}$/PDC$_{asym}$ ratio values from the numbers of ideal cell growths, the ratio values are plotted as a function of GT for given cell populations. The ideal PDC$_{sym}$/PDC$_{asym}$ curve for 100% symmetric self-renewal was derived from theoretical cell numbers of a cell population dividing with ideal exponential kinetics, Nt = N$_0$e^{kt}. The PDC$_{sym}$/PDC$_{asym}$ curve for each cell growth for the given F$_d$ value was derived from theoretical cell numbers of a cell population dividing with asymmetric self-renewal

27

As shown in Figure 5a, PDC_{asym} and PDC_{sym} can be calculated from conventional cell growth curve data. After cell growth data were transformed into PDC_{sym}/PDC_{asym} ratio curves, it was possible to estimate F_d of each self-renewing cell population by matching to the theoretical PDC ratio curves as a function of GT (Figure 5b). For example, the experimental PDC_{sym}/PDC_{asym} ratio data for Ind-8 cells inthe presence of Zn correlated well with the theoretical PDC_{sym}/PDC_{asym} ratio curve for the ideal asymmetric self-renewal ($F_d = 0.5$) from the second GT to the fourth GT of culture. The PDC_{sym}/PDC_{asym} ratio value at the time of the first GT is closer to the value where F_d is 0.3. In cell kinetics studies with asymmetrically self-renewing Ind-8 cells, the dividing fraction was variable during the first GT period and lower than those of later GT periods. From the second GT, the dividing fraction of asymmetrically self-renewing cell population maintains its steady state until the fourth or the fifth GT period. Therefore, the dividing fraction change of asymmetrically self-renewing cells with time was estimated by using the PDC_{sym}/PDC_{asym} ratio approximation.

F_d was sensitive to the both p53 expression level and cell density in model cell systems for asymmetric self-renewal

For ideal asymmetric self-renewal in adult stem cells, each cell division generates the same number of dividing and non-dividing daughter cells in culture, i.e., F_d is 0.5. The F_d of Ind-8 cells in culture was reported to be sensitive to subtle changes in zinc concentration (Liu et al., 1998a and 1998b; Rambhatla et al., 2001). In the previous cell division kinetic studies, F_d was not determined throughout culture times (Merok et al.,

2002). PDCs were calculated based on the assumption that F_d was invariant during the period of cell culture. Considering that F_d might vary with time of cell culture (Figure 5 and 6), the simple model with constant F_d for PDC determination might not be appropriate for the quality control for consistent cell kinetics behavior in microarray comparisons. Using the PDC_{sym}/PDC_{asym} ratio curve analyses, F_d was estimated as a function of culture time and other critical effectors of asymmetric cell kinetics like cell density and Zn concentration were evaluated (Rambhatla et al., 2001) (Figure 6 and 7). Figure 6 showed the effect of varying Zn concentration on F_d, which varys the level of p53 expression (Liu et al., 1998b). At the 65 μM of $ZnCl_2$, the F_d of cell culture approximately closed to the ideal asymmetric self-renewal until the fourth GT (Figure 6). By decreasing Zn concentration, the F_d of cell culture was gradually increased.

In addition to p53 expression level, F_d of Zn-responsive p53-inducing Ind-8 cell population was sensitive cell density (Liu et al., 1998a and 1998b; Rambhatla et al., 2001). The higher the initial cell density, the more F_d of a cell population increased under the conditions of Zn (Figure 7). The cell density dependent changes occurred primarily under conditions of p53 induction. However, a similar modest response was observed in the absence of Zn as well. Considering these effects when cells were cultured for microarray comparisons, both p53 induction conditions and input cell densities were carefully maintained to ideal populations with either symmetric self-renewal of aymmetric self-renewal.
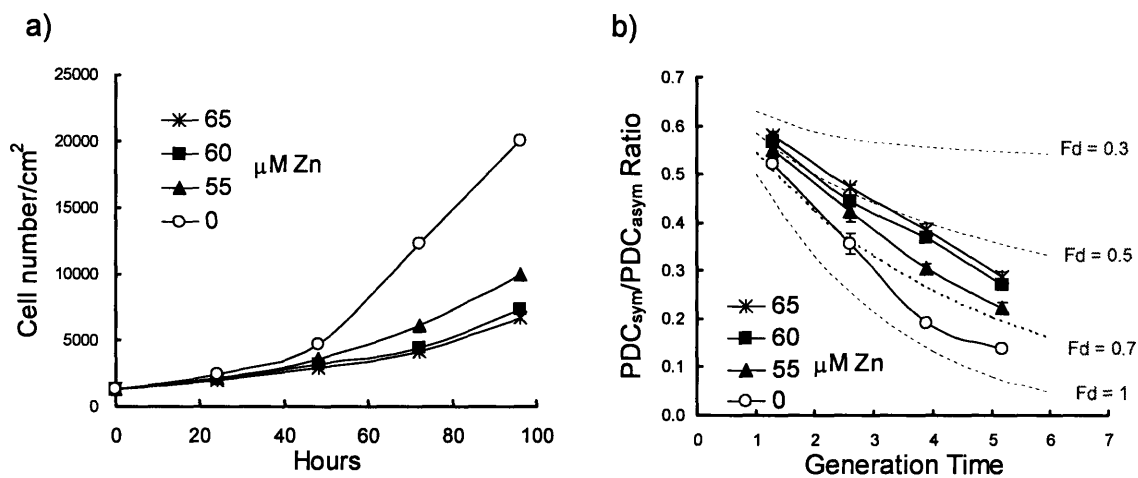
Figure 6. F_d **of Ind-8 cells changed with ZnCl$_2$ concentration.** a) typical cell growth curve analyses of Ind-8 cells in different concentrations of ZnCl$_2$. b) The PDC$_{sym}$/PDC$_{asym}$ ratio curves were transformed from the cell growth curves in the left panel. Input cell density was 1300 cells/cm^2.
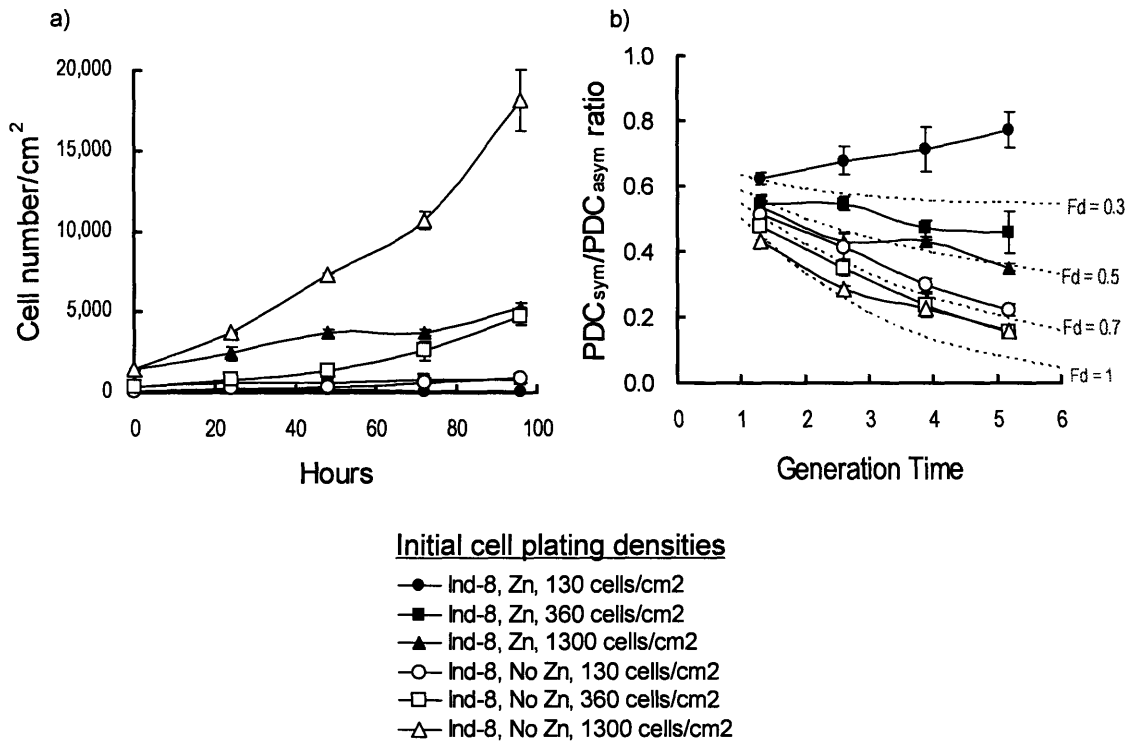
Figure 7. **F_d of Ind-8 cells was dependent on input cell density.** a) typical cell growth curve analyses of Ind-8 cells initiated at different cell densities in the presence or absence of 65 μM of $ZnCl_2$. b) The PDC_{sym}/PDC_{asym} ratio curves were developed with cell growth data shown in a).

31

The mathematical sign of the slope of the PDC_{sym}/PDC_{asym} ratio curve provided additional information regarding the condition of cell populations. A negative or positive slope for the PDC_{sym}/PDC_{asym} ratio curve indicated an increasing or decreasing dividing cell population, respectively. When the cell density was as low as 130 cell/cm^2 and the Zn concentration was as high as 65 μM, the slope of the PDC_{sym}/PDC_{asym} ratio curve was positive, indicative of cell death with increasing culture time (Figure 7b). Therefore, the PDC_{sym}/PDC_{asym} ratio curve analysis can be used not only as a quantitative tool to estimate dividing fraction in culture, but also as a qualitative tool for the viability of cultured cells with time in culture.

Dicussion and Conclusion:

The PDC_{sym}/PDC_{asym} ratio approximation approach provided the method to discern the cell kinetics pattern distribution of a cell population using only the assumption that individual cells adopt either symmetric or asymmetric cell kinetics characterized by the production of one non-dividing daughter each cycle. The PDC_{sym}/PDC_{asym} ratio allowed quantitative comparison of the self-renewal kinetics pattern of cell cultures without experimental determination of F_d. An initial dividing fraction (F_d) value of 1.0 corresponds to ideal exponential cell kinetics, which is the result of complete symmetric self-renewal in all cell divisions. The value of 0.5 for F_d represented ideal asymmetric stem cell kinetics. Intermediate F_d values indicated cell cultures with mixtures of cells in either symmetric state. The positive slope value of the PDC ratio curves indicated cell senescence and/or cell death. F_d value of ASCs *in vivo* may deviate from 0.5 depending

32

on physiological states (Matioli et al., 1970; Loeffler et al., 1997; Punzel et al., 2002). The PDC_{sym}/PDC_{asym} ratio calculated at given time of culture can be used to predict the self-renewal pattern of cell cultures.

Therefore, the PDC ratio provided a method for quality control of the self-renewal pattern of examined cells, i.e. whether asymmetric or symmetric. In addition, the PDC_{sym}/PDC_{asym} ratio might be used to estimate the F_d in primary tissue cell cultures. If the average GT of an isolated cell population is measured, the percentage of cells dividing with asymmetric self-renewal (i.e. adult stem cells) can be estimated by a PDC ratio analysis of simple growth data. Thus, further development and application of this approach may provide novel insights to the self-renewal properties of mammalian tissues.

Chapter 2. In Search of a Self-Renewal Pattern Associated Gene Expression Signature I: Orthogonal-intersection cDNA microarray analyses

Recently, transcription profiles have been reported that attempted to describe common genetic features of stem cells by identifying differentially expressed genes in ESCs and in ASC-enriched preparations (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003). In these studies, ASCs included hematopoietic stem cell (HSC)-enriched fractions, cultured neural stem cells (NSCs), and cultured retinal progenitor cells (RPCs). Some of stem cell specific genes reported in these "stemness" studies have been supported by independent investigators (Mitsui et al., 2003; Etheridge et al., 2004), however, the lack of gene overlapping among the virtually same transcriptional profiling studies has been criticized (Burns and Zon, 2002; Zipori, 2004; Cai et al., 2004). ASC-preparations used in the studies contain a significant fraction of non-stem cell lineage-committed progenitors and differentiating progeny cells that limited their utility for identifying genes whose expression was unique to stem cells, *i.e.,* stemness genes (Ivanova et al., 2003; Cai et al., 2004). Gene expression profiles based on specific expression in both ESCs and ASC-enriched populations will exclude genes whose expression is specific to either of these distinctive stem cell classes. One essential difference is that ESCs propagate in culture by symmetric self-renewal, whereas ASCs are defined by asymmetric self-renewal (Merok and Sherley, 2001; Sherley, 2002).

A novel strategy was used to identify genes whose expression levels are related to ASC function based on targeting their unique asymmetric self-renewal. Because

asymmetric self-renewal is unique to ASCs, some genes whose expression profiles are associated with asymmetric self-renewal may specify adult stemness and also identify ASCs. It was possible to pursue this strategy because of the availability of cultured cell lines whose self-renewal symmetry can be reversibly converted between symmetric and asymmetric self-renewal, conditionally (Sherley, 1991; Sherley et al., 1995a; Liu et al., 1998a and 1998b). Restoration of normal wild-type p53 protein expression induces these lines to undergo asymmetric self-renewal like ASCs (Rambhatla et al., 2001; Merok et al., 2002; Rambhatla et al., 2005). When p53 expression is reduced, the cells switch to symmetric self-renewal, resulting in exponential proliferation. *In vivo*, symmetric self-renewal by ASCs is regulated to increase tissue mass during normal adult maturation and to repair injured tissues (Sherely, 2002). When controls that constrain ASCs to asymmetric self-renewal are disrupted (*e.g.,* by p53 mutations), the risk of proliferative disorders like cancer may increase (Sherley, 2002; Cairns, 2002).

In this research, a novel strategy was applied to identify genes whose expression levels are related to ASC function by targeting asymmetric self-renewal with the model cell systems for self-renewal symmetry. Some genes whose transcription profiles are associated with asymmetric self-renewal may specify adult stemness and also help to identify ASCs. In microarray analyses, cultured cell lines whose self-renewal symmetry is clearly defined were used in order to avoid the complications of the heterogeneous cell populations that characterize ASC-enriched preparations.

Materials and Methods:

Cell Culture

For microarray analyses, the model cell lines for the regulation of self-renewal symmetry were used. In addition to the zinc-responsive p53-inducible Ind-8 cell line and its respective vector-transfected control Con-3 cell line, an isogenic p53-inducible cell line constitutively expressing the inosine-monophosphate dehydrogenase (IMPDH) gene regardless of cellular p53 level was derived from mouse embryo fibroblasts and called tI-3. Its respective control tC-2 cells were an isogenic p53-inducible line derived with an IMPDH-deleted vector. The cell lines whose self-renewal symmetry is temperature dependent were derived from non-tumorigenic, immortalized cells that originated from mouse mammary epithelium cells, 1h-3 and respective non-p53 inducible 1g-1 cells. The Ind-8, tC-2, and 1h-3 cells for asymmetric self-renewal and Con-3, tI-3, and 1g-1 cells for respective control symmetric self-renewal were maintained as described (Sherley, 1991; Sherley et al., 1995a; Liu et al., 1998a and 1998b). For analyses of cell division kinetics of Ind-8 and Con-3 cells, cells were grown over a 3-day period to about 50 % confluency, trypsinized, and replated in zinc-free medium (DMEM, 10% dialyzed fetal bovine serum (DFBS), 5 μg/ml puromycin) at a cell:plating area:medium volume ratio of 1×10^5:75 cm^2:20 ml. This ratio was held constant for all experiments, unless specified otherwise. Sixteen to 24 hours later, the culture medium was replaced with the same volume of zinc-free medium or medium containing the specified concentration of $ZnCl_2$. This time was designated as 0 hour in the analyses.

For the 1h-3 and 1g-1 cell growth, cells were grown to 50 % confluency at 37 °C and replenished with fresh growth medium (DMEM, 10% DFBS, 1 mg/ml G418 sulfate). Sixteen hours later, the cells were replated at either 37 °C or 32.5 °C. This time was 0 hour in the analyses. In case of the growth analyses of tI-3 and tC-2 cells, cells were grown and replated as the same way for those of Ind-8 and Con-3 cells except for the medium (DMEM, 10% DFBS, 1 mg/ml G418 sulfate plus 5 μg/ml puromycin). Cells were harvested by trypsinization and counted with a Model ZM Coulter Counter to determine cell kinetics parameters. The GT for each asymmetric cell growth model was determined from time lapse microscopic observation or from the average doubling time of the symmetrically dividing counterpart (Sherley et al., 1995a; Rambhatla et al., 2001).

cDNA Microarray and Normalization

Cells for total RNA extraction were harvested at the time of the second PDC (36 hours for Ind-8 and Con-3, 48 hours for tC-2, tI-3, 1h-3, and 1g-1), and were selected for microarray experiments based on the actual PDC ratio curves in parallel cell cultures. The same PDC ratio values were maintained to obtain similar fractions of cycling stem-like cells and non-cycling transient-like cells for all compared asymmetric models. Total RNA was extracted with the Trizol reagent (Invitrogen, Carlsbad, CA) and impurities were removed with the Qiagen RNeasy kit (Qiagen, Valencia, CA). Fifty to seventy μg of total RNA was used for cDNA syntheses. *Arabidopsis thaliana* mRNAs (Stratagene, La Jolla, CA) were introduced as internal probe standards into reverse transcription reactions to normalize data between different arrays. Cy3- or Cy5-fluorescently labeled cDNAs were

hybridized onto the NIA 15K mouse cDNA prefabricated arrays (Tanaka et al., 2000; Kargul et al., 2001), supplied by the MIT-BioMicro Center, using the procedure provided by the MIT-BioMicro Center (http:// web. mit. edu/biomicro/). Hybridized microarrays were scanned with the *arrayWoRx*^{eTM} Biochip Reader (Applied Precision LLC, Northwest Issaquah, WA). The fluorescence intensity of each spot was analyzed from the scanned tiff images by using the DigitalGenome^{TM} software (MolecularWare, Inc. Cambridge, MA). The Cy3 and Cy5 fluorescence intensities were normalized by calculating the normalization factor from total intensity normalization (Quackenbush, 2001 and 2002). Each comparison for each asymmetric cell kinetics model was performed as duplicate independent experiments. For each experiment, we performed two chip hybridizations with reciprocally labeled Cy3 or Cy5 target cDNAs to each biological sample. Thus, this entire first analysis incorporated data from 16 independent chip analyses. With the same statistical procedure above, we also generated a "noise" gene set from pair-wise comparisons of normalized expressed values from experimentally equivalent samples.

Statistical Determination

A gene was selected for data analyses only if the mean value of foreground pixels of the spot was greater than the sum of the mean and two standard deviations of the background pixels. For individual spots, the expression levels of Cy5 and Cy3 channels were estimated by subtracting mean backgrounds from mean foregrounds. The average ratio and the average difference of gene expression levels between the asymmetrically dividing cells and the symmetrically proliferating cells were calculated. Histograms for

the data based on the ratio and difference of gene expressions were obtained to analyze the distribution of the microarray data by using the StatView™ software (SAS, Cary, NC). Empirically, based on noise data developed from pair-wise comparisons of experimentally equivalent samples, the upper bound for confidence is $p < 0.0004$. The noise data were also used as a comparison control for Fisher's exact test analyses.

Results and Discussion:

Model cell systems in orthogonal-intersection gene microarrays

Previously, we derived cell lines with conditional self-renewal symmetry from non-tumorigenic, immortalized cells that originated from mouse mammary epithelium (MME) cells and mouse embryo fibroblasts (MEFs) (Sherley, 1991; Sherley et al., 1995a; Liu et al., 1998a and 1998b). The self-renewal symmetry of these cells can be reversibly switched between symmetric and asymmetric by varying either culture temperature or Zn concentration, as a consequence of controlling p53 expression with respectively responsive promoters. These diverse properties allowed a 2 X 4 orthogonal-intersection microarray analysis to identify genes whose expression consistently showed the same pattern of change between asymmetric versus symmetric self-renewal (Figure. 8).

2 X 4
orthogonal
intersection

1 Asymmetric
Self-Renewal

2 Symmetric
Self-Renewal

● : Dividing

▨ : Non-dividing

Model 1. Fibroblasts	p53-induced, Zn	vs.	p53 null, Zn
Model 2. Fibroblasts	p53 ON, Zn	vs.	p53 OFF
Model 3. Epithelial cells	p53 Normal, 32.5 °C	vs.	p53 Low, 32.5 °C
Model 4. Fibroblasts	p53 ON	vs.	p53 ON, IMPDH ↑
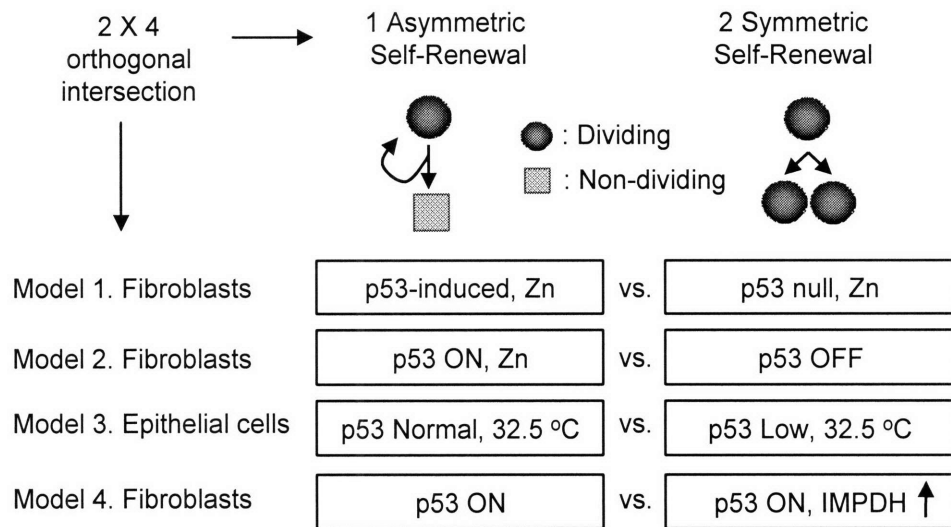
Figure 8. **A 2 X 4 orthogonal-intersection microarray analysis.** The first comparison was made between asymmetrically dividing Ind-8 cells and symmetrically dividing Con-3 cells both in the presence of zinc chloride (Model 1). In the model 2, gene expression between asymmetrically self-renewing Ind-8 cells by p53 induction and symmetrically self-renewing Ind-8 cells, in the absence of p53 induction, was compared. As the third model for asymmetric self-renewal, p53-inducible mouse mammary epithelial cell lines were tested. When p53 expression is elevated 1.5-fold due to low temperature (32.5 °C)-dependent gene promoter, these lines (1h-3) initiate asymmetric self-renewal ($F_d \sim 0.5$). Gene expression of 1h-3 cells after induction of p53 was compared to that of isogenic control lines (1g-1) that lack a p53 transgene and divide with symmetric self-renewal at 32.5 °C. Finally, gene expression was evaluated with p53-inducible-inosine-5'-monophosphate dehydrogenase (IMPDH)-transfected cell lines (tI-3) and isogenic p53-inducible vector-transfected cell lines (tC-2) (Model 4). Asymmetric self-renewal associated genes were selected by finding genes consistently changed in all 4 self-renewal pattern models. For statistical testing, we also selected p53-responsive genes showing differential expression in the model 1-3 but not in the model 4.

40

Four different pair-wise comparisons were developed in which a transcription profile of asymmetric self-renewal was compared to that of symmetric self-renewal. The first 3 comparisons were based on a difference in p53 expression, but each had a different biological context. For asymmetric versus symmetric, respectively, these comparisons were: 1) Zn-responsive p53-inducible MEFs (Ind-8 cells) versus p53-null control MEFs (Con-3 cells), both cultured in Zn-supplemented medium (Liu et al., 1998b; Rambhatla et al., 2001); 2) Zn-responsive p53-inducible MEFs in Zn-supplemented medium versus in Zn-free medium; 3) and low temperature (32.5°C)-dependent p53-inducible MME cells (1h-3) versus control MME cells (i.e., no p53 transgene; 1g-1 cells), both cultured at 32.5°C (Sherley et al., 1995a).

The fourth comparison had a special purpose. It provided a comparison of asymmetric versus symmetric self-renewal that was not based on a difference in p53 expression. Two previously described derivatives of the Zn-responsive p53-inducible MEFs were used to make this comparison. One line (tI-3 cells) is stably transfected with a constitutively expressed inosine monophophate dehydrogenase (IMPDH) gene (Liu et al., 1998a and 1998b). IMPDH is the rate-limiting enzyme for guanine nucleotide biosynthesis. Its down-regulation by p53 is required for asymmetric self-renewal (Liu et al., 1998a). Therefore, even in Zn-supplemented medium, which induces normal p53 expression, cells derived with a stably expressed IMPDH transgene continue to undergo symmetric self-renewal. This abrogation of p53 effects on cell division frequency occurs even though other p53-dependent responses remain intact (Liu et al., 1998a and 1998b). Under the same conditions, control vector-only transfectants (tC-2 cells) continue to

exhibit asymmetric self-renewal. Thus, this fourth comparison could be used to exclude genes whose change in expression was primarily due to changes p53 expression and not specifically transitions in self-renewal symmetry.

As the quality control of cell cultures for replicate microarray analyses, we used PDC_{sym}/PDC_{asym} ratio curve analyses discussed in Chapter 1. In microarrays, we harvested the cells in culture at the two population doubling (GT = 2) for total RNA isolation. The ideal PDC_{sym}/PDC_{asym} ratio at the second GT in culture is 0.5 for asymmetric self-renewal (F_d = 0.5) and 0.33 for symmetric self-renewal (F_d = 1). Microarray experiments were performed by using cell cultures whose experimental PDC ratio curves in parallel cell cultures correlated with ideal ratio curves for 0.5 of F_d (Figure 9, Table 2). For each of the 4 experimental comparisons, two independent total RNA samples were isolated. Thus, data from 4 microarrays, representing 2 independent experiments, was available for each of the four contexts for asymmetric versus symmetric self-renewal.
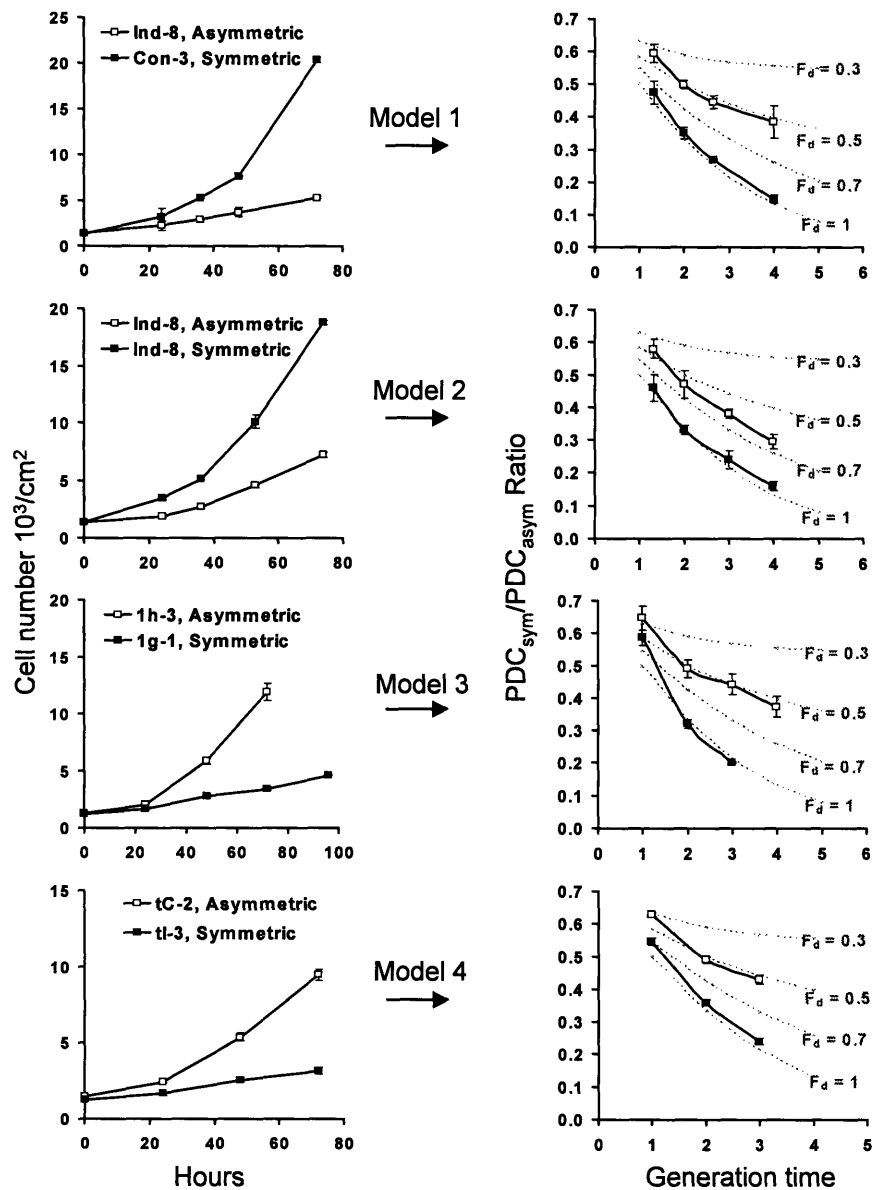
Figure 9. Cell growth curves and PDC$_{sym}$/PDC$_{asym}$ ratio analyses of the model cell systems used in 2 X 4 orthogonal-intersection microarray analysis. Cells for the isolation of total RNA samples for the microarrys were harvested at the time of 2 GT. The ideal PDC$_{sym}$/PDC$_{asym}$ ratio at 2 GT in culture is 0.5 for asymmetric self-renewal (F$_d$ = 0.5) and 0.33 for symmetric self-renewal (F$_d$ = 1).

	Asymmetric growth	Symmetric growth
	PDC_{sym}/PDC_{asym} ratio	PDC_{sym}/PDC_{asym} ratio
Model 1	0.50 ± 0.01	0.35 ± 0.02
Model 2	0.50 ± 0.01	0.33 ± 0.01
Model 3	0.49 ± 0.01	0.34 ± 0.01
Model 4	0.50 ± 0.02	0.35 ± 0.00

Table 2. Experimental values of PDC_{sym}/PDC_{asym} ratio values for cDNA microarray for each cell population (n = 2)

a) Data (Asym/Sym) **b)** Noise (Cy3/Cy5)
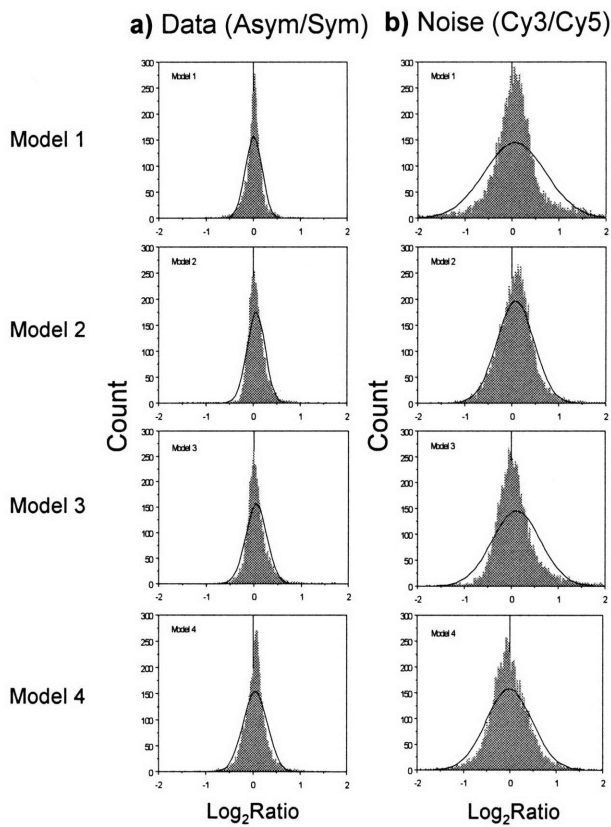
Figure 10. Data distributions for the expression-ratio values of the four self-renewal pattern comparisons.
(a) Data distributions: ratio values of asymmetric to symmetric expression were calculated. (b) Noise distributions: noise data were developed from pair-wise comparisons between biologically equivalent samples. The ratio values were calculated with Cy3 and Cy5 channel expression values of each gene from the identical RNA samples.

44

a)

b)
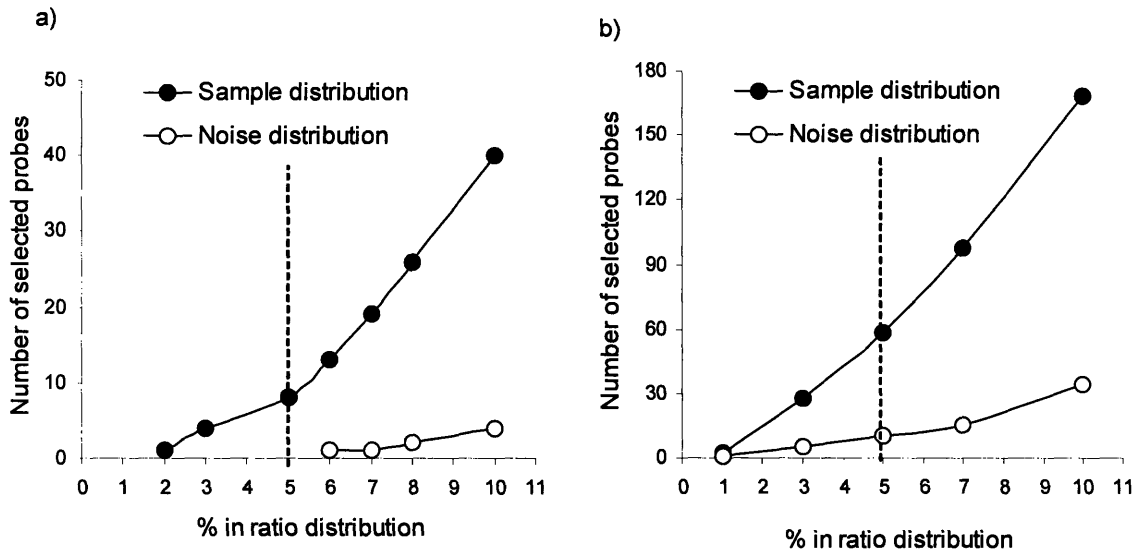
Figure 11. Number of selected genes at the each % tail area of sample distribution and noise distributions. (a) expression-ratio data selection. (b) expression-difference data selection

Selection of self-renewal pattern-associated genes

Prefabricated cDNA microarrays constructed with the National Institute for Aging (NIA) mouse 15K mouse clone set were used for the analysis (Kargul et al., 2001). For each of the 4 comparisons, frequency distributions were constructed for the number of genes with a given magnitude of change in terms of normalized gene expression data which are based on either expression-ratio data or expression-difference data (Figure 10a). To develop these distributions, gene-specific determinations from each of the 4 microarrays in each comparison type were averaged. These distributions were normal in character and had less than a 2-fold range for either up-regulation or down-regulation of gene expression between the two self-renewal symmetry states (Figure 10a). "Noise" data distributions were also developed from pair-wise comparisons of experimentally equivalent RNA samples that were labeled with different fluorescent dyes. One different feature of noise data sets was that they were based entirely on inter-chip comparison, whereas the study data were based on only intra-chip comparisons (Figure 10b).

We identified genes that were consistently present in the right (up-regulated) or left (down-regulated) tails of the distributions for all 4 comparisons. For each comparison, we selected genes located in the tails of these distributions that corresponded to 5% of the total distribution area. As results of this orthogonal-intersection analysis, we found 26 cDNA probes representing 21 genes that were consistently up-regulated in all 4 contexts of asymmetric self-renewal and 36 cDNA probes for 31 genes that were consistently down-regulated, based on comparisons to symmetric self-renewal. Twenty one genes were named as asymmetric self-renewal associated (ASRA) genes and 31 genes as

symmetric self-renewal associated (SSRA) genes. Collectively, we referred to these 52 genes as self-renewal pattern associated (SRPA) genes (Table 3). In the noise data set, no gene was identified in the expression-ratio distributions and 9 probes were found in the expression-difference analyses (Figure 11). No genes from noise analyses overlapped with the SRSA genes. Using a Poisson approximation based on setting the average value from the noise data set, the upper bound for confidence of this analysis is $p < 0.0004$.

A p53-related gene set, unrelated to self-renewal pattern regulation was identified from the same microarray data (Table 4). This p53-related gene set showed differential expression in the three comparisons whose self-renewal patterns were regulated in response to p53 gene induction but not in the fourth comparison with IMPDH-expressing tI3 cells (Figure 8). A similar noise gene set was selected for the p53-related genes.

Table 3 summarizes cellular functions of self-renewal symmetry associated (SRSA) genes. Some SRSA genes, like cyclin G and survivin, were already known for their association with p53 expression, although their exact cellular functions are still unclear (Mirza et al., 2002 and 2003). In contrast, a well-known p53-responsive gene, *mdm2*, was discovered in the 5% tails for the first three contexts, but not in comparison 4 (Table 4). This analysis does not exclude the possibility that these p53-responsive genes are involved in asymmetric self-renewal functions. However, it indicates that the observed p53-dependent changes in their expression do not require asymmetric self-renewal; whereas this possibility holds for SRSA genes at this stage of their characterization.

	ASRA genes (21)	SSRA genes (31)
Cell cycle/ growth control	*Ccng* (Mm.2103, H3077H03), *Btg1* (Mm.16596, H3114D11)	*Birc5* (Mm.8552, H3013G10), *Pcna* (Mm.7141, H3021F12), *Ncl* (Mm.117581, H3014B10), *Anp32b* (Mm.154985, H3026E06), *Cdk6* (Mm.31672, H3135E11)
GTPase/ GTP-GDP exchange	*Gnb2-rs1* (Mm.5305, H3083H02)	*Ranbp1* (Mm.3752, H3006B07)
Transcription factor/ signal transduction	*Cstd* (Mm.231395, H3156A10)	*Mxd4* (Mm.38795, H3131B07), *Shcbp1* (Mm.37801, H3057D08), *Calm1* (Mm.34246, H3019D01), *Usp28* (Mm.21630, H3146F08)
Plasma membrane receptor protein	*Kcnn4* (Mm.9911, H3054H04), *Cd63* (Mm.4426, H3075G01), *Tde2* (Mm.29344, H3014H10)	
Energy/ metabolism	*Gsta4* (Mm.2662, H3119G08), *Ftl1* (Mm.7500, H3023H10), *Npc1* (Mm.3484, H3055G06), *Fn1* (Mm.193099, H3116A10)	*Cycs* (Mm.35389, H3032C01), *Asns* (Mm.2942, H3154F02)
Chromosome organization	*Hmga2* (Mm.3953, H3040E06)	*Hmgb1* (Mm.16421, H3027D07), *Hmgb2* (Mm.1693, H3113B06), *Hmgb3* (Mm.340, H3030H10), *H2afz* (Mm.916, H3009D10)
Matrix/ structure	*Mglap* (Mm.243085, H3134C05)	*Tra1* (Mm.4526, H3026G05)
Translation/ RNA binding/ RNA processing	*Pabpc1* (Mm.2642, H3012E01)	*Eif2s2* (Mm.29859, H3039F04), *Eif4g2* (Mm.525, H3113E10), *Hnrpa1* (Mm.27927, H3034F01), *Hnrpa2b1* (Mm.155896, H3015F09), *Hnrpa3* (Mm.213045, H3057B02), *Sfrs2* (Mm.6787, H3025G02), *Rpa3* (Mm.29073, H3090H03), *Rpl4* (Mm.359301, H3011G02)
Unknown or ESTs	*Suv39h2* (Mm.14012, H3064A10), AU021895 (Mm.11780, H3061E07), AU018794 (Mm.178455, H3075B03), D1Ertd564e (Mm.155160, H3065B06), 2410003B16Rik (Mm.29933, H3104E03), Ttc3 (Mm.3679, H3121G11), *Atp1a7* (Mm.44101, H3140A11)	*Sumo2* (Mm.29073, H3144H03), BB631262 (Mm.102992, H3130E11), BG079363 (Mm.216374, H3038H11), *Btbd9* (Mm.218280, H3142F10), *Lsm5* (Mm.25642, H3108H10), *Tomm20* (Mm.6932, H3152B10)

Table 3. Self-renewal pattern associated genes and their functions or proposed functions. Genes are listed with a gene symbol (UniGene, NIA 15K mouse cDNA clone number).

	p53-induced non-SRSA genes (18)	p53-repressed non-SRSA genes (44)
Cell cycle/ growth control	*S100a6* (Mm.100144, H3087H09)	*Tmpo* (Mm.124, H3028A08), *Chek1* (Mm.16753, H3042B08), *Top2a* (Mm.4237, H3139A05), *Cdc2a* (Mm.4761, H3024E04)
Transcription factor/ signal transduction	Mdm2 (Mm.22670, H3051D04), Emd (Mm.18892, H3125F11)	*Ttk* (Mm.1904, H3137A12), *Ube2d2* (Mm.168915, H3018G09), *Calm2* (Mm.18041, H3006H06), *Ptp4a2* (Mm.193688, H3008A11), *Irak1* (Mm.38241, H3005F06)
Heat shock/ stress Chaperone activity		*Hspd1* (Mm.1777, H3023B11), *Cct5* (Mm.1813, H3023F08), *Vbp1* (Mm.8294, H3020D11)
Plasma membrane receptor protein	*M6pr* (Mm.1358, H3092C05), *B2m* (Mm.163, H3014C06)	*Slc23a1* (Mm.22702, H3097H06), *Ssfa1* (Mm.29504, H3088H11), *Lbr* (Mm.4538, H3044H11)
Energy/ metabolism	*Grp58* (Mm.709, H3097B08)	*Oxct1* (Mm.13445, H3058A06), *Decr2* (Mm.35760, H3069H02), Eno1 (Mm.90587, H3027G08
Matrix/structure	*Nup153* (Mm.260781, H3010H02), *Dstn* (Mm.28919, H3122E03)	*Neu3* (Mm.103703, H3100E03), *Kpna2* (Mm.12508, H3043A09)
Translation/ RNA binding/ RNA processing	*Eef1a1* (Mm.196614, H3040C11), *Eef1b2* (Mm.2718, H3013D05), *Rpl7* (Mm.102941, H3125G11), *Rpl37a* (Mm.21529, H3131H12)	*Tardbp* (Mm.22453, H3080F07), *Snapc1* (Mm.25937, H3023C01), *Hnrpa3* (Mm.316306, H3014F08)
Unknown or ESTs	AU018794 (Mm.188455, H3075B03), BG080759 (Mm.22482, H3056D05), 6720475J19Rik (Mm.100378, H3099E06), 1700066C05Rik (Mm.182434, H3147D11), 1700022L09Rik (Mm.25410, H3131C01), 6720475J19Rik (Mm.259177, H3125A02)	*Atp5h* (Mm.371641, H3019E12), *Ddx3x* (Mm.18459, H3018F11), *Wrd43* (Mm.257762, H3027G08), 2610208E05Rik (Mm.1182, H3106G09), 2810433K01Rik (Mm.25546, H3074B08), NM_019828 (Mm.172202, H3121D11), 15 more ESTs.

Table 4. p53-related non SRSA genes and their functions or proposed functions. Genes are listed with a gene symbol (UniGene, NIA 15K mouse cDNA clone number).

ASRA genes specify ASC-enriched cell populations

Due to their specific relationship to asymmetric self-renewal, it was postulated that SRPA genes may specify the genetic features of ASCs. This postulation was tested by analyzing overlaps of SRSA genes with various stem cell-specific genes reported in three stemness studies (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003). Both the UniGene identification (ID) and Entrez Gene ID were used to evaluate overlap among the datasets. For statistical evaluation, Fisher's exact test was performed using the noise gene set as the comparison basis (Table 5).

For the first comparison, the SRPA genes were evaluated for overlapping genes with other stemness genes that were commonly expressed in both ASC-enriched populations and ESCs (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003). There was no significant overlap between ASRA or SSRA genes and stemness genes (Table 5a). When the ASC-specific and ESC-specific gene sets used for previous stemness gene signature determinations were evaluated independently, a significant overlap was observed only between the ASRA genes and ASC-specific genes (80.0%; p = 0.048). As a further evaluation of ASRA gene specificity for adult stemness, the same overlap was analyzed with p53-responsive non-SRPA genes (Table 5b). No significant overlap was observed for any comparison between p53-responsive genes and stem cell specific genes defined by the other groups. Therefore, the ASRA gene set may specify the molecular features of ASC-enriched populations.

a)

Self-renewal pattern associated (SRPA) genes	Stemness Data Set		
	ASCs	ESCs	Stemness genes
ASRA genes	**80.0 % (p = 0.048)**	45.0 % (p = 0.53)	5.0 % (p = 0.34)
SSRA genes	67.7 % (p = 0.098)	61.3 % (p = 0.27)	25.8 % (p = 0.16)

b)

p53-related, non-SRSA genes	Stemness Data Set		
	ASCs	ESCs	Stemness genes
p53-induced genes	44.4 % (p = 0.99)	38.9 % (p = 0.33)	5.6 % (p = 0.60)
p53-repressed genes	40.9 % (p = 0.99)	36.4 % (p = 0.61)	18.2 % (p = 0.47)

Table 5. Overlap between the SRPA, p53-related genes, and reported stem cell specific gene sets. Overlap was determined based on the UniGene ID number for each SRSA gene. SRPA genes whose UniGene ID was unavailable were not included in analysis. The adult stem cell comparison set was created by HSC-specific, NSC-specific, and RPC-specific genes described previously (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003).

Discussion and Conclusion:

A significant barrier to study ASC function is the lack of reliable molecular markers to identify and isolate ASCs (Sherley, 2002; Zipori, 2004; Eckfeldt et al., 2005). Currently, there are no available tools to identify and isolate ASCs as pure populations from mammalian tissues (Eckfeldt et al., 2005). These limitations account for the confounded data in recent transcriptional profiling studies, which have been criticized for the lack of any significant overlap between gene sets selected by virtually the same methods (Burns and Zon, 2002; Cai et al., 2004). Instead of using currently available stem cell-enriched populations, model cell lines were used that mimic self-renewal pattern regulation to find genes associated with two different self-renewal patterns, asymmetric vs. symmetri. Fifty two SRPA genes from the 2 X 4 orthogonal-intersection microarray analysis suggested that they, especially up-regulated genes during asymmetric self-renewal, specify molecular features of ASCs. Over 80 % of up-regulated ASRA genes exist in specific genes expressed in ASC-enriched populations defined by other laboratories. Interestingly, cyclin G which had been well known as a p53-responsive gene was reported as a highly expressed gene in ASC-enriched preparations from mouse hair follicles and human bone marrow by the independent transcription profiling studies (Morris et al., 2004; Wagner et al., 2004).

There are previous reports of gene microarray studies to identify genes associated with p53 expression. Among other differences, these studies are based on states of uniform cell division arrest (Zhao et al., 2000; Mirza et al., 2003; Robinson et al., 2003). In contrast, our analyses is based on cells cycling in two different cell kinetics states, one

52

producing only cycling cells (*i.e.,* symmetric self-renewal) and the other producing both cycling and arrested cells (*i.e.,* asymmetric self-renewal). Asymmetric self-renewal by our model cell lines is characterized by stem-like cells that cycle to produce a viable non-cycling sister at each division (Rambhatla et al., 2001). In addition, when cycling asymmetrically, these lines exhibit non-random chromosome segregation, another ASC-unique property (Merok et al., 2002; Potten et al., 2003; Rambhatla et al., 2005; Karpowicz et al., 2005). Based on these features and the fourth comparison, our analysis was predicted to reveal adult stemness genes that had not been identified in previous p53-associated gene profiles.

In statistical testing, we found that only the ASRA gene set had significantly overlap with the specific genes in previously described ASC-enriched populations. No significance was observed with previously describe stemness genes and ESC-specific genes. Previously defined stemness genes were based on commonly up-regulated genes in embryonic, neuronal, and hematopoietic stem cell populations compared to unfractionated adult tissues or stem cell depleted cell preparations (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003). Based on human studies, the HSC-enriched populations used to define stemness genes are likely to consist of both asymmetrically and symmetrically self-renewing stem cells (Punzel et al., 2002 and 2003), whereas ESCs predominantly proliferate by symmetric self-renewal in culture (Wiles and Johansson, 1999; Schuldiner et al., 2000). The self-renewal pattern of neurospheres used as a source of NSCs in stemness studies is more complex (Carpenter et al., 1999; Gottlieb 2002). Overall, the cell division kinetics of neurosphere cultures appears exponential, indicative

of symmetric self-renewal. About, only 0.1 ~ 3 % of the cells in each neurosphere cluster regenerate secondary neurospheres (Reynolds and Rietze, 2005). This rare population is considered to represent the NSC fraction. The self-renewal pattern of this rare population is unclear, but whatever it might be, it would contribute only a small component to the overall symmetric cell kinetics after neurosphere cultures. Therefore, there were clear differences in the self-renewal patterns among stem cell enriched-populations in the stemness studies. These differences in self-renewal patterns among comparing stem cell-enriched populations may have resulted in confounding of the analyses to find common stem cell genes. These studies confirmed the precept that self-renewal pattern differs significantly between ASCs and ESCs.

In vivo, symmetric self-renewal by ASCs is proposed to be regulated to increase tissue mass during normal adult maturation and to repair injured tissues (Sherley, 2002). When control mechanisms that constrain ASCs to asymmetric self-renewal are disrupted (*e.g.,* by p53 mutations), the risk of cancer may increase (Cairns, 1975 and 2002; Sherley, 2002). Thus, defects in SRPA genes may be associated with human cancer. At least 8 SRPA genes were previously associated with human cancers. For instance, the two SSRA genes, *tra1* and *birc4*, are highly expressed in human tumor tissues compared to normal tissues (Claudio et al., 2002; Altieri, 2003); and the two ASRA genes, *hmga2* and *btg1*, were reported to have chromosomal loci that were rearranged in human leukemia (Rodier et al., 2001; Hisaoka et al., 2002). These observations implicate the disruption of self-renewal pattern regulation as a potentially important event in human.

Recently, the concept of cancer stem cells has been supported by serial transplantation experiments with rare malignant cells isolated from human blood, breast, and brain cancers (Reya et al., 2001; Pardal et al., 2003; Radtke and Clevers, 2005). In NOD/SCID mice repopulation experiments, the self-renewal and hierarchical progeny of cancer cells have been similar to those of normal ASCs. However, the underlying characteristics of the self-renewal and hierarchical progeny between them may be distinct. These cancer-associated SRSA genes may provide key clues to understanding molecular differences in the self-renewal pattern between ASCs and cancer stem cells.

Many ASRA genes are implicated in regulation of chromatin function and chromosome segregation. For instance, the expression of *hmgb1*, *hmgb2*, *hmgb3*, *birc5*, and *h2afz*, were consistently down-regulated in asymmetrically self-renewing cells. The high mobility group B (HMGB) protein family is known to modulate chromatin architecture (Muller et al., 2001; Agresti and Bianchi, 2003). The *birc5* gene product, survivin, is a component of the chromosome passenger protein complex (Altieri, 2003; Carmena and Earnshaw, 2003). *H2afz* may function in chromatin remodeling during embryogenesis, especially when totipotent ESCs begin to differentiate into specialized cells (Rangasamy et al., 2003). These genes may also function in non-random chromosome segregation mechanisms that are unique to cells undergoing asymmetric self-renewal by affecting chromosome structure or regulating mitotic machinery during mitosis.

In conclusion, asymmetric self-renewal by ASCs is a well-established precept in stem cell biology that is often overlooked (Cairns, 2002; Sherley, 2002). Though there may be molecular features common to embryonic and ASCs, they will also have stem-cell

specific properties that distinguish them because of their different developmental roles. This result showed that genes whose expression changes specify asymmetric self-renewal were highly represented among genes up-regulated in natural ASC-enriched cell populations. A similar association was not found with ESC-specific genes. It was proposed that asymmetric self-renewal was such a property, and these findings support this hypothesis. It may also be the case that some SRPA genes specify other unique attributes of ASC that are associated with asymmetric self-renewal but are not critical for execution of changes in self-renewal pattern. In particular, such genes may be present but undetectable in the p53-responsive gene set.

Chapter 3. In Search of a Self-Renewal Pattern Associated Gene Expression Signature II: Extending the SRPA Gene Signature by Whole Genome Transcriptional Profiling

Rationale:

The cDNA microarray analysis for studying self-renewal symmetry associated (SRSA) genes supported the precept that asymmetric self-renewal is an important properties to consider in research to elucidate adult stem cell function and character (Chapter 2). However, the orthogonal-intersection gene array had limitations for finding ideal molecular markers for ASCs. The genetic coverage of the cDNA microarrays used in the study was less than 13,000 genes. So, many significant genes may have been overlooked. Also, there were few plasma membrane protein-encoding genes in the SRPA gene set. One desired property for a candidate molecular markers for ASCs is plasma membrane integration, because this feature has great practical potential for identifying and purifying ASCs by generating antibodies for it (Weissman et al., 2001). There was also the possibility that many genes with low copy number were excluded because of the less than optimal detection sensitivity of the cDNA microarray analyses. For instance, differential expressions of p53 and IMPDH were not detected in the cDNA microarrays after normalization with the total 16 arrays in the study, although the mRNAs from these two genes showed differential expression in Northern blotting analyses (Liu et al, 1998a and 1998b).

As a unique marker, it woulds be ideal to have a gene set expressed in just one condition of self-renewal symmetry state and not expressed in another state, i.e. an

exclusively expressed gene set. However, there was no gene that was exclusively expressed in either state of self-renewal pattern in the cDNA microarray study. In the expression-ratio based analysis, information on the absolute expression level of genes might be lost (Quackenbush, 2002). Therefore, the expression-difference analysis was aimed to search for exclusively expressed genes in the cDNA microarray analysis (Chapter 2). However, in the expression-difference analysis there was evidence for a high false-positive rate. Many of the SRPA genes identified by this basis were highly expressed genes that were more likely to exhibit large difference due to statistical variation between large numbers. The fluorescent signals of many genes that have low transcript copy number may not be detected by the co-hybridization of two different fluorescent labeled comparing sample pair. In the two color microarray system with Cy3 and Cy5 dyes, the intensities of Cy3-labeled samples are far stronger than those of Cy5-labeled ones. During the normalization procedure between Cy3 and Cy5 fluorescent intensities, any small difference can be canceled out by noise signal or global normalization in cDNA microarray data. In general, the normalization procedure itself removes noise in the data, but it also may decrease the sensitivity for signal detection (Quackenbush, 2001 and 2002; Yeung and Ruzzo, 2001).

To address many of the shortcomings of the cDNA array analyses, whole genome transcripts was evaluated with Affymetrix GeneChip® mouse whole genome 430 2.0 arrays that contain over 42,000 genes. In contrast to the cDNA microarrays, the Affymetrix oligonucleotide microarrays measured gene expressions from only one biological sample per each array. In the Affymetrix array, a single probe set is composed

of 22 oligonucelotides, 11 perfect match (PM) and 11 mismatch (MM) cell pairs. Softwares to analyze Affymetrix GeneChip® image files calculate expression values of probe sets and also decide whether each gene is 'present' or 'absent' in a sample by using the PM/MM statistical model (Li and Wong, 2001). By the Affymetrix whole genome array analysis, the genetic coverage was increased by at least 4-fold compared to the cDNA analyses. Importantly, exclusively expressed gene sets associated with each self-renewal pattern were identified.

Methods

Affymetrix oligonucleotide microarray

Whole genome expression profiles of p53-induced Ind-8 cells, p53 null Con-3 cells, and p53 induced/IMPDH transfected tI-3 cells were compared by analyzing Affymetrix mouse whole genome GeneChip® 430 2.0 arrays (Affymetrix, Inc. Santa Clara, CA). Three independent cell cultures for Ind-8 and Con-3 cells and two for tI-3 cells were collected by the same procedure described in the cDNA microarray experiements (Chapter 2). Total RNA samples were extracted using the Trizol reagent (Invitrogen, Carlsbad, CA) and followed by a clean-up step with the Qiagen RNeasy kit (Qiagen, Valencia, CA). Total RNA quality was tested with the Agilent 2100 BioAnalyzer (Agilent Technologies, Palo Alto, CA). Five μg of total RNA was used for cDNA synthesis and then cDNA was used to make biotinylated cRNA. cRNA was fragmented and hybridized onto the Affymetrix mouse whole genome GeneChip® 430 2.0 array. The arrays were washed and

qaantified with a fluorescence array scanner. After scanning the chips, quantification and statistics were performed using model-based expression and the perfect match (PM) minus mismatch (MM) method in the G-COS® software (version 1.0) and/or the dChip software version 2005. Data across 8 arrays were originally normalized by setting target intensity at 500 in the G-COS.

Microarray data analysis and statistical analysis

An exclusively expressed gene set for asymmetric self-renewal was selected from probe sets that were called as 'present' in all Ind-8 cells and called as 'absent' in both Con-3 and tI-3 cells by the PM/MM statistical model in the G-COS® software. The converse analysis was performed to identify a set of genes expressed exclusively during symmetric self-renewal. Expression-ratio values were calculated with the average expression values for each self-renewal symmetry model (Chapter 2). Histrograms for the data distributions were generated with the expression-ratio values of probe sets called as 'present' in all GeneChip® arrays. Histograms were obtained to analyze the expression-ratio distribution by using the StatView™ software (SAS, Cary, NC). Like the cDNA microarray analysis, genes that consistently existed in the 5% tail areas of the ratio distributions of the two comparisons were selected as differentially expressed genes. In order to compare the data between cDNA microarray and Affymetrix whole genome array results, principal component analysis was performed to test the clustering gene expression data to test the biological samples. Parameters for principal component analysis were calculated by the CLUSTER 3.0 software (Eisen *et al.*, 1998).

Results:

Selection of self-renewal symmetry-associated genes from whole genome arrays

Two or three replicate microarrays were evaluated for gene expression of three cultured cell models for self-renewal pattern, p53-induced Ind-8 cells with asymmetric self-renewal, p53-null Con-3 cells with symmetric self-renewal, and p53-induced cells with symmetric self-renewal as a consequence of forced IMPDH gene expression, tI-3 cells. First, exclusively expressed gene sets were found by using 'present' and 'absent' calls decided by the PM/MM algorithm in Affymetrix GCOS® software. One hundred and thrity two genes that were exclusively expressed in p53-induced Ind-8 cells undergoing asymmetric self-renewal and simultaneously not present in symmetric self-renewal states. Sixty six genes were discovered exclusively in the cells dividing with symmetric self-renewal (Table 6). Second, two expression-ratio data histograms were generated with 17133 probe sets called as 'present' in all microarrays in the analyses. One comparison was made between p53-induced Ind-8 cells and p53-null Con-3 cells. The other comparison was performed between p53-induced Ind-8 cells and p53-induced/IMPDH expressed tI-3 cells in order to exclude p53-reponsive genes but not associated with self-renewal pattern regulation. Differentially expressed gene sets were selected from the genes which existed in the 5% tail areas of both distributions. As results, 178 genes were identified as up-regulated in asymmetrically self-renewing Ind-8 cells, and 167 genes were identified in symmetric self-renewal (Table 6).

Self-renewal symmetry	Gene set	Number of probe sets	Number of genes
Asymmetric	Exclusive	141	132
	Differential	203	178
Symmetric	Exclusive	74	66
	Differential	186	167
	Total	604	543

Table 6. Self-renewal symmetry associated genes identified in Affymetrix mouse whole genome arrays.

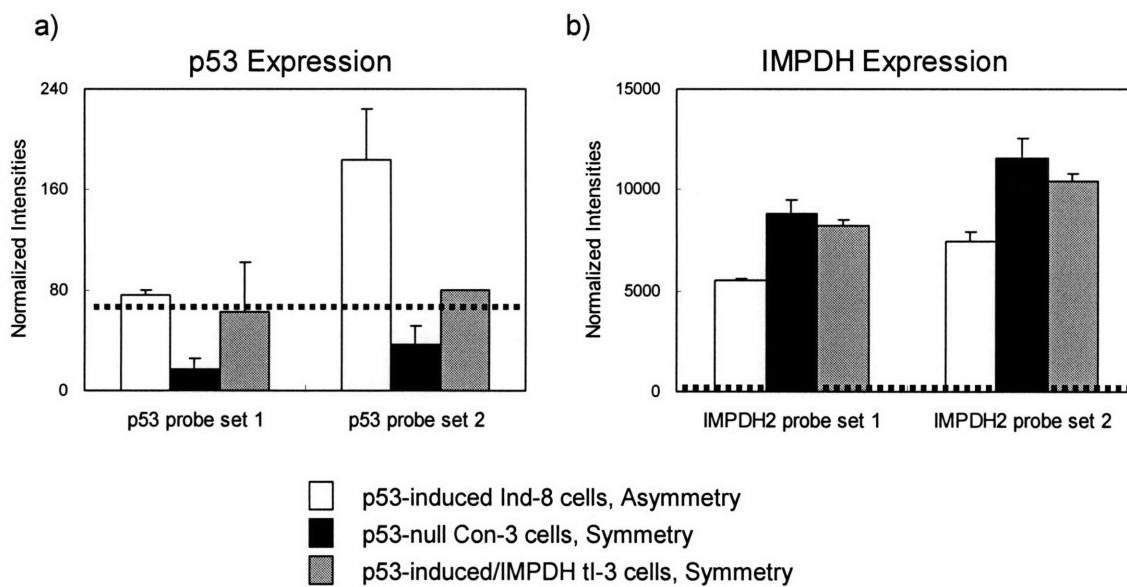**Figure 12. Quality and sensitivity evaluation of the Affymetrix whole genome array analysis for p53 and IMPDH genes**. Expression levels for each gene were normalized with all 8 oligonucleotide arrays by the G-COS® software. The line (⋯) indicated the average background noise level in arrays. a) Normalized intensities of p53 gene probe sets. b) Normalized intensities of p53-repressed gene *Impdh* probe sets.

Quality and detection sensitivity of the Affymetrix whole genome microarrays

For evaluating the quality and sensitivity of the Affymetrix whole genome array data, the expression of p53 and IMPDH genes on the arrays was evaluated. These two genes are major molecular components that regulate the self-renewal pattern in model cell systems (Figure 12). In the earlier cDNA microarray analyses, it was not possible to detect differential expression of p53 and IMPDH genes. In the Affymetrix array analysis, p53 gene expression with the corresponding probe sets was detected from the p53-induced Ind-8 samples. In the p53-induced states, a significant level of p53 gene expression was detected from only one p53 probe set, $p < 0.05$ (Figure 12a). On average, the expression level for p53 probe sets in the p53-induced samples was just above the background noise level (Figure 12a). This could be due to a low copy number for p53 transcripts or to lower efficiency of Affymetrix probe sets for the p53 gene. Expression of p53 gene in p53-induced Ind-8 cells dividing with asymmetric self-renewal was higher than those of both symmetric self-renewal conditions (Figure 12a).

The expression levels of the probe sets for the IMPDH gene were significantly greater than background noise level ($p < 0.001$ for all microarrays). Compared to the expression values in p53-null Con-3 cells and p53-induced/IMPDH transfectant tI-3 cells, the expression level was consistently lower in asymmetric self-renewing Ind-8 cells consistent with published Northern and Western blotting (Figure 12b) (Liu et al., 1998a and 1998b). The IMPDH cDNA sequence used for constitutive expression was originated from the Chinese hamster and the sequence homology between two species is 93.1 % (Liu et al., 1998a). Therefore, the expression level of IMPDH gene in tI-3 cells may be

affected by the interactions of the Chinese hamster cDNA with the mouse IMPDH gene sequences in the probe sets in the Affymetrix arrays.

Comparison of orthogonal-intersection cDNA microarray analyses and Affymetrix mouse whole genome array analyses

Only 46 out of 52 SRPA genes identified from the cDNA microarray analysis could be compared with data from the Affymetrix GeneChip mouse whole genome 430 2.0 array, because the UniGene ID number for 6 SRPA genes in the cDNA microarrays was not available. When the 543 SRPA genes from the Affymetrix whole genome array analysis were compared with the 46 cDNA SRPA genes, only 5 genes, two ASRA genes, *ccng1* and *gsta4*, three SSRA genes, *hmgb3*, *mxd4*, and *h2afz*, overlapped between the two microarray analyses. The low overlap could have been due to differences in the genetic coverage between the cDNA and the Affymetrix oligonucleotide arrays. The Affymetrix whole genome GeneChip® array covers 4 times more genes than the cDNA microarray. Therefore, application of the same 5 % tail-based selection for the Affymetrix analysis might have resulted in exclusion of genes scored in from smaller cDNA array distributions. For the selection of SRPA genes, the expression-difference based analysis was not used in the Affymetrix analysis because the expression-difference based analysis in the cDNA microarray analysis was designed to detect exclusively expressed genes that would not appear in expression ratio-based analyses. In the Affymetrix analysis, the PM/MM statistical model was used to detect exclusively expressed genes. When the expression-difference based analysis was performed with the Affymetrix array data, 32 SRPR genes

from the cDNA microarray analyses overlapped with the genes in the Affymetrix whole genome array analysis.

In order to test whether the 46 SRSA genes from the cDNA microarray analysis can classify the self-renewal pattern of cell populations, they were evaluated as a signature for clustering the data sets form the Affymetrix analysis by principal component analysis (PCA) (Figure 13b). The 543 SRSA genes from the Affymetrix whole genome array analysis clearly separated asymmetric self-renewing populations from symmetrically self-renewing populations (Figure 13a). When PCA was performed with the expression values of the 46 SRSA genes extracted from the Affymetrix array data, there were three separate clusters, each of which represented one of three self-renewal symmetry cell culture models (Figure 13b). Although the cluster of p53-induced/IMPDH expressing tI-3 cells was separated from both cell culture conditions, the center of the symmetrically dividing tI-3 cell cluster is closer to the center of the symmetrically dividing Con-3 cell group than that of the asymmetrically dividing cell populations. With the 32 overlapping SRPA genes identified between cDNA microarray and Affymetrix microarray analyses based on both expression-ratio and expression-difference, the clustering pattern of PCA was almost same as that of 46 cDNA SRPA genes (Figure 13c).
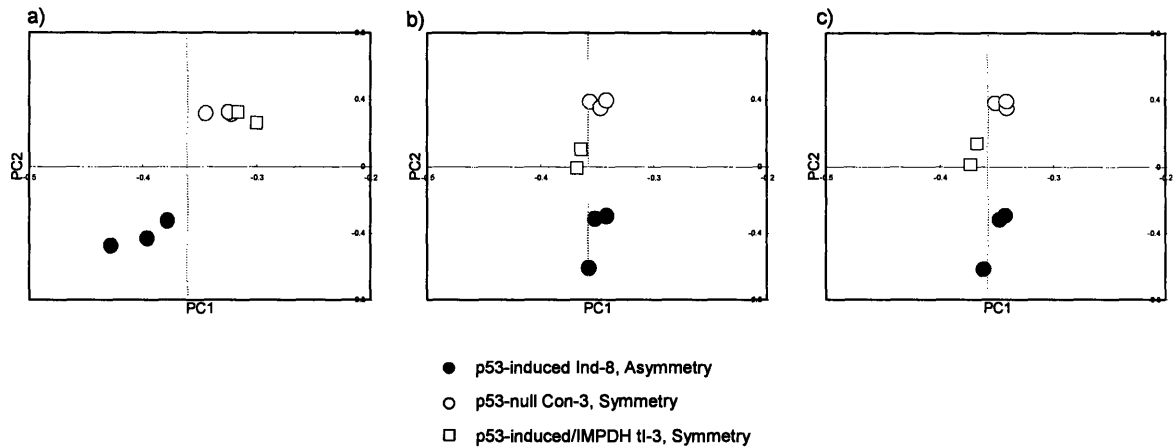
Figure 13. Principal component analysis for the comparison of two SRPA gene sets from the cDNA and Affymetrix microarray analyses. a) Principal component analysis (PCA) was performed with 543 SRSA genes identified from the Affymetrix whole genome analysis with the expression values form the Affymetrix data (CLUSTER 3.0). b) The expression values of 46 SRPA genes in the cDNA microarray analysis were extracted from the Affymetrix data and parameters for PCA were calculated. c) PCA using 32 SRPA genes overlapped between the cDNA and the oligonucleotide microarray analyses.

66

ASRA genes specify ASC-enriched cell populations

Next, the overlap was evaluated between the 543 SRSA genes and the genes reported in the stemness literature by using array comparison databases in the Netaffy center at the Affymetrix Inc. The probe set ID numbers and the Entrez Gene ID numbers among the data sets were used for comparison. The degree of overlap with the 543 SRSA genes and reported stem cell specific genes was not as great as observed in the cDNA microarray analyses (Table 7). However, the genetic coverage of Affymetrix GeneChip® Mouse U74v2 arrays used in the stemness literature is approximately a third of the Affymetrix whole genome 430 2.0 array (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003). One stemness study used the full set of the Affymetrix U74v2 series (Ivanova et al., 2002). However, only 309 probe sets out of 543 SRSA genes (604 probe sets) exist in the 430 2.0 version array. In the 309 probe sets, 211 probe sets for ASRA genes and 98 probe sets for SSRA genes could be compared with the stemness literature.

The orthogonal-intersection cDNA microarray results showed that ASRA genes were more highly associated with previously described ASC-enriched cell populations. This supported the hypothesis that ASRA genes specify adult stem cell identity and potentially function (Chapter 2). In the Affymetrix whole genome array analysis, the unique relationship between ASRA genes and specific genes of ASC-enriched populations was also observed (Table 7). Thirty three percent of ASRA genes were reported as differentially expressed in the ASC-enriched populations whereas 18 % SSRA genes in those populations.

	Overlaps	ASRA (211 genes)	SSRA (98 genes)	% ASRA/SSRA ratio
Adult stem cell specific genes	HSC only	24 (11 %)	4 (4 %)	**2.8**
	NSC only	33 (16 %)	12 (12 %)	**1.3**
	Both HSC and NSC	12 (6 %)	2 (2 %)	**2.8**
	Sum	**69 (33 %)**	**18 (18 %)**	**1.8**
Embryonic stem cell specific genes	ESC only	29 (14 %)	31 (32 %)	**0.4**
Stemness genes (665 genes)	ESC + NSC + HSC	11 (5 %)	8 (8 %)	**0.6**
Total overlaps		150 (71 %)	91 (93 %)	**0.8**

Table 7. Overlap between the self-renewal pattern associated (SRPA) genes and the stemness gene data sets. Only 211 probe sets in the 310 ASRA genes and 98 probe sets in the 233 SSRA genes could be compared because of the difference in the Affymetrix arrays used in the different studies. Percent overlap was calculated based on the number of total comparable genes in the respective SRPA gene sets.

In particular, HSC-associated genes showed a 3-fold greater overlap with ASRA genes compared to SSRA genes. However, the NSC-associated genes overlapped with both ASRA and SSRA genes at almost same level (12~16 %). In terms of cell kinetics, NSCs divide with symmetric self-renewal, which might result in the same level of overlap with the SSRA genes as with the ASRA genes. Interestingly, the genes associated with both HSCs and NSCs in the stemness studies showed the same overlap pattern of HSC-associated genes (Table 7). The adult stemness in NSCs may be appeared as the genes associated with both NSCs and HSCs. This overlap pattern suggested that the genes specific for both HSCs and NSCs may be associated with a common ASC feature, especially in asymmetric self-renewal. In contrast to the ASC-enriched populations, ESC-specific genes showed a 2-fold greater overlap with the SSRA genes. The relationship between the ASRA genes and the ASC-specific genes and between the SSRA genes and the ESC-specific genes were highly significant (Fisher's exact p-value, $p < 0.001$). Therefore, two independent microarray studies supported the hypothesis that asymmetric self-renewal is an essential feature of ASCs. Moreover, since the HSCs in the compared reported studies were not cultured, this observation is the first molecular evidence that mammalian cells in vivo undergo asymmetric self-renewal.

Discussion and Conclusion

A new SRPA gene set was identified, which consists of 310 ASRA genes and 233 SSRA genes, by an Affymetrix whole genome oligonucleotide array analysis. In particular, 132 genes were identified that showed exclusive expression in asymmetrically

69

self-renewing cells. These genes are the focus of further studies as potential adult ASC-specific markers. In the Affymetrix whole genome microarray analysis, sensitivity, consistency, statistical power, and gene coverage was improved compared to the cDNA microarray analyses. As an example, the expression of p53 and IMPDH, the two essential factors to regulate self-renewal pattern, were not detected in the cDNA microarray studies. The Affymetrix GeneChip® arrays successfully detected p53 and IMPDH gene expression. In addition, the SRPA genes in the whole genome array analysis included more plasma membrane encoding genes in the gene ontology (GO) cellular component classification. Such genes have potential to serve as exclusive markers for ASCs (Table 8, Supplementary Table 1).

The ASRA gene that encodes a plasma membrane protein may be a good candidate as an adult stem cell marker because of its great potential for application. In the 132 exclusive ASRA genes, at least 17 genes are predicted to encode plasma membrane proteins or extracellular proteins attached to the plasma membrane (Table 8). Interestingly, *notch1* was reported as playing roles in the asymmetric fate decision during embryonic development, in the self-renewal of HSCs or their progenitor cells, and in the regulation of epithelial stem cell niches (Artavanis-Tsakonas et al., 1999; Stier et al., 2002; Dallas et al., 2005; Radtke and Clevers, 2005). Endoglin, one of the exclusive ASRA genes, has been reported for expression in the long-term HSCs, although its expression is observed in other tissue (Chen et al., 2002). The functions of several exclusive ASRA genes like *plxdc2*, *thsd1*, and *9130213B05Rik* remain unknown to date.

Probe set ID, Gene title (Gene symbol)	Gene Ontology Biological Process Description
1417271_a_at, endoglin (Eng)	angiogenesis, cell adhesion, regulation of transforming growth factor beta receptor signaling pathway etc
1418051_at, Eph receptor B6 (Ephb6)	transmembrane receptor protein tyrosine kinase signaling, etc.
1418205_at, thrombospondin, type I, domain 1 (Thsd1)	Unknown
1418393_a_at, integrin alpha 7 (Itga7)	cell adhesion, cell-matrix adhesion, etc
1418633_at, Notch gene homolog 1 (Drosophila) (Notch1)	cell fate specification, branching morphogenesis, epithelial to mesenchymal transition, regulation of transcription, etc.
1418912_at, plexin domain containing 2 (Plxdc2)	Development
1419238_at, ATP-binding cassette, sub-family A (ABC1), member 7 (Abca7)	Unknown
1419671_a_at, interleukin 17 receptor C (Il17rc)	Unknown
1419758_at, ATP-binding cassette, sub-family B (MDR/TAP), member 1A (Abcb1a)	transport, drug transport
1422812_at, chemokine (C-X-C motif) receptor 6 (Cxcr6)	G-protein coupled receptor protein signaling pathway
1424478_at, Bardet-Biedl syndrome 2 homolog (human) (Bbs2)	cell adhesion, cell-matrix adhesion, protein localization, adult behavior, negative regulation of body size, etc
1426082_a_at, solute carrier family 16, member 4 (Slc16a4)	Transport
1426443_at, rhomboid, veinlet-like 7 (Drosophila) (Rhbdl7)	Unknown
1428891_at, RIKEN cDNA 9130213B05 gene	Unknown
1432826_a_at, CD80 antigen (Cd80)	defense response
1447602_x_at, sulfatase 2 (Sulf2)	sulfur metabolism, heparan sulfate proteoglycan metabolism
1449957_at, protein tyrosine phosphatase, receptor type, V (Ptprv)	protein amino acid dephosphorylation

Table 8. Example of the exclusively expressed ASRA genes encoding or having potential to encode a plasma membrane protein. At least 17 genes in the 132 exclusive ASRA genes may encode plasma membrane proteins or extracellular proteins anchored to plasma membrane by the gene ontology (GO) cellular component descriptions.

Many genes in the exclusive SRPA gene sets have low expression levels and some are close to the background level. The gene selection method that is simply based on the 'present' or 'absent' calls by the PM/MM statistical modeling has rarely been used in microarray studies. This analysis was used to capture genes that cannot be analyzed with expression-ratio based methods because 0 denominator values are not allowed. So far, the feasibility of the Affymetrix PM/MM statistical model to decide 'present' or 'absent' for each probe set has been rarely evaluated by independent experiments. The expression of these genes must be independently verified before testing as a potential molecular marker for ASCs. Analyses with this goal are described in Chapter 4.

The initial gene overlap analysis between the 46 SRPA gene signature from cDNA microarray analysis and the 543 SRPA gene signature from the Affymetrix whole genome array analysis was not significant. However, by evaluating signatures derived by both expression-ratio and expression-difference comparisons, the SRPA genes from the Affymetrix whole genome array matched with 70 % of the SRPA genes from the cDNA microarray analysis. In the PCA, the asymmetrically self-renewing samples were clustered and this asymmetric cluster was separated from the symmetric clusters by the 46 SRSA genes from the cDNA microarray analysis. In addition, when both SRPA sets were compared with the stemness data (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003), the two independent ASRA gene sets are specifically related to the genes expressed in ASC-enriched populations (Table 5 and 7).

In particular, the 543 SRPA genes could capture the bi-symmetic self-renewal pattern of NSC-enriched populations that showed overall symmetric cell kinetics in

neurosphere cultures (Reynolds and Rietze, 2005). However, only 0.1 ~ 3% cells that may self-renew asymmetrically in the neurosphere can regenerate secondary neurospheres, suggesting that most of cells in the neurosphere may be symmetrically dividing transient amplifying cells. The NSC-specific genes didn't have any preferential overlap with either ASRA or SSRA genes wherease the genes overlapped NSCs and HSCs reported by the stemness gene studies showed the 3-fold ASRA gene preferential overlap as the same level observed with HSC-specific genes (Table 7). Therefore, the common molecular feature of NSCs and HSCs captured by the overlap pattern of the 543 SRPA genes might be asymmetric self-renewal, an essential molecular feature of ASCs.

In conclusion, the microarray studies for finding the genes associated with self-renewal symmetry states strongly supported the proposal that asymmetric self-renewal is an essential feature of adult stemness *in vitro*. Future studies focused on molecular and cellular aspects of ASC asymmetric self-renewal may deliver the long sought phenotypic markers that uniquely identify ASCs.

Exclusive expression: 132 asymmetric self-renewal associated genes

Blood coagulation: 1423285_at (Coch)

Cell adhesion: 1421279_at (Lamc2), 1418393_a_at (Itga7), 1418450_at (Islr), 1422977_at (Gp1bb), 1424595_at (F11r), 1422253_at (Col10a1), 1424478_at (Bbs2)

Cell cycle regulation: 1427739_a_at (Trp53), 1441880_x_at (MGC30332), 1421679_a_at (Cdkn1a), 1458849_at (Asah2)

Cell death/apoptosis: 1418626_a_at (Clu), 1452863_at (1700003F12Rik)

Chromatin modulation: 1425868_at (N/A)

Cytokine/growth factor: 1429546_at (Ecgf1)

Cytoskeleton/cellular motility: 1447364_x_at (Myo1b), 1427485_at (Lmod1), 1456346_at (Dnm1), 1434264_at (Ank2)

Development: 1418633_at (Notch1), 1454830_at (Fbn2), 1434917_at (Cobl)

Electron transport: 1456638_at (Wdr59), 1448426_at (Sardh), 1422534_at (Cyp51), 1419435_at (Aox1)

Immune response: 1418638_at (H2-DMb1), 1432826_a_at (Cd80)

Metabolism: 1447602_x_at (Sulf2), 1421987_at (Papss2), 1435708_at (Gls), 1416645_a_at (Afp), 1418519_at (Aadat), 1444723_at (6530418L21Rik)

Mitosis: 1448001_x_at (Cdca3)

Proteolysis: 1443689_at (Usp32), 1418632_at (Ube2h), 1425743_at (Trim7), 1417682_a_at (Prss2), 1453836_a_at (Mgll), 1451019_at (Ctsf), 1417009_at (C1r), 1425170_a_at (Adam15)

Receptor/signal transduction: 1436978_at (Wnt9a), 1419671_a_at (Il17rc), 1418051_at (Ephb6), 1417271_a_at (Eng), 1454931_at (Cri2)

Signal transduction: 1440844_at (Tob1), 1437012_x_at (Rapgef3), 1449957_at (Ptprv), 1452127_a_at (Ptpn13), 1417801_a_at (Ppfibp2), 1426926_at (Plcg2), 1428025_s_at (Pitpnc1), 1422994_at (Pip5k3), 1419389_at (Pde10a), 1449630_s_at (Mark1), 1433553_at (Garnl3), 1422812_at (Cxcr6), 1437613_s_at (AW456874)

Transcriptional regulation: 1450929_at (Zfp57), 1425816_at (Zfp287), 1455165_at (Rora), 1422647_at (Ring1), 1418912_at (Plxdc2), 1427142_s_at (Jarid1b), 1422024_at (Fli1), 1437820_at (Fkhl18), 1423845_at (Csdc2), 1443227_at (Bzw2), 1452322_a_at (Brwd1), 1453959_at (1700065O13Rik)

Transporter, membrane: 1434793_at (Wdr78), 1417392_a_at (Slc7a7), 1426568_at (Slc2a9), 1434015_at (Slc2a6), 1421924_at (Slc2a3), 1426082_a_at (Slc16a4), 1451224_at (Scamp5), 1426225_at (Rbp4), 1425391_a_at (Osbpl5), 1444418_at (Itpr5), 1455396_at (Atp8b1), 1438431_at (Abcd2), 1419758_at (Abcb1a), 1419238_at (Abca7)

Unknown: 1426563_at (Zfp553), 1417310_at (Tob2), 1451479_a_at (Tmem53), 1418205_at (Thsd1), 1425217_a_at (Synj2), 1441906_x_at (Syap1), 1421668_x_at (Speer3), 1426443_at (Rhbdl7), 1420578_at (Optc), 1434877_at (Nptx1), 1427015_at (LOC380969), 1453317_a_at (Khdrbs3), 1452875_at (Hddc3), 1423091_a_at (Gpm6b), 1434645_at (C530008M17Rik), 1439194_at (C030048H21Rik), 1456287_at (BB236558), 1457459_at (AU014973), 1448034_at (AI842396), 1437366_at (AI608492), 1434762_at (A730041O15Rik), 1433358_at (A230102O21Rik), 1428891_at (9130213B05Rik), 1430097_at (8430436C05Rik), 1435744_at (6720401G13Rik), 1434277_a_at (6430570G24), 1429899_at (5730414N17Rik), 1430766_at (5033403F01Rik), 1432112_at (4930589L23Rik), 1437641_at (4930535B03Rik), 1457415_a_at (4930513N10Rik), 1451653_a_at (4930430E16Rik), 1451287_s_at (2810003C17Rik), 1434327_at (2610020H08Rik), 1429098_s_at (1700029B21Rik), 1428705_at (1700007K13Rik), 1437451_at (N/A), 1439011_at (N/A), 1443686_at (N/A), 1443687_x_at (N/A), 1446155_at (N/A), 1455970_at (N/A), 1432438_at (4930597L12Rik), 1458894_at (N/A)

Supplementary table 1. Functional classification of self-renewal symmetry associated genes identified in the Affymetrix mouse whole genome analysis. Functions of genes were classified based on the gene ontology (GO) biological process description.

Differential expression: 178 asymmetric self-renewal associated genes

Blood coagulation: 1417732_at (Anxa8)

Cell adhesion: 1438672_at (Parvb), 1459749_s_at (Fat4), 1454015_a_at (Cdh13), 1450757_at (Cdh11), 1435345_at (2600006K01Rik)

Cell cycle regulation: 1419015_at (Wisp2), 1438470_at (Socs2), 1437132_x_at (Nedd9), 1423605_a_at (Mdm2), 1423100_at (Fos), 1450241_a_at (Evi2a), 1420827_a_at (Ccng1), 1434745_at (Ccnd2)

Cell death/apoptosis: 1451596_a_at (Sphk1)

Chromatin modulation: 1435866_s_at (Hist3h2a), 1418072_at (Hist1h2bc), 1416101_a_at (Hist1h1c)

Cytokine/growth factor: 1419123_a_at (Pdgfc), 1419662_at (Ogn), 1424050_s_at (Fgfr1)

Cytoskeleton/cellular motility: 1436986_at (Sntb2), 1456028_x_at (Marcks), 1457670_s_at (Lmna), 1451891_a_at (Dysf), 1416499_a_at (Dctn6)

Development: 1427231_at (Robo1), 1451475_at (Plxnd1), 1439794_at (Ntn4), 1434795_at (Disp1), 1416455_a_at (Cryab)

DNA repair/replication: 1449483_at (Polk)

Electron transport: 1435469_at (Qscn6l1), 1433783_at (Ldb3), 1418996_a_at (4930469P12Rik)

Immune response: 1453304_s_at (Ly6e), 1423754_at (Ifitm3), 1445897_s_at (Ifi35), 1416978_at (Fcgrt)

Ion transport: 1440355_at (Kctd12b)

Metabolism: 1424167_a_at (Pmm1), 1424431_at (MGI:1926002), 1435203_at (Man2a2), 1418061_at (Ltbp2), 1448239_at (Hmox1), 1418172_at (Hebp1), 1416368_at (Gsta4), 1418483_a_at (Ggta1), 1423554_at (Ggcx), 1417951_at (Eno3), 1419456_at (Dcxr), 1435281_at (Cpt1c), 1428902_at (Chst11), 1418509_at (Cbr2), 1425704_at (BC022224), 1421149_a_at (Atn1), 1416239_at (Ass1), 1422184_a_at (Ak1), 1428236_at (Acbd5)

Post-transcriptional regulation: 1421063_s_at (Snrpn)

Protein folding: 1418888_a_at (Sepx1), 1422943_a_at (Hspb1), 1416803_at (Fkbp7), 1433481_at (Fkbp14)

Protein synthesis: 1452890_at (Ttll5), 1422603_at (Rnase4), 1449531_at (Leprel2), 1449076_x_at (AL024210)

Protein targeting: 1452330_a_at (Mxra8)

Proteolysis: 1423959_at (Ropn1l), 1422139_at (Plau), 1418269_at (Loxl3), 1422438_at (Ephx1), 1416871_at (Adam8), 1416315_at (Abhd4)

Receptor/signal transduction: 1421296_at (Tnfrsf10b), 1421917_at (Pdgfra), 1418674_at (Osmr), 1420382_at (MGI:2176230), 1418476_at (Crlf1), 1451446_at (Antxr1)

Signal transduction: 1426037_a_at (Rgs16), 1449124_at (Rgl1), 1417333_at (Rasa4), 1424648_at (Rabl4), 1420842_at (Ptprf), 1418181_at (Ptp4a3), 1417323_at (Psrc1), 1429005_at (Mfhas1), 1421365_at (Fst), 1415834_at (Dusp6), 1421425_a_at (Dscr1l1), 1416600_a_at (Dscr1), 1456226_x_at (Ddr1), 1417311_at (Crip2), 1454642_a_at (Commd3), 1455106_a_at (Ckb), 1424842_a_at (Arhgap24)

Transcriptional regulation: 1447432_s_at (Zfp263), 1416935_at (Trpv2), 1419082_at (Serpinb2), 1416638_at (Sall2), 1416505_at (Nr4a1), 1417483_at (Nfkbiz), 1458299_s_at (Nfkbie), 1436188_a_at (Ndrg4), 1434378_a_at (Mxd4), 1434900_at (Mkl1), 1429088_at (MGI:1925139), 1417780_at (Lass4), 1436763_a_at (Klf9), 1419355_at (Klf7), 1417394_at (Klf4), 1448890_at (Klf2), 1418517_at (Irx3), 1436050_x_at (Hes6), 1436329_at (Egr3), 1449152_at (Cdkn2b), 1416250_at (Btg2), 1428454_at (Bcas3), 1433742_at (Ankrd15), 1450637_a_at (Aebp1)

Transporter, membrane: 1439433_a_at (Slc35a2), 1417902_at (Slc19a2), 1422684_a_at (Sec8l1), 1417963_at (Pltp), 1426714_at (D11Ertd18e), 1426534_a_at (Arfgap3)

Unknown: 1429722_at (Zbtb4), 1450418_a_at (Yipf4), 1444012_at (Yipf3), 1447966_a_at (Tmem69), 1422587_at (Tmem45a), 1416261_at (Tmem19), 1419073_at (Tmeff2), 1420124_s_at (Tcta), 1454646_at (Tcp1l2), 1455812_x_at (Slitl2), 1434743_x_at (Rusc1), 1449002_at (Phlda3), 1448958_at (MGI:1929890), 1418454_at (Mfap5), 1455978_a_at (Matn2), 1435628_x_at (LOC277193), 1428326_s_at (Hrsp12), 1424927_at (Glipr1), 1425503_at (Gcnt2), 1435750_at (Gchfr), 1428306_at (Ddit4), 1458148_at (D230007K08Rik), 1419978_s_at (D10Ertd610e), 1426669_at (C530044N13Rik), 1452351_at (C030027K23Rik), 1417153_at (Btbd14a), 1451344_at (BC025600), 1451533_at (BC022687), 1424652_at (BC014699), 1424726_at (BC014685), 1438035_at (AW061290), 1433453_a_at (Abtb2), 1428909_at (A130040M12Rik), 1448251_at (9030425E11Rik), 1442002_at (7030402D04Rik), 1429909_at (4833411O04Rik), 1434240_at (4632434I11Rik), 1434150_a_at (3300001H21Rik), 1429089_s_at (2900026A02Rik), 1424186_at (2610001E17Rik), 1436729_at (2600003E23Rik), 1424239_at (2310066E14Rik), 1417030_at (2310028N02Rik), 1454224_at (2010300F17Rik), 1430596_s_at (1700110N18Rik), 1429027_at (0610007N19Rik), 1434115_at (N/A), 1434585_at (N/A), 1435588_at (N/A), 1436330_x_at (N/A)

Supplementary table 1. Continued.

75

Exclusive expression: 66 symmetric self-renewal associated genes

Blood coagulation: 1451790_a_at (Tfpi)

Cell adhesion: 1423606_at (Postn), 1437932_a_at (Cldn1)

Cell cycle regulation: 1425072_at (Skp2), 1416211_a_at (Ptn), 1415923_at (Ndn)

Cell death/apoptosis: 1418648_at (Egln3)

Cytokine/growth factor: 1420458_at (Tac4)

Cytoskeleton/cellular motility: 1433147_at (Cald1), 1455213_at (4930488E11Rik)

DNA repair/replication: 1438453_at (Rad51c), 1443172_at (Orc1l)

Immune response: 1452544_x_at (H2-D1), 1419597_at (Eda)

Metabolism: 1416798_a_at (Nme4), 1431004_at (Loxl2), 1450387_s_at (Ak3l1), 1428695_at (9130227C08Rik)

Post-transcriptional regulation: 1431505_at (Ppih)

Protein folding: 1416940_at (Ppif)

Protein synthesis: 1430100_at (Mrps15), 1436178_at (Leprel1)

Proteolysis: 1422671_s_at (Naalad2), 1431210_at (Agbl3)

Receptor/signal transduction: 1427165_at (Il13ra1), 1425574_at (Epha3), 1460661_at (Edg3)

Signal transduction: 1455304_at (Unc13c), 1449730_s_at (Fzd3), 1443162_at (N/A)

Transcriptional regulation: 1430651_s_at (Zfp191), 1454007_a_at (Zfp142), 1422836_at (Mbnl3), 1447393_at (Ifrd2), 1453291_at (Hmgb2l1), 1442109_at (Fubp1), 1446820_at (6330583I20Rik)

Transporter, membrane: 1422583_at (Rab3b), 1448326_a_at (Crabp1)

Unknown: 1430304_at (Wdr76), 1423874_at (Wdr33), 1416355_at (Rbmx), 1458919_at (Mkln1), 1456767_at (Lrfn3), 1456284_at (Gm905), 1442487_at (D14Ertd24e), 1423286_at (Cbln1), 1449680_at (C80678), 1440513_at (C80258), 1443978_at (Ankrd41), 1440083_at (A430061O12Rik), 1429846_at (9030411K21Rik), 1429810_at (4921505C17Rik), 1427979_at (4732418C07Rik), 1423266_at (2810405K02Rik), 1430586_at (2700007P21Rik), 1457999_at (2410005O16Rik), 1454031_at (2310029O18Rik), 1439363_at (1200014J11Rik), 1421837_at (N/A), 1430581_at (N/A), 1438245_at (N/A), 1445210_at (N/A), 1456840_at (N/A), 1458025_at (N/A), 1460138_at (N/A)

Supplementary table 1. Continued.

Differential expression: 167 symmetric self-renewal associated genes

Cell adhesion: 1452391_at (Cxadr)

Cell cycle regulation: 1455834_x_at (Tacc3), 1441677_at (Smc4l1), 1429658_a_at (Smc2l1), 1445178_at (Sh3md2), 1437179_at (Rif1), 1425166_at (Rbl1), 1438239_at (Mid1), 1437244_at (Gas2l3), 1442340_x_at (Cyr61), 1422252_a_at (Cdc25c), 1460549_a_at (Cdc23), 1416076_at (Ccnb1-rs1), 1417911_at (Ccna2), 1427192_a_at (Brd8), 1445928_at (Mar6)

Chromatin modulation: 1432236_a_at (Suv39h1), 1416155_at (Hmgb3), 1438091_a_at (H2afz)

Cytokine/growth factor: 1450413_at (Pdgfb), 1458277_at (Ccl25)

Cytoskeleton/cellular motility: 1456623_at (Tpm1), 1420172_at (Myh9), 1440924_at (Mphosph1), 1423520_at (Lmnb1), 1451642_at (Kif1b), 1452314_at (Kif11), 1437491_at (Bicd2)

Development: 1439650_at (Rtn4), 1440037_at (Pbx1), 1419078_at (Nin), 1425907_s_at (Amot)

DNA repair/replication: 1442454_at (Top2a), 1430294_at (Ssbp1), 1448369_at (Pola2), 1436454_x_at (Fen1), 1432097_a_at (Dclre1a), 1421081_a_at (Banf1)

Immune response: 1419132_at (Tlr2), 1456403_at (Pag1), 1422802_at (Defcr3), 1417268_at (Cd14)

Ion transport: 1440168_x_at (Kctd7)

Metabolism: 1436727_x_at (Sptlc1), 1422198_a_at (Shmt1), 1418768_at (Opa1), 1425050_at (Isoc1), 1426620_at (Chst10), 1455991_at (Ccbl2), 1450110_at (Adh7)

Mitosis: 1448113_at (Stmn1), 1423093_at (Incenp), 1442453_at (Fcho2), 1439040_at (Cenpe)

Post-transcriptional modification: 1434966_at (Sfrs8), 1436898_at (Sfpq), 1437322_at (Rbm4), 1455831_at (Fus), 1421882_a_at (Elavl2), 1455523_at (Cstf2)

Protein folding: 1431274_a_at (Hspa9a), 1453233_s_at (Calr3)

Protein synthesis: 1426958_at (Rps9), 1450838_x_at (Rpl37), 1419158_a_at (Harsl), 1438510_a_at (Hars), 1453359_at (Exosc1), 1420024_s_at (Etf1)

Protein synthesis/modification: 1435303_at (Taf4b), 1455341_at (2010003J03Rik)

Protein targeting: 1446234_at (Utx)

Proteolysis: 1439201_at (Usp14), 1417109_at (Tinagl), 1422966_a_at (Tfrc)

Receptor/signal transduction: 1421905_at (Ncoa6ip), 1429871_at (Hmmr)

Signal transduction: 1427467_a_at (Rpgr), 1419553_a_at (Rabggtb), 1452788_at (Ppp2r5e), 1443798_at (Pik3cd), 1448627_s_at (Pbk), 1448154_at (Ndrg2), 1421055_at (Lats2), 1418665_at (Impa2), 1424142_at (Ikbkap), 1453582_at (Chka), 1435807_at (Cdc42)

Transcriptional regulation: 1438714_at (Zfp207), 1435136_at (Whsc1), 1421033_a_at (Tcerg1), 1442764_at (Suv420h1), 1441253_at (Rfx3), 1456898_at (Pura), 1423325_at (Pnn), 1455836_at (Papola), 1422646_at (Mga), 1451739_at (Klf5), 1453596_at (Id2), 1437372_at (Cpsf6), 1434767_at (C79407), 1420628_at (6330411E07Rik)

Transporter, membrane: 1416110_at (Slc35a4), 1438545_at (Slc25a5), 1426341_at (Slc1a3), 1449207_a_at (Kif20a), 1452013_at (Atp10a)

Unknown: 1444717_at (Zwint), 1434433_x_at (Wdr61), 1451087_at (Wdr36), 1437426_at (Wac), 1435452_at (Tmem20), 1419361_at (Ss18), 1419967_at (Seh1l), 1454030_at (Saps3), 1456964_at (Rbm12), 1431287_at (Pcm1), 1430343_at (Nup205), 1432216_s_at (Mpp7), 1441272_at (Matr3), 1448720_at (Lrrc40), 1456659_at (LOC552902), 1415968_a_at (Kap), 1456880_at (Hpvc2), 1440559_at (Hmga2-ps1), 1423878_at (Gypc), 1431120_a_at (Golga1), 1424843_a_at (Gas5), 1438700_at (Fnbp4), 1441178_at (Dtwd2), 1416467_at (Ddx3x), 1436581_at (D3Ertd789e), 1442280_at (D2Ertd750e), 1458941_at (D130016B08Rik), 1444318_at (Chchd7), 1440332_at (Cdv3), 1427953_at (BC025462), 1434426_at (B130055D15Rik), 1433935_at (AU020206), 1457900_at (Ascc3), 1435584_at (AI662791), 1437570_at (AI503301), 1451456_at (6430706D22Rik), 1457218_at (6430510M02Rik), 1442933_at (6230415M23Rik), 1440163_at (6030490B17Rik), 1431347_at (5730407M17Rik), 1429537_at (5730406M06Rik), 1435728_at (5230400J09Rik), 1419612_at (4632415L05Rik), 1432218_a_at (4632412I24Rik), 1439726_at (4432406C05Rik), 1442421_at (2900083I11Rik), 1453050_at (2700085M18Rik), 1438429_at (2610319H10Rik), 1429268_at (2610318N02Rik), 1453067_at (2610040C18Rik), 1429882_at (2610005L07Rik), 1442083_at (1500011J06Rik), 1453683_a_at (1200008O12Rik), 1431235_at (1110061A14Rik), 1416033_at (1110006I15Rik), 1433640_at (N/A), 1441177_at (N/A), 1456077_x_at (N/A), 1456145_at (N/A), 1458902_at (N/A), 1459302_at (N/A)

Supplementary table 1. Continued.

77

Chapter 4. Validation of Self-Renewal Pattern-Associated (SRPA) Gene Expression

Rationale:

Normalization is required for most microarray platforms to extract information on differential gene expression among compared biological samples (Tseng et al., 2001; Quackenbush, 2001 and 2002). Normalization adjusts for the differences caused by assay technical variability due to hybridization, washing, and fluorescent dye-labeling efficiencies. Many normalization algorithms assume that the majority of the genes in the array express similar transcript levels among samples (Quackenbush, 2001 and 2002). Based on this assumption, the global normalization approach has been widely used. Global normalization even may correct for different amounts of total RNA samples used in the labeling. However, in special conditions that relatively large numbers of transcripts are changing among samples, the global normalization may skew the actual amount of transcripts. Therefore, the gene expression values of microarray studies should be carefully interpreted. Ideally, microarray expression of the special genes of interest for following studies should be confirmed by independent experiments like quantitative real-time reverse transcriptase polymerase chain reaction (Q-RT-PCR), Northern blotting, or other RNA quantitation method.

The state of a cell can be evaluated by measuring mRNA expression. However, many functions of a cell are ultimately regulated by the activity or modulation of proteins. Not surprising, changes in gene expression are not always correlated with the changes in the corresponding protein level (Gygi et al., 1999; Washburn et al., 2003). Gene

expression by itself is not sufficient to predict protein expression level and activity. Thus, it is essential to measure the level of expression of proteins of interest directly when evaluating the functional significance of the genes discovered by microarray studies.

In this research, Q-RT-PCR was performed to confirm the gene expression of the exclusive ASRA genes predicted to encode plasma membrane proteins. The correlation between microarray data and Q-RT-PCR results was evaluated as a test of the reliability of the PM/MM statistical modeling. The expression of particular proteins of interest predicted to be encoded by SRPA genes were verified with Western blotting.

Materials and Methods

Quantitative real-time reverse transcription polymerase chain reaction (Q-RT-PCR)

To validate the microarray data, Q-RT-PCR was performed with the TaqMan® gene expression assay kit (Applied Biosystems, Foster City, CA). The kit consists of a FAM™ dye-labeled TaqMan® MGB probe and two primers for detecting genes of interest. The gene expression of *Plxdc2* (Mm00470649_m1), *Eng* (Mm00468256_m1), *Notch1* (Mm00435245_m1), *9130213B05Rik* (Mm00523913_m1), *Cxcr6* (Mm00472858_m1), *Rhbdl7* (Mm00524533_m1), *Sulf2* (Mm00511193_m1), *Ephb6* (Mm00432456_m1), *Fst* (Mm00514982_m1), *Robo1* (Mm00803879_m1), *Trpv2* (Mm00449223_m1), and *Kcnn4* (Mm00464586_m1) was evaluated. The expression of each gene was normalized against the expression level of glyceraldehydes-3-phosphate dehydrogenase (*Gapdh*). One µg of total RNA from the p53-induced asymmetrically self-renewing Ind-8 cells and p53-null

symmetrically dividing Con-3 cells was used for reverse transcribing cDNA with the Superscript Reverse Transcriptase (RT) II kit (Invitrogen, Carlsbad, CA). The cDNA from the reverse transcription reaction was amplified by PCR to measure the FAM fluorescence of each PCR cycle using the Applied Biosystems 7700 Sequence Detection System. The reactions were performed using the manufacturer's instruction after adjusting the reaction volume as 25 µl.

Immunoblotting

Western blot analyses were performed with protein extracts prepared from cells cultured for 36-48 hrs, corresponding to two PDCs after asymmetric self-renewal was induced as described previously (Liu et al., 1998a; Rambhatla et al., 2001). Survivin (1:100), cyclin G1 (1:300), PCNA (1:1000), EPHB6 (1:1000), and EDG (1:1000) proteins were detected with the indicated dilutions of the corresponding antibodies (Santa Cruz Biotech, CA). Anti-HMGB1 (1:250), anti-HMGB2 (1:250), and anti-notch1 (1:400) antibodies (BD Biosciences, CA) were used according to the manufacturer's protocols. Anti-pleiotrophin (Abcam Inc. Cambridge, MA) and anti-proliferin (R&D Systems, Minneapolis, MN) antibodies were used at a 1:500 dilution. Detection of primary antibodies was performed with a 1:5000 dilution of peroxidase-conjugated anti-mouse, anti-rabbit, or anti-goat IgG polyclonal antibodies (Amersham Biosciences, UK). Specific protein bands were detected with ECL western blotting detection reagents according to manufacturer's protocols (Amersham Biosciences, UK). Band intensities were quantified with NIH-image J software.

Results

Validation of the microarray expressions of the ASRA genes

The expression of seven exclusive ASRA genes, *9130213B05Rik*, *Plxdc2*, *Cxcr6*, *Notch1*, *Ephb6*, *Rhbdl7*, and *Eng*, all of which encode plasma membrane proteins, were evaluated by Q-RT-PCR. Four ASRA genes, *Robo1*, *Trpv2*, *Kcnn4*, and *Fst*, that were in the set of genes up-regulated in asymmetric self-renewal were also evaluated. Eight of the evaluated 11 ASRA genes showed statistically significant up-regulated expression in asymmetrically self-renewing cells compared to their symmetric counterparts (Table 9). In the Q-RT-PCR study, one exclusive ASRA gene, *plxdc2*, was only detected exclusively during asymmetric self-renewal. The transcripts of the other 6 exclusive ASRA genes were detected at a significant level in symmetrically dividing cells.

Although the exclusive ASRA genes were called as 'absent' in the Affymetrix arrays, they have expression values for the probe sets based on their fluorescent intensities. The expression-ratio values of the asymmetric and symmetric samples were calculated after reading the Affymetrix cel files with the dChip software by setting the PM only model (Li and Wong, 2001; www. dChip. org). The correlation between the Affymetrix array analysis and the Q-RT-PCR analysis was affected by genes that showed lower expression level (Figure 14). When the correlation was re-evaluated with the ASRA genes whose gene expression was verified, better linear correlation was observed between the Affymetrix array analysis and the Q-RT-PCR analysis (Data not shown).

	Mean C_t, (STD)		Fold change (Asym/Sym) Mean (STD)	Paired t-test
	Asymmetry Ind-8 cell	Symmetry Con-3 cell		
18S rRNA (n = 6)	15.19 (1.37)	14.53 (0.50)	0.99 (0.82)	0.38
Gapdh (n = 6)	16.44 (0.48)	16.63 (0.84)	1.24 (0.57)	0.48
9130213B05Rik (n = 6)	25.93 (0.45)	27.88 (1.01)	4.75 (3.13)	0.006
Plxdc2 (n = 3)	23.53 (0.19)	Not detected	Exclusive	0.001
Notch1 (n = 3)	24.47 (1.25)	25.50 (1.18)	2.08 (0.47)	0.03
Eng (n = 3)	29.07 (0.51)	30.59 (0.81)	3.22 (1.62)	0.11
Ephb6 (n = 3)	25.10 (0.20)	27.16 (0.55)	4.60 (2.56)	0.04
Cxcr6 (n = 6)	27.26 (0.51)	30.21 (0.73)	8.59 (4.40)	0.002
Rhbdl7 (n = 3)	22.95 (0.38)	23.47 (0.57)	1.66 (1.16)	0.44
Fst (n = 3)	24.53 (0.43)	27.20 (0.59)	6.48 (1.56)	0.007
Kcnn4 (n = 3)	26.99 (0.47)	29.58 (0.50)	6.90 (4.59)	0.004
Robo1 (n = 3)	26.52 (0.80)	29.05 (0.76)	8.26 (8.82)	0.10
Trpv2 (n = 3)	25.10 (0.20)	25.83 (0.63)	1.80 (0.86)	0.23

Table 9. Quantitative real time-RT PCR analyses of the ASRA genes. For validation of microarray expression data, Q-RT-PCR was performed. Gene expression was normalized by adjusting the threshold cycle (C_t) for each gene to the C_t of gapdh gene.

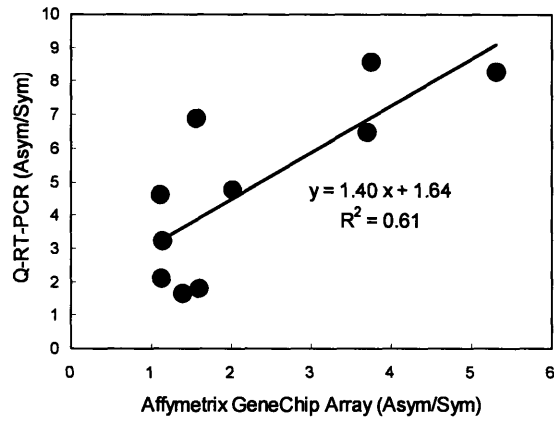Figure 14. Correlation between the Affymetrix GeneChip® array data and the Q-RT-PCR analyses. For calculating the ratio values of the exclusive ASRA genes, the expression values from the PM only model of the cChip software were used. Plxdc2 was not included in the analysis because its expression was not detected in symmetrically self-renewing Con-3 cells in Q-RT-PCR analysis. Ten ASRA genes that were evaluated by Q-RT-PCR.

The expression of five exclusive and 3 differentially expressed ASRA genes were validated by the Q-RT-PCR. The gene selection method for finding exclusively expressed ones detected genes that were disregarded in the conventional expression-ratio analyses. In addition, the PM data analysis indicated that in large part differences in gene expression determined by microarray were quantitatively similar to Q-RT-PCR measurements.

Validation of SRPA protein expression

The SRPA genes for the validation of protein expression were selected based on the following criteria: 1) availability of antibodies for immunoblotting; 2) identification in both from the cDNA and Affymetrix arrays; and 3) potential functional relevance to ASC-specific functions. Cells from the comparison models 1, 2, and 4 were used in these studies (Figure 8, Chapter 2). The expression of five ASRA gene-encoded proteins were evaluated in immunoblot analyses (Figure 15). Proliferin, a member of prolactin cytokine/growth hormone family, was selected for its potential application to ASC expansion (Choong et al., 2003). Cyclin G and notch1 were selected for analysis, because they have been reported for the differential expression in HSCs and for the association with the asymmetric cell divisions during embryonic development, respectively (Artavanis-Tsakonas et al., 1999; Stier et al., 2002; Morris et al., 2004; Wagner et al., 2004). In Western blotting analyses, cyclin G, proliferin, and notch 1 proteins were up-regulated in asymmetrically self-renewing cells compared to symmetrically renewing cells (Figure 15; compare lanes 3 to 4 and 7 to 8).
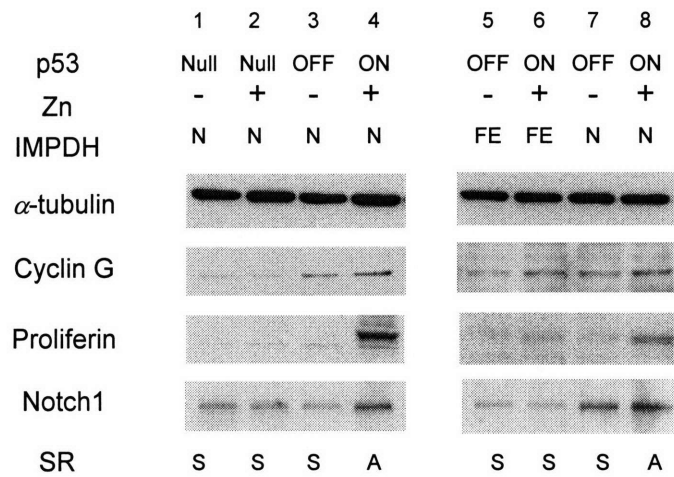
Figure 15. Evaluation of predicted ASRA gene protein expression. Immunoblottings for cyclin G1, proliferin, and notch 1 proteins were performed with total cell extracts harvested from the same conditions as used for the model 1 and 4 in the cDNA microarray analyses (*i.e.,* lane 2 versus 4 and lane 6 versus 8, respectively). In addition, the Zn-free control state of each compared condition was evaluated. Lanes 1 and 2, p53-null Con-3 cells; lanes 3 and 4, Zn-responsive p53-inducible Ind-8 cells (*i.e.,* Comparison 2); lanes 5 and 6, Zn-responsive p53-inducible IMPDH-transfected tI-3 cells (N, normal IMPDH expression; FE, forced expression); lanes 7 and 8, Zn-responsive p53-inducible vector-transfected control tC-2 cells. For all analyses, 20 μgs of extract protein were analyzed. α-tubulin intensity was used as a normalization standard to quantify the changes in protein levels. SR: self-renewal, S: symmetric self-renewal, A: asymmetric self-renewal, N: normal level, FE: forced expression
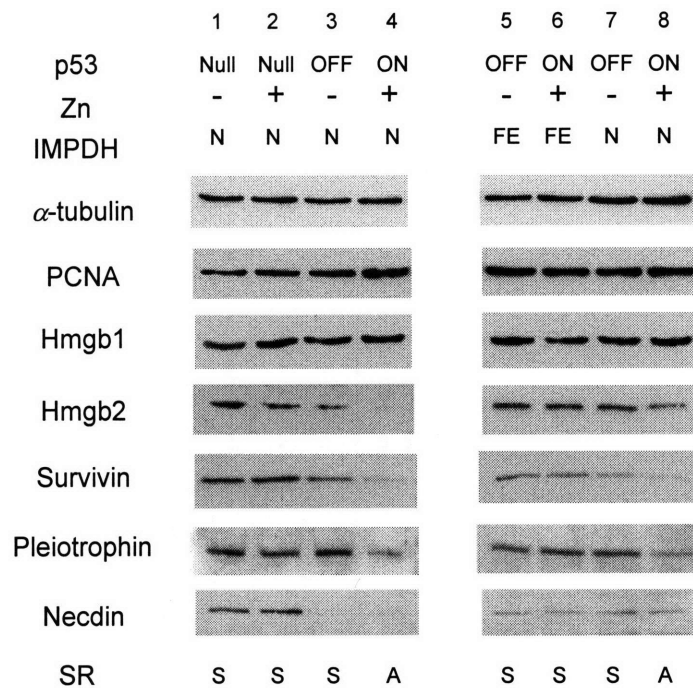
Figure 16. Evaluation of predicted SSRA gene protein expression. Western blotting analyses for PCNA, HMGB1, HMGB2, survivin, pleoitrophin, and necdin proteins were performed with total cell extracts prepared from cells grown as the same conditions as used for the model 1 and 4 in the cDNA microarray analyses (*i.e.,* lane 2 versus 4 and lane 6 versus 8, respectively). In addition, the Zn-free control state of each compared condition was evaluated. Lanes 1 and 2, p53-null Con-3 cells; lanes 3 and 4, Zn-responsive p53-inducible Ind-8 cells (*i.e.,* Comparison 2); lanes 5 and 6, Zn-responsive p53-inducible IMPDH-transfected tI-3 cells (N, normal IMPDH expression; FE, forced expression); lanes 7 and 8, Zn-responsive p53-inducible vector-transfected control tC-2 cells. For survivin, pleoitrophin and necdin analyses, 20 μgs of extract protein were analyzed. For PCNA, HMGB1, and HMGB2 analyses, 2, 0.5, and 1 μgs of extract protein were used, respectively. α-tubulin intensity was used as a normalization standard to quantify the protein levels. PCNA intensity, which was determined to be constant, was used as a normalization standard to quantify HMGB2 protein level.

As an ideal ASRA protein, protein expression must be sensitive to self-renewal pattern changes, not to p53 gene expression in the model cell systems. Interestingly, cyclin G protein still showed modest up-regulation in response to p53 expression without a shift to asymmetric self-renewal (Figure 15, FE; compare columns 5 and 6). Cyclin G was not expressed at a higher level in the asymmetrically self-renewing cells of the model 4 comparison in the cDNA microarrays (Figure 15; compare columns 6 and 8), suggesting that it underwent a p53-dependent up-regulation independent of asymmetric self-renewal state. The transcripts of cyclin G showed the expression pattern of ASRA genes, the protein expression was not an ASRA pattern. These relationships highlight the importance of independent criteria for refining interpretations of microarray analyses.

Six SSRA gene encoded proteins were evaluated in immunoblotting analyses (Figure 16). Survivin and HMGB proteins were reported to be associated with chromosome segregation and chromatin architecture regulation (Altieri, 2003; Muller et al., 2001; Agresti and Bianchi, 2003). These proteins could potentially play a role in the unique non-random chromosome co-segregation of ASCs. Three predicted SSRA proteins, HMGB2, survivin, and pleiotrophin, were down-regulated in asymmetrically self-renewing cells compared to symmetrically self-renewing cells, which correlated with the microarray data (Figure 16; compare lanes 3 to 4 and 7 to 8). Importantly, the changes in the protein expression of HMGB2, survivin, and pleoitrophin did not occur when p53 was expressed but asymmetric self-renewal was prevented by the forced expression of IMPDH proteins (Figure 16, FE; lanes 5 and 6). This finding illustrates that the protein

expressions of these three SSRA genes are associated with the state of self-renewal pattern, not directly with p53 expression. Therefore, they are SSRA proteins.

In contrast, HMGB1 and PCNA proteins showed no significant change between the states of self-renewal pattern. Such disagreements between the microarray and protein analysis results can occur due to translational and post-translational mechanisms (Gygi et al., 1999; Washburn et al., 2003). The expression of the necdin gene was also not self-renewal pattern specific (Figure 16). Necdin was identified as an SSRA gene from the Affymetrix whole genome arrays that evaluated gene expressions of the three self-renewal patterns, p53-induced Ind-8 cells for asymmetric, p53-null Con-3 cells for symmetric, and p53-induced/IMPDH transfectant tI-3 cells for symmetric self-renewal. The pattern of necdin expression in these three conditions was the same as its gene expression in the microarrays (Figure 16; lanes 3 and 4, lanes 4 and 6). The changes of necdin expression are more likely to be associated with the cell line specific properties, not with the conditions of self-renewal pattern. Therefore, the comparison between two different cell culture conditions with the same isogenic cell lines (Model 2 in the orthogonal-intersection cDNA microarray analysis) was useful to exclude the genes like necdin. These results indicated that 543 SRPA gene signature contained false-positives that require independent analyses for detection.
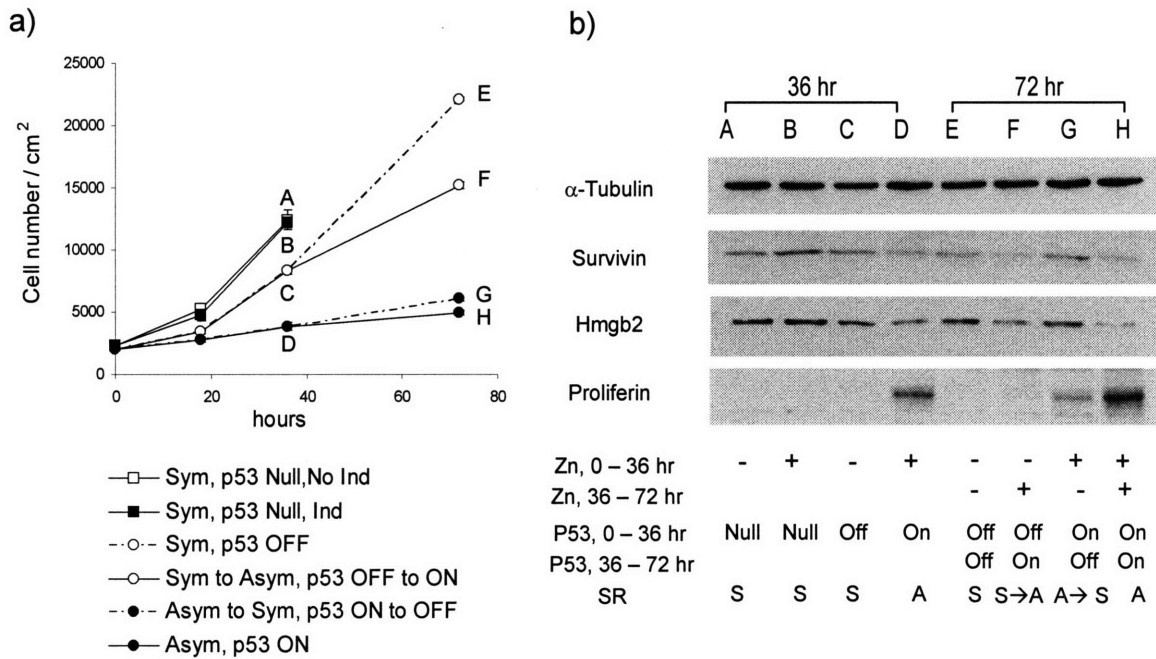
Figure 17. Reversible changes of SRPA protein expressions in response to the self-renewal pattern changes. Western blotting analyses for survivin, HMGB2, and proliferin were performed with total cell extracts harvested from the Zn-responsive p53-inducible Ind-8 cells and p53-null Con-3 cells with time in Zn-free (-) and Zn-supplemented (+) medium which induces transition from symmetric self-renewal to asymmetric self-renewal by p53-inducible cells but not p53-null cells. Twenty μgs of protein extracts were used for survivin and proliferin immunoblot analysis and 1 μgs for HMGB2 analysis. SR, self-renewal; S, symmetric self-renewal; A, asymmetric self-renewal.

SRPA proteins respond reversibly to self-renewal symmetry conditions

Previous change in SRPA protein expression was compared between static states of asymmetrically and symmetrically self-renewing cells. Next, changes in expression were evaluated when cells were shifted from symmetric to asymmetric self-renewal or vice versa (Figure 17). When the cell culture condition for self-renewal pattern was changed from non-inducing condition (symmetric) to p53-inducing condition (asymmetric) after two PDCs (36 hrs), survivin and HMGB2 protein levels were down-regulated by two PDCs (72 hrs) (Figure 17; F). These two SSRA proteins were also evaluated after cells were shifted from conditions of asymmetric self-renewal to conditions of symmetric self-renewal (Figure 17; G). Therefore, the expression of suvivin and HMGB2 proteins changes reversibly with the self-renewal pattern.

When the self-renewal pattern was switched from symmetric to asymmetric after two PDCs (36hrs), the expression of proliferin was barely detected after 72 hrs (Figure 17; F). This effect was reproducible. When cells have undergone symmetric self-renewal for at least two PDCs, proliferin gene may adopt an irreversible state. This also may be related to the cell density effect on self-renewal symmetry (Chapter 1. Figure 7). Even in the p53-inducing conditions, as the number of cells increased over the certain density in culture, the small fraction of cells in population started to self-renew symmetrically and cell growth kinetics were skewed toward exponential cell growth kinetics (Chapter 1). However, expression of proliferin, the ASRA protein, responded to the self-renewal symmetry transition from asymmetric to symmetric very well (Figure 17; G). Thus, proliferin may regulate the maintenance of asymmetric self-renewal in cell culture.

The pattern of changes in the expression level of survivin, HMGB2, and, to lesser extent, proliferin proteins can specify transitions from symmetric to asymmetric self-renewal in culture. As more ASRA and SSRA proteins are evaluated and selected for inclusion in this phenotypic signature for self-renewal symmetry regulation, its specificity and sensitivity for indentifying the self-renewal pattern of any given cell population will increase.

Discussion and Conclusion

The microarray expression of the SRPA genes were validated with Q-RT-PCR and immunoblotting analyses. The five ASRA genes in the exclusive gene set showed differential expression by Q-RT-PCR. Three differentially expressed ASRA genes, *fst*, *robo1*, and *kcnn4*, were validated in Q-RT-PCR analysis. Especially, Q-RT PCR analysis confirmed that the transcripts of the *plxdc2* gene were exclusively expressed in the p53-induced asymmetrically self-renewing cells. In general, PCR based detection has greater sensitivity than hybridization-based detection methods. Also, the hybridization efficiency of the oligonucleotide sequences in the Affymetrix probe sets may be differ from those of the primer and probe sets in the TaqMan® detection assay. For instance, the threshold cycle numbers (C_t) of *notch1* and *rhbdl7* genes in the Q-RT-PCR were same or even smaller than those of differentially expressed ASRA genes like *robo1* and *trpv2* (Table 9). These results can be explained with the hybridization efficiency of oligonucleotide sequences in the Affymetrix probe sets.

With the 7 verified ASRA genes except *plxdc2* gene, the ratio values of the asymmetric to symmetric expressions from the Affymetrix whole genome array data well correlated with those of the Q-RT-PCR analysis (Figure 14). Importantly, 5 genes in the 7 examined exclusive ASRA genes showed up-regulation or exclusive expression in asymmetric self-renewal by independent experiments. This result suggested that the present/absent call-based method used for selecting the exclusive gene sets was effective for identifying genes that the expression ratio based analyses overlook.

In the immunoblotting analyses, translation of the 12 SRSA genes was evaluated. The predicted protein expression was validated for only 50% of the SRSA genes. The expression of survivin, HMGB2, pleoitrophin, proliferin and notch 1 proteins were dependant on the self-renewal pattern in cell culture specifically and not directly associated with p53 gene expression. Proteins like cyclin G responded not only to the self-renewal symmetry condition but also to p53 gene expression directely. Survivin, HMGB2, and proliferin expression changed sensitively and reversibly with the self-renewal pattern transitions. Because cellular states and functions are mostly determined by protein activity and modulation, the protein expression signature for given self-renewal state of cells may often digress from the gene signature (Gygi et al., 1999). In order to understand the regulation on self-renewal pattern and determine the self-renewal states of a specific cell population, more SRPA proteins should be verified to increase the pool of proteins associated with both asymmetric and symmetric self-renewal.

Through the cDNA and Affymetrix microarray analyses with model cell systems, hundreds of genes associated with self-renewal pattern regulation were identified. The

microarray expression of only a small set of SRPA genes was validated with Q-RT-PCR and Western blotting analyses with the originating model cell systems. Although the genetic signature for asymmetric self-renewal well represented the genes differentially expressed in ASC-enriched populations already reported in literature (Chapter 2 & 3), the gene expression profiles must be verified with the well-defined adult stem cells like those in small intestinal epithelium and skin epithelium or with well-characterized adult stem cell enriched populations like the bone marrow-Hoechst 33342 dye effluxing side-population that HSCs are enriched (Uchida et al., 2004; Fuchs et al., 2004). Therefore, the SRPA genes and proteins whose microarray expression was confirmed should be evaluated in these tissues or cell populations in order to demonstrate the thesis that asymmetric self-renewal associated genes may provide clues to find molecular markers for ASCs and also to understand ASC function.

Chapter 5. Self-Renewal Pattern Associated Genes As a Genetic Signature for Identification of Adult Stem Cell-Enriched Populations

<u>**Rationale**</u>:

Since genome-wide transcriptional profiling analysis was introduced, many gene expression signatures have been reported to diagnose specific disease states and their subtypes, as well as to define the functional states of various cells or tissues (Liotta and Petricoin, 2000; Ramaswamy et al., 2003; Segal et al., 2005; Eckfeldt et al., 2005; Ludwig and Weinstein, 2005). In previous gene expression signature studies, tens or hundreds of genes were proposed as signatures that were used to classify cancer subtypes, predict cancer progression, and select appropriate treatment regimens for diseases. In order to identify gene expression signatures, different data classification techniques have been applied such as hierarchical clustering, self-organizing maps, and principal component analysis (Eisen et al., 1998; Quackenbush, 2001; Yeung and Ruzzo, 2002; Cobb et al., 2005). Each method has its own advantages and limitations, and therefore their appropriate application is critical for the analytical quality of the gene-expression signature analysis (Quackenbush, 2001).

Five hundreds and fourty three SRSA genes were identified from the Affymetrix whole genome array analysis. This gene set contained many genes whose expression was previously associated with ASC-enriched populations (Chapter 3). Because asymmetric self-renewal is a unique ASC property, it was postulated that the 543 SRPA gene set would distinguish ASCs from other cell populations in appropriate clustering analyses. Since the stemness gene signatures were proposed, there have been controversies about

their poor correlation among similar studies in different laboratories (Burns and Zon, 2002; Cai et al., 2004; Zipori, 2004). In the microarray analyses for studying various cancers, principal component analysis was successfully applied to discover and predict cancer classes with a selected group of genes, i.e. cancer gene signature (Yeoh et al., 2002; Pomeroy et al., 2002). Therefore, principal component analysis was applied to evaluate the 543 SRPA gene set and the stemness gene set for their ability to specify ASC containing cell populations as distinct from non-adult stem cell populations. These evaluations were conducted by principal component analysis with published microarray data sets for different types on stem cell and non-stem cell populations.

Principal component analysis (PCA) is a multivariate statistical tool to reduce the dimensionality of large data sets and simplify evaluation of common features of the data (Alter et al., 2000, Raychaudhuri et al., 2000, Yeung and Ruzzo, 2001). The mathematical procedure of PCA transforms the large data into a small number of variables called principal components. Briefly, the symmetric covariance matrix that is calculated from the data matrix that consists of expression values by genes x arrays space. The covariance matrix is used to solve the "eigenvalues" and "eigenvectors" by eigen analysis. Then, the eigenvector with the largest eigenvalue becomes the first principal component. The first principal component contains the largest amount of variance information in a given data set. Usually, the first two or three principal components capture most of the variation in the data set. With these first few principal components, PCA can classify both genes and experiments from microarray data analysis. In other words, PCA can not only identify genes that have correlated expression patterns, but also cluster biological samples that

have similar expression patterns of any given gene set (Alter et al., 2000, Quackenbush 2001).

In this study, PCA was performed to test whether the 543 SRPA gene set and the stemness gene sets could be used as gene expression signatures to detect ASCs by analyzing publicly available microarray data sets. First, these gene sets from the transcriptional profiling studies for epidermal stem cell-enriched samples were evaluated. Two reports described the use of the same microarray platform as stemness signature studies to analyze isolated fluorescent label-retaining hair follicle cells enriched for epidermal stem cells (Tumbar et al., 2004; Morris et al., 2004). Next microarray data for CD34+-hair follicle stem cell-enriched populations derived with an Agilent microarray platform (manuscript in preparation with Dr. Carol Trempus) were analyzed. In order to test the functional relevance of the SRPA gene set with human stem cell population, a public microarray database for human CD34+ bone marrow cells and human bone marrow cells was also evaluated by PCA using gene expression data for human orthologues of the SRPA genes (Su et al., 2004; Ge et al., 2005).

Materials and Methods:

Data analysis from the public gene expression profiling databases

The raw data files for the Ivanova *et al* stem cell molecular signature study were available from the RNA Abundance Database (RAD: www.cbil.upenn.edu /RAD/php/index.php) (Ivanova et al., 2002). The cel files from the RAD were

mBM_LTHSC_2_U74Av2 (LT-HSC1), mBM_LTHSC_3_U74Av2 (LT-HSC2), mBM_LTHSC_4_U74Av2 (LT-HSC3), mBM_LTHSC_5_U74Av2 (LT-HSC4), ESC_1_U74Av2 (ESC1), ESC_2_U74Av2 (ESC2), NSC_1_U74Av2 (NSC1), NSC_2_U74Av2 (NSC2), and NSC_3_U74Av2 (NSC3). The raw data from Tumbar et al. and Morris et al. were acquired from www.rockefeller.edu/labheads/fuchs/database.php and www.uphs.upenn.edu/dermatol/re-search/HFGenedatafile.html, respectively (Tumbar et al., 2004; Morris et al., 2004). These two studies reported the gene expression profiling of hair follicle label-retaining stem cell-enriched populations. The Affymetrix raw data cel files for the human and mouse gene atlas were provided by the Genomics Institute of the Novartis Research Foundation (Su et al., 2004). The cel files for normal human tissue were also downloaded from the supplementary information web-site for the Ge et al. study (www2. genome.rcast.u-tokyo.ac.jp/normal/) (Ge et al., 2005). For the analytic groups of Affymetrix cel files for caculating PCA parameters, all cel files were analyzed as a group and normalized together with dChip to the median intensity using the PM-only model (www. dchip. org).

Microarray data analysis of CD34$^+$-hair follicle stem cell-enriched population study

Dr. Carol Trempus at the National Institute of Environmental Health Sciences (NIEHS) kindly provided the microarray expression data for the CD34$^+$ hair follicle stem cell-enriched populations. The microarray platform for this study was the Agilent Mouse Oligo Microarray (V2) G4121B array (Agilent Technologies, Palo Alto, CA). Genes in this data were identified that constantly existed in the 5% tail areas of triplicate microarray

comparisons between CD34+ hair follicle cells and CD34- hair follicle cells (Chapter 2). For PCA, genes of interest in the Agilent array were found by using Entrez Gene ID. Normalized expression values for genes of overlaps were used without any modification of the original data.

Principal component analysis

For each PCA, expression values were normalized and extracted from the comparing group of the Affymetrix cel files simultaneously by using dChip software 2005. The normalized expression values from the same analysis were used to calculate the parameters of PCA. The calculation of PCA was performed with the Cluster 3.0 program (Eisen et al., 1998). The first two eigenvecters were used for clustering evaluations.

<u>**Results**</u>:

Stemness gene sets do not separately cluster asymmetrically self-renewing cell populations from symmetric self-renewing cell populations

Because few overlaps existed among stemness signatures produced by similar methods and similar cell preparations (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003), they have not been utilized to evaluate other cell populations for evidence of stem cell character. In addition to discussed technical limitations, previous stemness signatures may have inadvertently excluded ASC-specific genes.
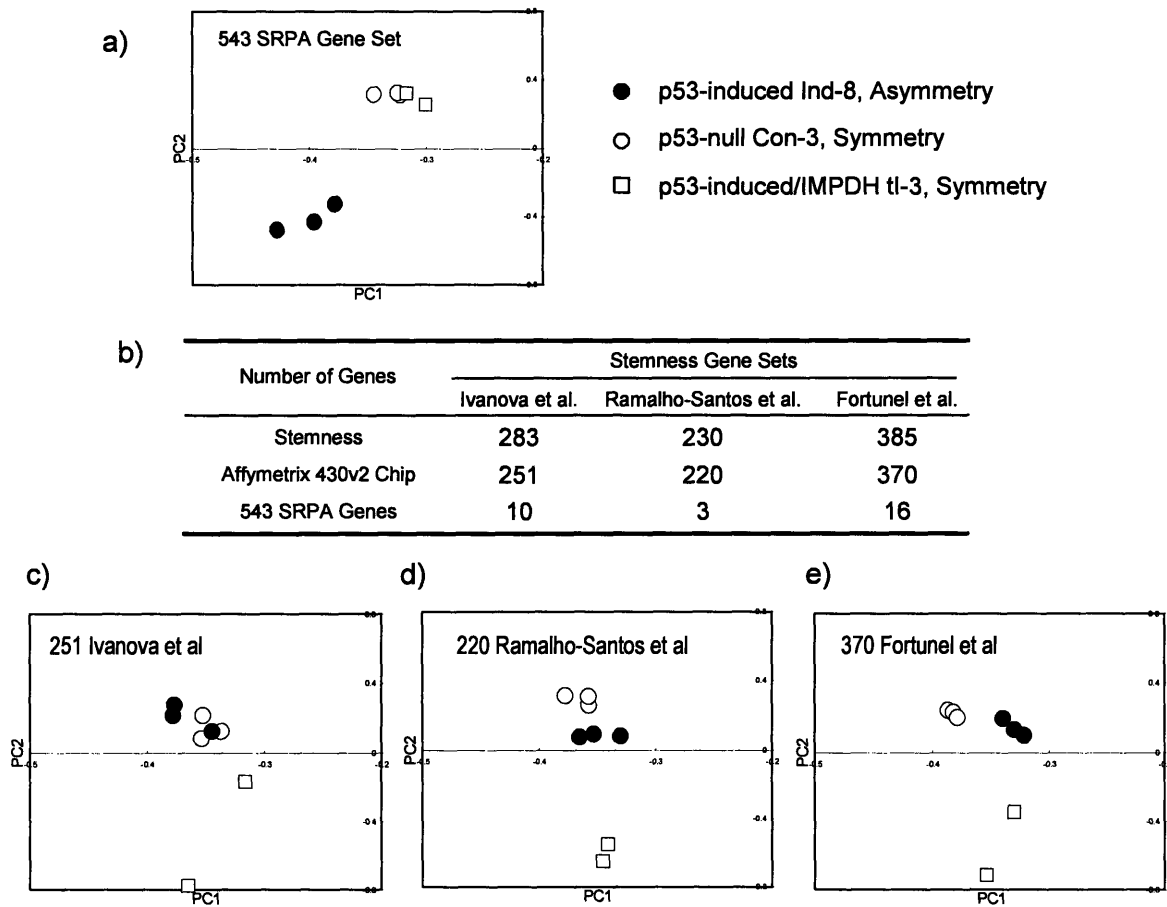
Figure 18. Principal component analyses of reported stemness gene sets for evaluating the clustering patterns of model cell lines for self-renewal pattern regulation. The parameters for PCA were calculated from the expression values of the Affymetrix whole genome mouse 430 2.0 arrays that contained gene expression information of p53-induced Ind-8 (closed circle), p53-null Con-3 (open circle), and p53-induced/IMPDH transfected tl-3 (open square) cell populations used for defining the 543 SRPA genes. a) PCA with the 543 SRPA genes (same as Figure 13a). b) The first row and second rows show the number of the stemness genes from each stemness study, and the number of stemness genes exist in the Affymetrix 430 2.0 array, respectively. The numbers in the third row are the overlapping genes between the stemness gene sets and the 543 SRPA genes. c) PCA with the 251 stemness genes proposed by Ivanova et al. 2002. d) PCA with 220 stemness genes by Ramalho-Santos et al. 2002. e) PCA with 370 stemness genes by Fortunel et al. 2003.

As shown in Figure 18, the stemness gene sets from the three reported studies had no power to distinguish asymmetrically self-renewing cell populations (asymmetric cluster) from symmetrically self-renewing populations (symmetric cluster). Of the 283 stemness genes reported by Ivanova et al., 251 genes existed on the Affymetrix mouse whole genome 430 2.0 array. PCA using the 251 gene set was unable to separate different self-renewal pattern clusters (Figure 18b,c). The stemness genes reported by Ramalho-Santos et al. also could not separate the asymmetric cluster from the symmetric cluster (Figure 18d). The PCA with the 370 stemness gene set by Fortunel et al. resulted in three separated clusters. The individual clusters corresponded to the three model cell populations for self-renewal pattern analyses. However, both symmetric clusters were widely separated by the 370 gene set (Figure 18e). These analyses showed that the stemness gene sets lack of the genetic features associated with differences in self-renewal pattern.

The SRPA gene expression signature distinguishes HSC-enriched cell populations

In the stemness transcriptional profiling studies, genes commonly up-regulated in three different stem cell preparations were identified (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003). ESCs in culture undergo symmetric self-renewal that manifest as ideal exponential cell accumulation kinetics (Evans and Kaufmann, 1981). In neurosphere cultures that contain NSCs, the overall cell growth kinetics are close to exponential, although all cells in the neurosphere cannot give rise to a secondary neurosphere (Carpenter et al., 1999; Gottlieb, 2000; Reynolds and Rietze, 2005).

Long-term functional HSCs (LT-HSCs) have characteristics of asymmetric self-renewal (Morrison and Weissman, 1994; Punzel et al., 2002 and 2003).

The SRPA gene signature was evaluated for its ability to classify stem cell-enriched populations based on their predicted self-renewal pattern. Affymetrix raw cel files were only publicly accessible for the Ivanova et al. study (Ivanova et al., 2002). Only 151 of the 543 SRPA genes were available on the Affymetrix mouse U74Av2 chip. PCA using the 151 SRPA genes with the model cell line microarray data indicated that it was a somewhat better signature than the full 543 gene set (Figure 18a, Figure 19a). This 151 SRPA gene signature generated two clusters by PCA from the three stem cell-enriched samples in the Ivanova et al. study (Figure 19b). The samples for ESCs and NSCs clustered together, and the samples for LT-HSCs were separated from the ESC/NSC cluster. Considering that *in vivo* sample preparations for LT-HSC from mouse bone marrow are usually variable (Cheshier et al., 1999), the less densely structured cluster of LT-HSCs was not surprising. These PCA results showed that the partial 151 SRPA gene signature was able to classify adult stem cell-enriched populations. Moreover, the classification matched predictions for the self-renewal pattern of examined stem cell containing cell populations.
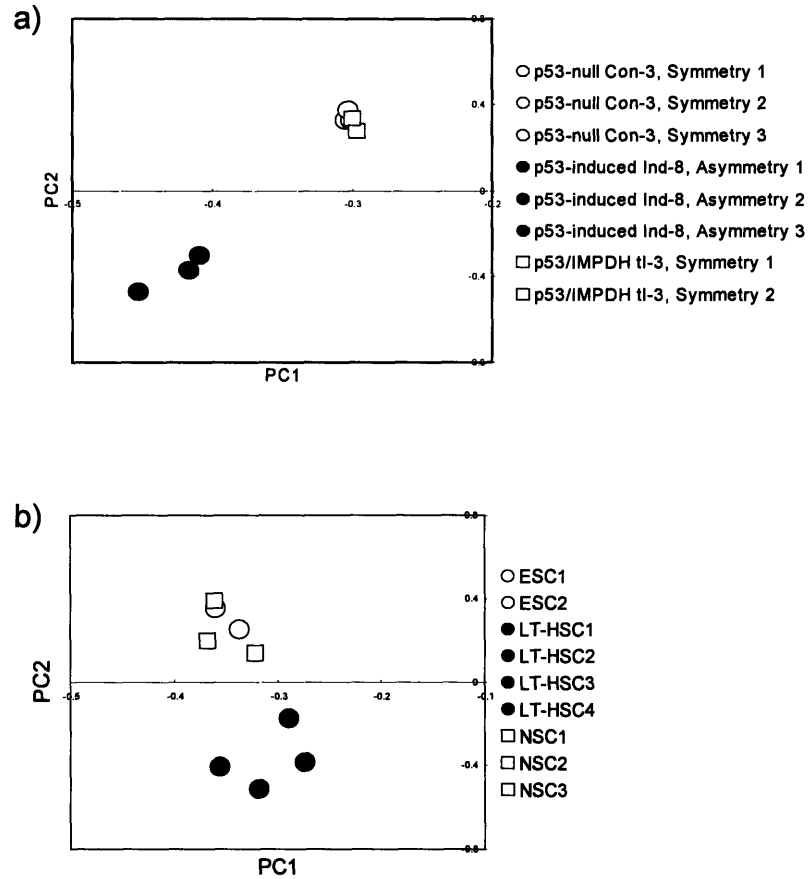
Figure 19. Principal component analysis with a 151 gene subset of SRPA gene signature for evaluating the clustering patterns of stem cell-enriched populations. a) PCA using 151 SRPA genes with the expression values from the self-renewal symmetry transcriptional profiling study with indicated model cell line data b) The parameters for PCA were calculated with the expression values of the 151 SRPA gene set from the Affymetrix U74Av2 arrays that contain gene expression data of Ivanova et al. The expression values of the 151 SRPA genes were extracted from the Affymetrix cel files by the dChip software. PCA parameters were calculated with Cluster 3.0 (Eisen et al., 1998). ESC: embryonic stem cell, LT-HSC: Long-term hematopoietic stem cell, NSC: neuronal stem cell neurosphere culture.

Evaluation of SRPA and stemness genes for identifying epidermal label-retaining cell populations

Label-retaining cells (LRCs) in the adult hair follicle bulge region are widely believed to be epidermal stem cells (Taylor et al., 2000; Trempus et al., 2003). Microarray data for the LRCs in the adult mouse hair follicle bulge region were analyzed with the SRPA gene and stemness gene signatures (Tumbar et al., 2004; Morris et al., 2004). The transcriptional profiling studies for LRCs were performed with Affymetrix mouse U74Av2 chips, the same microarray platform used in the three stemness studies. For the Tumbar et al study, both the SRPA gene set and the stemness gene sets generated two distinctive clusters by PCA (Figure 20). One cluster was the LRC-cluster and the other was the non-LRC populations. However, for the Morris et al study, it was difficult to decide the clustering pattern without prior sample information, especially for the two stemness gene sets proposed by Ramalho-Santos et al. and Fortunel et al. (Figure 21). Both the SRPA gene set and the stemness gene set by Ivanova et al. showed marginal separation of the stem cell-enriched population by PCA. PCA using the 100 differentially expressed genes (> 3 fold) from the Morris et al. study showed that the EGFP-label retaining populations were highly variable (Figure 21b). In spite of the sample variation for the LRCs, the SRPA gene signature showed better classification of epidermal LRC populations compared to the other stemness gene signatures.

Figure 20. The clustering patterns of epidermal label-retaining hair follicle bulge cells by the SRPA and stemness gene sets. PCAs were performed with the expression values in the epidermal label-retaining hair follicle cell populations in the Tumbar et al study. a) PCA by the SRPA and stemness genes. b) PCA using the 152 genes differentially expressed in LRCs from hair follicle bulge region (Tumber et al., 2004).

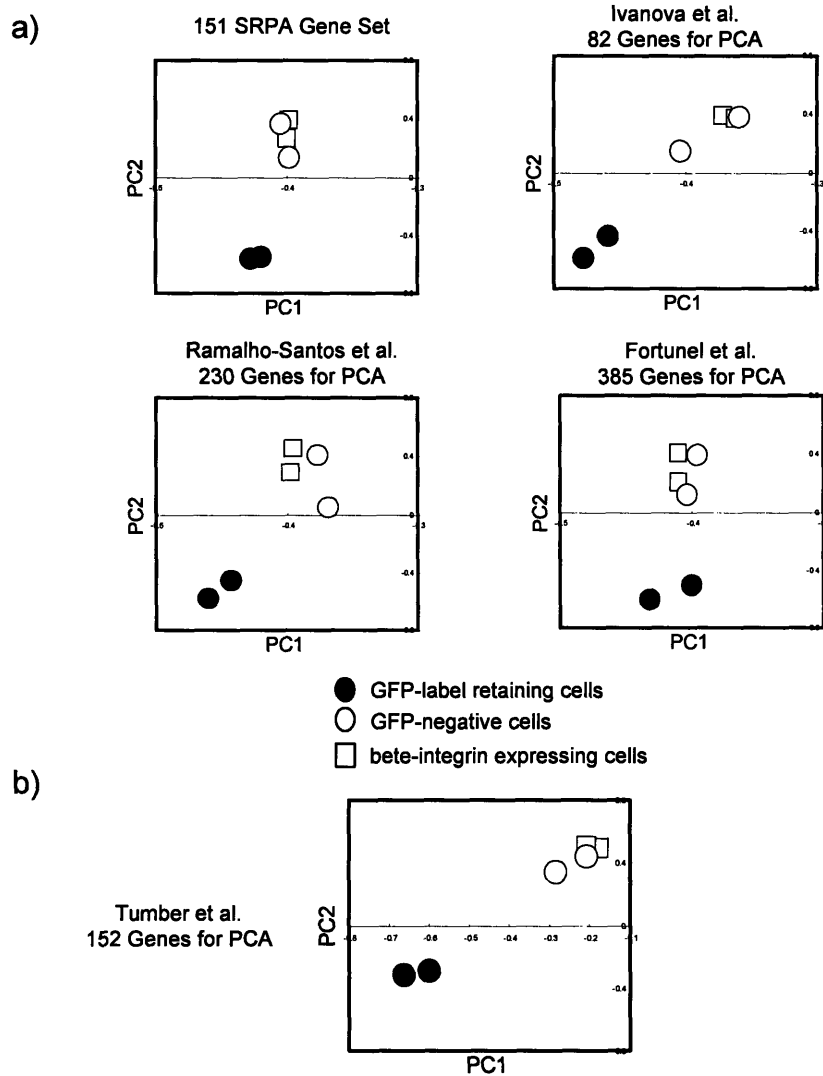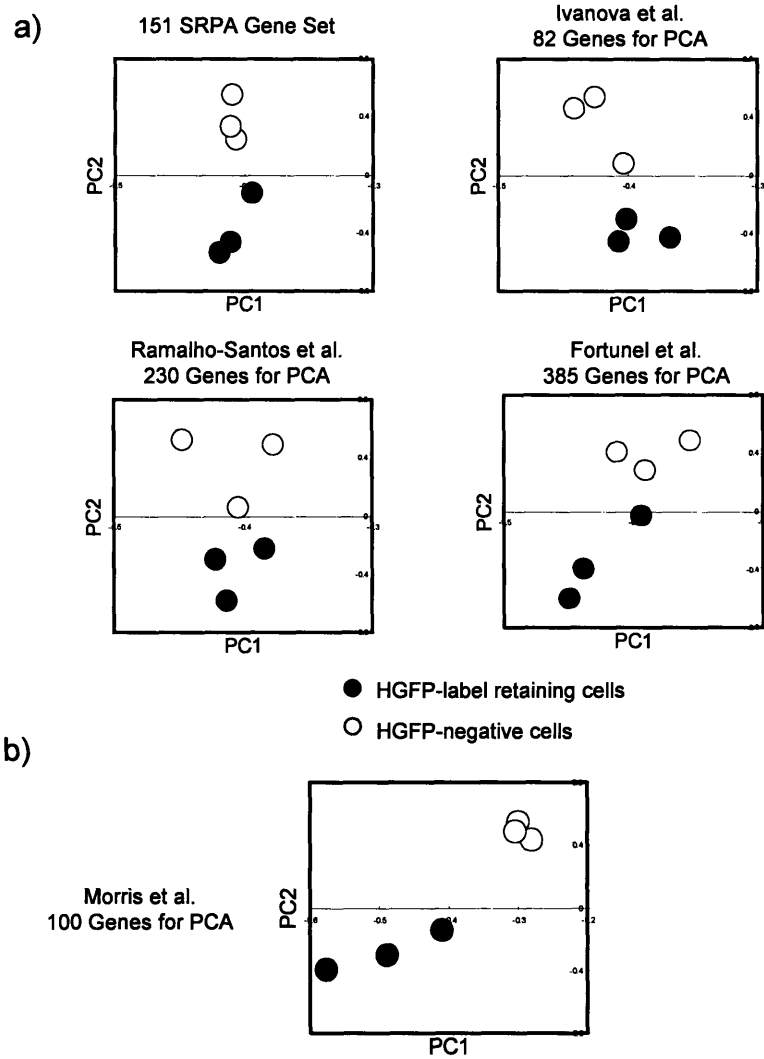Figure 21. The clustering patterns of epidermal label-retaining hair follicle bulge cells by the SRPA and stemness gene sets. PCAs were performed with the expression values in the epidermal label-retaining hair follicle cell populations reported by Morris et al. a) PCA by the SRPA and stemness genes. b) PCA using the 100 genes that showed over a 3-fold difference between LRCs and non-LRCs (Morris et al. 2004).

The association of the ASRA genes with CD34$^+$-hair follicle stem cell-enriched populations

Although LRCs in hair follicle bulge have been shown to be an epidermal stem cell-enriched population (Taylor et al., 2000; Oshima et al., 2001), currently, few surface markers are available for detecting and isolating epidermal stem cells (Toni et al, 2000; Trempus et al., 2003; Schneider et al., 2003). Recently, the putative hair follicle bulge epidermal stem cell populations were enriched and characterized based on the CD34$^+$/α6 integrin immunostaining (Trempus et al., 2003). The transcriptional profiling with the CD34$^+$ hair follicular bulge cells compared to CD34- cells was studied with an Agilent Mouse Oligo G4121B array platform. The normalized expression data (provided by Dr. Carol Trempus at the NIEHS) were analyzed for testing the SRPA gene signature. Genes were identified that consistently existed in the upper or lower 5% tail areas of all three of the triplicate CD34$^+$/CD34$^-$ ratio comparisons (Chapter 2, expression-ratio analysis). As results, 454 genes were identified as up-regulated in CD34$^+$ cells and 412 genes as up-regulated in CD34- cells.

	Self-renewal pattern associated genes	
	Asymmetric (ASRA)	Symmetric (SSRA)
Up-regulation in the CD34+ hair follicular bulge cells	*C1r, Ctsf, Jcam1, Myo1b, Tmtsp, Sall2, Rgl1, Ltbp2, Klf4, Hsp25, Ddr1, 1200013A08Rik, 1110020P09Rik*	*Osf2-pending, Nfib, Idb2, Ndr2*
Up-regulation in the CD34- hair follicular bulge cells	*Adam8, Dda3-pending, Klf2, Robo1, Serpinb2, 5830413E08Rik*	*Cldn1, Hmgb2, Ccna2, Cdc25c, Cxadr, Impa2, Incenp, Klf5, Lag, Rab6kifl*

Table 10. Overlap between SRPA gene and CD34 phenotype-associated gene sets

Table 10 shows the genes that overlapped between the 543 SRPA genes and the CD34+ or CD34⁻ associated genes identified for cell from the hair follicular bulge region. Thirteen genes overlapped between the ASRA genes and the up-regulated genes in the $CD34^{+}$ cells, and 11 overlapped between the SSRA genes and the up-regulated genes in the CD34- cells. This overlap pattern of ASRA genes among genes up-regulated in $CD34^{+}$ cells compared to SSRA genes among genes up-regulated in CD34- cells was significant (p = 0.037, Fisher's exact test). This overlap result implicated that epidermal stem cell-enriched population to contain molecular features of asymmetric self-renewal.

The SRPA gene set can distinguish $CD34^{+}$-hair follicle epidermal stem cell-enriched populations

The transcriptional profiling data of the $CD34^{+}$/CD34- epidermal cell populations were evaluated by PCA. Gene sets from the SRPA genes, the stemness genes, and also the differentially-expressed genes from the two previously reported hair follicle bulge LRC studies were evaluated. The gene symbols or Entrez Gene IDs were used to find the corresponding probes in the Agilent Mouse Oligo G4121B array (Figure 22a). PCA with the specific gene sets for epidermal LRCs gave rise to two clusters (Tumbar et al., 2004; Morris et al., 2004) (Figure 22b,c). One cluster consisted of data sets for six CD34- hair follicle bulge cells (open circles) and the other was for CD34+ hair follicle bulge cell data sets (closed circles).

108

a)

Number of Gene Overlap	SRPA Genes	Ivanova et al.	Ramalho-Santos et al.	Fortunel et al.	Tumbar et al.	Morris et al.
Gene Set	543	283	230	385	152	100
Agilent array	460	228	195	354	130	88

b)

Tumbar et al.
130 Genes for PCA

c)

Morris et al.
88 Genes for PCA

d)

460 SRPA Gene Set

e)

Ivanova et al.
228 Genes for PCA

f)

Ramalho-Santos et al.
195 Genes for PCA

g)

Fortunel et al.
354 Genes for PCA

● CD34+ hair follicle bulge cells

○ CD34- hair follicle bulge cells
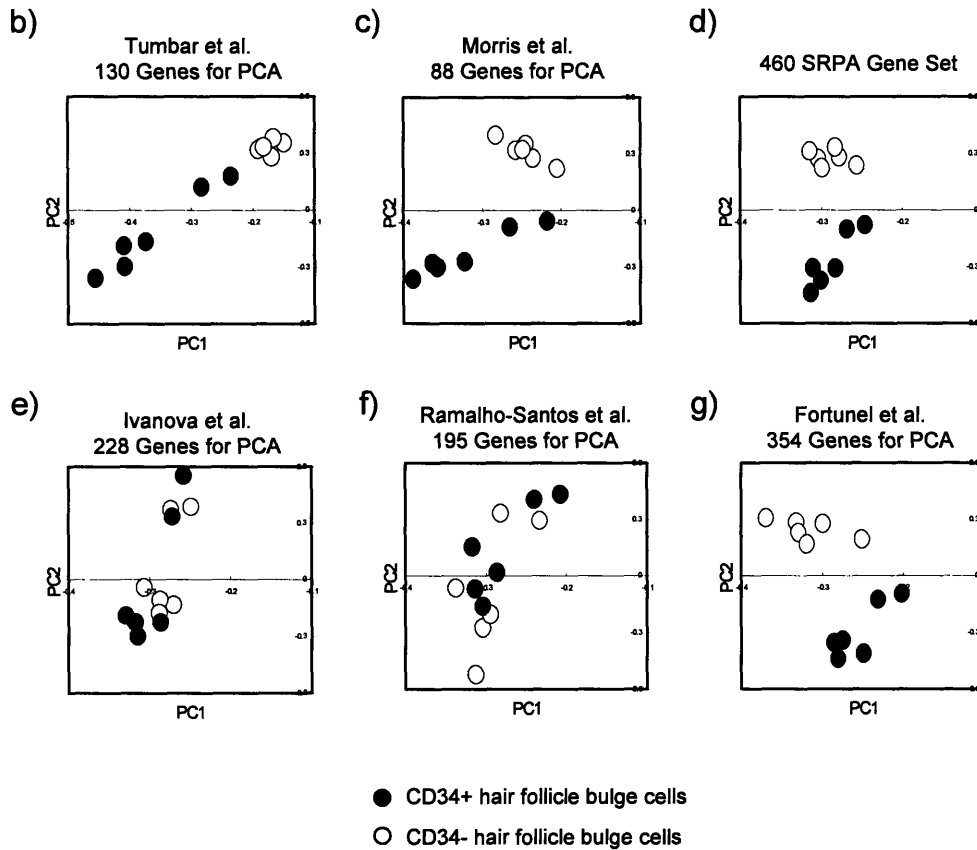
Figure 22. The clustering patterns of CD34[+] and CD34- hair follicle bulge cell populations. PCAs were performed with the expression values of the SRPA, stemness, and hair follicle LRC-specific genes from replicate Agilent Oligo Mouse arrays. a) The array comparison b-g) PCA by the SRSA, stemness, and hair follicle LRC-specific genes genes.

Figure 23. PCA for human bone marrow samples with the SRPA gene set.
PCAs using 320 SRSA genes generated the distinctive cluster of CD34[+] bone marrow samples separated from the normal BM samples. All cel files were processed and normalized together by one process with dChip.

110

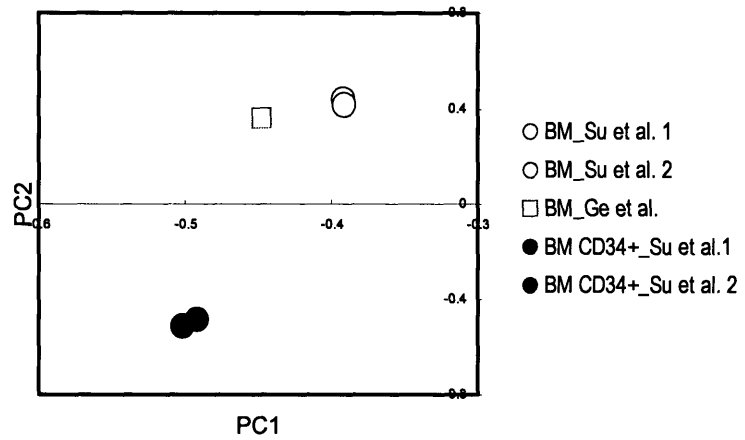Next, the SRPA gene signature was evaluated. In the Agilent array, 460 of the 543 SRPA genes were available for comparison (Figure 22a). PCA using the 460 SRPA gene signature clearly showed the distinctive clusters between the $CD34^+$ and CD34- hair follicle bulge cell populations (Figure 22d). However, separation was not obtained between $CD34^+$ and $CD34^-$ hair follicle cells by PCA with the two stemness gene sets identified from ESCs, NSCs, and HSCs (Ivanova et al., 2002; Ramalho-Santos et al, 2002). In contrast, the stemness gene set from the study with ESCs, NSCs, and retina progenitor cells (RPCs) separated CD34+ hair follicle bulge cells from CD34- cells (Fortunel et al., 2003).

The SRPA gene set can distinguish human CD34+ cells from unfractionated human bone marrow cells.

The feasibility of using the SRPA gene expression signature to evaluate human stem cell-enriched populations was tested. Recently, the transcriptional profiling studies for normal human tissues were published with Affymetrix GeneChip® human U133A arrays (Su et al., 2004; Ge et al., 2005). Three hundreds and twenty genes of the 543 SRPA genes were available for analysis. PCA using the 320 SRPA genes separated the cluster of bone marrow (BM) $CD34^+$ cells from total bone marrow samples (Figure 23). The $CD34^+$ cell fraction is highly enriched for human HSCs (Hess et al., 2006). Two transctiptional profiling studies for human samples were publicly available (Su et al., 2004; Ge et al., 2005). Although it is too early to make a clear statement, the SRPA gene set may contain the genetic feature for ASC self-renewal in human. Further studies and

analyses should be followed to find the common gene signature that may be evolutionary conserved among diverse species for which ASC self-renewal pattern regulation is critical for survival.

<u>Discussion and Conclusion</u>:

The 543 SRPA gene expression signature identified from microarray studies with cell culture models for self-renewal pattern successfully classified specific ASC-enriched populations by PCA. Compared to previously described stemness gene signatures (Ivanova et al, 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003), the SRPA gene signature has more power to distinguish molecular features of the epidermal LRCs and CD34$^+$ cells in mouse hair follicle bulge region. These two cellular phenotypes are highly associated with hair follicle stem cell activity. Therefore, the PCA results using the SRPA gene signature provides more evidence that knowledge of ASC self-renewal pattern regulation is essential to understanding the function of ASCs.

Several studies have supported the functional correlation of a few genes in the stemness gene sets (Mitsui et al., 2003; Etheridge et al., 2004). However, the PCA results showed that their significance as a genetic signature for analyzing adult stem cells may be limited (Figure 21 and 22). In addition, the stemness gene set identified only from symmetrically self-renewing stem cell populations was relatively better for epidermal stem cell clustering compared to two other stemness studies (Figure 22g). The difference for the Fortunel et al signature was the substitution of RPCs in palce of HSCs. Considering that RPCs were cultured like NSCs, the stem cell-enriched populations used

in the Fortunel et al. study would predominantly undergo symmetric self-renewal (Fortunel et al., 2003). Fortunel et al. identified stemness genes by comparing the gene expression of these actively cycling stem cell-enriched populations to adult brain tissues adjacent to the lateral ventricles where most of cells are expected to be quiescent though some may divide with asymmetric cell kinetics (Fortunel et al., 2003). Thus, the 370 stemness genes identified by Fortunel et al. may be more homogenous in terms of self-renewal pattern compared to those by Ivanova et al. and Ramalho-Santos et al. Difference in self-renewal pattern reflected in stemness gene signatures may contribute to the clustering patterns by PCA (Figure 22e-g).

This study also showed that gene signatures are powerful to evaluate and find sample variability. Two stemness gene signatures failed to distinguish CD34$^+$ epidermal hair follicle bulge cells from CD34- cells. In the statistical analysis of gene sets identified from the Affymetrix array applied to the Agilent platform, there were two clusters in PCA using the stemness gene set by the Ivanova et al. (Figure 22e). One cluster located in the upper half of the graph consisted of the four arrays, all of which were originated from the one comparing set of the triplicate experiment. In the Agilent platform, compared biological samples were labeled with two different fluorescent dyes and competed for binding a probe, so-called two color assay. In the two color assay, there is dye-labeling bias, usually caused by differences in fluorescent photochemistry and labeling efficiency (Quackenbush, 2002; Yang et al., 2002). The dye-labeling bias is corrected by using another microarray hybridized with duplicate reciprocal dye-labeled samples, called a dye-swapping array. The stemness gene sets by Ivanova et al. and Ramalho-Santos et al.

clustered these dye-swapped arrays together. Therefore, PCA using the stemness gene signatures by Ivanova et al. and Ramalho-Santos et al. may have generated clusters representing microarray features instead of biological features. This might occur for signatures that contain a significant number of highly expressed genes in two color assay. These results support the explanation that the few overlaps among very similar transcriptional profilings for stemness may be due in part to the inadvertent comparison of ASCs and ESCs, which differ in self-renewal pattern.

Q-RT-PCR result showed that some SRPA genes are false positives. In addition, their protein expression and cellular activities may not be directly proportioned to changes in their transcriptional level. Nonetheless they are effective at distinguish ASC populations by virtue of their asymmetric self-renewal pattern.

For cell replacement therapy, quality evaluation of expanded cells or genetically engineered cells may be indispensable before application to patients. A gene expression signature for a specific purpose and its application to proper analysis like PCA may provide a tool to monitor the safety and efficacy of therapeutic cell populations.

Chapter 6. ASRA and SSRA Gene Clusters Associated with Chromosome Rearrangements in Human Cancers

Rationale:

Recently, the concept of cancer stem cells have been well supported by the serial transplantation experiments of malignant cells in human blood, breast, and brain cancers (Pardal et al., 2003; Huntley and Gilliland, 2005). In the nonobese diabetic-severe combined immunodeficiency (NOD/SCID) mice repopulation experiments, the self-renewal and hierarchical progeny lineages of cancer cells were found to be similar to those of normal ASCs. Both ASCs and cancer stem cells self-renew and give rise to differentiated progeny cells. However, they differ in the property of self-renewal kinetics pattern. ASCs have the unique property of asymmetric self-renewal and give rise to differentiated progeny cells specific for a particular tissue (Sherley, 2002; Cairns, 2002). In contrast, the self-renewal pattern of cancer stem cells is still unclear, and non-stem cell progeny of cancer stem cells may have extensive proliferating potential (Reya et al., 2001; Pardal et al., 2003; Huntley and Gilliland, 2005).

Many human cancers may originate from mutated stem cells in adult tissues (Cairns, 1975 and 2002). However, repopulation studies with human malignant cells in NOD/SCID mice have not provided any direct information to understand the relationship between stem cells in adult tissues and cancer stem cells. Even in the human hematopoietic system, in which both normal ASCs and cancer stem cells have been better studied than in any other tissue, no specific hematopoietic cancer has been related to the hematopoietic stem cell (HSC). Many human blood malignancies are attributed to

differentiated lineages from HSCs, but not the HSC itself (Sandberg, 1981). In spite of the similarities between ASCs and cancer stem cells in NOD/SCID mice-repopulation studies, the characteristics of the self-renewal pattern of the two types of stem cells are predicted to differ.

Study of the regulation of normal ASC self-renewal may provide key clues to understanding the self-renewal pattern of cancer stem cells and carcinogenic mechanisms of human cancers. In addition, unique molecular markers for ASCs and cancer stem cells may be identified by studying asymmetric self-renewal. Unique markers to detect cancer stem cells *in vivo* may help in the quest to discover ideal targets for future cancer therapy. Some of the SRPA genes are expected to have implications in cancer because the disruption of self-renewal pattern regulation may be directly or indirectly associated with carcinogenesis. In the microarrays, many SRPA genes showed association with human cancer (Chapter 2 and 3). For instance, the chromosome loci of ASRA genes like *hmga2*, *btg1*, and *Notch1*, rearranged in human cancers (Rodier et al., 2001; Hisaoka et al., 2002; Ellison et al., 1991). Some SSRA genes like *tra1* and *birc5* are up-regulated in human cancer tissues (Claudio et al., 2002; Altier, 2003). Pleoitrophin was reported to be involved in carcinogenesis and neural development (Kadomatsu and Muramatsu, 2004).

In this study, chromosome mapping of the 543 SRPA genes was performd. Two SRPA gene clusters were found, having significant implications for human cancers. The expression of SRPA genes in the two significant clusters was analyzed *in silico* in human leukemias with public transcriptional profiling databases.

116

Materials and Methods:

Chromosome mapping of the 543 SRPA genes

Chromosome mapping of the 543 SRSA genes was performed by using a gene proximity analysis in the dChip software (www. dChip. org). In mapping analysis, p values were calculated for less than 20 gene stretches as the measure of the proximity of a stretch of n genes against that of n genes randomly put on the chromosome. Thus, the significance of the gene stretch was decided by how much the genes in a stretch on a chromosome differ from the random occurrence probability. With the 543 SRPA genes, 6140 p value assessments were performed by the dChip software.

Cancer expression analysis in silico

For comparative expression analyses for the SRPA genes, the transcriptional profiling study on pediatric acute lymphoblastic leukemia (ALL) were evaluated (Yeoh et al., 2002; Ross et al., 2003). The Affymetrix raw data for 132 pediatric ALL patients were available from the St. Jude Children's Research Hospital (www. stjuderesearch. org / data / ALL3) (Ross et al., 2003). In order to find corresponding human genes for the 543 SRPA genes, the comparison spreadsheets for human and mouse ortholog targets available from the Affymetrix NetAffx[TM] Analysis Center (www. affymetrix. com/ analysis/ index.affx) were used. All Affymetrix cel files were analyzed and normalized with the dChip software to the median intensity using the PM-only model.

<u>**Results**</u>:

Chromosome mapping of the 543 SRPA genes

Figure 24 shows the locations of the 543 SRPA genes on mouse chromosomes (Figure 24a). The 16 gene cluster on mouse chromosome 18 was highly significant (p < 10^{-9}). The second most significant cluster was found on chromosome 2 (p < 10^{-8}). The next two significant gene clusters on chromosome 14 and 17 had 100-fold lower statistical confidence (p < 10^{-6}) (Figure 24b). Interestingly, the 16-gene cluster on chromosome 18 consisted of 15 SSRA genes and 1 ASRA gene. This cluster was referred to as the SSRA chromosome cluster (Table 11). In contrast, the gene cluster on chromosome 2 was comprised of 15 ASRA genes and 1 SSRA gene (p < 10^{-8}). Two additional ASRA genes were closely adjacent to this region (p < 10^{-7}). This 18 gene cluster was referred to as the ASRA chromosome cluster (Table 12). When analyzed by PCA, the 34 SRPA gene chromosome signature clearly distinguished asymmetric self-renewing cells from symmetrically self-renewing cells (Figure 24c).

The synteny between human and mouse for the two SRPA chromosome clusters

The dominant gene expression by one type of the SRPA genes in the specific regions of chromosomes suggested that these chromosome locations might be essential for the regulation of self-renewal pattern. The two SRPA chromosome cluster regions were evaluated for association with human cancers by comparing synteny between human and mouse chromosomes.

a) Mouse chromosomes

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
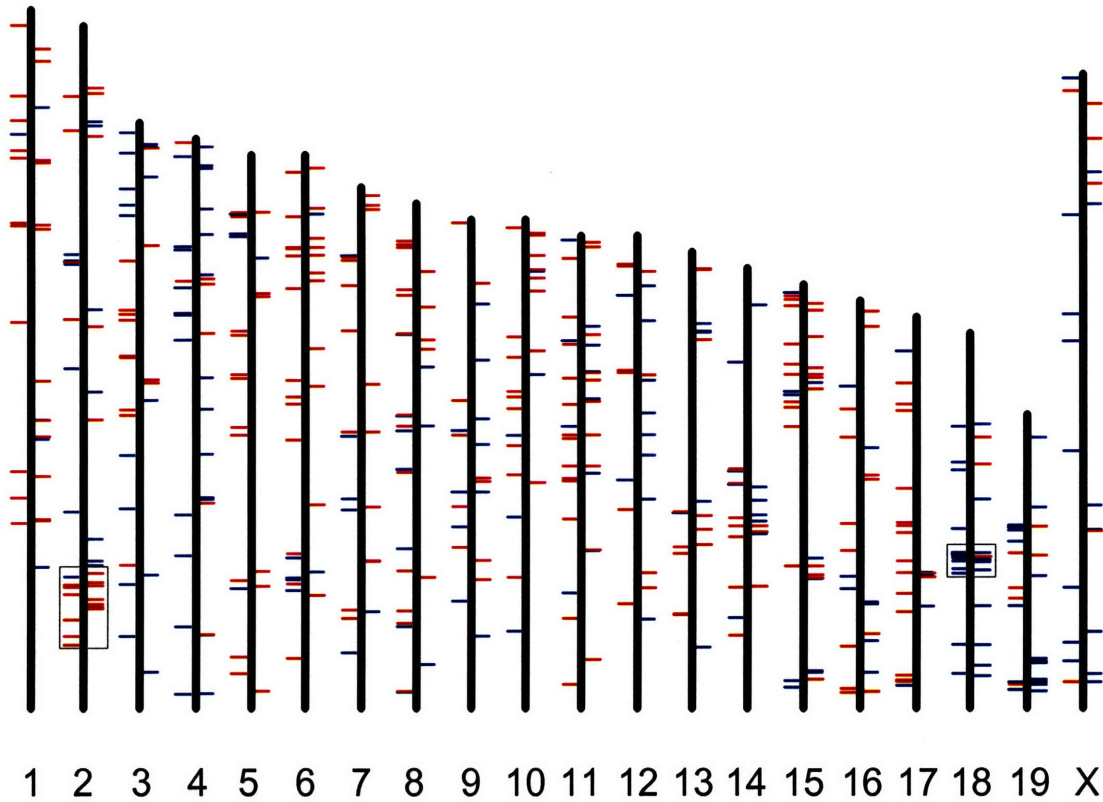
b)

p value	Number of significant gene stretches (Total 6140 p value assessments)
$<10^{-9}$	1 (Ch. 18)
$<10^{-8}$	2 (Ch.18, 2)
$<10^{-7}$	2 (Ch. 18, 2)
$<10^{-6}$	4 (Ch. 18, 2, 14, 17)
$<10^{-5}$	5 (Ch. 18, 2, 14, 17, 1)
$<10^{-4}$	8 (Ch. 18, 2, 14, 17, 1, 5, 6, 19)

c)

PC2 / PC1

● p53-induced Ind-8, Asymmetry
○ p53-null Con-3, Symmetry
□ p53-induced/IMPDH tl-3, Symmetry
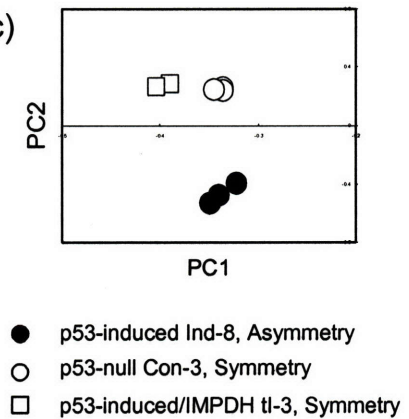
Figure 24. Mapping of the 543 SRPA genes on mouse chromosomes. a) Gene mappling was performed with dChip. The horizontal bars represent genes as the transcription start site for gene position. The genes in the left are on the forward strand and the genes on the right are on the reverse strand. The boxes on the chromosome 2 and 18 located the two significant SRPA gene stretches at the level of $p < 10^{-8}$. Red horizontal bars represent ASRA genes and blue bars are SSRA genes. b) Number of significant gene clusters by degree of statistical confidence. c) PCA using 34 genes in the ASRA and SSRA chromosome clusters.

Affymetrix ID	Gene Title	Gene Symbol	SRSA
1416110_at	solute carrier family 35, member A4	Slc35a4	SSRA
1417268_at	CD14 antigen	Cd14	SSRA
1419158_a_at	histidyl-tRNA synthetase-like	Harsl	SSRA
1420024_s_at	eukaryotic translation termination factor 1	Etf1	SSRA
1420628_at	purine rich element binding protein A	Pura/6330411E07Rik	SSRA
1422252_a_at	cell division cycle 25 homolog C (S. cerevisiae)	Cdc25c	SSRA
1423874_at	WD repeat domain 33	Wdr33	SSRA
1423878_at	glycophorin C	Gypc	SSRA
1427192_a_at	bromodomain containing 8	Brd8	SSRA
1431235_at	RIKEN cDNA 1110061A14 gene	1110061A14Rik	SSRA
1431274_a_at	heat shock protein 9A	Hspa9a	SSRA
1437451_at	RIKEN cDNA 1110006O17 gene	1110006O17Rik	*ASRA*
1441272_at	Matrin 3	Matr3	SSRA
1449207_a_at	kinesin family member 20A	Kif20a	SSRA
1451087_at	WD repeat domain 36	Wdr36	SSRA
1460549_a_at	CDC23 (cell division cycle 23, yeast, homolog)	Cdc23	SSRA

Table 11. The 16 genes in the SSRA cluster on mouse chromosome 18

Affymetrix ID	Gene Title	Gene Symbol	SRSA
1417153_at	BTB (POZ) domain containing 14A	Btbd14a	ASRA
1417271_a_at	endoglin	Eng	ASRA
1418483_a_at	glycoprotein galactosyltransferase alpha 1, 3	Ggta1	ASRA
1418633_at	Notch gene homolog 1 (Drosophila)	Notch1	ASRA
1418912_at	plexin domain containing 2	Plxdc2	ASRA
1422184_a_at	adenylate kinase 1	Ak1	ASRA
1428236_at	acyl-Coenzyme A binding domain containing 5	Acbd5	ASRA
1428705_at	RIKEN cDNA 1700007K13 gene	1700007K13Rik	ASRA
1431347_at	RIKEN cDNA 5730407M17 gene	5730407M17Rik	*SSRA*
1433553_at	GTPase activating RANGAP domain-like 3	Garnl3	ASRA
1434015_at	solute carrier family 2 (facilitated glucose transporter), member 6	Slc2a6	ASRA
1435345_at	RIKEN cDNA 2600006K01 gene	2600006K01Rik	ASRA
1435469_at	quiescin Q6-like 1	Qscn6l1	ASRA
1439194_at	RIKEN cDNA C030048H21 gene	C030048H21Rik	ASRA
1448426_at	sarcosine dehydrogenase	Sardh	ASRA
1451287_s_at	RIKEN cDNA 2810003C17 gene	2810003C17Rik	ASRA
1454642_a_at	COMM domain containing 3	Commd3	ASRA
1456346_at	Dynamin 1, mRNA	Dnm1	ASRA

Table 12. The 18 genes in the ASRA cluster on mouse chromosome 2

The 5.1 mega base pair (Mbp) region for the SSRA cluster on mouse chromosome 18 was syntenic to three human chromosome bands, 2q14-q21, 5q22, and 5q31. These regions in human chromosomes contain tumor suppressor genes like *BIN1* (2q14-q21) and *APC* (5q22) (Sakamuro et al., 1988; Fearnhead et al., 2001). The 5q31 region is known as one of the most frequently deleted loci in human leukemia (Giagounidis et al., 2006). Chromosome translocation in 5q31 has also been reported in infant acute lymphoblastic leukemia (ALL) with ins(5;11)(q31;q13q23) (Taki et al., 1999). It is likely that the loss or mutation of SSRA chromosome cluster genes may play role in mechanisms of leukocyte carcinogenesis.

The ASRA gene chromosome cluster encompassing a 16 Mbp region on mouse chromosome 2 corresponded to human chromosome bands, 2q14-q21, 9q33, and 9q34. The human chromosome band 2q14-q21 region for the ASRA cluster was separated as 0.3 Mbp (113.5-113.8) and 0.9 Mbp (138.4-139.3) regions (separated by 4.6 Mbp) in the NCBI human genome sequence contigs, whereas the SSRA cluster for this region covered 0.8 Mbp (127.7-128.5) sequence. Most interestingly, the region of the ASRA gene cluster included synteny to human chromosome band 9q34, the most frequent chromosome rearrangement locus in chronic myeloid leukemia (CML) and acute lymphoblastic leukemia (ALL) (Sandberg, 1981; Calabretta and Perrotti, 2004). Also, *Notch1* in the ASRA cluster has been known to function in T cell acute lymphoblastic leukemia (T-ALL) as one of frequent chromosome translocation loci (Ellisen et al., 1991). Both the ASRA and SSRA clusters are associated with chromosome abnormalities in human ALL and some ASRA genes like *Notch1* are directly associated with human cancers.

Therefore, this association of the SRPA genes with human cancers supported the proposal that the SRPA genes may provide essential clues to cancer stem cells, origin and carcinogenic mechanisms, related to self-renewal pattern regulation.

In silico investigation for alterations in expression of SRPA chromosome cluter gene in human ALL

The expression profile of SRPA chromosome cluster genes were analyzed in transcriptional profiling studies for pediatric ALL patients (Yeoh et al., 2002; Ross et al., 2003). Pediatric ALL is a heterogeneous disease having different leukemia subtypes. The ALL subtypes can be classified based on their molecular phenotype, especially related to chromosome abnormalities, for instance, t(9;22)[*BCR-ABL*], t(1;19)[*E2A-PBX1*], t(12;21)[*TEL-AML1*], and rearrangements in the *MLL* gene, etc. (Yeoh et al., 2002; Ross et al., 2003). The raw data files for ALL cells were compared with those for normal, healthy human tissues (Su et al., 2004; Ge et al., 2005). In Affymetrix human UG-133A GeneChip® arrays, only 6 ASRA and 11 SSRA genes (17 total) from total 34 SRPA chromosome cluster genes were available for comparative data analysis. Gene expression *per se* was evaluated by counting the number of present or absent calls for each sample (Table 13a). Although the number of normal bone marrow (BM) cell samples was small, SSRA chromosome cluster genes showed similar detection frequencies between whole BM cells and CD34[+] BM cells which are enriched for HSCs. Only two SSRA chromosome cluster genes showed a different frequency of detection in ALL cells. CD14 was expressed less frequently, and WD repeat domain 33 was expressed more often. Two

ASRA chromosome cluster genes, endoglin and sarcosine dehydrogenase, were not detected in whole BM, but were detected in several types of ALL cells and $CD34^+$ BM cells.

Next the expression level of each gene in the ASRA and SSRA clusters was analyzed. Gene expression of each ALL sample was compared to that of normal whole BM and $CD34^+$ BM cells. Three ASRA genes, endoglin, adenylate kinase 1, and *Notch1*, and one SSRA gene, glycophorin C (4 out of 17 evaluable genes; 24%), showed significant changes in gene expression in one or more different types of ALL tumors (Figure 25a). Endoglin, also known as CD105, was significantly up-regulated in four different types including the BCR-ABL ALLs. In the *TEL-AML* subtype of ALL patients, expression of adenylate kinase 1, and *Notch1* genes were significantly increased. As reported, *Notch1* gene was highly up-regulated in T-ALL patients (Grabher et al., 2006). Interestingly, the expression of the SSRA gene, glycophorin C, was down-regulated in several types of ALL patients (Figure 25a).

In addition to SRPA chromosome cluster genes, microarray expression of two human HSC markers, CD34 and CD133, was evaluated to compare their expression with that of SRPA genes whether there was a similar pattern of expression (Figure 25b). Athough CD34 and CD133 expression level was relatively higher in some of ALL subtypes than in normal BM cells, the change only in the up-regulation of CD133 in the MLL type of ALL was statitstically significant. This may be due to huge variabilities existed among patients for each type of ALL existed (Figure 25b)

| | | Acute Lymphoblastic Leukemia | | | | | | | | |
| | | | | | | | Karyotype | | | |
	BM_CD34+	BM	T_ALL	BCR-ABL	E2A-PBX1	TEL-AML	MLL	Normal	Hyper	Hypo	Pseudo
ASRA gene cluster					**Fractions of Present Calls**						
Adenylate kinase 1	1	1	0.5	0.9	1	1	0.9	0.9	1	1	1
Dynamin 1	0.5	0	0.1	0.1	0	0.1	0.4	0.1	0	0	0.3
Endoglin (Osler-Rendu-Weber syndrome 1)	1	0	0.1	1	0.2	0.9	0.7	0.7	0.9	1	1
Notch1	1	0.3	1	1	0.7	1	1	0.9	0.8	1	1
Sarcosine dehydrogenase	0.5	0	0.4	0.4	0.2	0.1	0.2	0.3	0.3	0.3	0.3
Solute carrier family 2, member 6	0	0	0.1	0.1	0	0	0.1	0.1	0	0	0
SSRA gene cluster					**Fractions of Present Calls**						
Bromodomain containing 8	1	1	1	1	1	1	1	1	1	1	1
CD14 antigen	0.5	1	0.2	0.3	0	0.1	0.2	0	0.1	0.3	0.2
CDC23 (cell division cycle 23, yeast, homolog)	1	1	1	1	1	1	1	1	1	1	1
Cell division cycle 25C	0.5	0.3	0.6	0.5	0.8	0.4	0.6	0.7	0.6	0.5	0.3
Eukaryotic translation termination factor 1	1	1	1	1	1	1	1	1	1	1	1
Glycophorin C (Gerbich blood group)	1	1	1	1	0.9	1	0.9	1	1	1	1
Histidyl-tRNA synthetase-like	1	0.3	0.9	1	0.9	1	1	0.9	1	1	0.9
Kinesin family member 20A	1	1	0.9	0.6	0.8	0.9	0.8	0.6	0.6	0.8	1
Matrin 3	1	1	1	1	1	1	1	1	1	1	1
Purine-rich element binding protein A	1	1	1	1	1	1	1	1	1	1	1
WD repeat domain 33	0	0	0.6	0.1	0	0.4	0.3	0.4	0.3	0.3	0.2
Number of Samples	2	3	14	15	18	20	20	10	20	4	11

Table 13. Frequency of present calls for ALL samples and normal bone marrow (BM) samples. The present/absent calls for samples were acquired by using the dChip software. a) Fractions of present calls. BM: healthy bone marrow samples (Su et al., 2004: Ge et al., 2005); BM_CD34+: healthy BM CD34+ cells (Su et al., 2004); T-ALL: T-cell lineage ALL; BCR-ABL, E2A-PBX1, and TEL-AML1 are B-cell lineage ALL subtype t(9;22)(*BCR-ABL*), t(1;19)(*E2A-PBX1*), t(12;21)(*TEL-AML1*), respectively; MLL: B-cell lineage ALL subtype with rearrangement in the *MLL* gene on chromosome 11q23, and B-cell lineage ALL with hyperdiploid, hypodiploid, or normal karyotype (Ross et al., 2003).
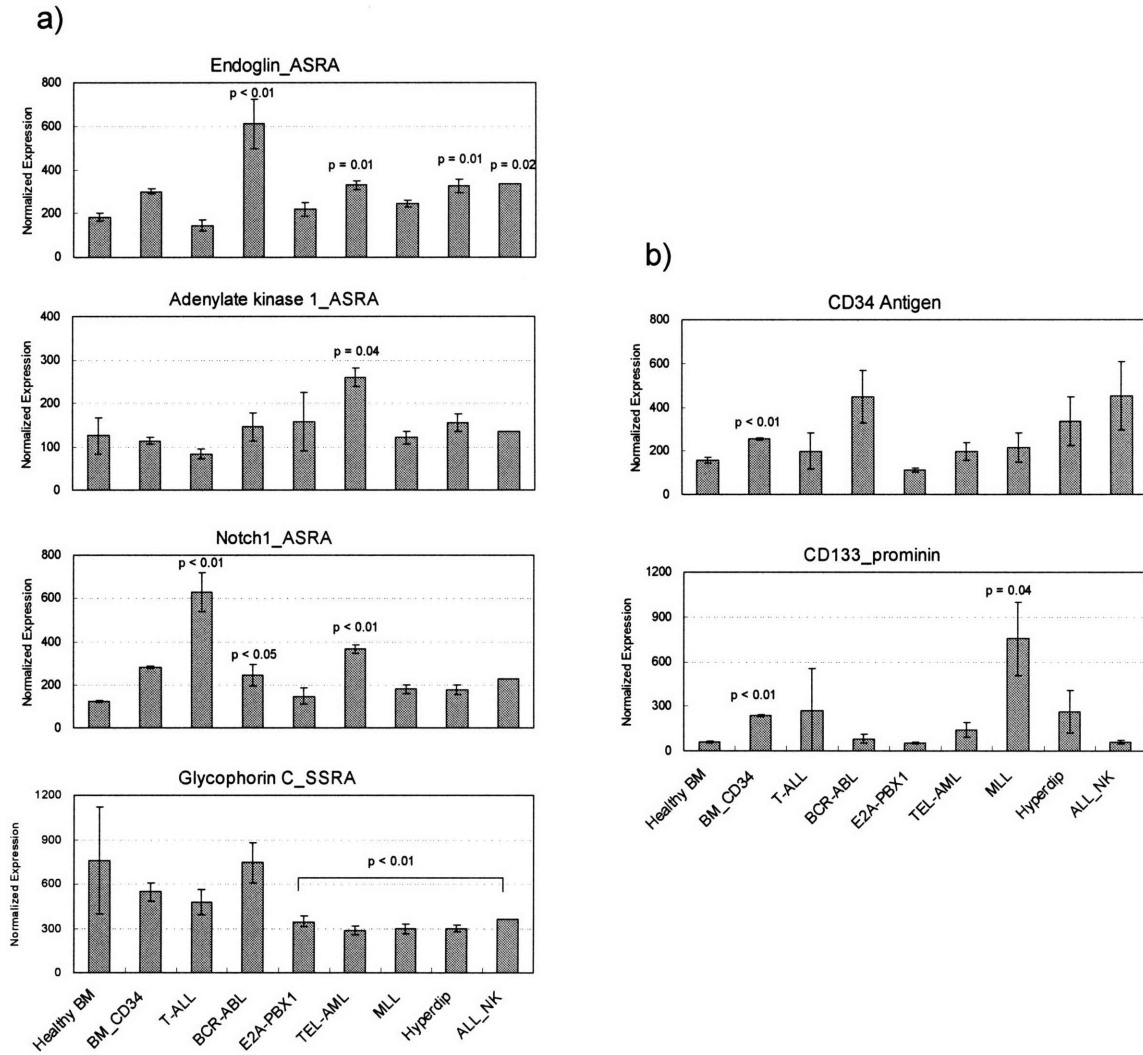
124

Figure 25. Microarray expression of the SRPA chromosome cluster genes in ALL cells. Gene expression was detected by using the PM-only model of the dChip software. Student t-tests were performed between normal BM samples and each kind of ALL sample. BM: healthy bone marrow samples; T-ALL: T-cell lineage ALL; BCR-ABL, E2A-PBX1, and TEL-AML1 are B-cell lineage ALL subtype t(9;22)(*BCR-ABL*), t(1;19)(*E2A-PBX1*), t(12;21)(*TEL-AML1*), respectively; MLL: B-cell lineage ALL subtype with rearrangement in the *MLL* gene on chromosome 11q23; Hyperdip: B-cell lineage ALL having > 50 chromosomes; ALL_NK: acute lymphoblastic leukemia with normal karyotype. a) Expression of SRPA genes in ALL cells. b) Microarray expression of two human HSC-enrichment markers, CD34 and CD133, in ALL cells

125

Discussion and Conclusion:

The ASRA gene cluster on mouse chromosome 2 and the SSRA gene cluster on mouse chromosome 8 were found by the chromosome mapping of the 543 SRSA genes. Both 18 ASRA gene and 16 SSRA gene clusters are associated with chromosome abnormalities in human cancers (Taki et al., 1999; Calabretta and Perrotti, 2004; Giagounidis et al., 2006). The close proximity of one type of a SRPA gene in the specific regions and their association with human cancers implicate these regions as critical for regulation of the self-renewal pattern of ASCs and cancer stem cells. In analyses of the transcript levels of SRPA chromosome cluster genes in the different type of ALL cells (Ross et al., 2003), three ASRA genes, endoglin, *Notch1*, and adenylate kinase 1, in the ASRA chromosome cluster were significantly up-regulated in the cancers of several types of ALL patients. In particular, endoglin was highly up-regulated in ALL with the t(9;22)(*BCR-ABL*) translocation (Figure 25).

In addition to ALL, endoglin is highly detectable in the microvasculature of solid tumors and has been reported for its association with tumor angiogenesis (Cho et al., 2001; Fonsatti et al., 2003). Endoglin also exists as a soluble form in circulation of cancer patients (Takahashi et al., 2001; Calabro et al., 2003). In normal tissues, endoglin has been detected mainly in microvascular endothelium and also in hematopoietic progenitor cells (Cho et al., 2001; Fonsatti et al., 2003). Endoglin was reported to play a role in the signal transduction of multiple members of transforming growth factor-β receptor superfamily. The complex association with various membrane receptors suggested that endoglin may function differently depending on cellular and tissue contexts (Barbara et al.,

126

1999). In the microarray study, endoglin transcript was increased during inducing asymmetric self-renewal. It is still unclear how endoglin overexpression is functionally related to the carcinogenic mechanisms of human cancers. The detection of this chromosome cluste gene in ALL cells may signify and origin from HSCs that self-renew asymmetrically. If so, endoglin might not play a direct role in regulating asymmetric self-renewal pattern, but instead co-vary in cells with this potential. It is also of interest whether endoglin up-regulation in the BCR-ABL subtype of ALL patients is directly associated with the chromosome aberration at t(9;22)(*BCR-ABL*).

Notch1 gene transcripts were increased in both TEL-AML subtype of ALL and T-ALL compared with that of normal whole BM samples (Figure 25a). The ASRA gene, *Notch1*, has been suggested as major carcinogenic factor for T-ALL and cervical cancer (Ellisen et al., 1991; Zagouras et al., 1995). However, there was a report that *Notch1* may function as a tumor suppressor in mouse skin (Nicholas et al., 2003). Considering that the *Notch1* gene is involved in different functions in embryogenesis and hematopoiesis, the role of *Notch1* may be oncogenic or tumor suppressive depending on developmental and tissue context (Stier et al., 2002; Grabher et al., 2006). *Notch1* was initially identified as an exclusively expressed gene in the asymmetrically self-renewing cells, and its expression was verified with both Q-RT-PCR and Western-blotting (Chapter 4). The association of *Notch1* with asymmetric self-renewal of stem cells was not recognized previously. In order to understand the contradictory association of the Notch1 gene in carcinogenesis and to elucidate its possible role in adult stem cell self-renewal, its

127

molecular function should be investigated in the context of the asymmetric self-renewal of ASCs and cancer stem cells.

Interestingly, 10~15 percent of chronic myelogeneous leukemia (CML) patients have deletions of the derivative chromosome 9 (der(9)) with the loss of BCR-ABL fusion gene. These cancers progress more rapidly to acute blast crisis (Sinclair et al., 2000; Huntley et al., 2003). The ASRA genes in the gene cluster on mouse chromosome 2 (human 9q34) may play crucial roles in regulating asymmetric self-renewal of ASCs. Their loss may lead to more rapid blast crisis in CML. Deletion of any region of the ASRA gene cluster at the 9q34 region in humans may affect the self-renewal of ASCs and cancer stem cells. The ASRA gene expression in the 16 Mbp cluster may explain why CML patients with der(9) deletion have a more severe clinical outcome and a shorter survival. In the preliminary analysis of 11 evaluable SRPA chromosome cluster genes, including endoglin, no significant change was observed in CML cells (Crossman et al., 2005). However, in this analysis, only 6 CML samples from a total of 28 samples were called as 'present' for endoglin, whereas endoglin in several subtypes of ALL was almost alwasys present (See Table 13; Ross et al., 2003; Crossman et al., 2005). Based on the earlier reasoning that endoglin gene expression in human hematopoietic tumors signifies an HSC origin, this finding suggests that CML tumors arise from a more differentiated cell without asymmetric self-renewal properties.

In conclusion, the ASRA cluster and SSRA cluster identified from the chromosome mapping of the 543 SRPA genes may be critical in the molecular pathogenesis of human cancers. The ASRA gene endoglin was up-regulated in the BCR-

ABL subtype of ALL. However, it is unknown whether chromosome translocation t(9;22)(*BCR-ABL*) affects its transcription. Although no SSRA genes have been found that may be strongly involved in ALL, there is still a chance that the genes in the SSRA cluster region on mouse chromosome 18 may be involved in carcinogenesis of other human cancers. In addition to the 34 genes in the ASRA and SSRA chromosome cluster regions on chromosomes 2 and 18, other SRPA genes may be important for the self-renewal of cancer cells and for attempts to identify cancer stem cells. Therefore, by studying the SRPA genes, novel molecular factors regulating asymmetric self-renewal of ASCs and pathological mechanisms related to the disruption of self-renewal symmetry control of ASCs in human cancers may be elucidated in the future. Furthermore, unique molecular markers to identify ASCs and/or cancer stem cells may be discovered by studying self-renewal pattern regulation *per se*.

Chapter 7. Evaluation of Asymmetric Self-Renewal Associated Protein Modulation for Finding New Approaches for the Suppression of Asymmetric Cell Kinetics

Rationale:

Genes responsible for the regulation of asymmetric self-renewal by ASCs have many potential applications such as molecular markers for ASC identification and isolation, diagnostic and therapeutic targets for stem cell-related human cancers, and engineering tools for ASC expansion (Merok and Sherley, 2001; Sherley, 2002). However, mRNA expression alone is not enough to predict rate determining gene function in ASCs because protein expression is not always correlated with changes in mRNA expression (Gygi et al., 1999; Washburn et al., 2003). Importantly, changes in gene expression may also be the result of secondary responses to the primary transcriptional changes (Kannan et al., 2001).

As an ASC application, the concept of the suppression of asymmetric cell kinetics (SACK) was recently introduced for expanding functional ASCs (Sherley et al., 2002). The first described SACK trials were based on p53 regulated cellular metabolic changes (Sherley, 2002). The p53 protein controls cellular guanine ribonucleotide level by regulating IMPDH gene expression (Liu et al., 1998a; Rambhatla et al., 2005). In the first SACK studies, the guanine ribonucleoside xanthosine (Xs) was supplemented in cell culture media to overcome IMPDH down-regulation. Using Xs as a SACK agent, rat liver hepatic adult stem cell strains were successfully expanded and propagated (Lee et al., 2003).

Several SRPA genes have been validated by Q-RT-PCR and Western blotting. The validated SRPA genes are candidate targets for modulation of self-renewal pattern. Proliferin, the autocrine or paracrine growth factor, is a good example. Proliferin is a member of the prolactin-growth hormone gene family. Proliferin in the adult mouse is expressed in placenta, hair follicles, and small intestine (Fassett and Nilsen-Hamilton, 2001; Choong et al., 2003). Proliferin-2 was also shown to modestly expand mouse HSCs (Choong et al., 2003). However, no role for proliferin in self-renewal pattern regulation was recognized previously.

The function of another ASRA gene, *Kcnn4*, can be pharmacological modulated by using its antagonist, clotrimazole. The *Kcnn4* gene encodes the intermediate conductance Ca^{2+}-activated potassium channel (IKCa1). Mouse knockout study for *Kcnn4* gene showed that volume regulation of T-lymphocytes and red blood cells was mildly impaired (Begenisich et al., 2004). Except for those modest changes, the healthy phenotype of *Kcnn4* gene knock-out mice indicated that pharmacological blocking of the potassium channel IKCa1 may not be toxic in clinical applications (Begenisich et al., 2004). Therefore, it was interesting to investigate the potential for a functional connection between the potassium channel IKCa1 with self-renewal pattern regulation.

In this study, the blocking of proliferin with specific antibodies was evaluated as a potential SACK approach. Binding of secreted proliferin in culture medium by anti-proliferin antibody treatment was performed to see if it altered in self-renewal pattern regulation. In addition, the potassium channel IKCa1 was pharmacologically blocked by

clotrimazole. These analyses provided opportunities to find novel SACK approaches for expanding ASCs in culture.

Materials and Methods:

Cell growth and colony formation analyses

The culture conditions and maintenance of p53-inducible Ind-8 cells and p53-null Con-3 cells have been described (Liu et al., 1998b; Chapter 2). Cell growth analyses were performed as described in Chapter 2. Colony formation studies were initiated as cell growth curve analyses. Cells were grown over a 2~3 day period to about half confluency, trypsinized, and replated in zinc-free medium. For routine colony formation assay, three hundred cells were plated per each well of a 6 well-culture plate (300 cells/9.6 cm^2; Becton Dickinson, Franklin Lakes, NJ). Sixteen to twenty four hours after plating, cell culture media were changed into either Zn free medium or medium containing 45 μM ZnCl$_2$. After 12 days of incubation, cell colonies were stained with crystal violet and counted by gross inspection.

Acetone precipitation of culture supernatants for proliferin analysis

Cell culture media were harvested, and cell debris in culture supernatants were removed by centrifugation at 5000 rpm for 5 minutes in 4 °C with a table top centrifuge. The supernatants were mixed with the 5 x volumes of ice-cold acetone, and the mixtures were precipitated overnight at -20 °C. The mixtures were then centrifuged at 13,000 rpm

for 30 min at 4 °C. The supernatants were decanted and the precipitate was dried and re-dissolved in a buffer containing 1 % sodium dodecyl sulfate, 1 % Nonidet P-40, and 0.5 % sodium deoxycholate in phosphate buffered saline (PBS). Immunoblotting for proliferin was performed as described in Chapter 4.

Cell culture treatment

For testing effects of blocking proliferin in culture, anti-proliferin antibody (R&D Systems, Minneapolis, MN) was used. Anti-proliferin antibody solutions that were reconstituted with distilled water were sterilized with 0.2 μm Whatman syringe filters (VWR, Batavia, IL). Anti-proliferin antibody solutions were added to cell culture media when asymmetric self-renewal was induced. For colony formation assay, the same amounts of antibody solutions were added a second time to the same cell culture wells 6 days after the initial treatment. In order to evaluate the effect of *Kcnn4* antagonism on cell growth kinetics, clotrmazole (Sigma-Aldrich, St. Louis, MO) was used. Clotrimazole was dissolved in 70% ethanol as a 1000x stock solution. For cell growth and colony formation analyses, clotrimazole was added to cell culture media when asymmetric self-renewal was induced.

Results:

The effect of anti-proliferin antibodies on asymmetric self-renewal kinetics

As results of Western blotting, the production and secretion of proliferin into the cell culture supernatants was identified, especially from the asymmetrically self-renewing Ind-8 cells (Figure 26). With the same antibody used in Western blotting, the neutralization of proliferin was evaluated for possible effects on self-renewal kinetics. Up to 10 µg/ml of anti-proliferin antibody (the highest concentration practical), there was no change in cell growth kinetics. Next, the effect of anti-proliferin antibody was evaluated by colony formation assay. Colony formation assay has been used as a more sensitive analysis to evaluate the effect of test compounds on self-renewal kinetics (Sherley 1991; Sherley et al., 1995; Liu et al., 1998a). However, no significant change in colony formation efficiency was detected upon addition of anti-proliferin antibody.

There were several possible explanations why the anti-proliferin antibody had no effect on cell growth kinetics, including proliferin does not play a rate-determining role. Also, the amount of neutralizing antibody might not have been enough to block all secreted proliferin in cell culture supernatants. The anti-proliferin antibody might have bound to proliferin without affecting its function. Proliferin proteins were exclusively detected in all cell culture model systems dividing with asymmetric self-renewal. This feature makes the explanation that the antibody would not work well. Thus, it may be worthwhile to pursue alternative methods to effectively block proliferin function.
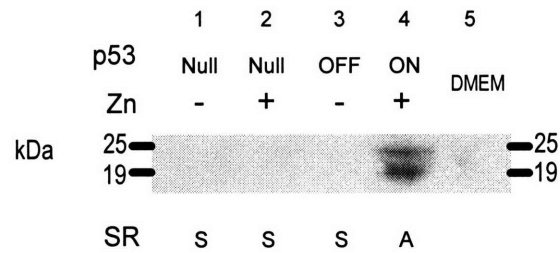
Figure 26. Evaluation of proliferin secretion with different patterns of self-renewal. Western blotting analyses for proliferin were performed with cell culture supernatants. Cell culture supernatants were concentrated 5- fold by acetone precipitation. Lanes 1 and 2, p53-null Con-3 cells; lanes 3 and 4, Zn-responsive p53-inducible Ind-8 cells; lanes 5, culture medium control for acetone precipitation. SR: self-renewal, S: symmetric self-renewal, A: asymmetric self-renewal.

Antagonism of the potassium channel IKCa1 by clotrimazole

Clotrimazole has been widely used to antagonize the potassium channel IKCa1 in the voltage clamping studies or to measure the cellular Ca^{2+}-induced fluorescence fluctuation in cells (Ghanshani et al., 2000; Elliott and Higgins, 2003). However, the effective concentration of clotrimazole has a wide range, from 20 nM to 10 μM for each experiement. In the presence of 65 μM $ZnCl_2$, clotrimazole showed no apparent cellular toxicity up to 10 μM for p53-inducible Ind-8 cells. In the cell growth analyses with conventional 25 cm^2-cell culture flasks, no significant change was observed in the doubling times or PDCs by clotrimazole treatment. However, consistently, more cells were presetn in cell cultures treated with less than 10 nM clotrimazole after 72 hours of culture compared to vehicle treated cells. In the cell cultures treated with more than 1 μM clotrimazole, cell counts were consistently lower than those of vehicle-treated cells. These observations were not statistically significant (data not shown).

In the colony formation analysis, statistically significant increase in colony formation were observed with 5 and 10 nM clotrimazole treatment of cells induced for asymmetric self-renewal (Figure 27). There was no change in colony formation observed for p53-null symmetrically self-renewing Con-3 cells. Therefore, the potassium channel IKCa1 may be involved in the regulation of self-renewal pattern in the p53 inducible Ind-8 cells.
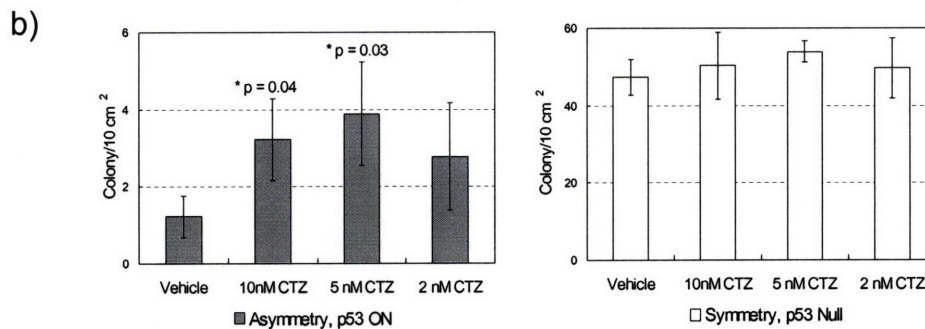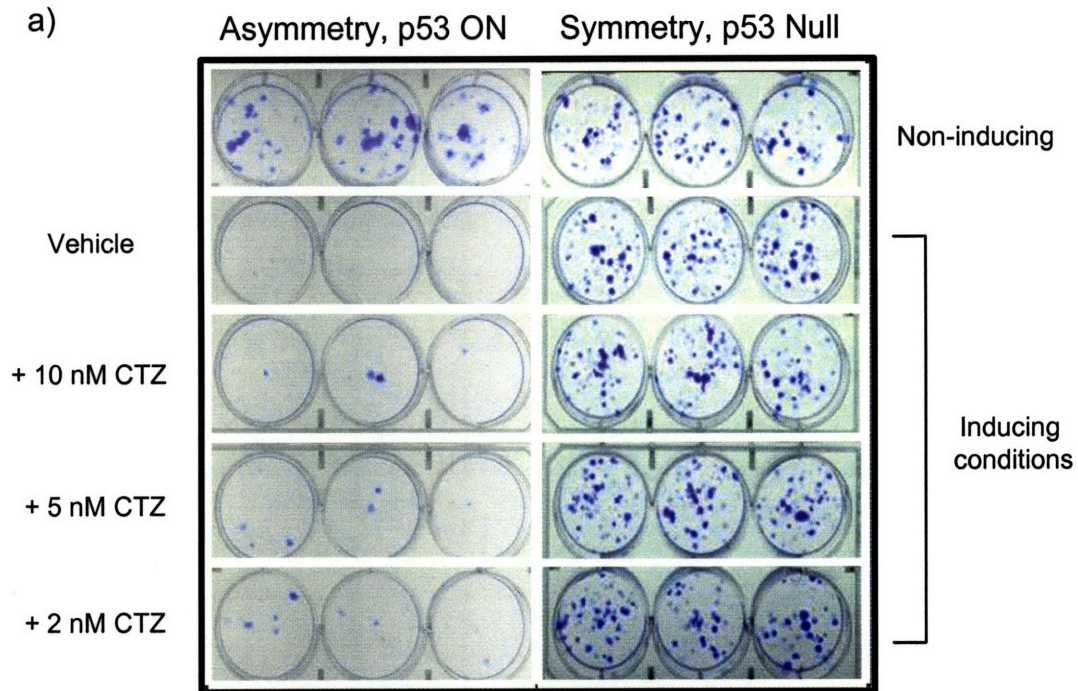
Figure 27. Effects of clotrimazole on colony formation by asymmetrically self-renewing cells. a) Colony formation analyses. Three hundred cells were plated per each well. Cell culture media were changed to either Zn free medium for symmetric self-renewal or medium containing 45 μM $ZnCl_2$ for asymmetric self-renewal. Clotrimazole (CTZ) was added with Zn-containing media. After 12 days of cell growth, cell colonies were stained with crystal violet and counted. b) The data from three independent experiments were averaged. Error bars = S.D. Un-paired t-tests were performed between vehicle-treated wells and CTZ-treated wells.

Discussion and Conclusion:

The purpose of this study was to investigate whether particular SRPA genes were functionally involved in self-renewal pattern regulation. These studies also had potential to reveal novel approaches to SACK expansion of ASCs. It has been proposed that when ASCs asymmetrically self-renew in culture, they become diluted among their non-stem progeny cells and lost on successive passages (Sherley, 2002; Lee et al., 2003). Asymmetric self-renewal is a barrier to expand ASCs. SACK shifts self-renewal pattern of ASCs to symmetric cell kinetics which promotes their exponential expansion in culture. Two ASRA proteins were evaluated as SACK targets for immunological or pharmacological agents. Clotrimazle increased colony formation by asymmetrically self-renewing cells indicating SACK activity, but no effect were observed from the anti-proliferin antibody treatment.

The ASRA gene, *Kcnn4*, was identified in the cDNA microarray studies (Chapter 2) and a similar expression pattern was observed in the Affymetrix whole genome analysis. Q-RT-PCR studies verified that its transcript was up-regulated in asymmetrically self-renewing Ind-8 cells (Table 9 in chapter 4). The potassium channel IKCa1, encoded by the *Kcnn4* gene, has been detected in various tissues including the hematopoietic system, pancreas, lung, and colon (Begenisich et al., 2004). The potassium channel IKCa1 has also been postulated to play a role in transduction of mitogenic signals. A Ca^{2+}-activated potassium channel, IKCa1 is up-regulated during T-cell activation by mitogens (Ghanshani et al., 2000). Interestingly, clotrimazole treatment made T-cells more resistant to cell shrinkage following apoptosis (Elliott and Higgins, 2003). The activated IKCa1

hyperpolarized cells like T-lymphocytes, and this change in plasma membrane potential modulated intracellular Ca^{2+} concentration to help signal activation of mitogenic cytokines (Fanger et al., 1999; Begenisich et al., 2004). Recently, *Kcnn4* gene was reported to be more expressed in proliferating smooth muscle cells (Cheong et al., 2005). The relationship between *Kcnn4* gene expression and cell proliferation is unclear at the molecular level. Although its expression is not exclusive for asymmetric self-renewal, the modulation of the potassium channel IKCa1 may be tested in the future as a novel SACK approach for expansion of ASCs.

Recently, genetic modification was applied to expand pancreatic β-cells (Narushima et al., 2005). Genetic engineering for adult stem cell expansion may not be applicalbe for all adult cells for clinical applications, because such genetic changes raise controversial safety issues (Asahara et al., 2000). The SACK method allows for reversible conversion of self-renewal pattern from asymmetric to symmetric and *vice versa* in cell culture (Sherley, 2002). In this sense, metabolic modulations are the primary levers used to expand ASCs. Among the SRPA genes, those encoding growth factors, plasma membrane receptors for cellular signaling, and enzymes in the metabolic pathways may be components of new SACK pathways (Table 14). Going forward, the microarray expression of these genes should be verified with appropriate orthogonal mRNA analyses. Verified ASRA genes should be targeted for down-regulation using specific neutralizing antibodies, pharmacological antagonists, and molecular regulators (e.g. siRNAs) as the specific molecular features dictate (Table 14a). In contrast, SSRA genes should be targeted for up-regulation to induce shifts to symmetric self-renewal (Table 14b). IMPDH,

139

which is up-regulated during symmetric self-renewal, is the prototype for such investigation.

In conclusion, pharmacological antagonism against the ASRA protein IKCa1, encoded by *Kcnn4*, increased the proliferation rate of cells undergoing asymmetric self-renewal, consistent with a shift to symmetric self-renewal. More exacting assays will be required to confirm this mechanism (Lee et al., 2003). These results show that the SRPA genes may be good candidates for developing new SACK methods for ASC expansion. The verification of the candidate SRPA gene and protein expression should be performed before the evaluation of their functional relevance with model cell systems for self-renewal pattern regulation. More importantly, new methods should be tested for whether they effectively expand ASCs in culture. Clotrimazole or other specific antagonists against the potassium channel IKCa1 may be the first new SACK agents for this purpose.

a)

ASRA genes	Gene ID	Affymetrix ID (Gene Symbol)	Gene Title
Cytokine/ growth factor	14017	1450241_a_at (Evi2a)	Ecotropic viral integration site 2a
	14182	1424050_s_at (Fgfr1)	Fibroblast growth factor receptor 1
	18295	1419662_at (Ogn)	Osteoglycin
	54635	1419123_a_at (Pdgfc)	Platelet-derived growth factor, C polypeptide
	72962	1429546_at (Ecgf1)	Endothelial cell growth factor 1 (platelet-derived)
Receptor/ Signal transduction	12931	1418476_at (Crlf1)	Cytokine receptor-like factor 1
	13805	1417271_a_at (Eng)	Endoglin
	13848	1418051_at (Ephb6)	Eph receptor B6
	18414	1418674_at (Osmr)	Oncostatin M receptor
	18595	1421917_at (Pdgfra)	Platelet derived growth factor receptor, alpha polypeptide
	21933	1421296_at (Tnfrsf10b)	Tumor necrosis factor receptor superfamily, member 10b
	22403	1419015_at (Wisp2)	WNT1 inducible signaling pathway protein 2
	56229	1418205_at (Thsd1)	Thrombospondin, type I, domain 1
	67448	1418912_at (Plxdc2)	Plexin domain containing 2
	69538	1451446_at (Antxr1)	Anthrax toxin receptor 1
	171095	1419671_a_at (Il17rc)	Interleukin 17 receptor C
	216795	1436978_at (Wnt9a)	Wingless-type MMTV integration site 9A
Metabolism	11636	1422184_a_at (Ak1)	Adenylate kinase 1
	11898	1416239_at (Ass1)	Argininosuccinate synthetase 1
	12409	1418509_at (Cbr2)	Carbonyl reductase 2
	13498	1421149_a_at (Atn1)	Atrophin 1
	13808	1417951_at (Eno3)	Enolase 3, beta muscle
	14594	1418483_a_at (Ggta1)	Glycoprotein galactosyltransferase alpha 1, 3
	14860	1416368_at (Gsta4)	Glutathione S-transferase, alpha 4
	29858	1424167_a_at (Pmm1)	Phosphomannomutase 1
	54447	1458849_at (Asah2)	N-acylsphingosine amidohydrolase 2
	56316	1423554_at (Ggcx)	Gamma-glutamyl carboxylase
	58250	1428902_at (Chst11)	Carbohydrate sulfotransferase 11
	67880	1419456_at (Dcxr)	Dicarbonyl L-xylulose reductase
	72043	1447602_x_at (Sulf2)	Sulfatase 2
	74159	1428236_at (Acbd5)	Acyl-Coenzyme A binding domain containing 5
	78070	1435281_at (Cpt1c)	Carnitine palmitoyltransferase 1c
	78752	1424431_at (MGI:1926002)	Chondroitin sulfate GalNAcT-2
	140481	1435203_at (Man2a2)	Mannosidase 2, alpha 2

b)

SSRA genes	Gene ID	Affymetrix ID (Gene Symbol)	Gene Title
Cytokine/ growth factor	16007	1442340_x_at (Cyr61)	Cysteine rich protein 61
	18591	1450413_at (Pdgfb)	Platelet derived growth factor, B polypeptide
	19242	1416211_a_at (Ptn)	Pleiotrophin
	20300	1458277_at (Ccl25)	Chemokine (C-C motif) ligand 25 (Ccl25), mRNA
	93670	1420458_at (Tac4)	Tachykinin 4
Receptor/ Signal transduction	13610	1460661_at (Edg3)	Endothelial differentiation, sphingolipid GPCR, 3
	13837	1425574_at (Epha3)	Eph receptor A3
	15366	1429871_at (Hmmr)	Hyaluronan mediated motility receptor (RHAMM)
	16164	1427165_at (Il13ra1)	Interleukin 13 receptor, alpha 1
Metabolism	11639	1450387_s_at (Ak3l1)	Adenylate kinase 3 alpha-like 1
	66307	1425050_at (Isoc1)	Isochorismatase domain containing 1
	94352	1431004_at (Loxl2)	Lysyl oxidase-like 2
	98388	1426620_at (Chst10)	Carbohydrate sulfotransferase 10
	229905	1455991_at (Ccbl2)	Cysteine conjugate-beta lyase 2
	268656	1436727_x_at (Sptlc1)	Serine palmitoyltransferase, long chain base subunit 1

Table 14. List of the SRPA genes predicted to encode potential targets for new SACK approaches. Cytokine, growth factor, membrane receptors, and enzymes in metabolic pathways were listed from the 543 SRPA genes. a) Genes exclusively or differentially expressed in asymmetric self-renewal. b) Genes up-regulated with symmetric self-renewal.

Chapter 8. Implications of Symmetric Self-Renewal Associated Genes in the Immortal DNA Strand Co-segregation of Adult Stem Cells

<u>Rationale</u>:

Recent experimental evidence supported the hypothesis that ASCs co-segregate all chromosomes having the oldest parental DNA strands to their stem cell daughters during asymmetric self-renewal (Sherely, 2002; Potten et al., 2002; Merok et al., 2002; Karpowicz et al., 2005). This unique non-random chromosome co-segregation may provide ASCs with the advantage of protecting themselves from DNA replication errors (Cairns, 1975 and 2002). The cellular mechanism responsible for the the non-random chromosome co-segregation of ASC is poorly understood at the molecular level.

The chromosome passenger complex containing survivin, Aurora B kinase, borealin, and INCENP regulates chromosome segregation during mitosis. Aurora B kinase regulates the kinetochore/microtubule interactions during mitosis (Maiato et al., 2004). Survivin, originally suggested as an anti-apoptotic protein, also actively modulates the microtubule organization during mitosis (Chen et al., 2003; Rosa et al., 2006). INCENP has multiple functional domains that may modulate the other passenger complex protein functions (Maiato et al., 2004). Borealin is a recently characterized member of the complex that may play a role in the correction of kinetochore attachment errors with the other chromosome passenger complex proteins (Gassmann et al., 2004; Sampath et al., 2004).

Among SRPA genes, two chromosome passenger complex proteins, survivin and inner centromere protein (INCENP), were up-regulated in the symmetrically self-

renewing cells. The protein expression of survivin was also sensitive to the self-renewal pattern and reversibly responded to the changes in the self-renewal pattern of dividing cells, indicative of the existence of a highly regulated mechanism (Chapter 4). The activity of chromosome passenger complex proteins is known to be synergistically activated by survivin (Chen et al., 2003; Lens et al., 2003). In budding yeast, the deletion mutants for the *Ipl1-Sli15* kinase complex resulted in segregation defects for which all chromosomes co-segregated with the old spindle pole body, a caricature of non-random chromosome co-segregation (Tanaka et al., 2002). The mammalian orthologues of the yeast kinase complex are Aurora B-kinase and INCENP. In mammalian cells, however, the effect of mutations or knock-outs of the chromosome passenger complex genes has not been fully evaluated, especially for association with non-random chromosome co-segregation mechanisms.

There was a great possibility that survivin and the other chromosome passenger complex proteins like Aurora B kinase and INCENP might be involved in the process of non-random chromosome co-segregation during asymmetric self-renewal. In this study, the functional associations of the SRPA genes related to the chromosome passenger complex were evaluated by using the model cell systems which also exhibit non-random chromosome co-segregation in association with asymmetric self-renewal (Merok et al., 2002; Rambhatla., 2005). The SRPA chromosome passenger complex protein expression and subcellular localization in the cells undergoing non-random chromosome co-segregation were compared with those in the symmetric counterparts by *in situ* immunofluorescent studies.

Materials and Methods:

Cell culture and fixation for in situ immuno-fluorescence microscopy

The culture conditions and maintenance of p53-inducible Ind-8 cells and p53-null Con-3 cells were described in Chapter 2. Cells were grown over a 2~3 day period to about half confluency, trypsinized, and replated in zinc-free medium. Cells were plated at 2000 cells/cm^2 in 4-well LAB-TEK$^®$ chamber slides (Nunc, Inc., Naperville, IL). After 24 hours of culture, cell culture media were replaced with either Zn free medium or medium containing 65 µM ZnCl$_2$ to induce asymmetric self-renewal and non-random chromosome co-segregation. Thirty six hours after the induction of asymmetric self-renewal in the p53-inducible Ind-8 cells, the media was aspirated form each well, and the cells were washed with ice cold PBS filtered with a 0.2 µm membrane filter (Millipore, Billerica, MA). Cells were immediately fixed in 3.7% formaldehyde in filtered PBS for 15 minutes at room temperature. After removing the 3.7% formaldehyde/PBS solution, 0.5% formaldehyde in filtered PBS was added at least 30 minutes at 4 °C. After fixation, cells were washed with filtered PBS at room temperature three times for 5 minutes each. Cells were blocked with 2% fatty acid free-bovine serum albumin (BSA) in filtered PBS solution (2 % BSA-PBS) at least two hours at 4 °C.

Immunostaining

The 2% BSA-PBS solution was removed and cells were permeabilized using 0.4% Triton X-100, 0.2 % non-fat dried skim milk, and 2% BSA in filtered PBS solution (blocking buffer) for 10 minutes at room temperature. Primary antibodies were diluted in the blocking buffer. Mouse monoclonal anti-survivin antibody (D-8, sc-17779) and anti-α-tubulin (B-7, sc-5286) were diluted at a concentration of 1:30 (Santa Cruz Inc., Santa Cruz, CA). Anti-Aurora A kinase (IAK1) and anti-Aurora B kinase (AIM-1) antibodies were used at a 1:50 dilution (BD Bioscience, Inc., San Jose, CA). All antibody solutions were centrifuged at 13,000 rpm for 5 minutes in order to remove interfering particulates. After overnight incubation with the fixed, washed, blocked cells in chamber slides at 4 °C, the primary antibody solution was removed and the cells were washed 5 times for 5 minutes each with 0.2 % Tween 20 in filtered PBS (washing buffer). Goat anti-mouse IgG-FITC conjugated antibody in the blocking buffer (1:200 dilution) was used as the secondary antibody (BD Bioscience, Inc., San Jose, CA). The secondary antibody was incubated for at least 1 hour at room temperature. The antibody treated cells were washed 5 times with the washing buffer and cell nuclei were counterstained with the 4'-6-diamido-2-phenylindole (DAPI) containing-VectaShield® mounting media (Vector Laboratories, Inc., Burlingame, CA). Immunofluorescent images were observed using a Zeiss Axioskop MOT microscope and images were captured with a Zeiss Axiocam camera. Images that were captured with different filters for the same field were merged using Adobe Photoshop software (Adobe, Inc. San Jose, CA).
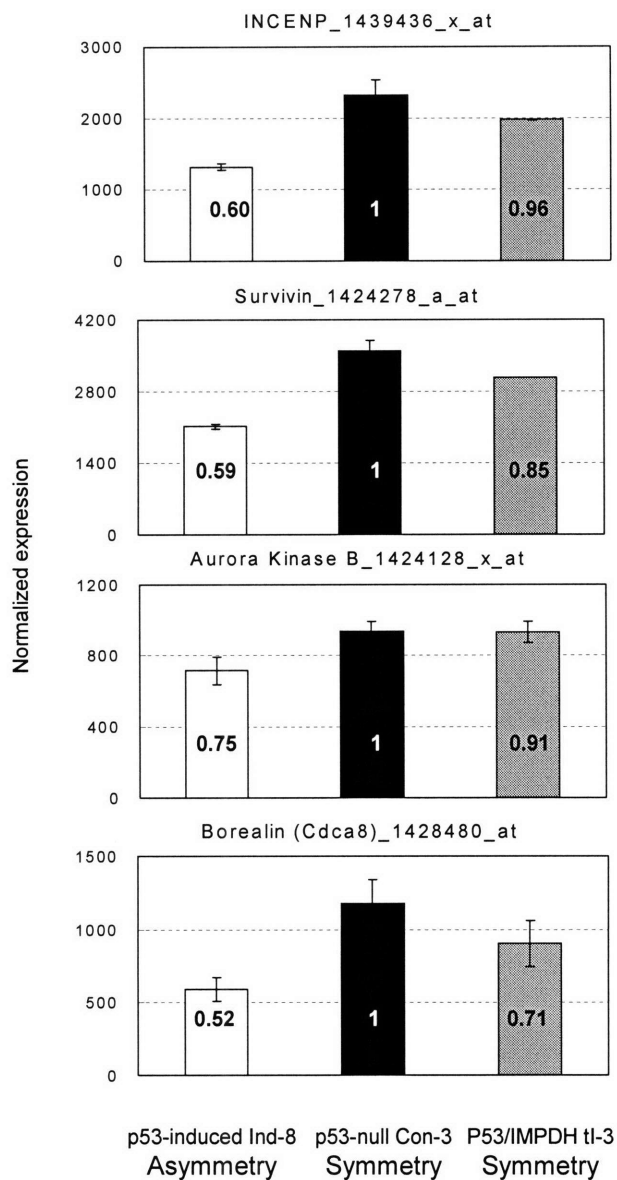
Figure 28. Microarray expression of the chromosome passenger complex genes with SSRA properties. Expression levels for each gene were normalized and averaged for replicaate oligonucleotide arrays. There were 3 probe sets for INCENP and Borealin (Cdca8), 2 probe sets for Aurora B kinase, and 1 probe set for survivin. All probe sets showed similar results for each gene. The graph for each gene was selected for one probe set. The numbers in the bar graphs represented the averaged relative expression level of all probe sets for each gene including those not used for graph after setting the expression values in the symmetrically self-renewing Con-3 cells as 1.
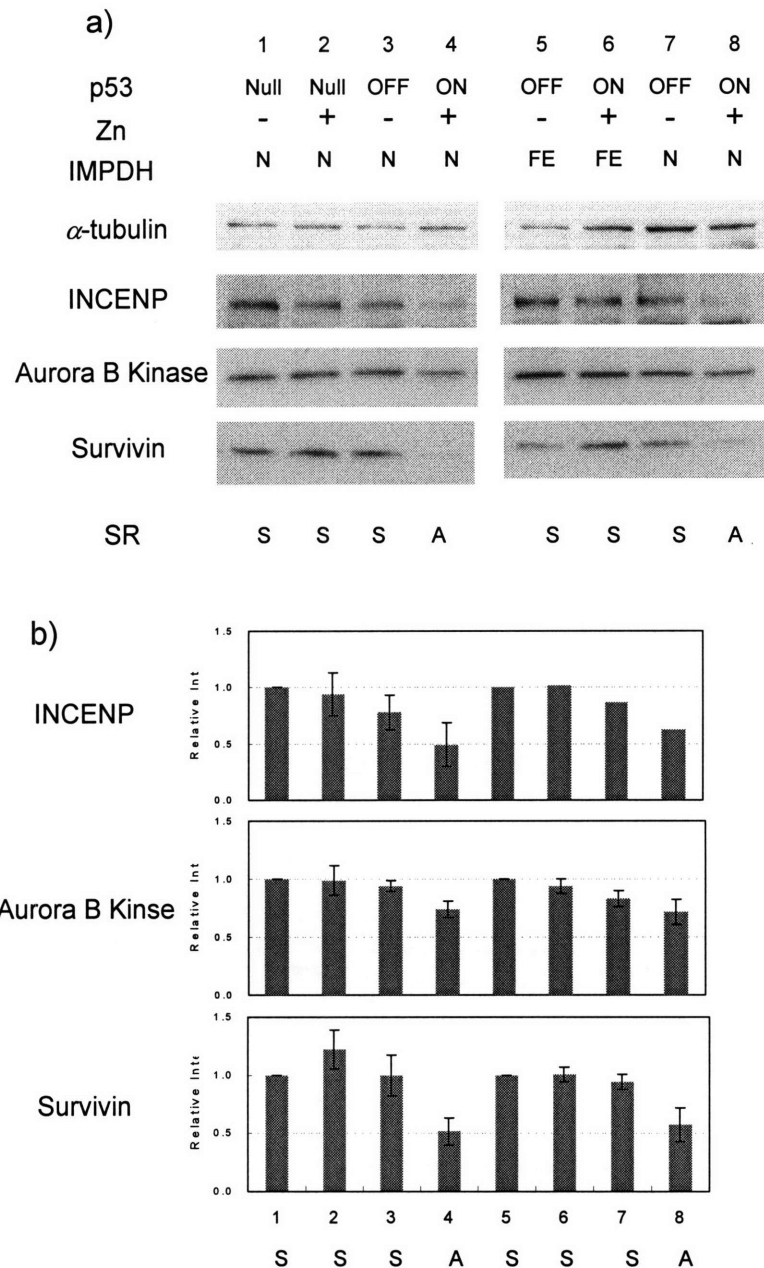
146

Figure 29. Evaluation of chromosome passenger complex protein expression in model cell systems for regulated self-renewal pattern. a) Western blotting analyses for INCENP, Aurora B kinase, and survivin proteins were performed with total cell extracts harvested from the same conditions as used for the the Affymetrix whole genome array analyses (*i.e.,* lane 2, lane 4 and lane 6, respectively). The other lanes were the same as in the Figure 16 in Chapter 4. b) The expression levels were measured by using the NIH-image J software after normalized with α-tubulin band intensities. Lane 1-4, n=3; Lane 5-8, INCENP, n=1; Aurora B kinase, n=2; survivin, n=3. SR: self-renewal, S: symmetric self-renewal, A: asymmetric self-renewal, N: normal level, FE: forced expression of IMPDH gene.

147

Results:

Validation of the microarray expression profiles of the chromosome passenger complex proteins

In the Affymetrix whole genome arrays, two chromosome passenger complex proteins, survivin and INCENP, were identified as the SSRA genes. Two other passenger proteins, borealin and Aurora B kinase, were not included in the 5% tail areas of the ratio-expression distribution histograms. However, they also had expression properties of SSRA genes. The normalized expression levels of the genes encoding chromosome passenger complex proteins were analyzed (Figure 28). Similar to survivin and INCENP, the microarray expression of Aurora B kinase and borealin in the cells dividing with symmetric self-renewal were significantly higher than in cells undergoing asymmetric self-renewal (t-test, $p < 0.05$). Therefore, the encoded proteins were predicted to be down-regulated during asymmetric self-renewal and associated immortal DNA strand co-segregation.

Immunoblotting analyses were performed to investigate the expression of the chromosome passenger complex proteins. Three of the proteins, INCENP, Aurora B kinase, and survivin, were up-regulated in symmetrically self-renewing cells compared to asymmetrically self-renewing cells (Figure 29a). Importantly, the changes in the chromosome passenger complex proteins did not occur when p53 was expressed but asymmetric self-renewal was prevented by the forced expression of IMPDH proteins

148

(Figure 29a, lane 6). Therefore, there were correlations between protein and microarray expressions for three chromosome passenger complex SRPA proteins.

The subcellular localization of Aurora Kinases

The subcellular localization of chromosome passenger complex proteins has been well characterized (Honda et al., 2003; Gassmann et al., 2004; Meraldi et al., 2004). However, there have been no reports on the structure and function of the chromosomal passenger complex in the context of non-random chromosome co-segregation. *In situ* immunofluorescence studies for survivin and Aurora B kinase (AurkB) were performed to study whether there was any difference in the subcellular localization of these chromosome passenger complex proteins between asymmetric and symmetric self-renewal.

The subcellular localization of AurkB was correlated with its known function in regulation of centromere-kinetochore interactions (Figure 30a). The immunofluorescence of AurkB was observable in prophase, reached maximal levels in the metaphase plate, and concentrated in the mid-body in telophase. However, there was no distinctive difference between asymmetrically and symmetrically dividing cells in terms of its localization pattern and intensity (Figure 30a).
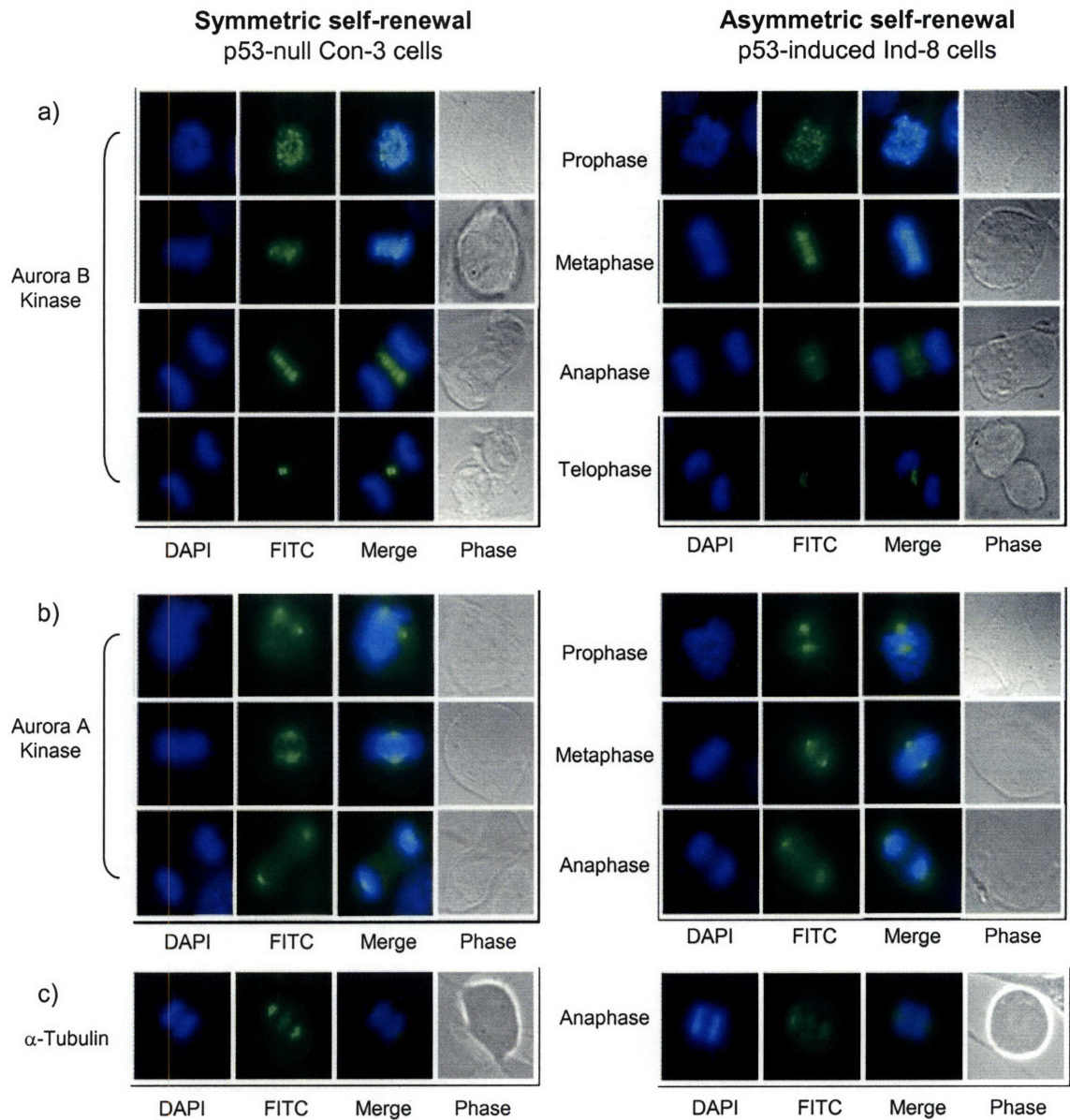
149

Figure 30. Subcellular localization of Aurora B kinase, Aurora A kinase and α-tubulin in mitotic cells by in situ immunofluorescence. Cells were plated at 2000 cells/cm^2 in 4-well chamber slide. After 24 hours, cell culture media were replaced with either Zn-free medium or medium containing 65 μM ZnCl$_2$. At the 36 hours after the induction of asymmetric self-renewal in the p53-inducible Ind-8 cells, the cells were fixed and in situ immunofluorescent studies were performed. a) Localization of Aurora B kinase. b) Aurora A kinase. c) α-tubulin. DNA was visualized by DAPI fluorescence. FITC-conjugated antibodies were used to visualize protein locations
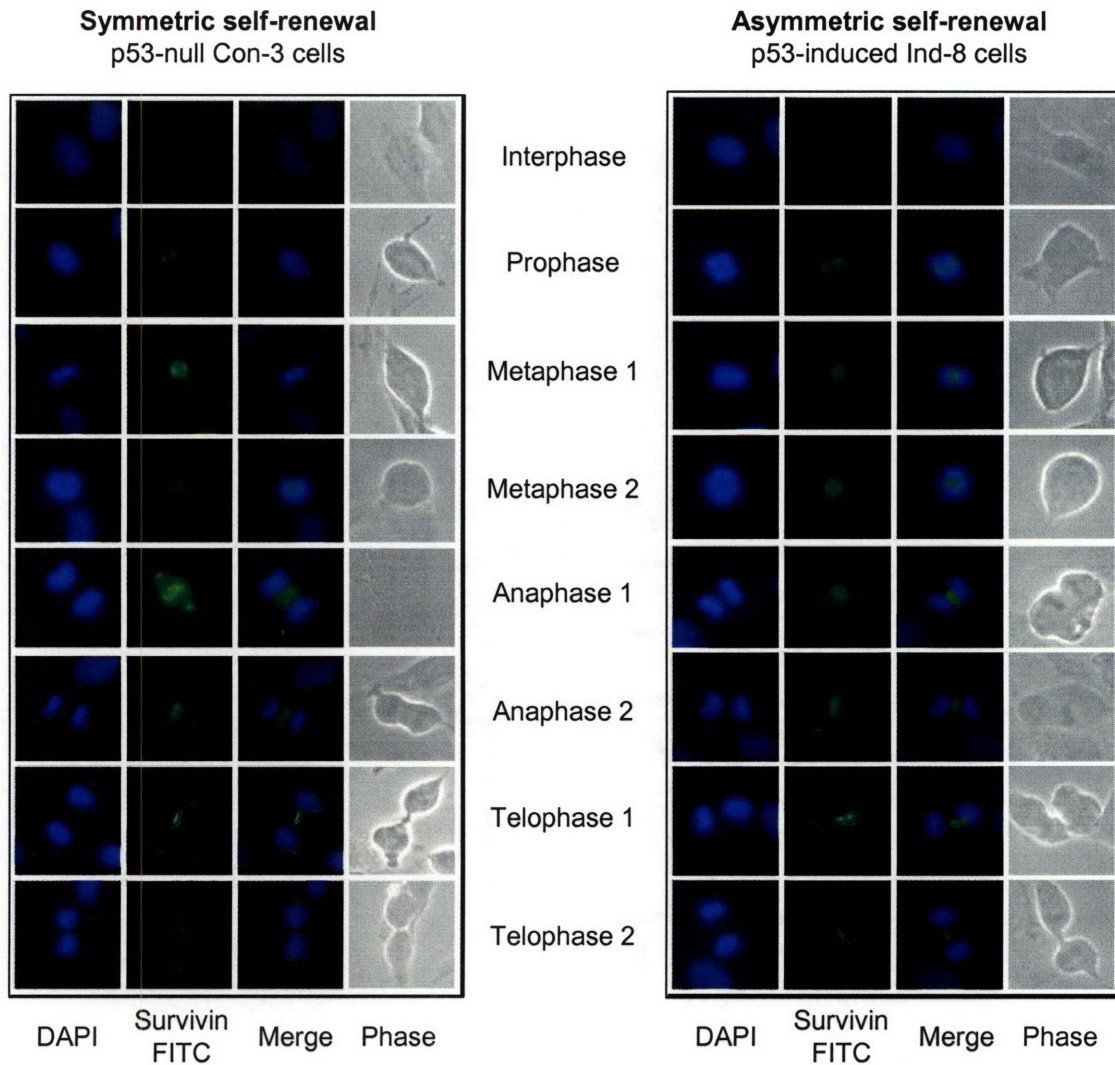
150

Symmetric self-renewal
p53-null Con-3 cells

Asymmetric self-renewal
p53-induced Ind-8 cells

Interphase

Prophase

Metaphase 1

Metaphase 2

Anaphase 1

Anaphase 2

Telophase 1

Telophase 2

DAPI   Survivin FITC   Merge   Phase

DAPI   Survivin FITC   Merge   Phase

Figure 31. Survivin subcellular localization in cells undergoing symmetric or asymmetric self-renewal. Cells were plated at 2000 cells/cm^2 in 4-well chamber slides. After 24 hours, cell culture media were replaced with either Zn-free medium or medium containing 65 μM ZnCl$_2$. At 36 hours after the induction of asymmetric self-renewal in the p53-inducible Ind-8 cells, the cells were fixed and in situ immunofluorescence studies were performed. DNA was visualized by DAPI fluorescence. FITC-conjugated antibodies were used to visualize protein locations

151

At the same time, the localizations of Aurora A kinase (AurkA) and α-tubulin were also evaluated throughout mitosis. In budding and fission yeast, there is a single Aurora kinase, whereas mammalian cells have multiple Aurora kinases, A, B, and C (Tanaka et al., 2005). The AurkA was used as a control when compared the subcellular localization of AurkB in the asymmetrically dividing cells with their symmetric counterparts. The protein α-tubulin is a structural element of microtubules that form mitotic spindles and plays roles in kinetochore dynamics in mitosis (Meraldi et al., 2004; Tanaka et al., 2005). In Western blotting studies, like α-tubulin, there was no significant change in the AurkA protein expression level with changes in self-renewal pattern (data not shown). AurkA immunofluorescence was observed at centrosomes during mitosis and there was no difference noted in its localization pattern between the two patterns of the self-renewal (Figure 30b). Strong α-tubulin signals were detected around centrosomes and mitotic spindles in the metaphase plates where microtubules are required for the mitotic chromosome dynamics and segregation (Figure 30c).

The subcellular localization of survivin

In the interphase cells, there was no significant survivin immunofluorescence. Centrosomal survivin in immunofluorescence was detected in prophase cells and was remained later in mitosis (Figure 31). Survivin immunofluorescence reached maximal levels in metaphase and was strongly observable inside the annular array of metaphase chromosomes (Figure 32a). In contrast, the AurkB signal was aligned parallel to the metaphase chromosome annulus (Figure 32b). This was a clear distinction in the

152

subcellular localization of survivin versus AurkB (Figure 32c). The survivin-AurkB-INCENP chromosome passenger complex proteins together regulate the chromosome segregation (Maiato et al., 2004). *In situ* immunofluorescence studies showed that survivin also localized to centrosomes and also distributed along with the mitotic spindles, survivin may have additional roles with mitotic spindles, independent of AurkB. The survivin localization patterns between asymmetrically and symmetrically self-renewing cells were similar until anaphase (Figure 31). At the telophase, this difference in the centrosomal survivin localization was increased.

In telophase, the immunofluorescence intensities of centrosomal survivin in the cells undergoing with asymmetric self-renewal were generally weaker than those of symmetrically self-renewing cells. There were two patterns of the survivin distribution in mitotic cells at the telophase. In the first pattern, survivin immunofluorescence was detected from three distinctive locations, the two centrosomes, one in each sister cell, and the midbody (Figure 31, Figure 33a). This pattern was predominantly observed in the cells undergoing symmetric self-renewal (Figure 33c). In asymmetrically self-renewing cells, survivin was predominantly detected only in the midbody region between dividing cells (Figure 31, Figure 33b). The difference in the survivin centrosomal localization in telophase was significant in Fisher's exact test (p = 0.0027; Figure 33c). In addition, to this major difference, metaphase chromosome annuli is asymmetrically self-renewing cells showed an open-"C" structure at a higher frequency than symmetrically self-renewing cells. Insufficient observations were made for a meaningful statistical analysis of the significance of this observation (data not shown).
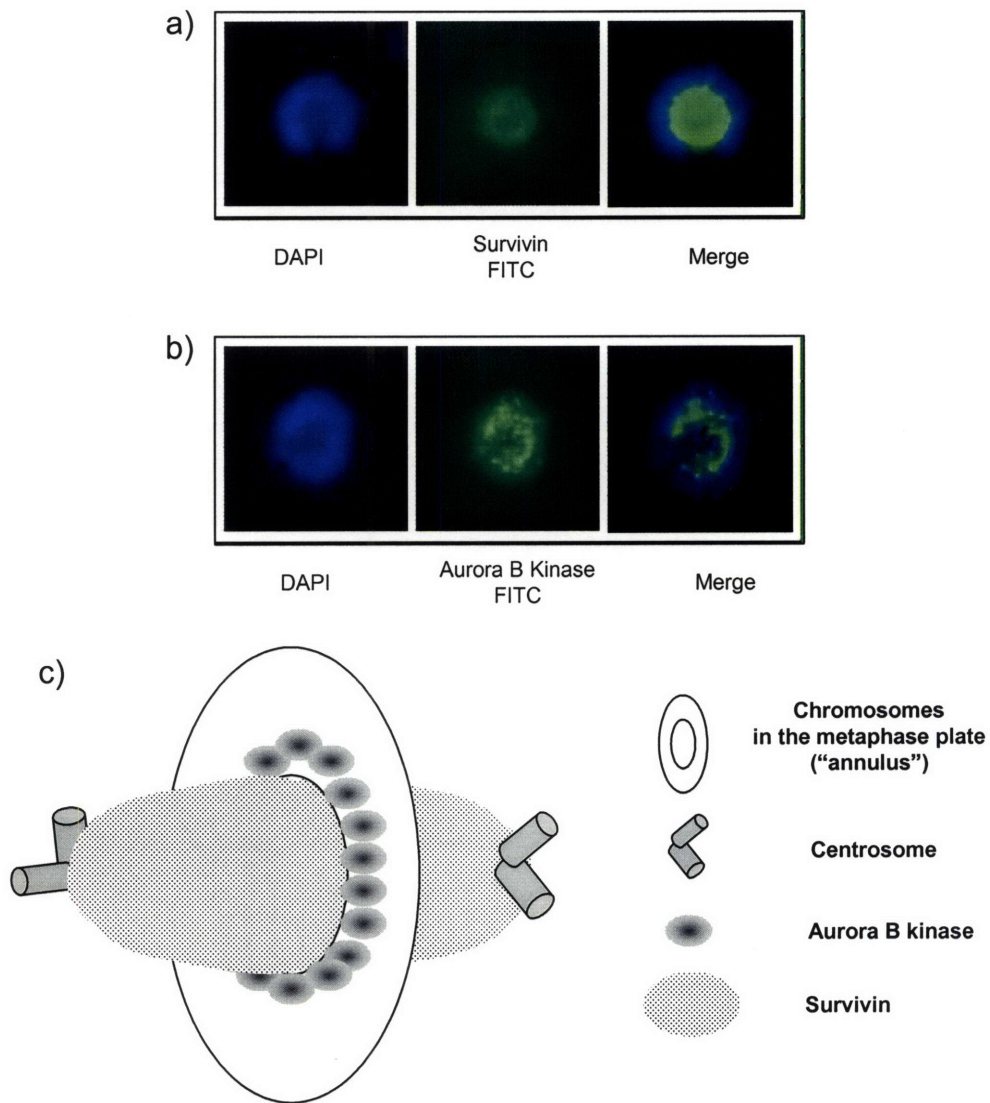
153

Figure 32. Survivin and Aurora B kinase localization in metaphase. Cells were plated at 2000 cells/cm^2 in 4-well chamber slide. After 24 hours, cell culture media were replaced with either Zn free medium or medium containing 65 μM ZnCl$_2$. At 36 hours after the induction of asymmetric self-renewal in the p53-inducible Ind-8 cells, the cells were fixed and in situ immunofluorescent studies were performed. a) survivin immunofluorescence. b) Aurora B kinase immunofluorescence. c) Schematic of basic topology of survivin and Aurora B kinase in metaphase.
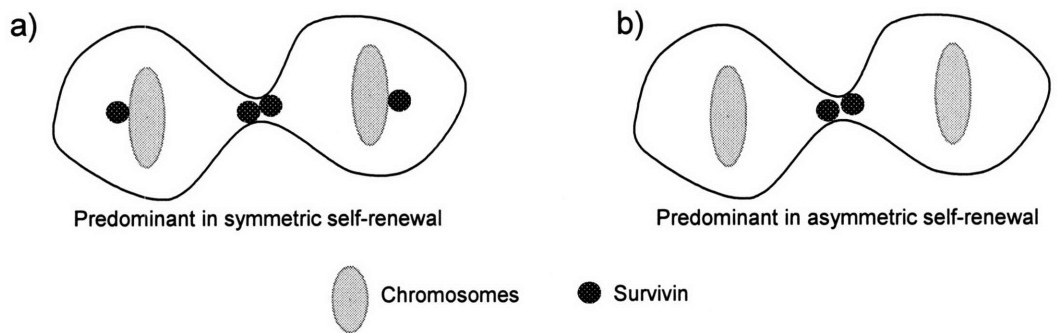
	Symmetric Self-Renewal			Asymmetric Self-Renewal			
	Centrosomal survivin immunostaining		Observed (212)	Centrosomal survivin immunostaining		Observed (52)	Fisher's Test
	Yes	No		Yes	No		
Prophase	14	0	14	5	1 (17 %)	6	
Metaphase	108	0	108	22	0	22	
Anaphase	17	2 (11 %)	19	5	2 (29 %)	7	p = 0.28
Telophase	46	25 (35 %)	71	4	13 (77 %)	17	p = 0.0027

Figure 33. Reduced centrosomal survivin localization in asymmetrically self-renewing cells. a) The centrosomal and midbody survivin localizations in telophase cells that were predominantly observed in symmetrically self-renewing cells. b) The midbody-only survivin localization in telophase cells undergoing asymmetric self-renewal. c) Difference in centrosomal survivin localization related to self-renewal pattern. Cells were fixed at 36 hours after induction of asymmetric self-renewal. Approximately 10,000 cells for each self-renewing condition were observed. Only mitotic cells were counted.

Discussion and Conclusion:

In this study, the functional relevance of SSRA proteins associated with chromosome segregation to the asymmetric self-renewal was evaluated by the *in situ* immunofluorescence localization in cell with different self-renewal patterns. Because the SSRA proteins, survivin, INCENP, and AurkB, form the chromosome passenger complex that regulates kinetochore-centromere association and chromosome segregation during mitosis, the molecular mechanism for the non-random chromosome co-segregation may be discovered by studying these SSRA proteins.

Western blotting analyses for evaluating correlation between microarray and protein expression showed that these chromosome passenger complex protein expression was sensitive to self-renewal pattern. The down-regulation of AurkB after the induction of asymmetric self-renewal was not as much as those of INCENP and survivin (Figure 29). Because Aurora B kinase is fully activated during mitosis by the presence of survivin, the small change in the protein level may not reflect the change in Aurora B kinase activity (Chen et al., 2003; Lens et al., 2003).

In symmetrically self-renewing cells, survivin exists at centrosomes in prophase, on mitotic spindles and centrosomes in metaphase, in midbodies and centrosomes in anaphase, and at centrosomes and midbodies in telophase (Figure 31). However, in telophase of cells undergoing asymmetric self-renewal, the frequency of centrosomal survivin localization was significantly lower than that of symmetrically self-renewing cells (Figure 33). AurkB forms the chromosome passenger complex with surviving and other proteins. However, its cellular localizations in both self-renewing states was almost

156

identical (Figure 30). In addition, there was no noticeable difference in the immunological localization of Aurora A kinase, which may play an essential role at centrosomes during mitosis (Figure 30).

In spite of its difference in gene and protein expression, there was no detected difference in the AurkB localization pattern between asymmetric and symmetric self-renewal. The changes in AurkB protein expressions in the Western blotting may be due to a difference in the number of mitotic cells between the two conditions of self-renewal pattern. However, previous unpublished work in this laboratory showed that the same changes in expression were noted in pure mitotic cell preparations. *In situ* immunofluorescence study with the polyclonal anti-INCENP antibody was also performed. However, so far, these analyses have insufficient resolution for evaluation. In the Affymetrix array results, borealin, another chromosome passenger complex protein, was also significantly down-regulated in asymmetrically self-renewing cells (Figure 28). Further analyses of borealin and INCENP with the asymmetrically dividing cells will be required to establish whether these two chromosome passenger complex proteins show alterations in localization.

During mitosis, survivin was reported as the modulator for microtubule dynamics and nucleation (Giodini et al., 2002; Rosa et al., 2006). These reports showed that survivin increased microtubule turnover around mitotic kinetochores. In asymmetrically self-renewing cells, the lower expression of survivin compared to those in symmetrically self-renewing cells may increase the stability of microtubule attachment to kinetochores. The stable attachment of microtubules to kinetochores may be a mechanism used in non-

random chromosome co-segregation of ASCs. Microtubules in nuclei of cells undergoing asymmetric self-renewal may be stably attached to old parental chromosomes. Because survivin is involved in various cell functions like apoptosis (Altieri, 2003), specialized independent modulators may function regulate this role for survivin. Therefore, there may be additional factors in ASCs that stabilize the kinetochore-microtubule attachment. In studies with *S. cerevisiae* mutants, down-regulation of analogues of two other SSRA-passenger complex proteins, AurkB and INCENP, is predicted to have a similar effect of centrosome-microtubule-kinetochore stability (Tanaka et al., 2002). Thus, continued research of SRPA genes may lead to discovery of the molecular basis for non-random chromosome co-segregation in ASCs.

Chapter 9. Expressions of Self-Renewal Pattern Associated Genes in HSC-Enriched Populations: Towards the Discovery of Novel Molecular Markers for Adult Stem Cells

<u>Rationale</u>:

The SRPA genes were identified from the microarray analyses of the model cell systems whose self-renewal patterns were reversibly controlled. The comparative analyses between the SRPA gene signature and the stemness gene signatures showed that the SRPA gene sets were more effective as a gene expression signature to classify ASC populations (Chapter 5). Importantly, the ASRA genes from the two independent microarray analyses represented the gene expression characteristics of ASC-enriched populations (Chapter 2 and 3). These supportive analyses increased the possibility that novel molecular markers for ASCs might be discovered from studying SRPA genes. However, in order to evaluate potential markers of ASCs, the expression and function of the SRPA genes needed to be investigated in the well-defined uncultured ASC populations.

Several ASC niches have been well characterized. These include gastrointestinal epithelial crypts, epidermal hair follicles, and bone marrow (Fuchs et al., 2004). The demonstration of the existence of ASCs has been usually evaluated with repopulation studies (Loeffler and Potten, 1997; Weissman et al., 2001). Several molecular markers for each ASC compartment have been reported and widely studied. Musashi-1 was suggested as a putative intestinal crypt stem cell marker (Potten et al., 2003). Epidermal hair follicle stem cells were successfully enriched by anti-CD34 and anti-α6-integrin immunoselection (Trempus et al., 2003). The various combinations of CD34, CD133, and

other lineage markers have been used for studying mouse and human HSCs (Weissman et al., 2001; Fuchs et al., 2004). However, many molecular markers like CD34 have been used without knowing their exact molecular function in ASCs. The SRPA genes were identified based on their association with ASC self-renewal pattern. Therefore, molecular markers discovered based on SRPA genes may have a greater likelihood of identifying ASCs exclusively.

ASCs have been enriched by CD-based immunoselection, laser capture microdissection (LCM), and EGF/FGF-based tissue-sphere cultures (e.g. neurosphere cultures for neural stem cells) (Weissman et al., 2001; Stappenbeck et al., 2003; Reynolds and Rietze, 2005). It has been difficult to apply these methods to the large-scale isolation of ASCs. Although neurosphere type-enrichment may produce millions of cells relatively easy and inexpensively, a recent study showed that the stem cell fraction in neurosphere culture was less than 0.1% (Reynolds and Rietze, 2005).

Recently, Hoechst 33342 dye-effluxing side population (SP) cells have been shown to be enriched for stem cells from several adult tissues (Challen and Little, 2006). Originally, this method was applied to the isolation and characterization of the HSCs in mouse bone marrow (Goodell et al., 1996). Mouse bone marrow repopulation assays showed that the Hoechst 33342 dye-effluxing cell population in bone marrow contains 1000~3000-fold more HSCs than whole bone marrow cells. The transplantation efficiency of SP cells was almost the same level as other HSC enrichment methods using the combination of several cell surface markers (Challen and Little, 2006). Practically

speaking, the Hoechst 33342 dye method may provide cheaper and easier ASC enrichments than any other methods currently available.

In order to study candidate ASRA genes as a putative ASC markers, especially identified from the whole genome transcriptional profiling studies, large scale ASC enrichments were indispensable for high-throughput gene expression evaluation. In this study, the plasma membrane encoding ASRA genes were evaluated for their expression distribution between the HSC-enriched Hoechst 33342 dye-effluxing side population of mice and SP cell-depleted bone marrow.

Materials and Methods:

Isolation of mouse bone marrow Hoechst 33342 dye-effluxing side population cells

Bone marrow cells were isolated from the femurs and tibias of FVB mice at 3~6 months of age. In average, 30 million bone marrow cells were collected from one three month old mouse. After centrifugation, the bone marrow cells were resuspended at 1 X 10^6 cells/ml in pre-warmed DMEM with 2% dialyzed fetal bovine serum (DFBS) (Invitrogen, Carlsbad, CA). Hoechst 33342 was added to a final concentration of 5 μg/ml (Goodell et al., 1996). Verapamil was added to an additional Hoechst 33342-treated bone marrow cell suspension for verifying Hoechst 33342 dye exclusion. After mixing, the cells were incubated for 90 minutes at 37 °C. After Hoechst 33342 staining, cells were centrifuged and resuspended in ice-cold Hanks balanced salt solution (HBSS) containing 2% DFBS. The Hoechst 33342 stained cells were maintained at 4 °C until FACS analysis.

The bone marrow cell suspension was filtered through a 70 μM nylon mesh filter before flow cytometry and fluorescent-activated cell sorting (FACS) analysis. Propidium iodide (PI) was added to a final concentration of 2 μg/ml for dead cell discrimination during cell sorting. The mouse bone marrow Hoechst 33342 dye effluxing side population (BM-SP) cells were isolated by the BM-SP isolation protocol for the Cytomation MoFlo high speed sorter at the MIT Center for Cancer Research Flow Cytometry Core Facility. From the each mouse, on average, 50,000 BM-SP cells were isolated. In addition, the SP-depleted BM cells were collected to 2 million cells during parallel cell sorting.

Total RNA isolation and Q-RT-PCR.

Total cellular RNA was isolated from the BM-SP cell population and the BM Hoechst stained cell population (SP-depleted BM). The total RNA quality was evaluated by using the Agilent BioAnalyzer 2100. Q-RT-PCR was performed with the TaqMan® gene expression assay kits as decribed in Chapter 4. Gene expression of *Plxdc2*, *Eng*, *Notch1*, *9130213B05Rik*, *Cxcr6*, and *kcnn4* was evaluated and normalized against the expression of *Gapdh* as a housekeeping mRNA standard. Five hundred ng of total RNA from each cell population was used for transcribing cDNA with the Superscript Reverse Transcriptase (RT) II kit (Invitrogen, Carlsbad, CA). The cDNA from the reaction was amplified by PCR measuring the FAM fluorescence of each PCR cycle using the Applied Biosystems 7700 Sequence Detection System. The reactions were performed according to the manufacturer's instruction after adjusting the reaction volume as 25 μl.

Results:

Isolation of the bone marrow Hoechst 33342 dye-effluxing side population cells and investigation of ASRA gene expression

The Hoechst 33342 dye-effluxing side population (BM-SP) cells were isolated from 3~6 month old adult FVB mice. On average, 50,000 BM-SP cells were sorted from the 30 million bone marrow cells from each mouse by the cell sorter (R4 in Figure 34a). The overall yield was relatively lower compared with the values in literature mainly because of using older adult mice. SP-depleted BM Hoechst-stained cells were also isolated for comparing gene expression in Q-RT-PCR assays (R6 in Figure 34a).

Six ASRA genes, *9130213B05Rik*, *Plxdc2*, *Cxcr6*, *Notch1*, *Eng*, and *Kcnn4*, were evaluated for gene expression in BM-SP by Q-RT-PCR (Figure 34b). Three ASRA genes, *Plxdc2*, *Cxcr6*, and *Epbh6*, were differentially expressed in the BM-SP cells to a significant degree ($p < 0.05$). The *Plxdc2* transcripts in the BM-SP cells were up-regulated by 3.8 fold compared to those in the Hoechst 33342 stained non-SP cells (Figure 34b). The *Plxdc2* transcript was exclusively detected only in the asymmetrically self-renewing cells in the Q-RT-PCR analysis (Table 9 in Chapter 4). Therefore, this result suggests that more asymmetrically dividing cells were in the BM-SP cells than in the SP-depleted Hoechst 33342-stained cells. Because the BM-SP population was enriched for HSCs, this result indirectly suggests that HSCs undergo with asymmetric self-renewal.

a)
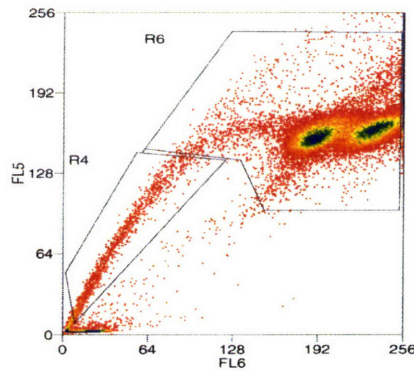
b)

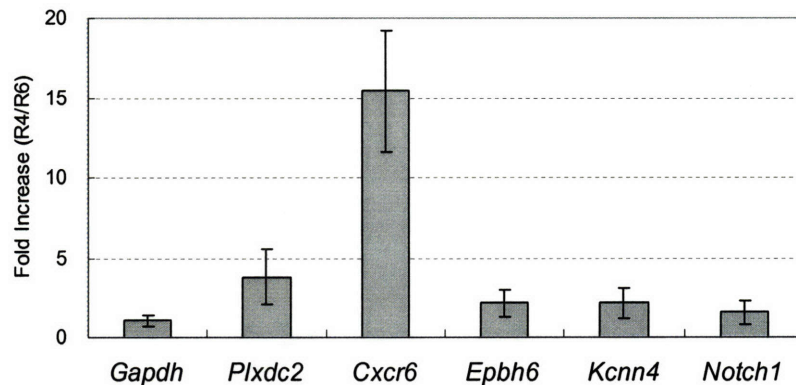|	Average (R4/R6)	STD	p, (ΔCt_{norm})
Gapdh	1.06	0.35	0.99
9130213B05Rik	Undetectable		
Plxdc2	**3.8**	1.8	0.01
Cxcr6	**15**	3.8	0.01
Epbh6	**2.1**	0.83	0.04
Kcnn4	2.1	0.92	0.07
Notch1	1.6	0.71	0.34

Figure 34. ASRA gene expression in bone marrow Hoechst 33342 dye-effluxing side population cells. a) The Hoechst 33342 fluorescent-activated cell sorting profile of female FVB mouse bone marrow cells. The BM-SP cells (R4) were collected with the Cytomation MoFlo sorter. The cell sorting area is marked (R4). The Hoechst 33342-stained non-SP cell population (R6) was also isolated for the Q-RT-PCR comparison. FL5: Hoechst blue, FL6: Hoechst red. b) Q-RT-PCR results of the ASRA genes. Gene expression was normalized by adjusting the threshold cycle (C_t) for each gene to the C_t of *Gapdh* gene. Student's t-test was used to estimate confidence in observed differences between the BM-SP and the Hoechst-stained non-SP BM cells using the normalized C_t values. In graph, bar height equals mean fold increase for n = 2 independent experiments. Error bars = S. D.

Interestingly, there was 15 fold difference in *Cxcr6* gene expression between the BM-SP and Hoechst 33342-stained cells. This change was much bigger than that observed the Q-RT-PCR studies with model cells (Table 9). *Epbh6* gene transcripts were 2-fold enriched in the BM-SP cells. However, the transcript levels for *Notch1* and *Kcnn4* genes were not significantly different between the two bone marrow cell populations. There was no detectable transcript for *9130213B05Rik* in bone marrow cells of either type (Figure 34c). In this study, the total RNA used for cDNA synthesis was as half as the amount used in Q-RT-PCR analysis with model cell systems. Thus, there was limitation in sensitivity for detecting *9130213B05Rik* gene expression.

Discussion and Conclusion:

The Q-RT-PCR analysis of the plasma membrane encoding ASRA genes in the BM-SP cells showed that three ASRA genes, *Plxdc2*, *Cxcr6*, and *Epbh6*, were differentially expressed in the HSC-enriched population to a significant level. The differential expression of the *Plxdc2* gene in the BM-SP cells suggested that there were more asymmetrically self-renewing cells in the BM-SP, corresponding to the self-renewal pattern of HSCs. Again, this result underscores why exclusively expressed genes in asymmetric self-renewal will be important for the identification of ASCs *in vivo*. There was a previous report that *Plxdc2* transcripts were highly detected in normal lung and skeletal muscle tissues in the mouse (Carson-Walter et al., 2001). This report also showed that the *Plxdc2* gene was expressed in the tumor endothelium in colon cancer tissues. One possible implication of this finding is that the endothelium of colon cancer may acquire

the feature of ASC self-renewal and thereby show the increased expression of *Plxdc2* gene. Currently, there is no available tool to detect the protein encoded by the *Plxdc2* gene. In the future, antibodies against the protein may help better understand its role in HSCs and also possible tumor associated endothelial cells.

The transcript of the *Cxcr6* gene was highly up-regulated in the BM-SP cells. The *Cxcr6* gene, also called BONZO, STRL33, has been under focus as a co-receptor in lymphocytes for human immunodeficiency virus (HIV) infection (Deng et al., 1997). In preliminary studies, protein expression of the *Cxcr6* gene was detected in human CD34$^+$ and CD133$^+$ HSC-enriched populations, and a significant fraction of CXCR6$^+$ cells are CD38- and CD33- (data not shown). This result suggested that *Cxcr6* may be highly associated with HSCs. In mouse studies, *Cxcr6* gene expression was detected in various tissues like spleen, thymus, small intestine, and placenta (Deng et al., 1997). In the peripheral blood, the low level of *Cxcr6* gene expression was also detected in the polarized memory/effector Th1 cells (Kim et al., 2001). Even though *Cxcr6* gene expression is not exclusive in HSCs, in combination with traditional human HSC-associated markers, the CXCR6$^+$ phenotype may provide much more exclusive identification.

Outside of marking human HSC, the CXCR6 protein may play a role in asymmetric self-renewal by HSCs. CXCL16, a transmembrane ligand for CXCR6, is expressed in a subset of CD4, CD8, and natural killer T-cells (Matloubian et al., 2000). Because T-lymphocytes actively interconvert from symmetric expansion to stable maintenance as memoary T-cells in order to respond to foreign agents, self-renewal

control may be more critical than in any other tissues. In addition, when the self-renewal symmetry state of putative pancreatic stem cell lines was changed from symmetric to asymmetric, the exclusive *Cxcr6* gene expression was observed (in collaboration with Dr. Jean-Francois Pare, data not shown). Therefore, the function of *Cxcr6* should be studied in the context of self-renewal pattern regulation of HSCs as well as other types of ASCs.

In this study, three ASRA gene expressions were confirmed in the HSC-enriched BM-SP cells by the evaluation of the 9 candidate genes from the microarray analyses. One third of the ASRA genes were actually expressed in the HSC-enriched population, although some of them, like *Cxcr6,* are reported to be also expressed in the lineage committed cells that may conditionally divide with asymmetric self-renewal. Conclusively, the approach to discover novel molecular markers of ASCs by studying asymmetric self-renewal itself provided the great chance of finding genes actually functioning in the ASCs and led to the discovery of novel molecular markers for ASCs.

Summary of Conclusions on the Study of Self-Renewal Pattern-Associated (SRPA) Genes

The genome-wide transcriptional profiling studies for the identification of self-renewal pattern associated (SRPA) genes and subsequent functional analyses of SRPA gene-encoded proteins showed that asymmetric self-renewal is an important consideration for the effective characterization of adult stem cells (ASCs). This conclusion was supported by the following findings in this thesis research:

Thesis hypothesis: Investigations based on the defining property of ASCs, asymmetric self-renewal, will provide key clues not only to understanding their functions but also to improving tools for their identificaiton.

Unique molecular markers for ASCs have been in high demand for their research and clinical applications (Sherely, 2002; Fuchs et al., 2004). The development of exclusive molecular markers of ASCs has been difficult mainly due to tough technical barriers in their identification and isolation from adult tissues. Again, the lack of molecular markers themselves has been major barrier to purify extremely rare ASCs in adult tissues. To date, although molecular markers have been described that are highly associated with specific types of ASCs (e.g., hematopoietic stem cells), none that are exclusively expressed by adult stem cells have been reported.

Previously reported genetic profiles for stemness have included genes commonly up-regulated in both ASCs and embryonic stem cells (ESCs). (Ivanova, et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003). However, the poor correlations

among the stemness transcriptional profiling studies have brought more controversy (Burns and Zon, 2002; Ivanova 2003; Cai et al., 2004; Zipori et al., 2004; Eckfeldt et al., 2005). Since ESCs undergo symmetric self-renewal, genes associated with asymmetric self-renewal were excluded from these stemness gene profiles. Moreover, all previous ASC gene expression profiles, whether or not limited to ASCs, were compromised by the use of impure cell preparations in which ASCs were obscured by the actively proliferating committed progeny cells.

In this research, the defining property of ASCs, asymmetric self-renewal, was targeted to identify essential genes involved in ASC function. It was hypothesized that asymmetric self-renewal held essential information for understanding ASC functions. In addition, the discovery of novel molecular markers of ASCs might have been possible by identifying and studying genes associated with the regulation of self-renewal pattern, especially those up-regulated in asymmetric self-renewal.

Two sets of self-renewal pattern associated (SRPA) genes were identified in gene microarray studies with model cell systems for self-renewal pattern regulation

The described *in vitro* cell culture models were used to find genes whose expression was associated with the self-renewal patterns of ASCs. The self-renewal model cell approach was designed to avoid expected complications that occur when using heterogeneous stem cell-enriched preparations like other stemness studies (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Fortunel et al., 2003). For replicate microarray results, population doubling cycle (PDC) ratio modeling was developed to estimate the dividing

cell fraction in cell culture as a quality control metrix (Chapter 1). With the model cell cultures whose self-renewal pattern was verified by PDC ratio modeling, a 4x4 orthogonal-intersection cDNA microarray analysis was performed. From this analysis, 21 *asymmetric self-renewal associated* (ASRA) and 31 *symmetric self-renewal associated* (SSRA) genes were identified. The 52 genes were collectively referred to as *self-renewal pattern associated* (SRPA) genes (Chapter 2). In order to expand the SRPA gene pool to the whole genome level, the gene expression of three essential self-renewal patterns in the 4x4 orthogonal-intersection analysis was comparatively evaluated with Affymetrix Mouse Genome 430 2.0 arrays (Chapter 3). The whole genome transcript evaluation resulted in the 543 SRPA genes in which there were 310 ASRA and 233 SSRA genes. Seventy percent of the 52 SRPA genes from the cDNA microarray analyses overlapped with those from the whole genome arrays based on the same analytical procedure. In the SRPA gene sets from both microarray analyses, the asymmetric self-renewal associated (ASRA) genes highly represented the genetic feature of ASC-enriched populations. Therefore, the two independent microarray studies supported the proposal that asymmetric self-renewal is an essential feature of adult stemness; and for the first time this statement could be made for primary, uncultured ASC-enriched populations.

The SRPA gene signature can classify ASC-enriched populations based on their self-renewal pattern.

In principal component analyses (PCA) using the SRPA gene expression signature, asymmetrically self-renewing ASC-enriched populations were clustered distinctly from

the symmetrically self-renewing cells (Chapter 5). Importantly, the SRPA gene expression signature was more effective than the stemness gene signatures for detection and delineation of epidermal hair follicle bulge $CD34^+$ cell populations by PCA. In addition, PCA using the SRPA gene signature clearly distinguished human bone marrow $CD34^+$ HSC-enriched cell populations from whole bone marrow cells. Therefore, the SRPA genes identified by the microarray studies with model cell systems can be used for the detection and classification of ASC-enriched populations isolated from adult tissues.

The SRPA genes may provide clues to understanding cancer stem cells.

The chromosome mapping of the 543 SRPA genes identified the two gene clusters on mouse chromosomes ($p < 10^{-8}$) (Chapter 6). The ASRA cluster region on mouse chromosome 2 was syntenic to the human chromosome 9q34 region, the most frequent chromosome rearrangement locus in human leukemia. The SSRA cluster on mouse chromosome 18 was also associated with the chromosome abnormalities in human cancers (Taki et al., 1999; Calabretta and Perrotti, 2004; Giagounidis et al., 2006). The predominant frequency of either ASRA or SSRA genes in their respective chromosome clusters and their association with human cancers strongly implicated these regions as important in carcinogenesis. Their mutation in normal ASC might lead to cancer stem cells that develop into tumors. Two genes in the ASRA cluster, endoglin and *Notch1*, were frequently expressed in the BCR-ABL subtype of pediatric acute lymphoblastic leukemia (ALL). However, how chromosome translocation t(9;22)(*BCR-ABL*) may affect the transcriptions of these ASRA genes remains unknown. Answering this question might

lead to a better understanding of ALL carcinogenesis. Therefore, the SRPA genes can be useful to study pathological mechanisms related to the disruption of self-renewal pattern regulation of ASCs. In addition, unique molecular markers to identify cancer stem cells may be discovered by studying the SRPA genes, especially those present in the ASRA and SSRA gene clusters.

SRPA protein modulation as a new suppression of asymmetric cell kinetics (SACK) approach.

SRPA proteins whose function can be modulated by metabolic, immunological, or pharmacological agents may be useful for application to ASC expansion (Chapter 7). These SRPA genes may encode cytokines and growth factors, signal transducing plasma membrane receptors, and rate determining enzymes for metabolic pathways. To investigate a potential approach for expansion of ASCs by suppression of asymmetric cell kinetics (SACK) approaches, pharmacological antagonism of the protein encoded by the ASRA gene *Kcnn4* was evaluated in colony formation assays. The IKCa1 potassium channel antagonist clotrimazole induced an increase in growth rate by asymmetrically self-renewing cell model cells. This result could be explained by a shift of cells to symmetric self-renewal (i.e. by SACK). Therefore, this analysis not only supported the notion that SRPA genes were functionally involved in self-renewal pattern regulation, but also provided opportunities to find novel SACK approaches for expanding ASCs in culture.

The implications of SSRA genes, encoding chromosome passenger complex proteins, in the immortal DNA strand co-segregation mechanisms.

Two chromosome passenger complex proteins, survivin and inner centromere protein (INCENP), were present in the SSRA gene list; and two other proteins, Aurora B kinase (AurkB) and borealin, were identified to have SSRA profiles, retrospectively. Because the chromosome passenger complex regulates kinetochore-centromere attachment by microtubules and chromosome segregation during mitosis, the molecular mechanism for the non-random chromosome co-segregation may be explained by studying these SSRA proteins. The functional involvement of these SSRA proteins in immortal DNA stand co-segregation mechanisms was evaluated by *in situ* immunofluorescence localization. Reduced survivin centrosomal interaction was observed in asymmetrically self-renewing telophase cells compared to their symmetric counterparts. Reduction of the activity of the Aurora B kinase-INCENP-survivin complex has been shown to increase the stability of microtubule-kinetochore attachment in mitosis. This observation supports the hypothesis that the stable attachment of microtubules to kinetochores of immortal DNA strands throughout the cell cycle is a feature of the mechanism that controls immortal DNA strand co-segregation in ASCs (Chapter 8).

Towards discovery of novel molecular markers for ASCs by studying SRPA genes

To evaluate their potential as markers of ASCs, the expression of predicted plasma membrane protein-encoding SRPA genes was evaluated in well-defined, uncultured ASC-enriched populations. Initially, the microarray expression of 9 plasma membrane protein

encoding ASRA genes was verified by Q-RT-PCR (Chapter 4). The transcripts of three ASRA genes, *Plxdc2*, *Cxcr6*, and *Epbh6*, were confirmed as significantly up-regulated in the HSC-enriched mouse bone marrow Hoechst 33342 dye effluxing side population (Chapter 9). Therefore, one third of the evaluated ASRA genes were actually expressed and up-regulated in an HSC-enriched population. This result strongly supports the precept that asymmetric self-renewal is an important phenotype feature of ASCs that holds the secret to ASC function and identity.

References

Abkowitz, J. L., Linenberger, M. L., Persik, M., Newton, M. A., and Guttorp, P. Behavior of feline hematopoietic stem cells years after busulfan exposure. 82, 2096-2103 (1993).

Agresti, A. and Bianchi, M. E. HMGB proteins and gene expression. *Curr. Opin. Genet. Dev.* 13, 170-178 (2003).

Alter, O., Brown, P. O. and Botstein, D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc. Natl. Acad. Sci. USA* 97, 10101-10106, 2000.

Altieri, D. C. Validating survivin as a cancer therapeutic target. *Nature Rev. Cancer* 3, 46-54 (2003).

Artavanis-Tsakonas, S., Rand, M. D. and Lake, R. J. Notch signaling; cell fate control and signal integration in development. *Science* 284, 770-776 (1999).

Asahara, T., Kalka, C. and Isner, J. M. Stem cell therapy and gene transfer for regeneration. *Gene Ther.* 7, 451-457 (2000).

Begenisich, T., Nakamoto, T., Ovitt, C. E., Nehrke, K., Brugnara, C., Alper, S. L. and Melvin, J. E. Physiological roles of the intermediate conductance, Ca^{2+}-activated potassium channel Kcnn4. *J. Biol. Chem.* 279, 47681-47687 (2004).

Bjerknes, M. Simple stochastic theory of stem cell differentiation is not simultaneously consistent with crypt extinction probability and the expansion of mutated clones. *J. Theor. Biol.* 168, 349-365 (1994).

Burns, C. E., Zon, L. I. Portrait of a stem cell. *Dev. Cell* 3, 612-613 (2002).

Cai, J., Weiss, M. L., and Rao, M. S. In search of "stemness". *Exp. Hematol.* 32, 585-598 (2004).

Cairns, J. Mutation selection and the natural history of cancer. *Nature,* 255, 197-200 (1975).

Cairns J. Somatic stem cells and the kinetics of mutagenesis and carcinogenesis. *Proc. Natl. Acad. Sci. USA.* 99, 10567-10570 (2002).

Calabretta, B. and Perrotti, D. The biology of CML blast crisis. *Blood* 103, 4010-4022 (2004).

Calabro, L., Fonsatti, E., Bellomo, G., Alonci, A., Colizzi, F., Sigalotti, L., Altomonte, M. Musolino, C. and Maio, M. Differential levels of soluble endoglin (CD105) in myeloid malignancies. *J. Cell. Physiol.* 194, 171-175 (2003).

Carmena, M. and Earnshaw, W. C. The cellular geography of aurora kinases. *Nat. Rev. Mol. Cell Biol.* 4, 842-854 (2003)

Carpenter M. K., Cui, X, Hu, Z. Y., Jackson, J., and Sherman, S., et al. In vitro expansion of a multipotent population of human neural progenitor cells. *Exp. Neurol.* 158:265-278 (1999).

Carson-Walter, E. B., Watkins, D. N., Nanda, A., Vogelstein, B., Kinzler, K. W. and Croix, B. S. Cell surface tumor endothelial markers are conserved in mice and humans. *Cancer Res.* 61, 6649-6655 (2001).

Challen, G. A. and Little, M. H. A side order of stem cells: the SP phenotype. *Stem Cells* 24, 3-12 (2006)

Chen, C. Z., Li, M., Graaf, D., Monti, S., Gottgens, B., Sanchez, M.-J., Lander, E. S., Golub, T. R., Green, A. R., and Lodish, H. F. Identification of endoglin as a functional marker that defines long-term repopulating hematopoietic stem cells. *Proc. Natl. Acad. Sci. USA.* 99, 15468-15473 (2002).

Chen, J., Jin, S., Tahir, S. K., Zhang, H., Liu, X., Sarthy, A. V., McGonigal, T. P., Liu, Z., Rosenberg, S. H. and Ng, S. Survivin enhances Aurora-B kinase activity and localizes Aurora-B in human cells. *J. Biol. Chem.*, 278, 486-490 (2003).

Cheong, A., Bingham, A. J., Li, J., Kumar, B., Sukumar, P., Munsch, C., Buckley, N. J., Neylon, C. B., Porter, K. E., Beech, D. J., Wood, I. C. Downregulated REST transcription factor is a switch enabling critical potassium channel expression and cell proliferation. *Mol. Cell* 20, 45-52 (2005).

Cheshier, S. H., Morrison, S. J., Liao, X. and Weissman, I. L. In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. *Proc. Natl. Acad. Sci. USA.* 96, 3120-3125 (1999).

Cho, S. K., Bourdeau, A., Letarte, M., Zuniga-Pflucker, J. C. Expression and function of CD105 during the onset of hematopoiesis from Flk1$^+$ precursors. *Blood* 98, 3635-3642 (2001).

Choong, M. L., Tan, A. C. L., Luo, B. and Lodish, H. F. A novel role for proliferin-2 in the ex vivo expansion of hematopoietic stem cells. *FEBS Lett.* 550, 155-162 (2003).

Claudio, J. O., Masih-Kahn, E., Tang, H. et al. A molecular compendium of genes expressed in multiple myeloma. *Blood* 100, 2175-2186 (2002).

Cobb, J. P., Mindrinos, M. N., Miller-Graziano, C., Calvano, S. E. et al. Application of genome-wide expression analysis to human health and disease. Proc. Natl. Acad. Sci. USA 102, 4801-4806 (2005).

Dallas, M. H., Varnum-Finney, B., Delaney, C., Kato, K., and Bernstein, I. D. Density of the Notch ligand delta1 determines generation of B and T cell precursors from hematopoietic stem cells. *J. Exp. Med.* 201, 1361-1366 (2005).

Deng, H., Unutmaz, D., Kewalramani, V. N. and Littman, D. R. Expression cloning of new receptors used by simian and human immunodeficiency viruses. Nature 388, 296-300 (1997).

Eckfeldt, C. E., Mendelhall, E. M., and Verfaillie, C. M. The molecular repertoire of the 'almighty' stem cell. *Nat. Rev. Mol. Cell Biol.* 6, 726-737 (2005).

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA 95*, 14863-14868 (1998).

Elliott, J. I. and Higgins, C. F. IKCa1 activity is required for cell shrinkage, phosphatidylserine translocation and death in T lymphocyte apoptosis. *EMBO reports* 4, 189-194 (2003).

Ellisen, L. W., Bird, J., West, D. C., Soreng, A. L., Reynolds, T. C., Smith, S. D. and Sklar, J. TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromosome translocations in T-lymphoblastic neoplasms. Cell 66, 649-661 (1991).

Etheridge, S. L., Spencer, G. J., Heath, D. J. et al. Expression profiling and functional analysis of Wnt signaling mechanisms in mesenchymal stem cells. *Stem Cells* 22, 849-860 (2004).

Evans, M. J., and Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154-156 (1981).

Fassett, J. T. and Nilsen-Hamilton, M. Mrp3, a mitogen-regulated protein/proliferin gene expressed in wound healing and in hair follicles. Endocrinology 142, 2129-2137 (2001).

Fearnhead, N. S., Britton, M. P. and Bodmer, W. F. The ABC of APC. Human Mol. Genet. 10, 721-734 (2001).

Fonsatti, E., Altomonte, M., Nocotra, M. R., Natali, P. G. and Maio, M. Endoglin (CD105): a powerful therapeutic target on tumor-associated angiogenetic blood vessels. *Oncogene* 22, 6557-6563 (2003).

Fortunel *et al.* Comment on "'Stemness': transcriptional profiling of embryonic and adult stem cells" and "A stem cell molecular signature" (I). *Science*. 302, 393b (2003).

Fuchs, E. and Segre, J.A. Stem cells: a new lease on life. *Cell*, 100, 143-155 (2000).

Fuchs, E., Tumbar, T., and Guasch, G. Socializing with the neighbors: Stem cells and their niche. *Cell* 116, 769-778 (2004).

Gassmann, R., Carvalho, A., Henzing, A. J., Ruchaud, S., Hudson, D. F., Honda, R., Nigg, E. A., Gerloff, D. L. and Earnshaw, W. C. Borealin: a novel chromosome passenger required for stability of the bipolar mitotic spindle. J. Cell Biol. 166, 179-191 (2004).

Ge, X., Yamamoto, S., Tsutsumi, S., Midorikawa, Y., Ihara, S., Wang, S. M. and Aburatani, H. Interpreting expression profiles of cancers by genome-wide survey of breadth of expression in normal tissues. *Genomics* 86, 127-141 (2005).

Ghanshani, S., Wulff, H., Miller, M. J., Rohm, H., Neben, A., Gutman, G. A., Cahalan, M. D. and Chandy, K. G. Up-regulation of the IKCa1 potassium channel during T-cell activation. *J. Biol. Chem.* 275, 37137-37149 (2000).

Giagounidis, A. A. N., Germing, U. and Aul, C. Biological and prognostic significance of chromosome 5q deletions in myeloid malignancies. *Clin. Cancer Res.* 12, 5-10 (2006).

Giodini, A., Kallio, M. J., Wall, N. R., Gorbsky, G. J., Tognin, S., Marchisio, P. C., Symons, M. and Altieri, D. C. Regulation of microtubule stability and mitotic progression by survivin. *Cancer Res.* 62, 2462-2467 (2002).

Goodell, M. A., Brose, K., Paradis, G., Conner, A. S. and Mulligan, R. C. Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. J. Exp. Med.183, 1797-1806 (1996).

Gottlieb, D. I. Large scale sources of neural stem cells. *Annu. Rev. Neurosci.*, 25, 381-407 (2002).

Grabher, C., Boehmer, H. and Look, A. T. Notch1 activation in the molecular pathogenesis of T-cell acute lymphoblastic leukemia. *Nat. Rev. Cancer* 6, 347-359 (2006).

Gygi, S. P., Rochon, Y., Franza, R., and Aebersold, R. Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* 19, 1720-1730 (1999).

178

Hess, D. A., Wirthlin, L., Craft, T. P., Herrbrich, P. E., Hohm, S. A., Lahey, R., Eades, W. C., Creer, M. H., and Nolta, J. A. Selection based on CD133 and high aldhehyde dehydrogenase activity isolates long-term reconstituting human hematopoietic stem cells. *Blood.* 107, 2162-2169 (2006).

Hisaoka, M., Sheng, W. Q., Tanaka, A. et al. HMGIC alterations in smooth muscle tumors of soft tissues and other sites. *Cancer Genet. Cytogenet.* 138, 50-55 (2002).

Hoffman, L. M. and Carpenter, M. K. Characterization and culture of human embryonic stem cells. *Nat. Biotech.* 23, 699-708 (2005).

Honda, R., Korner, R. and Nigg, E. A. Exploring the functional interactions between Aurora B, INCENP, and surviving in mitosis. *Mol. Biol. Cell* 14, 3325-3341 (2003).

Huntly, B.J.P. and Gilliland, D. G.. Leukaemia stem cells and the evolution of cancer-stem-cell research. *Nat. Rev. Cancer* 5, 311-321 (2005).

Huntly, B. J. P., Guilhot, F., Reid, A. G., Vassiliou, G., Hennig, E., Franke, C., Byrne, J., Brizard, A. et al., Imatinib improves but may not fully reverse the poor prognosis of patients with CML with derivative chromosome 9 deletions. *Blood* 102, 2205-2212 (2003).

Ivanova *et al.*, Response to comments on "'Stemness': transcriptional profiling of embryonic and adult stem cells" and "A stem cell molecular signature" *Science.* 302, 393d (2003).

Ivanova, N.B., Dimos, J.T., Schaniel, C., Hackney, J.A., Moore, K.A. and Lemischka, I.R. A stem cell molecular signature. *Science.* 298, 601-604 (2002).

Jordan, J. D., Landau, E. M., and Iyengar, R. Signaling networks: the origins of cellular multitasking. *Cell* 103, 193-200 (2000).

Kadomatsu, K. and Muramatsu, T. Midkine and pleiotrophin in neural development and cancer. *Cancer Lett.* 204, 127-143 (2004).

Kannan, K., Amariglio, N., Rechavi, G., Jakob-Hirsch, J., Kela, I., Kaminski, N., Getz, G., Domany, E. and Givol, D. DNA microarrays identification of primary and secondary target genes regulated by p53. *Oncogene* 20, 2225-2234 (2001).

Kargul, G. J., Dudekula, D. B., Qian, Y., Lim, M. K., Jaradat, S. A., Tananka, T. S., Carter, M. G., and Ko. M. S. H. Verification and initial annotation of the NIA mouse 15K cDNA clone set. *Nat. Genet.* 28, 17-18 (2001).

179

Karpowicz, P., Morshead, C., Kam, A., Jervis, E., Ramunas, J., Cheng, V., and Kooy, D. Support for the immortal strand hypothesis: neural stem cells partition DNA asymmetrically in vitro. *J. Cell Biol.* 170, 721-732 (2005).

Lajtha, L. G. Stem cell concepts. Differentiation 14, 23-34 (1979).

Lee, H.S., Crane, G. G., Merok, J. R., Tunstead, J. R., Hatch, N. L., Panchalingam, K. P., Powers, M. J., Griffith, L. G., and Sherely, J. L. Clonal expansion of adult rat hepatic stem cell lines by suppression of asymmetric cell kinetics (SACK). *Biotech. Bioeng.* 83, 760-771 (2003).

Lenz, S.M.A., Wolthuis, R.M.F., Klompmaker, R., Kauw, J., Agami, R., Brummelkamp, T., Kops, G. and Medema, R.H. Survivin is required for a sustained spindle checkpoint arrest in response to lack of tension, *EMBO J.*, 22, 2934-2947 (2003).

Li, C. and Wong, W. H. Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. *Proc. Natl. Acad. Sci. USA* 98, 31-36 (2001).

Liotta, L. and Petricoin, E. Molecular profiling of human cancer. *Nat. Rev. Genet.* 1, 48-56 (2000)

Liu, Y., Bohn, S. A., and Sherley, J. L. Inosine-5'-monophosphate dehydrogenase is a rate-limiting factor for p53-dependent growth regulation. *Mol. Biol. Cell.* 9, 15-28 (1998a).

Liu, Y., Riley, L. B., Bohn, S. A., Stadler, P. B., Boyd, J. T., and Sherley, J. L. Comparison of bax, waf1, and IMP dehydrogenase regulation in response to wild-type p53 expression under normal growth conditions. *J. Cell Physiol.*, 177, 364-376 (1998b).

Loeffler, M., Birke, A., Winton, D., and Potten, C. Somatic mutation, monoclonality and stochastic models of stem cell organization in the intestinal crypt. *J. Theor. Biol.* 160, 471-591 (1993).

Loeffler, M., Bratke, T., Paulus, U., and Potten, C. S. Clonality and life cycles of intestinal crypts explained by a state dependent stochastic model of epithelial stem cell organization. *J. Theor. Biol.* 186, 41-54 (1997).

Loeffler, M. and Potten, C. S. Stem cells and cellular pedigrees – a conceptual introduction. In *Stem Cells*, C. S. Potten, ed. (San Diego, CA: Harcourt Brace & Co.), pp. 1-28 (1997).

Ludwig, J. A. and Weinstein, J. N. Biomarkers in cancer staging, prognosis and treatment selection. *Nat. Rev. Cancer* 5, 845-856 (2005).

Maiato, H., DeLuca, J., Salmon, E. D. and Earnshaw, W. C. The dynamic kinetochore-microtubule interface. *J. Cell Sci.* 117, 5461-5477 (2004).

Matioli, G., Niewisch, H., and Vogel, H. Stochastic stem cell renewal. *Rev. Europ. Etudes Clin. Et Biol.*, 15, 20-22 (1970).

Matloubian, M., David, A., Engel, S., Ryan, J. E. and Cyster, J. G. A transmembrane CXC chemokine is a ligand for HIV-coreceptor Bonzo. *Nat. Immunol.* 1, 298-304 (2000).

Meraldi, P., Honda, R. and Nigg, E. A. Aurora kinases link chromosome segregation and cell division to cancer susceptibility. *Curr. Opin. Genet. Dev.* 14, 29-36 (2004).

Merok, J.R. and Sherley, J.L. Breaching the kinetic barrier to in vitro somatic stem cell propagation. *J. Biomed. Biotech.* 1, 25-27 (2001).

Merok, J. R., Lansita, J. A., Tunstead, J. R., and Sherley, J. R. Cosegregation of chromosomes containing immortal DNA strands in cells that cycle with asymmetric stem cell kinetics. *Cancer Res.*, 62, 6791-6795 (2002).

Mirza, A., McGuirk, M., Hockenberry, T. N. et al. Human survivin is negatively regulated by wild-type p53 and participates in p53-dependent apoptotic pathway. *Oncogene* 2002, 21, 2613-2622 (2002).

Mirza, A., Wu, Q., Wang, L. et al. Global transcriptional program of p53 target genes during the process of apoptosis and cell cycle progression. *Oncogene* 22, 3645-3654 (2003).

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Maeda, M. and Yamanaka, S. The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* 113, 631-642 (2003).

Morris, R. J., Liu, Y., Marles, L., Yang, Z., Trempus, C., Li, S., Lin, J. S., Sawicki, J. A., and Cotsarelis, G. Capturing and profiling adult hair follicle stem cells. *Nat. Biotech.* 22, 411-417 (2004).

Morrison, S. J. and Weissman, I. L. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity* 1, 661-673 (1994).

Morrison, S. J. Shah, N. M. and Anderson, D. J. Regulatory mechanisms in stem cell biology. *Cell* 88, 287-298 (1997).

Muller, S., Scaffidi, P., Degryse, B. et al. The double life of HMGB1 chromatin protein: architectural factor and extracellular signal. *EMBO J.* 20, 4337-4340 (2001).

181

Muller-Sieburg, C. E., Cho, R. H., Thoman, M., Adkins, B., and Sieburg, H. B. Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. *Blood.* 100, 1302-1309 (2002).

Narushima, M., Kobayashi, N., Okitsu, T., Tanaka, Y., Li, S.-A., Chen, Y. et al. A human β-cell line for transplantation therapy to control type 1 diabetes. *Nat. Biotech.* 23, 1274-1282 (2005).

Nicolas, M., Wolfer, A., Raj, K., Kummer, J. A., Mill, P., Noort, M., Hui, C.-C., Clevers, H., Dotto, P. and Radtke, F. Notch1 functions as a tumor suppressor in mouse skin. *Nat. Genet.* 33, 416-421 (2003).

Oshima, H., Rochat, A., Kedzia, C., Kobayashi, K. and Barrandon, Y. Morphogenesis and renewal of hair follicle from adult multipotent stem cells. *Cell* 104, 233-245 (2001).

Pardal, R. Clarke, M.F. and Morrison, S. J. Applying the principles of stem-cell biology to cancer. *Nat. Rev. Cancer* 3, 895-902 (2003).

Pomeroy, S. L., Tamayo, P., Gaasebeek, M., Sturla, L. M., Angelo, M. et al. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* 415, 436-442 (2002).

Potten, C. S., Booth, C., Tudor, G. L., Booth, D., Brady, G., Hurley, P., Ashton, G., Clarke, R., Sakakibara, S. and Okano, H. Identification of a putative intestinal stem cell and early lineage marker; musashi-1. *Differentiation* 71, 28-41 (2003)

Potten, C.S., Owen, G., and Booth, D. Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J. Cell. Sci.,* 115, 2381-2388 (2002).

Punzel, M. and Ho, A.D. Divisional history and pluripotency of human hematopoietic stem cells. Ann. N. Y. Acad. Sci. 938, 72-81 (2001).

Punzel, M., Liu, D., Zhang, T., Eckstein, V., Miesala, K., and Ho, A.D. The symmetry of initial divisions of human hematopoietic progenitors is altered only by the cellular microenvironment. *Exp Hematol.* 31, 339-347 (2003).

Quackenbush, J. Computational analysis of microarray data. *Nat. Rev. Gen.,* 2, 418-427 (2001).

Quackenbush, J. Microarray data normalization and transformation. *Nat. Gen.* Supplement, 32, 496-501 (2002)

Radtke, F. and Clevers, H. Self-renewal and cancer of the gut: two sides of a coin. Science. 307, 1904-1909 (2005).

Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R.C. and Melton, D.A. Stemness: Transcriptional profiling of embryonic and adult stem cells. *Science.* 298, 597-600 (2002).

Ramaswamy, S., Ross, K. N., Lander, E. S., and Golub, T. R. A molecular signature of metastasis in primary solid tumors. *Nat. Genet.* 33, 49-54 (2003).

Rambhatla L., Bohn, S. A., Stadler, P. B., Boyd, J. T., Coss, R. A., and Sherley, J. L. Cellular senescence: *ex vivo* p53-dependent asymmetric cell kinetics. *J. Biomed. Biotech.* 1, 28-37 (2001).

Rambhatla L.,Ram-Mohan, S., Cheng, J. J., and Sherley, J. L. Immortal DNA strand cosegregation requires p53/IMPDH-dependent asymmetric self-renewal associated with adult stem cells. *Cancer Res.* 65, 3155- 3161(2005).

Rangasamy, D., Berven, L., Ridgway, P. et al. Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. *EMBO J.* 22, 1599-1607 (2003).

Raychaudhuri, S., Stuart, J. M. and Altman, R. B. Principal component analysis to summarize microarray experiments: application to sporulation time series. *Pac. Symp. Biocomput.* 455-466 (2000).

Reya, T. Morrison, S.J., Clarke, M. F. and Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature.* 414, 105-111 (2001).

Reynolds, B. A. and Rietze, R. Neural stem cells and neurospheres-reevaluating the relationship. Nat. Methods 2, 333-336 (2005)

Robinson, M., Jiang, P., Cui, J. et al. Global genechip profiling to identify genes responsive to p53-induced growth arrest and apoptosis in human lung carcinoma cells. *Cancer Biol. Ther.* 2, 406-415 (2003).

Rodier, A., Rochard, P., Berthet, C. et al. Identification of functional domains involved in BTG1 cell localization. *Oncogene* 20, 2691-2703 (2001).

Rosa, J., Canovas, P., Islam, A., Altieri, D. C. and Doxsey, S. J. Survivin modulates microtubule dynamics and nucleation throughout the cell cycle. *Mol. Biol. Cell* 17, 1483-1493 (2006).

Ross, M. E., Zhou, X., Song, G., Shurtleff, S. A., Girtman, K., Williams, W. K., Liu, H., Mahfouz, R., Raimondi, S. C., Lenny, N., Patel, A. and Downing, J. R. Classification of

pediatric acute lymphoblastic leukemia by gene expression profiling. *Blood* 102, 2951-2959 (2003).

Sakamuro, D., Elliott, K. J., Wechsler-Reya, R. and Prendergast, G. C. BIN1 is a novel MYC-interacting protein with features of a tumor suppressor. *Nat. Genet.* 14, 69-77 (1996).

Sampath, S. C., Ohi, R., Leismann, O., Salic, A., Pozniakovski, A., Funabiki, H. The chromosome passenger complex is required for chromatin-induced microtubule stabilization and spindle assembly. *Cell* 118, 187-202 (2004).

Sandberg, A.A. The Chromosome in Human Cancer and Leukemia, Elsevier, New York, (1981).

Schneider, T. E., Barland, C., Alex, A. M., Mancianti, M. L., Lu, Y., Cleaver, J. E., Lawrence, H. J. and Ghadially, R. Measuring stem cell frequency in epidermis: A quantitative in vivo functional assay for long-term repopulating cells. Proc. Natl. Acad. Sci. USA 100, 11412-11417 (2003).

Schuldiner, M., Yanuka, O., Itskovitz-Eldor, J., Melton, D.A., and Benvenisty, N. (2000). Effects of eight growth factors on the differentiation of cells derived from human embryonic stem cells. *Proc. Natl. Acad. Sci. USA* 97, 11307-11312.

Seery, J. P. and Watt, F. M. Asymmetric stem-cell divisions define the architecture of human oesophageal epithelium. *Curr. Biol.* 10, 1447-1450 (2000).

Segal, E., Friedman, N., Kaminski, N., Regev, A., and Koller, D. From signatures to models: understanding cancer using microarrays. *Nat. Genet.* 37, S38-S45 (2005).

Sharpless, N.E. and DePinho R.A. p53: Good Cop/Bad Cop. *Cell.* 110, 9-12 (2002).

Sherley, J. L. Guanine nucleotide biosynthesis is regulated by the cellular p53 concentration. *J. Biol. Chem.*, 266, 24815-24818 (1991).

Sherley, J. L., Stadler, P. B., and Johnson, D. R. Expression of the wild-type p53 antioncogene induces guanine nucleotide-dependent stem cell division kinetics. *Proc. Natl. Acad. Sci. USA*, 92, 136-140 (1995a).

Sherley, J.L., Stadler, P.B., and Stadler, J.S. A quantitative method for the analysis of mammalian cell proliferation in culture in terms of dividing and non-dividing cells. *Cell Prolif.*, 28, 137-144 (1995b).

Sherley, J.L. Asymmetric cell kinetics genes: the key to expansion of adult stem cells in culture. *Stem Cells*, 20, 561-572 (2002).

Sinclair, P. B., Nacheva, E. P., Leversha, M., Telford, N., Chang, J., Reid, A., Bench, A., Champion, K., Huntly, B. and Green, A. R. Large deletions at the t(9;22) breakpoint are common and may identify a poor-prognosis subgroup of patients with chronic myeloid leukemia. *Blood* 95, 738-744 (2000).

Stappenbeck, T. S., Mills, J. and Gordon, J. I. Molecular features of adult mouse small intestinal epithelial progenitors. *Proc. Natl. Acad. Sci. USA* 100, 1004-1009 (2003).

Stier, S., Cheng, T., Dombkowski, D., Carlesso, N., and Scadden, D. T. Notch1 activation increases hematopoietic stem cell self-renewal in vivo and favors lymphoid over myeloid lineage outcome. Blood 99, 2369-2378 (2002).

Su, A., Wiltshire, T., Batalov, S., Lapp, H., Ching, K. A., Block, D., Zhang, Z., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M. P., Walker, J. R. and Hogenesch, J. B. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. USA* 101, 6062-6067 (2004).

Takahashi, N., Kawanishi-Tabata, R., Haba, A., Tabata, M., Haruta, Y., Tsai, H. and Seon, B. K. Association of serum endoglin with metastasis in patients with colorectal, breast, and other solid tumors, and suppressive effect of chemotherapy on the serum endoglin. Clin. Cancer Res. 7, 524-532 (2001).

Tanaka, T. U., Stark, M. J. R. and Tanaka, K. Kinetochore capture and bi-orientation on the mitotic spindle. Nat. Rev. Mol. Cell. Biol. 6, 929-942 (2005).

Tani, H., Morris, R. J. and Kaur, P. Enrichment for murine keratinocyte stem cells based on cell surface phenotype. *Proc. Natl. Acad. Sci. USA* 97, 10960-10965 (2000).

Taylor, G., Lehrer, M. S., Jensen, P. J., Sun, T. T. and Lavker. R. M. Involvement of follicular stem cells in forming not only the follicle but also the epidermis. *Cell* 102, 451-461 (2000).

Tanaka, T. S., Jaradat, S.A., Lim, M. K., Kargul, G. J., Wang, X., Grahovac, M. J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H., Wood, W. H., Becker, K. G., and Ko. M. S. H. Genome-wide expression profiling of mid-gestation placenta and embryo using 15k mouse developmental cDNA microarray. *Proc. Natl. Acad. Sci. USA* 97, 9127-9132 (2000).

Tanaka, T. U., Rachidi, N., Janke, C., Pereira, G., Galova, M., Schiebel, E., Stark, M.J.R and Nasmyth, K. Evidence that the Ipl1-Sli15 (Aurora Kinase-INCENP) Complex Promotes Chromosome Bi-orientation by Altering Kinetochore-Spindle Pole Connections. *Cell.* 108, 317-329 (2002).

185

Taki, T., Kano, H., Taniwaki, M., Sako, M., Yanagisawa, M. and Hayashi, Y. AF5q31, a newly identified AF4-related gene, is fused to MLL in infant acute lymphoblastic leukemia with ins(5;11)(q31;q13q23). Proc. Natl. Acad. Sci. USA 96, 14535-14540 (1999).

Tokino, T., Satoh, H., Yoshida, M. C., Ochiya, T. and Matsubara, K. Regional localization of the LCA oncogene to human chromosome region 2q14q21. Cytogenet. Cell Genet. 48, 63-64 (1988).

Trempus, C. S., Morris, R. J., Bortner, C. D., Cotsarelis, G., Faircloth, R. S., Reece, J. M. and Tennant, R. W. Enrichement for living murine keratinoctes from the hair follicle bulge with the cell surface marker CD34. J. Invest. Dermatol. 120, 501-511 (2003).

Tseng, G. C., Oh, M. K., Rohlin, L., Liao, J. C. and Wong, W. H. Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. Nuc. Acid Res. 29, 2549-2557 (2001).

Tumbar, T., Guasch, G., Greco, V., Blanpain, C., Lowry, W. E., Rendl, M. and Fuchs, E. Defining the epithelial stem cell niche in skin. Science. 303, 359-363 (2004)

Uchida, N., Dykstra, B., Lyons, K., Leung, F., Kristiansen, M., and Eaves, C. ABC transporter activities of murine hematopoietic stem cells vary according to their developmental and activation status. Blood 103, 4487-4495 (2004).

Vogel, H., Niewisch, H., and Matioli, G. Stochastic development of stem cells. J. Theor. Biol. 22, 249-270 (1969).

Wagers, A. J. and Weissman, I. L. Plasticity of adult stem cells. 116, 639-648 (2004)

Wagner, W., Ansorge, A., Wirkner, U. et al. Molecular evidence for stem cell function of the slow-dividing fraction among human hematopoietic progenitor cells by genome-wide analysis. Blood 104, 675-686 (2004).

Washburn, M. P., Koller, A., Oshiro, G., Ulaszek, R. R., Plouffe, D., Deciu, C., Winzeler, E., and Yates, J. R. Protein pathway and complex clustering of correlated mRNA and protein expression analyses in Saccharomyces cerevisiae. Proc. Natl. Acad, Sci. USA 100, 3107-3112 (2003).

Weissman, I.L., Anderson, D.J., and Gage, F. Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. Annu. Rev. Cell Dev. Biol. 17, 387-403 (2001).

Wiles, M.V., and Johansson, B.M. (1999). Embryonic stem cell development in a chemically defined medium. Exp. Cell Res. 247, 241-248.

Yamashita, Y.M., Jones, D.L., and Fuller, M.T. Orientation of asymmetric stem cell division by the APC tumor suppressor and centrosome. *Science*. 301, 1547-1550 (2003).

Yang, Y. H., Dudoit, S., Luu, P., Lin, D., Peng, V., Ngai, J. and Speed, T. P. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nuc. Acids Res*. 30, e15 (2002).

Yatabe, Y., Tavare, S., and Shibata, D. Investigating stem cells in human colon by using methylation patterns. *Proc. Natl. Acad. Sci. USA*. 98, 10839-10844 (2001).

Yeoh, E. J., Ross, M. E., Shurtleff, S. A., Williams, W. K., Patel, D. et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell* 1, 133-143 (2002).

Yeung, K. Y. and Ruzzo, W. L. Principal component analysis for clustering gene expression data. *Bioinformatics* 17, 763-774 (2001).

Zagouras, P., Stifani, S., Blaumueller, C. M., Carcangiu, M. L. and Artavanis-Tsakonas, S. Alterations in Notch signaling in neoplastic lesions of the human cervix. *Proc. Natl. Acad. Sci. USA* 92, 6414-6418 (1995).

Zhao, R., Gish, K., Murphy, M. et al. Analysis of p53-regulated gene expression patterns using oligonucleotide arrays. *Genes Dev*. 14, 981-993 (2000).

Zipori, D. The nature of stem cells: state rather than entity *Nat. Rev. Gen*., 5, 873-878 (2004).

Acknowledgments:

I thank Prof. James L. Sherley, my advisor, for his patient guidance to all aspects of my life at the MIT. His creative scientific vision and wonderful personality have impacted me and made my great learning experience in both science and life. Throughout my thesis research, his creative idea and insight have never been burned out.

I would also like to acknowledge my thesis committee, Prof. John Essigmann, Prof. David Schauer, and Prof. Douglas Lauffenburger for their advice and support during the course of this work. I thank Prof. Schauer for having me to use many instruments in his laboratory.

I also appreciate Dr. Rouzbeh Taghizadeh for his excellent help for my graduate life at the MIT. He contributed to the most important part of my CXCR6 story. I am grateful for the wonderful help of Krisha Panchalingam from the beginning of my work. I also thank Dr. Jean-Francois Pare who encouraged me, taught me how to isolate adult stem cell-enriched preparations, and critically reviewed this thesis, Dr. Christopher Utzat who helped me for cell kinetics studies and also critically reviewed this thesis, and Dr. Samuel Boutin who provided me the first demonstration of using the dChip. Dr Samuel Boutin also contributed to find the asymmetric chromosome cluster region. I would like to acknowledge all other Sherley Lab members for my great moments in the Sherley laboratory.

Finally, thank you to my family. I specially thank Soohyon. I could not have completed this thesis without her encouragement and help.

List of Figures:

Figure 19. Principal component analysis with a 151 gene subset of SRPA gene signature for evaluating the clustering patterns of stem cell-enriched populations

Figure 20. The clustering patterns of epidermal label-retaining hair follicle bulge cells by the SRPA and stemness gene sets

Figure 21. The clustering patterns of epidermal label-retaining hair follicle bulge cells by the SRPA and stemness gene sets

Figure 22. The clustering patterns of CD34$^+$ and CD34- hair follicle bulge cell populations

Figure 23. PCA for human bone marrow samples with the SRPA gene set

Figure 24. Mapping of the 543 SRPA genes on mouse chromosomes

Figure 25. Microarray expression of the SRPA chromosome cluster genes in ALL cells

Figure 26. Evaluation of proliferin secretion with different patterns of self-renewal

Figure 27. Effects of clotrimazole on colony formation by asymmetrically self-renewing cells

Figure 28. Microarray expression of the chromosome passenger complex genes with SSRA properties

Figure 29. Evaluation of chromosome passenger complex protein expression in model cell systems for regulated self-renewal pattern

Figure 30. Subcellular localization of Aurora B kinase, Aurora A kinase and α-tubulin in mitotic cells by in situ immunofluorescence

Figure 31. Survivin subcellular localization in cells undergoing symmetric or asymmetric self-renewal

Figure 32. Survivin and Aurora B kinase localization in metaphase

Figure 33. Reduced centrosomal survivin localization in asymmetrically self-renewing cells

Figure 34. ASRA gene expression in bone marrow Hoechst 33342 dye-effluxing side population cells

Abbreviations:

ASC:	Adult Stem Cell
ASCs:	Adult Stem Cells
ASRA:	Asymmetric Self-Renewal Associated
BM:	Bone Marrow
BM-SP:	Bome Marrow Hoechst 33342 dye effluxing Side Population
CD:	Cluster of Differentiation
cDNA:	complementary DeoxyriboNecleic Acid
DAPI:	4',6-diamidino-2-phenylindole
ESC:	Embryonic Stem Cell
ESCs:	Embryonic Stem Cells
FACS:	Fluorescent-Activated Cell Sorting
FITC:	Fluorescein isothiocyanate
GT	Generation Time
HBSS:	Hanks Balanced Salt Solution
HSC:	Hematopoietic Stem Cell
HSCs:	Hematopoietic Stem Cells
INCENP:	Inner centromere protein
NSC:	Neural Stem Cell
NSCs:	Neural Stem Cells
PBS:	Phosphate Buffered Saline
PCA:	Principal Component Analysis
PDC	Population Doubling Cycle
Q-RT-PCR:	Quantitative real time Reverse Transcriptase Polymerase Chain Reaction
RNA:	RiboNucleic Acid
SACK	Suppression of Asymmetric Cell Kinetics
SRPA:	Self-Renewal Pattern Associated
SSRA:	Symmetric Self-Renewal Associated