From **Department of Microbiology, Tumor and Cell Biology**
Karolinska Institutet, Stockholm, Sweden

EXPLORING THE CANCER CELL ATTRACTOR IN THE EPSTEIN-BARR VIRUS INFECTION MODEL

Qin Li
李钦

Karolinska Institutet

Stockholm 2016

Published by Karolinska Institutet.
Printed by E-print AB

ISBN 978-91-7676-343-8

Exploring the cancer cell attractor in the Epstein-Barr virus infection model

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Qin Li

Principal Supervisor:
Professor Ingemar Ernberg
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology

Co-supervisor(s):
Dr. Jiezhi Zou
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology

Opponent:
Assistant Professor Sandro Santagata
Harvard Medical School
Department of Pathology

Examination Board:
Professor Pierre Åman
Gothenburg University
Department of Pathology, Sahlgrenska Cancer Center

Professor Tina Dalianis
Karolinska Institutet
Department of Oncology and Pathology

Professor Mats Nilsson
Stockholm University and SciLife Laboratory
Department of Biochemistry and Biophysics

To my family

献给我的家人

ABSTRACT

It has been proposed, based on the theory of complex gene regulatory networks, that cell types, including cancer cells, represent attractor states of the network dynamics. In this study, we proposed an Epstein - Barr virus (EBV) latency I to latency III switch model. Exploiting this EBV latency switch model, we characterized for the first time the detailed dynamics of a cancer cell attractor at single-cell-resolution and found that the edge cells from a non-malignant cell line could transiently and stochastically adopt some malignant characteristics due to biological noise. This work might impact design of future rational cancer therapies by taking into account the dynamic robustness and high volatility of a heterogeneous cancer cell population. We also evaluated the impact of EBV load as a biomarker for hematopoietic stem cell transplantation (HSCT) patients.

We established and validated a latency I to latency III switch model that is based on on-off status of the EBV C promoter during latent infection by competition between EBNA1 and Oct proteins together with co-regulators (Paper I).

The dynamics of gene expression space, owing to a balance between homeostatic forces and stochastic fluctuations, has led to the cancer cell attractor conceptual model. Using the immortalized and malignant carrying EBV B-cell lines, we characterized the detailed structure of cell attractors. Any subpopulation selected from a population of cells repopulated the whole original basin of attraction within days to weeks. Cells at the basin edges were unstable and prone to apoptosis. Cells continuously changed states within their own attractor, thus driving the repopulation. Perturbations of key regulatory genes induced a jump to a nearby attractor. Using the Fokker-Planck equation, this cell population behavior could be described as two virtual, opposing influences on the cells: one attracting towards the center and the other promoting diffusion in state space. Transcriptome analysis suggests that these forces result from high-dimensional dynamics of the gene regulatory network. The clonal cell population heterogeneity was investigated by single cell RNA sequencing method. We sequenced different subpopulations within the clonal cell populations and found that edge cells from the non-malignant cell line (non-malignant attractor), represent a distinct population more close to the malignant attractor. This was based on mRNA expression pattern, single cell imaging, clustering analysis, and functional studies. We propose that these findings can be generalized to all cancer cell populations and represented intrinsic behaviors of tumors, offering new perspectives in the study of cancer. The results provide quantitative knowledge on non-genetic intercellular heterogeneity and its dynamics within an isogenic cell population of cancerous cells, affording insights at a new level of resolution, between molecular pathways and macroscopic tumor behaviors (Paper II and III).

We evaluated the impact of EBV load on survival of 51 HSCT patients. Patients with very high or very low level of cell bound EBV-DNA levels had a shorter overall survival (OS) than those with moderate EBV load: OS at 5 years was 67% vs 90%, ($P < 0.03$). There was a conspicuous relationship between EBV load and the dynamics of reconstitution of total and EBV-specific T cells in a few patients. According to multivariate analysis, two other factors were also associated to early mortality: acute GVHD II-IV ($p<0.02$) and pre-transplant conditioning with total body irradiation (TBI) $\geqslant$6 Gy ($p<0.03$). All patients showing these three criteria died within two years after transplantation. This points to a subgroup of HSCT patients which deserve special attention aiming to improve their future treatment (Paper IV).

LIST OF SCIENTIFIC PAPERS

Articles included in this thesis:

1. JieZhi Zou, Jenny Almqvist, **Qin Li**, Ingemar Ernberg. Repression of Epstein-Barr virus enhancer Family of Repeats mediated transcription by Oct and Grg/TLE transcriptional regulators with implications for switching between latency programs. **Manuscript**.

2. **Qin Li**, Anders Wennborg, Erik Aurell, Erez Dekel, JieZhi Zou, Yuting Xu, Sui Huang, Ingemar Ernberg. Dynamics inside the cancer cell attractor reveal cell heterogeneity, limits of stability, and escape. **Proc Natl Acad Sci U S A.** 2016 Mar 8;113(10):2672-7.

3. **Qin Li**, Mtakai Ngara, Anders Wennborg, Andrii Savchenko , Laszlo Szekely, Sui Huang, Rickard Sandberg, Sten Linnarsson, Ingemar Ernberg. Transient phenotypic switch between non-malignant and cancer cell attractors. **Manuscript**.

4. **Qin Li,** Lalit Rane, Thomas Poiret, Jiezhi Zou, Isabelle Magalhaes, Raija Ahmed, Ziming Du, Nalini Vudattu, Qingda Meng, Åsa Gustafsson-Jernberg, Jacek Winiarski, Olle Ringdén, Markus Maeurer Mats Remberger, Ingemar Ernberg. Both high and low levels of cellular Epstein-Barr virus DNA in blood identify failure after hematologic stem cell transplantation in conjunction with acute GVHD and type of conditioning. **Oncotarget.** 2016 Apr 19. [Epub ahead of print]

Related publications and manuscripts:

1. Helen Vallhov, Cindy Gutzeit, Sara M. Johansson, Noémi Nagy, Mandira Paul, **Qin Li**, Sherree Friend, Thaddeus C. George, Eva Klein, Annika Scheynius and Susanne Gabrielsson. Exosomes containing glycoprotein 350 released by EBV-transformed B cells selectively target B cells through CD21 and block EBV infection in vitro. **J Immunol** 186(1): 73-82.

2. **Qin Li**, Jiezhi Zou, Ingemar Ernberg. Comparison of quantitative PCR and semi-quantitative PCR for detection of Epstein-Barr virus in patient blood**. Manuscript**

CONTENTS

1 INTRODUCTION....1
1.1 Brief overview of cancer 1
1.1.1 Cancer epidemiology and challenges in cancer therapy 1
1.1.2 The current view of cancer biology 3
1.1.2.1 Hallmarks of cancer....3
1.1.2.2 Models of tumorigenesis and their challenges 4
1.1.2.2.1 The clonal evolution model....4
1.1.2.2.2 The epigenetic clonal and epigenetic progenitor model....7
1.1.2.2.3 The cancer progenitor/stem cell model....8
1.2 Cell attractor 10
1.2.1 Origin of the idea....10
1.2.2 The conceptual development....11
1.2.2.1 Boolean network....11
1.2.2.2 The gene regulatory network and the attractor....12
1.2.2.3 The quasi-potential landscape, Waddington's epigenetic landscape, and network entropy....13
1.2.2.4 The cancer cell attractor 14
1.2.3 Potential applications of the cell attractor....15
1.2.3.1 The cell attractor as a tool to study cellular heterogeneity....15
1.2.3.2 The cell attractor as a model for cell differentiation and cell phenotypic switch....15
1.2.3.3 Cancer cell attractors as a tool to understand therapy resistance 16
1.3 Epstein-Barr virus biology 16
1.3.1 General introduction of EBV....16
1.3.2 EBV latency types in relation to tumors....17
1.3.2.1 Latency III and latency I/0 17
1.3.2.2 Latency II....19
1.3.2.3 CD10 and CD54 19
1.3.3 EBV load as a biomarker in stem cell transplantation 20
1.3.3.1 Overview of stem cell transplantation....20
1.3.3.2 EBV load together with T cell phenotyping as biomarkers in stem cell transplantation 20
2 AIMS OF THIS THESIS 22
3 MATERIAL AND METHODS 23
3.1 Clinical samples....23
3.2 Cell lines 23
3.3 Single-cell resolution methods....24
3.3.1 Flow cytometry....24
3.3.2 Single cell sequencing 24
3.3.3 Single cell imaging....25
3.4 Gene expression profiling by microarray 25

3.5 qPCR analysis 25
3.6 Malignancy test –agarose cloning 26
3.7 In silico modeling 26
3.8 Statistics 26
4 RESULTS AND DISCUSSION 27
4.1 Repression of Epstein-Barr virus enhancer Family of Repeats mediated transcription by Oct and Grg/TLE transcriptional regulators with implications for switching between latency programs (Paper I) 27
4.2 Dynamics inside the cancer cell attractor: cell population heterogeneity, stability and escape (Paper II) 28
4.3 Transient phenotypic switch between non-malignant and cancer cell attractors. Paper (III). 30
4.4 Both high and low levels of cellular Epstein-Barr virus DNA in blood identify failure after hematologic stem cell transplantation in conjunction with acute GVHD and type of conditioning (Paper IV) 31
5 UNSOLVED PROBLEMS AND HYPOTHESES FOR FUTURE WORK 34
5.1 What does the stronger negative fluorescent signal in FCM mean? 34
5.2 Subpopulation suicide as a population survival strategy? 35
5.2.1 Asymmetric cell division and self-suicide strategy 36
5.2.2 Delivered among cells or by cell-cell contact 36
6 ACKNOWLEDGEMENTS 37
7 REFERENCES 43

LIST OF ABBREVIATIONS

APC	Adenomatous polyposis coli
ONC	Oncogene
TSG	Tumor suppressor gene
TPG	Tumor progenitor gene
GKM	Gatekeeper mutation
CSC	Cancer stem cells
GRN	Gene regulatory network
iPSCs	Induced pluripotent stem cells
EBV	Epstein-Barr virus
IM	Infectious mononucleosis
BL	Burkitt lymphoma
EBNA	EBV nuclear antigens
LMP	Latent membrane proteins
LCLs	Lymphoblastoid cell lines
PTLD	Post-transplant lymphoproliferative disease
MME	Membrane metallo-endopeptidase
CALLA	Common acute lymphocytic leukemia antigen
ALL	Acute lymphocytic leukemia
ICAM-1	Intercellular adhesion molecule-1
HSCT	Hematopoietic stem cell transplantation
BMT	Bone marrow transplantation
aGVHD	Acute graft versus host disease
TBI	Total body irradiation
FCM	Flow cytometry
FACS	Fluorescence-activated cell sorting
MESF	Molecules of Equivalent Soluble Fluorochrome
STRT	Single cell tagged reverse transcription
FPE	Fokker-Planck equation
GEP	Gene expression profiles
OS	Overall survival
DN	Double negative

1 INTRODUCTION

1.1 Brief overview of cancer

1.1.1 Cancer epidemiology and challenges in cancer therapy

The original meaning of the word "cancer" is crab. This was based on the observation that conspicuous blood vessels feed the tumors and resemble the claws of a crab (1, 2). Hippocrates first used the term *Karkinos* for a swelling or ulcerous formation, even hemorrhoids, whereas he used *karkinoma* for non-healing "cancer".

Paleopathologic findings have demonstrated cancers already in animals of prehistoric times (3). The earliest known written description of cancer (a breast cancer) is from 3000 BC and was discovered in the Edwin Smith Papyrus (3). The Egyptians attempted to treat cancer with cautery, knives, salts, arsenic paste, and the Chinese attempted herbal remedies already more than 3000 years ago. Indians, Persians, Sumerians, and Hebrews also made early attempts to cure the cancer during the same period (3).

Although the long history with big efforts to understand and cure cancer, there are still major obstacles to achieve the goal and thus cancer is still a major challenge to public health. Cancer is a leading cause of death worldwide and this burden is expected to increase (4). In 2012, there were estimated 14.1 million new cancer cases and 8.2 million deaths of cancer worldwide (excluding non-melanoma skin cancers) (4). Breast and lung cancer show the highest incidence and mortality worldwide in women and men, respectively (4). Other frequently diagnosed cancers in women are those of colon and rectum, cervix and corpus uteri, stomach, ovary, thyroid, liver and non-Hodgkin lymphoma. The most common in men are those of prostate, colon and rectum, stomach, liver, esophagus, bladder, non-Hodgkin lymphoma, kidney and leukemia (**Figure 1**). The most common and lethal types of cancer differ between economically developed and developing countries, between gender, races and age. The estimated economic cost of cancer to society was approximately 1.16 trillion US dollar in 2010(5) excluding the substantial long-term costs for care-givers and families.

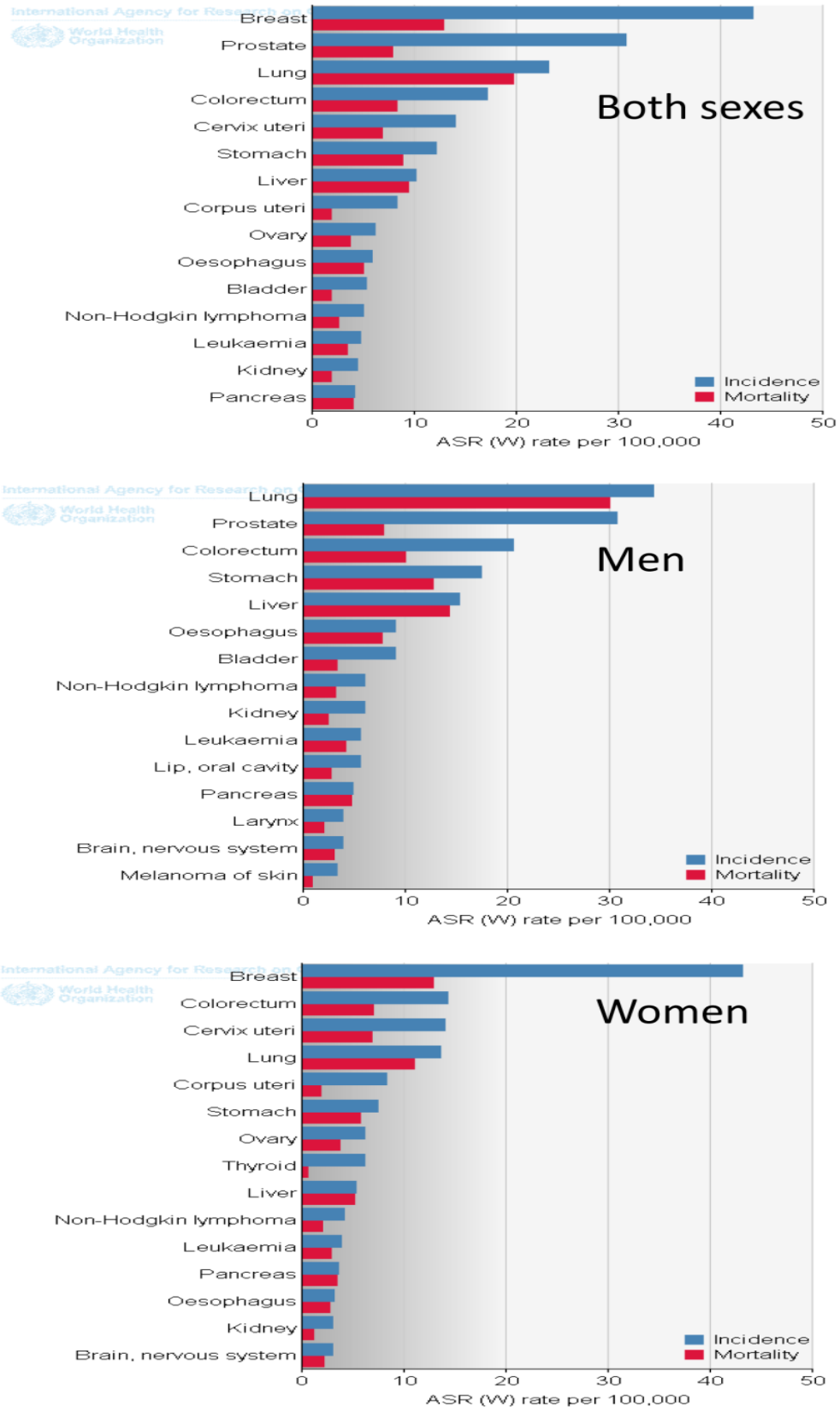

Figure 1. Estimated age-standardised incidence and mortality rates in men, women and both sexes worldwide, 2012. Source: GLOBOCAN 2012. (http://globocan.iarc.fr/Pages/fact_sheets_population.aspx)

During the last decades, the understanding of cancer has improved remarkably and a valuable body of knowledge of cancer biology has been built. However, several key features of tumors biology, such as intra-tumor heterogeneity, and therapy resistance are still controversial, while the understanding of tumorigenesis and tumor progression is incomplete. Although some success stories, cancer treatment is still inefficient. The global annual costs for cancer drugs was approximately 49 billon US dollars in 2011, of which $ 37 billion was linked to treatments resulting in adverse side effects without therapeutic benefit to patients (6). Cancer treatment and care are so expensive that not only developing countries, but also the developed countries, will struggle to manage the spiraling costs (5).

In summary, in spite of the major attempts to understand and cure cancer , the economic burden for and relative failures of cancer therapy, makes cancer one of the major challenges in public health, as well as a life-threatening disease to individual patient.

1.1.2 The current view of cancer biology

1.1.2.1 Hallmarks of cancer

Hanahan and Weinberg summarized the most common features or "programs" affected in cells undergoing tumorigenesis as the eight hallmarks and two enabling characteristics of cancer : sustaining proliferative signaling, resisting cell death, evading growth suppressors, enabling replicative immortality, inducing angiogenesis, avoiding immune destruction, deregulating cellular energetics, activating invasion and metastasis; acquiring genome instability and mutations, tumor promoting inflammation (**Figure 2**) (7). Currently new drugs are being developed to target each of the hallmarks and enabling characteristics(7).

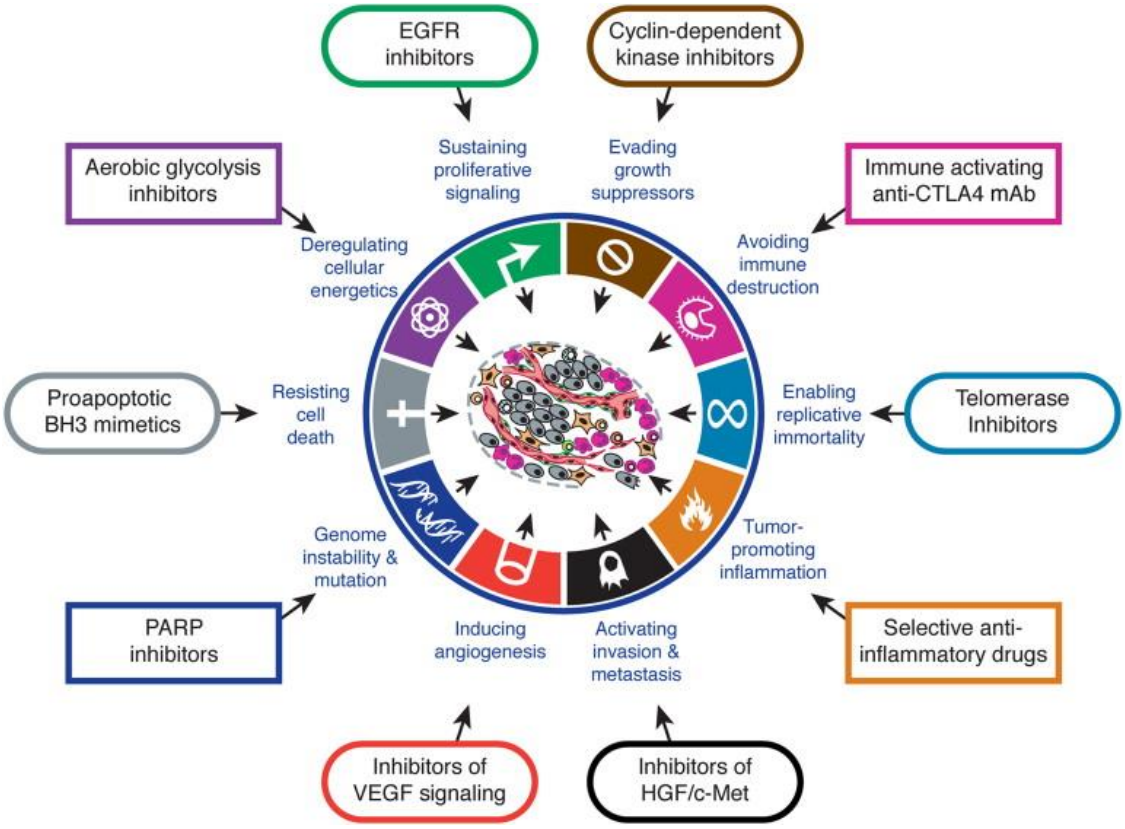

Figure 2. Eight hallmarks, two enabling characteristics of cancer and their corresponding targeting therapeutics (7).

1.1.2.2 Models of tumorigenesis and their challenges

Tumorigenesis is prevailingly understood as a multistep process driven by a sequence of irreversible genetic alterations, each conferring one or another type of survival and growth advantage, resulting in progressive transformation of normal cells into cancer cells. The process is thought to follow a Darwinian evolutionary process at the cell population level (8, 9). This is a DNA centric paradigm, in which the tumor suppressor genes and oncogenes play the crucial role. These key genetic aberrations are the main focus of current cancer drug development.

In line with this Darwinian evolutionary theory, (intra-)tumor heterogeneity is explained by different models, including the clonal evolution model (8) due to genetic and/or epigenetic events (2) and/or by the cancer progenitor/stem cell model.

1.1.2.2.1 The clonal evolution model

According to the clonal evolution model, mutations of genes starting in normal cells occur randomly and accumulate over time. Alterations which provide

advantages for survival and growth make such cells expand to compete out non-affected cells, and other cancer hallmarks as defined by Hanahan and Weinberg, are subsequently acquired (7, 9). By parallel selection more than one neoplastic subclone will be established (**Figure 3**).

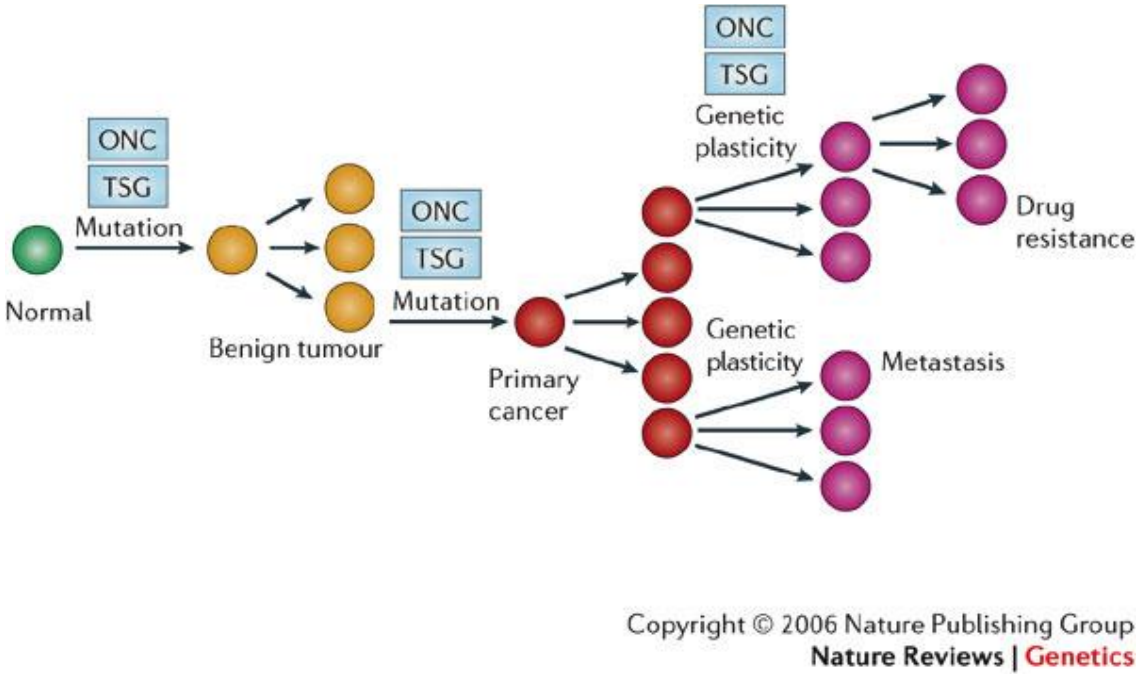

Figure 3. The clonal genetic model of cancer(2). ONC: dominantly acting oncogenes. TSG: tumor suppressor genes.

However, the mutation rate of a specific gene is so low, due to the efficient controls - the DNA repair system and the checkpoints in mitosis – so that the accumulation of mutations is inefficient. It is unlikely that all the mutations required from tumor initiation, to invasion and metastasis could be acquired during a human life span(10). To circumvent this problem, two concepts have been introduced: genome instability and clonal cooperation.

Genome instability enables and accelerates the evolution of subcolones to acquire the hallmarks of cancer (9). Mutations of key guardians of the genome, e.g., p53, is one cause of acquired genome instability and contributes to the mutator phenotype in cancer (10). According to the clonal cooperation model, cells of subclones interact with each other. While most of such interactions are neutral, positive interactions result in mutual benefits between cells by sharing hallmark properties and accelerating tumor progression by the Darwinian selection force (11, 12).

The intra-tumor heterogeneity is then explained by the genetic variation between the selected subclones. Cancer therapy can often only target some of the mutations, i.e., there are subclones with acquired resistant mutations or preexisting resistance, which cause failure of therapy.

Colon cancer is the prime illustration of this model (**Figure 4**). Vogelstein et al established the first linear progression scheme for colon cancer, based on familial polypus colon cancer (APC). During progression of colon cancer, phenotypic changes mirror the genetic (and also epigenetic) aberrations and the continued process follows a predictable order, starting with the loss of (adenomatous polyposis coli) APC, locus specific DNA hypomethylation, mutation and activation of K-ras, loss of a tumor suppressor gene in chromosome 18q and loss of p53. At the phenotypic level the normal epithelium becomes hyperplastic, then small benign adenomas appears, which develops into carcinoma, and eventually become invasive and metastatic (11-13). However, it turns out that almost no cancers really follow such an idealized linear progression model, not even colon cancer. This paradigm suggests a linear deterministic process caused by genes controlling complex phenotypes and that somatic mutations are the main drivers of tumor development. However, during the past ten years cell and tissue biology have undergone a revolution resulting in that new or largely neglected phenomena have to be adopted in tumor biology, such as epigenetic regulation, gene regulation by microRNAs, tumor cell heterogeneity including progenitor cells, the microenvironment including inflammatory processes and the rediscovery of the Warburg effect. This has resulted in new tools for description of the molecular biology of tumors, but also in the discovery of a much larger complexity than earlier anticipated.

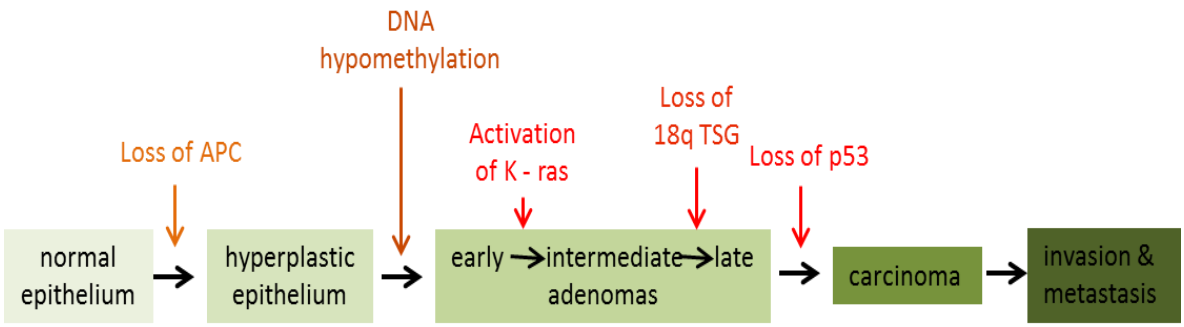

Figure 4. Linear progression scheme for colon cancer (APC). Adapted from 'the biology of cancer'(14).

There are limitations of this model:

1. The Darwinian evolutionary selection is based on phenotype, not genotype. The alterations in a gene do not necessarily cause phenotypic changes. Thus a phenotypic change does not strictly correspond to one specific event in the genome. Thus, there is no 1:1 correspondence between genotype and phenotype.

For example, even in the best defined and original example of this model, hereditary colon cancer, no common mutations responsible for development of invasion and metastasis have been identified. Thus some specific changes in gene expression has no underlying mutation resulting in invasion and metastasis (2).

2. In spite of the big amount of knowledge acquired on suppressor genes and oncogenes and on their signaling pathways, modern cancer drug development aiming at restoring mutated key gene functions to eliminate the tumor has not been convincingly successful, with an efficacy rate below 25% (15).

This linear paradigm seems to be an oversimplified view assuming that targeting some of the most important mutated genes can eliminate all the cancer cells, while ignoring gene expression stochasticity and gene regulatory networks involved.

1.1.2.2.2 The epigenetic clonal and epigenetic progenitor model

The epigenetic clonal model also follows a Darwinian evolutionary process (2). It is not in conflict with the clonal evolutionary model, but adds one more dimension to that model. In this model, epigenetic alterations could be surrogates for genetic changes, i.e., tumor suppressor gene could be silenced by epigenetic control rather than mutation, and oncogenes could be deregulated by epigenetic mechanisms as well. This can result in chromosomal instability and loss of imprinting, like classical mutations do, which also follows a Darwinian evolutionary process (2) (**Figure 5**).

The key difference between epigenetic and genetic mechanisms in tumorigenesis is that epigenetic alterations are reversible while genetic mutations are not (16). However, reversible epigenetic changes can serve as a substitute for irreversible genetic changes. As an example a treatment that reduced the level of heat shock protein 90 (HSP 90) caused morphological defects in flies when genetic modifications were strictly blocked and epigenetic alterations could be clearly confirmed (17, 18).

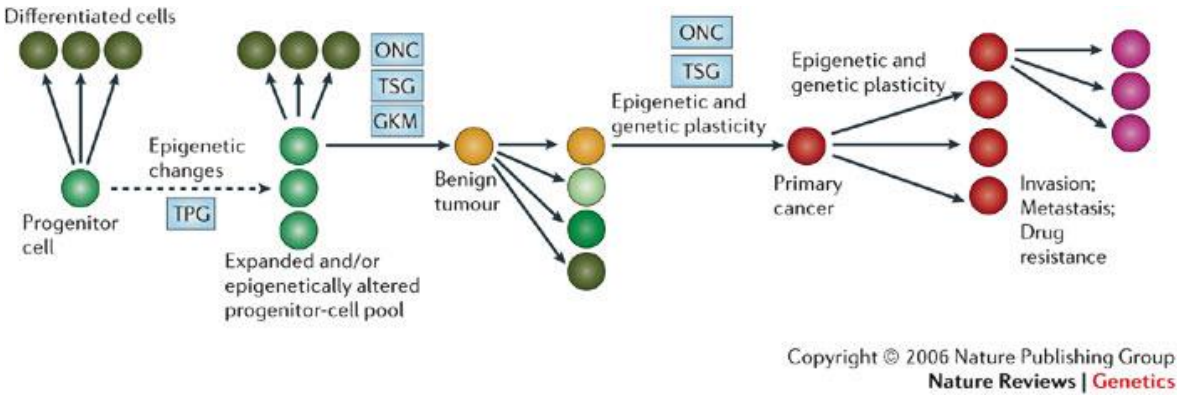

Figure 5. The epigenetic progenitor model of cancer (2). ONC: dominantly acting oncogenes. TSG: tumor suppressor genes. TPG: tumor-progenitor genes. GKM: gatekeeper mutation.

The epigenetic clonal model can also explain a large part of intra-tumor heterogeneity and therapy resistance, due to epigenetic variation in genome control (19).

A special case is the epigenetic progenitor model suggested by Feinberg et al, namely *the first event* in tumorigenesis are epigenetic due to perturbations in the local environment, instead of mutations(2). In fact it has been demonstrated that in rare tumors all the tumorigenic mechanisms are epigenetic, without any genetic mutations, like in ependymomas (20, 21).

Thus epigenetics can be involved at three levels of tumorigenesis: 1. Epigenetic reprogramming to cancer progenitor cells. 2. In subclonal evolution playing a similar role as mutations. 3. Generate epigenetic plasticity and heterogeneity within tumors(2).

1.1.2.2.3 The cancer progenitor/stem cell model

The concept and definition of cancer stem cells (CSC) is still controversial (22-27). The concept of stem cell is well established and clear. It refers to cells with pluripotent capacity to differentiate into cell types which comprise each organ and with the simultaneous ability to perpetuate themselves through self-renewal (27, 28). The differentiation of tissue specific stem cells is highly restricted to cell types within a specific organ, compared to the embryonic stem cells which are totipotent(28).

The idea that tumors might arise from a rare population of cells with stem cell-like properties was proposed about 150 years ago (28). The cancer stem cell

model posits that tumors are hierarchically organized including a small population of cells with the ability to seed new tumors and to spawn non-CSC populations lacking tumor initiating ability(**Figure 6**) (22, 23). CSCs express certain stem cell markers. In the CSC model, a small usually quiescent subpopulation of tumor cells-the CSCs- differentiate asymmetrically and generate both more tumor initiating CSC cells and non-tumor initiating cells. Great efforts have been made to identify specific CSC markers for a particular type of tumor. In some malignancies, such markers have been reported and validated (29, 30).

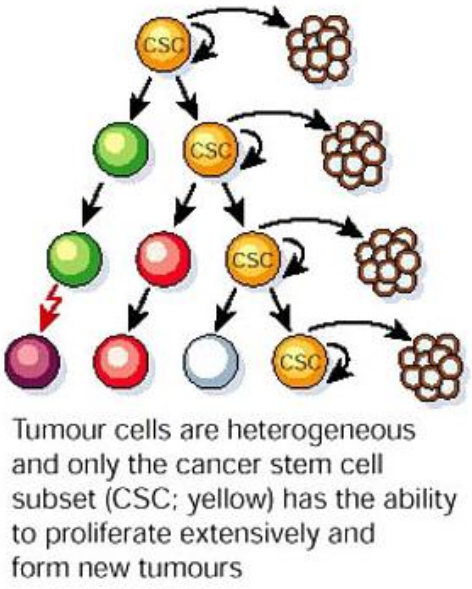

Figure 6. Cancer stem cell model. Adapted from (27).

Intra-tumor heterogeneity according to the CSC model is explained by asymmetric cell division and the ability to generate phenotypically and functionally different down-stream cell generations to the CSCs. CSCs are quiescent or slow proliferating cells. During systemic treatment they might therefore show a weak response to genotoxic damage and give rise to an even more malignant subpopulation of tumor cells during relapse. Once the selective pressure is withdrawn, the relatively therapy resistant CSCs would feed a new tumor (31-33).

The key concern about the CSC model is if such a differentiation hierarchy of CSCs is necessary for tumorigenesis. Although there are several examples which conform to the CSC model (30, 34), there is also good evidence that some types of tumors are not hierarchically organized and that the phenotypic

heterogeneity of the tumorigenic cells can be reversible (22, 35). This is a challenge to the CSC model. To pay tribute to caution, the term “cancer stem cell-like” has been applied recently.

In the model we have explored, the cancer attractor model, we found that the edge cells, which are at the extreme edge (in relation to a specific predetermined marker) of the basin of the attractor, showed slow, or even no proliferation. These cells are transiently at the edge, which means they would be able to leave this extreme state and acquire normal proliferation capacity once they are not at the edge any more. Thus we can conclude that a temporarily existing dormant and phenotypically different subpopulation of tumor cells does not necessarily have to be a product of a hierarchical organization, but can also arise due to biological noise and to variations in the gene regulatory network (36-39).

1.2 Cell attractor

Although efforts have been made to adapt the current “modern” cancer paradigm to the insights about complexity in cell and tissue biology, it is now more of a repair patchwork than a model with an integrated perspective (7, 9). The entry of tumor biology into the realm of complex (cellular) systems can not be neglected any more. It will be necessary to find new, additional novel approaches to understand cancer processes in order to improve clinical developments. This is a major challenge, as the science of complex system is lagging in providing useful models and tools. Even the scientific nomenclature and the definitions in this field of science are insufficiently developed.¨

1.2.1 Origin of the idea

Max Delbrück first pointed out in 1949 that ‘many systems in flux equilibrium are capable of several different equilibria under identical conditions. They can pass from one state to another under the influence of transient perturbations’ (40, 41). The ability of small gene regulatory circuits to produce more than one stable equilibrium state was also later proposed by Jacob and Monod, to explain that differentiation could result in a multitude of stable phenotypic states (42,

43). In the 1960s, Stuart Kauffman using Random Boolean Networks showed that when connected by Boolean functions, hundred thousands of mutually regulating genes would be able to establish hundreds of stable equilibrium state, termed attractors (44-46). Huang and Kauffman suggested that such an attractor would correspond to a gene expression pattern associated with a cell type (47).

There is another cell regulatory model relating to phenotypes, designated the 'independent attractor' model by Baverstock. It arises from 'the interactions between active gene products (mainly proteins) and independently of the genomic sequence'(41). The difference between these two 'attractor' models has been discussed in detail by Baverstock (48). The 'independent attractor' model is still very hypothetical and its biological meaning still awaits further exploration and validation. In this thesis, the 'independent attractor' model will not be discussed further.

1.2.2 The conceptual development

Though it has been decades since the concept of cell attractors was originally proposed, the development of this field has been slow and the impact of the idea has had a low impact in cell biology and tumor biology. Still at the stage of hypothesis, most of its applications have been based on *in silico* modelling while very few in vitro or in vivo experiments have been performed with the aim to explore and validate the theory. Most publications are reviews, opinions or perspectives while the original research papers are few.

To explore whether it is helpful to view cells as biological attractors would be an important step in conceptual consolidation. For this purpose network attractors would be applied as an ideal model to describe fundamental principles governing cell behavior and the appearance and maintenance of distinct cell types.

1.2.2.1 Boolean network

Using Boolean networks, the activation states of individual genes are described by a simple "on-or-off" parameter and studied in computer simulations with varying complexity. Given a set of simple rules, one gene network can assume a

very large number of theoretical positions in its n-dimensional state-space coordinate system. Using the reduction of levels of gene expression to two states (on-off;1-0) as digitals in Boolean algebra simulations, the number of stable states (termed attractors) that the network can assume is very limited, compared to the immense number of possibilities. This "conceptual model" represents fundamental and suggested features of the functionality of real gene regulatory networks.

1.2.2.2 The gene regulatory network and the attractor

A gene regulatory network (GRN) is established based on the fact that genes influences the expression of other genes through molecular regulatory interactions (49). It refers to a hard-wired architecture and is quite robust but might change as a consequence of mutations (47, 49). It is the wiring of the intracellular network that makes the system robust and compensates for the variations so that the cells can maintain their identity (cell type) and function in spite of the variations.

The expression level of the genes changes over time, collectively manifested as the change of the gene expression patterns, which is referred to as 'network dynamic'(49).

Based on the GRN concept, an attractor is seen as a point in space of the n-dimensional coordinate system to which objects in space seem to be attracted, where n is the total number of genes. An attractor of a GRN represents a virtual state of the expression level of all genes of a cell type. If there was no inter-cellular heterogeneity, the cells of the same type would be in exactly the same point, the attractor. However, in reality data shows that most individual cells representing one cell type will not be positioned at the point of the attractor, but rather in a more or less distributed cloud around the attractor due to a conspicuous heterogeneity of gene expression levels (**Figure 7**). The closer to the attractor, the higher density of cells there will be. The area of this cloud is designated the basin of attraction of the attractor. In this domain the gene network can operate in a relatively stable manner and is only perturbed by intrinsic fluctuations and responses to small, random external influences. In other positions the network is unstable. All these basin of attractors constitute together a virtual landscape with valleys and ponds separated by hills and ridges. So an attractor is a stable gene network state and represents one type of

cells, and all these basin of attractors representing the different cell types constitute together the virtual landscape (36, 43).

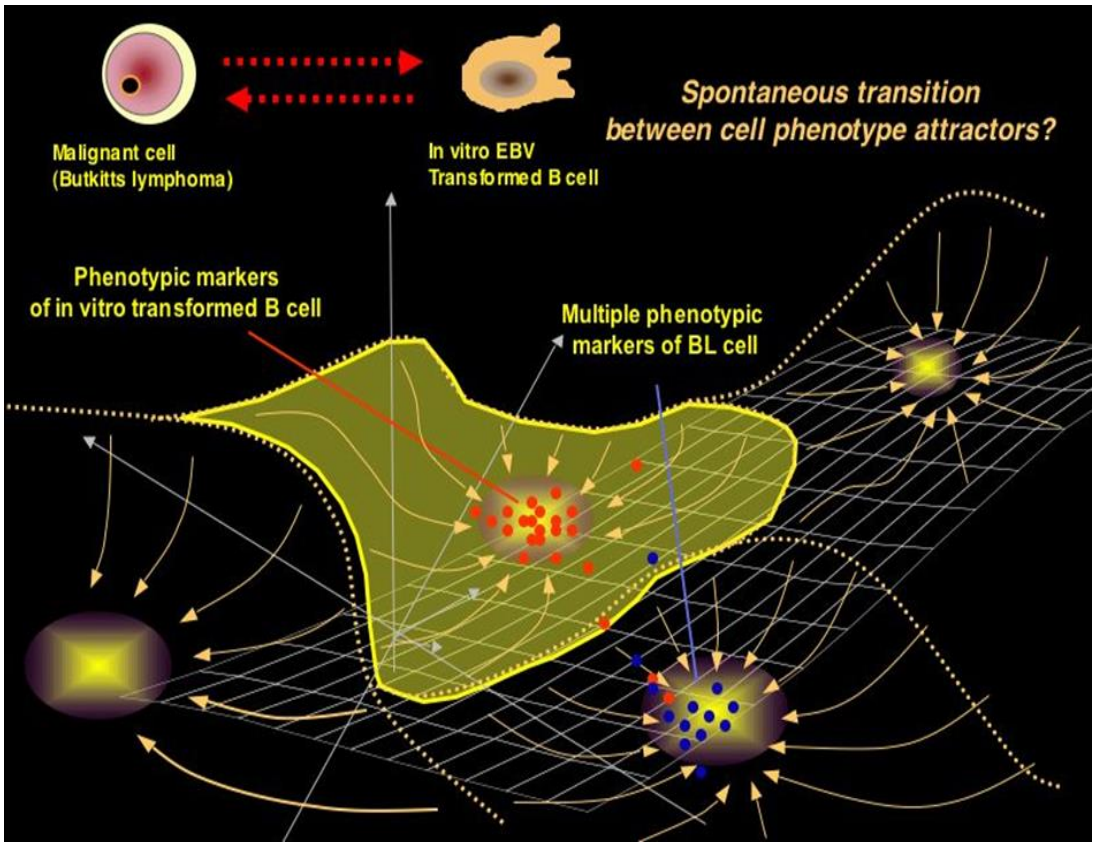

Figure 7. A schematic representation of the cancer cell attractor model.

1.2.2.3 The quasi-potential landscape, Waddington's epigenetic landscape, and network entropy

In the cell attractor model, each dimension in the n-dimensional space (n= number of genes in the genome) represents the expression level of one gene, respectively. Huang proposed that Waddington's landscape is more than a metaphor by computing and interpreting the quasi-potential landscape, in which n dimensions have been compressed into two and the quasi-potential U has been introduced as one dimension, based on his demonstration that the relative stability of a network state could be computed and represented as a quasi-potential energy *U* of each state *S (49, 50)*. The combined quasi-potential landscape of basins of attractors have superficial similarities with Waddington's epigenetic landscape suggested 60 years ago (**Figure 8**) (43, 51).

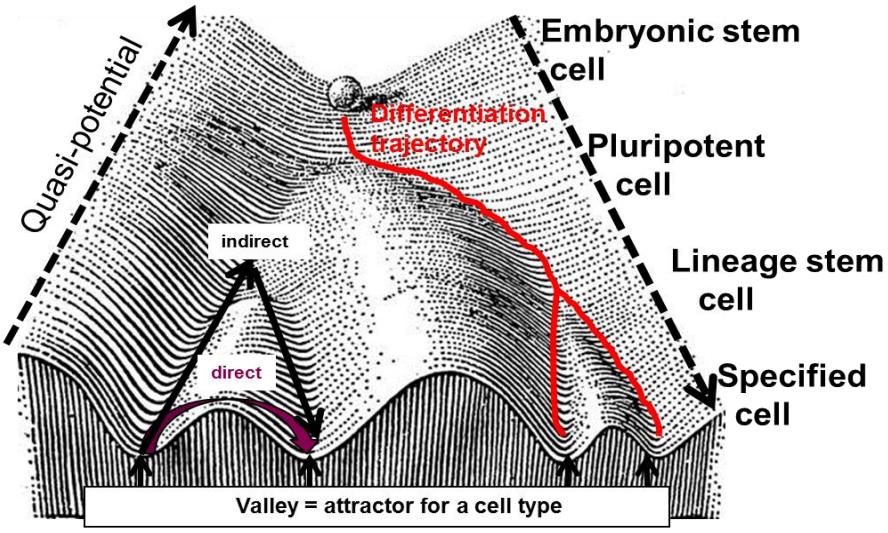

Figure 8. Waddington's epigenetic landscape, attractor and differentiation. Adapted from Waddington, 1957(51).

In the epigenetic landscape, pluripotent stem cells initially at the highest point in the landscape "roll down" though the valleys, to the bottom. This represents the development of the embryonic cells to the mature cell types. Development is the process when cells move to its 'lowest energy' attractors.

Erwin Schrödinger described living systems as 'feeding on negative entropy'(52). It has been demonstrated that entropy rate decreases during the differentiation and the cellular network entropy could reflect the energy potential of a virtual biological landscape like Waddington's differentiation landscape (53).

1.2.2.4 The cancer cell attractor

In this state space-landscape model of a gene network there might also be an unknown number of possible but normally unexploited positions. It has been suggested that such 'illegitimate' positions can be entered by cancer cells (43), hence the cancer cell attractor model. One ultimate question is whether cancer precursors can be derived from a normal heterogeneous cell population?

Edge cells- those at the edge of the basin of the attractor - according to the conceptual model are transitory, unstable states, and will either quickly find one of the adjacent attractors or disintegrate resulting in cell death. The characteristics of the edge cells, including higher mortality and significantly slower proliferation of the edge cells has been validated (36). We hypothesize that the edge cells may rarely be able to jump to a neighboring legitimate or

illegitimate attractor. The entering of a legitimate attractor results in cell type switch and to an illegitimate attractor results in tumor initiation.

In the cell attractor model, mutations are suggested to change the GRN and affect the 'height' of the ridges between the valleys so that access to normally inaccessible attractors, including malignant cell attractors, can occur more easily(49). Thus, mutations would neither be a sufficient nor a necessary condition for tumor initiation, but by modifying the local landscape, they could promote or accelerate tumorigenesis.

1.2.3 Potential applications of the cell attractor

Besides being one novel model of tumorigenesis, the 'cancer cell attractor model' as discussed above, the cell attractor model can also be applied as a tool to study aspects of complexity in cell biology:

1.2.3.1 The cell attractor as a tool to study cellular heterogeneity

The same GRN, due to biological fluctuations and regulation between genes, can generate different transcriptomes(43, 49) which correspond to the different cellular phenotype (54).

1.2.3.2 The cell attractor as a model for cell differentiation and cell phenotypic switch

In the epigenetic landscape based on GRN, the pluripotent cell moves down along the valley (i.e., trajectory of differentiation) to the bottom (i.e., the mature stable cell type). This is a non-reversible process like water falling down from a higher position to a lower. Thus, normally once the cell is committed to one valley at bifurcations (or intersections), it cannot go back and enter other(s). According to the cell attractor model, induced pluripotent stem cells (iPSCs) (55, 56) could be viewed as products of forcing the mature cell types upwards from the bottom by external perturbations (by adding Oct3/4, Klf4, Sox2 and c-Myc) and regain stemness. This forced process drives the mature cells back to an intersection in the landscape so that they may enter another valley and switch to another cell type; or it may lead them to a normally inaccessible illegitimate path, normally not being exploited by evolution, resulting in establishment of a malignant cell attractor. This hypothesis is supported by the fact that the process

of iPSC derivation shares many characteristics with cancer development and it has been reported that premature termination of reprogramming leads to cancer development through altered epigenetic regulation (20, 57-59). Tumor cells may mimic processes in iPSCs by alterations of key regulating transcription factors, which might be an explanation why many tumor cells show stemness.

Thus, the cell type switch could be achieved by either reversion to a more pluripotent state and entering another trajectory (indirect way, Figure 8), or by jumping over the ridge between two valleys (direct way, Figure 8) when its height seems reduced by mutation in the presence of sufficiently high levels of fluctuations or in the response to deterministic regulatory signals (36, 60, 61)

1.2.3.3 Cancer cell attractors as a tool to understand therapy resistance

First, according to the model, the inter-cellular heterogeneity is an inherent property of any type of cells, due to the fluctuations of the GRN. Secondly, the GRN is dynamic, so even within the same cell, its phenotype changes over time, within its basin of attraction. Thirdly, as a property of the attractor, there is a trend that the cells move to their own attractor and any selected subpopulation has the ability to re-establish the whole parental population (36). With this in mind anti-cancer treatment cannot efficiently kill all the subpopulations of the tumor cell population at the same time. Tumor cells which are not killed after treatment have the ability to re-establish the whole parental tumor population again, which leads to relapse. Also, the treatment itself could be a potential stress for the tumor cells and enhance inter-cellular heterogeneity(62). Thus, compound screening based on a network view (63), and tumor-specific therapy taking into account the cancer cell network concept (64) have been discussed. Network medicine, defined as 'combination drugs that interfere with disease network', has been suggested as a promising future direction for therapy (6, 36, 64, 65).

1.3. Epstein-Barr virus biology

1.3.1 General introduction of EBV

Epstein-Barr virus (EBV) (also known as Human herpesvirus 4), is a human-specific gamma-herpes virus. The EBV genome is a double-stranded, 172kb

linear DNA molecule (66, 67). It is one of the most common viral infections in humans and more than 90% of the adult population worldwide are EBV carriers. After primary infection with EBV, humans become EBV carriers for life time (68, 69). The primary infection by EBV in early childhood is often asymptomatic. Delayed first EBV infection may result in a benign, self-limiting disease, called infectious mononucleosis (IM), which is more common in developed countries than in developing areas.

EBV was first discovered in Burkitt lymphoma (BL), a B cell-derived lymphoma (70). Soon after its discovery, the in vitro transforming capacity of EBV was demonstrated (71). The in vitro EBV transforming system was proven to be useful in studies of the viral transforming mechanism and immune response against the virus carrying cell. In this thesis, I have exploited malignant vs. non-malignant EBV infected cell systems.

1.3.2 EBV latency types in relation to tumors

Based on the expression pattern of EBV nuclear antigens (EBNA) and latent membrane proteins(LMP) and also on the usage of different viral promoters, EBV latent infection is classified into latency III, latency II, latent I/0 and Wp-restricted type (found in some rare BL). EBV infection has been proposed to be an epigenetic driver of tumorigenesis(72), e.g., in gastric cancers(72-74).

1.3.2.1 Latency III and latency I/0

Upon in vitro EBV infection of B cells, the virus establishes a latent infection and the transformed B lymphocyte will proliferate continuously, giving rise to lymphoblastoid cell lines (LCLs)(71). Such LCL expresses six EBNAs (EBNA 1, EBNA2, EBNA 3 (3A), EBNA 4 (3B), EBNA 6 (3C) and EBNA 5 (-LP)) and 3 LMPs (LMP-1, LMP-2A and LMP-2B). This pattern of gene expression is referred to as latency III. In contrast, in latency I, only one virally encoded protein, EBNA1 is expressed, occasionally two (LMP2A). EBNA1 is essential for replication and maintenance of the viral episome (75). The different gene expression patterns are achieved by the use of different promoters to generate alternative primary transcripts: in latency III LCLs, one of two upstream promoters, either Cp or Wp, in the BamHI C or W region of the genome is exploited whereas in latency I, EBNA 1 expression is driven by Qp (75). The

latency I type does not induce proliferation and its phenotype corresponds to non-activated B cells (66). In the asymptomatic EBV persistence in healthy carriers, no EBV antigens may be expressed. This is referred as latency 0.

Latent EBV infection is associated with many kinds of malignancies and complication in immunosuppressed patients. Latency III EBV infection is associated with post-transplant lymphoproliferative disease (PTLD) (66). The latency III program is expressed only in B lymphocytes and latency III type cells can exist only during the acute phase of primary infection, before the EBV specific T cell response develops and in patients with impaired immune functions, e.g., immunosuppressed stem cell transplanted patients(66).

Latency I infection associated malignancy includes BL(75). BL seems to arise from the malignant transformation of germinal center B cells and it is classified into 3 types: the endemic BL (almost 100% EBV positive), the sporadic BL (around 30% EBV positive) and AIDS-associated BL (25-40% EBV positive). The etiology of BL is associated with EBV infection, but a key event in the pathogenesis of BL is a chromosomal translocation on chromosomes 14, 22 or 2 and the c-myc oncogene on chromosome 8, resulting in deregulated expression of the c-myc protein. Alteration in c-myc expression is a characteristic of all BL tumors including those that are not EBV associated (75, 76). It has been shown that MYC contributes to the phenotype of BL cells by upregulation of CD10 and CD38 and down regulation of activation markers (77). Although EBV carrying BL patients are immunocompetent, the BL cells can escape the immune system surveillance due to that they have only one EBV protein, EBNA1 expressed, but no other co-stimulatory surface molecules and EBNA 1 does not serve as target for CD8+ CTLs(66).

In our study, CBM1-Ral-Sto (CBM1), a typical LCL and Rael, derived from BL were used to represent the latency I and latency III switch model system. The characteristics of the two cell lines are as follows (**Figure 9**):

Figure 9. Comparison of latency III (CBM1) and latency I (Rael) EBV infected cells.

1.3.2.2 Latency II

EBV latency II cells express EBNA 1, LMP1, LMP 2A and LMP 2B, and are associated with the classical Hodgkin lymphoma (HL), nasopharyngeal carcinoma (NPC) and nasal NK/T lymphoma.

1.3.2.3 CD10 and CD54

CD10 and CD54 are two perfect markers to distinguish latency I and latency III cell phenotype. EBV latency I type cells are supposed to be CD10+CD54- and latency III cells are CD10-CD54+.

CD10 (membrane metallo-endopeptidase, MME), is a common acute lymphocytic leukemia antigen (CALLA). It is an important cell surface marker in the diagnosis of human acute lymphocytic leukemia (ALL). This protein is presented on leukemic cells of pre-B phenotype. However, it is not restricted to leukemic cells, but also found on a variety of normal tissues. It has been shown that CD10 expression relates to differentiation potential and stage(78).

CD54 (intercellular adhesion molecule-1, ICAM-1) is a cell adhesion molecule distributed among normal and neoplastic tissues(79). It plays a key role in cell-cell interaction leading to immune response. CD54 is predominantly expressed

in the early stages of normal hematopoiesis and during the activation of blood cells, and it is also expressed on neoplastic cells from 'stem cell derived' neoplasms(79).

1.3.3 EBV load as a biomarker in stem cell transplantation

1.3.3.1 Overview of stem cell transplantation

Hematopoietic stem cell transplantation (HSCT) or bone marrow transplantation (BMT) is a well-established treatment for hematologic malignancies. The survival after the HSCT is expected to be superior compared to use of only conventional chemotherapy (80). Patients receiving bone marrow (BMT) or hematologic stem cell transplants (HSCT) show a considerable risk to develop EBV-associated post-transplant lymphoproliferative disorder (PTLD) and lymphomas, which is a potentially fatal disorder associated with the EBV infection (81-83). Relapse and acute graft versus host disease (aGVHD) are two other major causes of death of patients undergoing HSCT. The most commonly applied strategy for ablative therapy in preparation for allogeneic HSCT has been cyclophosphamide combined with total body irradiation (TBI) (80, 84). Higher dose of TBI and /or chemotherapy could reduce relapse, but increase treatment toxicity, which could also be a cause of death. aGVHD is a frequent and at times severe inflammatory complication after HSCT. Efforts to diagnose, prevent and treat it are important for better survival of the HSCT patients(85).

1.3.3.2 EBV load together with T cell phenotyping as biomarkers in stem cell transplantation

EBV DNA levels in blood reflect the intricate and complex balance between EBV and the host. EBV load has been proven to be a useful biomarker associated with prognosis after HSCT and solid organ transplantation (81, 83, 86-95).

After transplantation the levels of EBV DNA-load in blood is affected both by immunosuppressive treatment and immune stimulatory mechanisms, like aGVHD(83, 91).

The importance of T cell mediated immune responses in maintaining balanced viral persistence is emphasized by the fact that patients with T cell dysfunction are at risk of developing EBV associated PTLD (96, 97). EBV specific T cells play an important role in the long term control of the EBV carrier state and it has been reported that PTLD may be prevented or even cured by administration of donor derived EBV-specific cytotoxic T lymphocytes(81, 82, 98).

2 AIMS OF THIS THESIS

The overall purpose of the studies presented in this thesis was to explore the cancer cell attractor using the EBV latent infection model.

More specifically, the thesis aimed to:

1. Study the molecular basis for switch between EBV latency I to latency III.

2. Examine the cell attractor model generated from in silico work and its biological relevance by in vitro experimental testing.

3. Study inter-cellular heterogeneity and cell type maintenance based on the cell attractor model.

4. Characterize the dynamics of cancer cell attractor at single cell resolution.

5. Explore the characteristics and fate of edge cells in the cell attractor.

6. Simulate the cell population dynamics of cell attractors in silico with Fokker-Planck equation.

7. Investigate EBV load together with T cell phenotyping and evaluate its prognostic value after hematopoietic stem cell transplantation.

3 MATERIAL AND METHODS

I will here primarily discuss the methods which are more specific for my studies. Otherwise I do refer to the Methods descriptions in the separate papers.

3.1 Clinical samples

The characteristics of the patients are summarized in table 1. Blood samples were collected from the patients at 1, 2, 3, 6 and 12 months after HSCT. Fifty-one patients were included. Seventeen were children (≤ 18 years old) and 34 were adults. The study was approved by the Stockholm Ethical Committee South 2010/760-31/1. In the case of children consent was also obtained from parents or legal guardians (on file at Center for Allogeneic Stem Cell transplantation, CAST, Karolinska University Hospital, Huddinge).

3.2 Cell lines

Six human Epstein-Barr virus (EBV) carrying B-cell lines were used in this study: DG75 is an EBV negative Burkitt´s lymphoma cell line(99). Rael is an EBV positive Burkitt´s lymphoma (BL) cell line with a latency I expression pattern (100). The CBMI-Ral-STO cell line was obtained by *in vitro* infection of cord blood cells with virus rescued from Rael and has a latency III expression pattern (101). Mutu I (clone 148) is a BL derived cell line with a type I phenotype. Mutu III (clone 99) was obtained by in vitro culturing of Mutu I and the cells have a type III phenotype. Akata is an EBV positive BL cell line that spontaneously loose EBV forming the EBV negative Akata-N. Fresh EBV transformed B cells (fresh LCL) was also applied (Paper III) : Resting human B-cells from “buffy coat” or from total blood was purified by affinity separation using magnetic microbeads loaded with specific antibodies against the B-cell marker CD19 and infected with the B95-8 laboratory strain of EBV.

The cell lines were maintained as suspension cultures in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS), L-glutamine, streptomycin and penicillin.

3.3 Single-cell resolution methods

3.3.1 Flow cytometry (FCM)

Multi-color FCM was applied to study the heterogeneity of cell surface proteins and their dynamics. With this method, expression of more than 12 proteins from each cell could be recorded at the same time. One key step in multi-color FCM is fluorescence signal compensation. It is required when two or more dyes are attached to cells and their ranges of color emission overlaps (spectral overlap). The more fluorescence dyes were applied, the more complicated and challenging the compensation would be. Fluorescence-activated cell sorting (FACS) were applied to select and isolate cells. Based on FCM, Quantum MESF (Molecules of Equivalent Soluble Fluorochrome) beads were applied to correct for the effect of day-to-day fluctuations of the flow cytometer. The cell apoptosis and death analysis was done with APC Annexin V/ Dead Cell Apoptosis Kit. The cell tracing and proliferation analysis was done by CFSE staining with the CellTraceTM CFSE Cell Proliferation Kit(102) (paper II, paper III). T cell phenotyping was also done by multicolor FCM (paper IV). FCM data was analyzed by Flowjo.

3.3.2 Single cell sequencing

In this study, the number of edge cells that could be collected is limited, thus, we applied single cell analysis. A single cell gene expression profiling, single cell tagged reverse transcription (STRT) method, combined with a microfluidic sample preparation, Fluidigm C1, was applied. In this STRT-C1 method, a 5 bp short random degenerate sequence was introduced to count the single cell molecule and remove the amplification bias(103).

The obtained sequence reads were mapped to a concatenated reference genome made up of the human reference genome (UCSF's Human Genome Assembly Release hg19) and ERCC's synthetic transcript controls using STAR version 2.3.0 (104). The quality of sample libraries were evaluated for complexity, sequencing error rates, reproducibility, exon/intron mapping ratio, sense/antisense ratio. For the high quality libraries (at least 10,000 mapped reads and 500 detected genes), gene expression levels were computed based on uniquely mapped reads in terms of read counts and normalized expression calculated in rpm (reads per million mapped reads). Samples were clustered

using principal component analysis based on the log-transformed expression values and the distances between different sub-samples were computed based on the Euclidean distances of the transcriptomes. Genes with significant PCA loading were identified using a jackstraws randomization approach (105).

The computational analyses were done using the STRT pipeline, in-house developed python and R scripts.

3.3.3 Single cell imaging

Cells isolated by FACS were placed in a30000 well nanochip. After 30 min of incubation most of the cells were trapped in the wells and forced to go down by gravity. Being distributed randomly most of them where single in a well. Minority of wells contained two, three or more cells. Those wells were excluded from analysis. The method in detail has been published(106).

Imaging of the cells in nanochips was performed with Hexascope. (Qantascope Biotech,Sweden).
(http://www.qantascope.com/products/instruments/hexascopehtp).

3.4 **Gene expression profiling by microarray**

Larger amount of RNA was required for microarray analysis, compared to single cell STRT-C1 method. In this study, we applied this method together with GEDI program to figure out the gene expression patterns of a cell population over time to study its dynamics.

The GEDI maps were generated using the program GEDI (http://www.childrenshospital.org/research/ingber/GEDI/gedihome.htm) which is applied for visual representation of global gene expression based on self-organizing maps. In this study, we were more interested to know the whole gene expression patterns than to find out differentiated genes, thus, GEDI is a useful tool.

3.5 qPCR analysis

In this study, qPCR analysis was applied to validate the microarray data (Paper II), and also to evaluate the EBV DNA load in HSCT patients (Paper IV).

3.6 Malignancy test –agarose cloning

The loss of contact inhibition and anchorage independence are nature of the tumor cells. Normal cells stop to grow due to cell-cell contact in many cases - tumor cells do not stop to grow when they meet neighboring cells. This was devised to test the breakdown of barriers and anchorage independence - *the cloning procedure in soft agarose.* In our study, we applied this method to test if the edge cells from non-malignant LCLs would transiently adopt malignant properties or not.

3.7 *In silico* modeling

The in silico modeling based on Fokker-Planck equation (FPE) describes the time evolution of a probability density function, PDF, f(X,t) under the combined influence of drift (corresponding to deterministic force of relaxation to the attractor basin center) and diffusion (corresponding to gene expression noise).The observed dynamics of relaxation of the edge cell fractions or the population segments was explored by FPE. The mathematical task was to determine the functions A(x) and B(x) in a one-dimensional FPE from observations of f(x, t):

$$\frac{\partial f(x,t)}{\partial t} = -\frac{\partial}{\partial x}(A(x)f(x,t)) + \frac{\partial^2}{\partial x^2}(B(x)f(x,t))$$

Several variations of the basic model have been investigated to exemplify that the complex dynamics of cell populations.

3.8 Statistics

Overall survival was analyzed using the Kaplan-Meier method and compared with the log-rank test. Uni- and multivariate analysis of factors associated to survival was performed with the Cox proportional hazards model. Factors with a p-value <0.2 in the univariate analysis were included in the backwards elimination multivariate analysis. Continuous variables were compared with the Mann-Whitney test and categorical variables with the Fisher exact test. Analysis was performed with the Statistica software (Statsoft, Tulsa, MN, USA). In cell viability analysis, paired student t test was performed to compare the differences between groups by GraphPad Prism 6.0.

4 RESULTS AND DISCUSSION

4.1 Repression of Epstein-Barr virus enhancer Family of Repeats mediated transcription by Oct and Grg/TLE transcriptional regulators with implications for switching between latency programs (Paper I)

We consider the switch between latency I and latency III to be crucial in vivo, as it controls the proliferative capacity of EBV-infected/-carrying B-cells and thus also the level of EBV infected B lymphocytes in blood and lymphoid tissues. In normal B cells latency I cells are resting and latency III cells are proliferating. We have earlier shown that the switch depends on two viral promoters Cp and Qp and their regulatory elements, especially the enhancer for Cp designated FR (107).

In this study, an EBV latency I to latency III switch was proposed: EBNA1 has higher affinity to FR than Oct-2 has, so in latency III where the EBNA1 expression is high and Oct-2 is low, EBNA1 occupy most of the binding sites to FR-enhancer, while no or few Oct-2 molecules bind to this DNA sequence and as a result the transcription from the C promoter active. In latency I where the concentration of Oct-2 is high enough to compete with EBNA1， it will occupy the binding sites for EBNA1, and then the protein Grg which converts Oct-2 to a repressor binds to Oct-2 and the C promoter is silenced, while “constitutive” C promoter takes over (**Figure 10**).

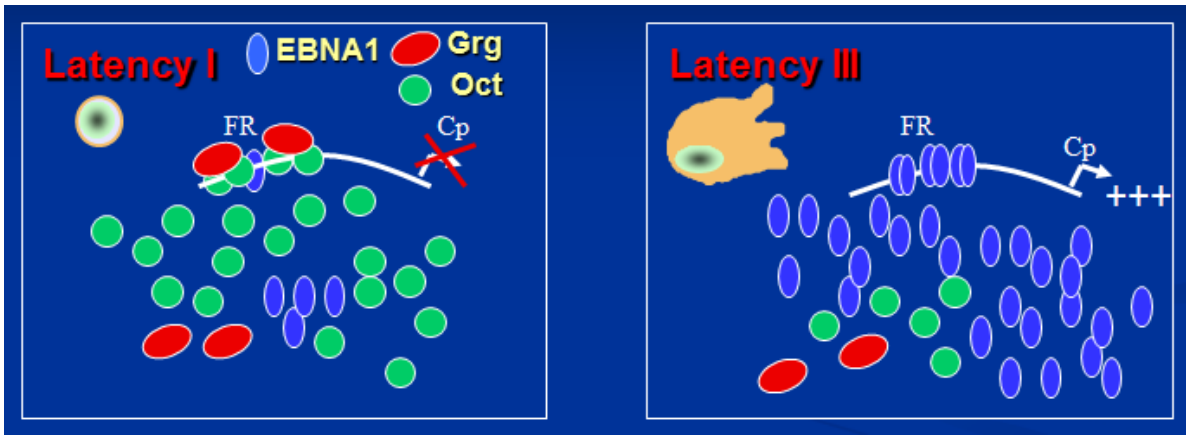

Figure 10. Latency I to latency III switch model.

We have earlier validated that Levels of Oct-2 vary between latency I and latency III cells: high in latency I and low in latency III, while EBNA 1 shows the reverse. We demonstrated that all members of Grg/TLE could repress Oct-2

induced promoter activity through FR. This repression could be reversed by EBNA1. We suggest that Oct-2 and Grg/TLE-proteins could be involved in promoter switching, and as a consequence in latency switch. This switch will in turn have dramatic consequences for the fate of the infected B-cell, determining whether to proliferate or rest.

4.2 Dynamics inside the cancer cell attractor: cell population heterogeneity, stability and escape (Paper II)

Exploiting EBV latency I cells (Rael cell line, Mutu I, malignant) and a latency III system (CBM1 cell line, non-malignant or Mutu III, malignant), we explored the cell attractor model and studied its characteristics.

Protein expression in clonal cell populations showed apparent heterogeneity and edge cells dynamically change their marker protein expression profile.

The expression level of surface markers from the Rael and CBM1 showed a characteristic bell-shaped histogram with a broad basis. Edge cells, from the tail of the histogram, showed their ability to reestablish the parental distribution after isolation. They also showed reduced proliferation and viability. Also, isolated edge cells of Rael show an mRNA expression profile distinctly different from that of parental Rael.

Our data suggest that edge cells may show ambivalence either transgressing the ridge between two basins or returning to their original attractor. In an organized microenvironment they might merely not survive. However, with tissue stress and disorder, such as during low grade chronic inflammation, they might have a larger chance to survive and thus could complete transition to neighboring attractors. This essentially would explain the adoption of malignant, cancer stem-cell like phenotypes due to non-genetic plasticity and not mutations, as others have suggested (2, 62, 108).

As from the modelling work of cell attractor and also from our in vitro experimental data, the transition of the edge cells to neighboring attractors would be an exceptional phenomenon as it would violate the ordered structure of cell phenotypes and tissues in a multicellular organism. Indeed, the

robustness of attractor states, as described here, suggests that such spontaneous transitions are extraordinarily rare.

Subpopulations representing different levels of marker expression can re-establish the parental marker distribution upon separate culturing.

The central finding – directly predicted from the concept of the attractor basin - is that any isolated subpopulation derived from the original spectrum of the expression levels of a marker can reestablish within days to a few weeks the original parental distribution with respect to not just that marker used to define the subpopulations but the entire transcriptome. Isolated edge cells reassumed properties of cells in the middle of the population distribution, whereas cells at the center (mode) of the distribution moved away from the middle towards the edges, thus, in both cases "diffusing" (in state space) throughout the entire population space.

One implication is that when the selective pressure of treatment is released, the original population can be reestablished by these few temporarily resistant edge cells, which may be one explanation for therapy resistance.

Single cells "move around" in the attractor basin over time.

We demonstrated that cells in the attractor are not static and they change their positon in the attractor dynamically over time. The wide distribution of single cells in a cancer cell attractor may result in that some cells in the population are resistant to a given drug at a certain time point, but not at another time point when they have shifted position in the state space (22, 26, 31, 63, 108, 109).

Spontaneous shift from one phenotype to another is demonstrated. A key regulatory gene can facilitate a switch between attractors.

Early during switching from Mutu I to III, some Mutu III cells resumed their original Mutu I phenotype as revealed by CFSE labelling. After apparent completion of the switch to the Mutu III phenotype, we occasionally detected a small subpopulation of CD10highCD54low (type I) cells at variable time points. Thus Mutu cells seem to oscillate spontaneously between the two "attractors".

By knock down Oct2, a small fraction of Rael cells transiently switch from CD10+CD54- (type I) to CD10-CD54+ (type III) phenotype.

Two virtual forces identified by mathematical modeling of the population re-establishment by Fokker-Planck equation.

The description of the cell population distribution relaxation using a simple FPE mathematical model suggests that the dynamics of cell states around a cancer cell attractor is reasonably represented by two virtual forces, a deterministic drift force and a diffusion term that captures the stochastic fluctuations due to molecular noise that affects gene/protein marker expression.

4.3 Transient phenotypic switch between non-malignant and cancer cell attractors (Paper III)

In line with our exploration of cancer cell attractors on Paper II, we deepened the analysis of gene expression profiles (GEP) in small distinct subpopulations representing a cross section of the basin of attraction covering both type I and type III cells and compared their biological functions.

We have demonstrated inter-cellular heterogeneity of the clonal cell population at protein level by measuring surface markers and at mRNA level by microarray in Paper II. Here we now explore the mRNA levels (transcriptome) by an RNAseq method designed for single cell analysis. Gene expression profiles show that the $CD10_{high}CD54_{low}$ edge cells from CBM1 is a distinct population and resemble Rael more than its parental population, CBM1, as analyzed by hierarchical clustering based on overall gene expression patterns. This was also evident by calculating the Euclidean distance between the edge cells to Rael and CBMI, and also between Rael and CBM1 cells, and finally also shown by principle component analysis (PCA). Also, longitudinal live single-cell imaging was performed to explore cell divisions of the edge cells and the capacity to proliferate varied dramatically between each single cell.

The key biological functions of the most differentiated genes in the CBM1 edge cells compared to parental CBM1 and Rael were found out by gene set enrichment analysis GSEA.

The edge cells of CBM1 also show higher cloning frequency in soft agarose. While type I BL cells clone well in soft agarose correlating to their malignant phenotype, in vitro EBV-transformed type LCLs show very low cloning

frequency, when freshly transformed 0,001 -0,01%. Here we choose to analyze recently infected B-cells instead of CBM1, because the long term transformed B cells do gradually increase their agarose cloning frequency. Thus, edge cell of a fresh LCL made colonies in soft agarose like BL cells do, while the mode cells did not.

In conclusion, our result suggests that the edge cell from a non-malignant cell line might adopt some of the characteristics of the cells of a nearby adjacent malignant cell attractor without genetic mutations. This will inspire further studies on non-genetic intercellular heterogeneity and the possibility that tumorigenesis could even initiate without genetic aberrations in the right adverse microenvironment.

4.4 Both high and low levels of cellular Epstein-Barr virus DNA in blood identify failure after hematologic stem cell transplantation in conjunction with acute GVHD and type of conditioning (Paper IV)

The number of EBV infected B cells in blood of EBV carriers – the EBV load - is ultimately a reflection of the control of the switch between latency I (non - proliferating cells) and latency III (proliferating cells). Exactly how and when this switch control operates in vivo is not clear. Most of the cells in the blood of healthy persons or in risk groups are of latency 0/I type and thus non-proliferating. Thus the expansion of the EBV-infected B-cell pool, when necessary, takes place elsewhere in vivo, and probably in lymphoid tissues. There is a small possibility also that the EBV+ B cells pool can be expanded by new infections of cells, probably in lymphoid tissues in the oropharynx where there also can be access of infectious virus.

In a clinical follow-up project we determined EBV load and immune parameters in HSCT patients by regular sampling during one year after transplantation.

Grouping according to the EBV load

The border values between the groups were set based on the EBV load data:

Grouping		EBV copy number of the three PBMC collected at 1st,2nd, 3rd month after transplantation (copies/10^6 PBMC)
$EBV_{high+low}$	EBV_{high}	> 90 000 in at least 1/3 samples(s), or > 60 000 in at least 2/3 samples
	EBV_{low}	< 6000 in all the 3 samples, or negative in at least 2/3 samples
$EBV_{intermediate}$		All the rest samples

Thirty of the 51 patients (60%) were assigned to the $EBV_{high+low}$ group, according to our definition. Twenty-one patients (40%) belong to the $EBV_{intermediate}$ group. There was no difference in clinical parameters between these two groups.

Overall survival of patients in the two EBV groups, and risk factors, survival and cause of death in the HSCT patients.

The $EBV_{high+low}$ patients had a lower overall survival (OS) rate than those in the $EBV_{intermediate}$ group (Figure 1; p =0.03). OS at 5 years was 67% vs 90%, (P < 0.03). In the combined multivariate analysis three factors were associated to mortality: high+low EBV DNA load, acute GVHD II-IV and conditioning with TBI ≥6 Gy.

The causes of death (n=11) in the $EBV_{high+low}$ group were specifically relapse (n=4), bacterial infection (n=3), organ failure (n=3, one also with bacterial infection), acute GVHD (n=1) and a secondary malignancy (n=1, not EBV related).

Dynamics of immune reconstitution in the two EBV load groups

In fourteen of the patients we performed follow up of cellular immune parameters. The analysis included frequency of CD3+ T cells, CD4+ cells, CD8+cells, Treg cells (CD3+/CD4+/CD4+CD25hi+/Foxp3+/CD127-) and CD4-CD8- cells (double negative, DN). Seven of these patients belonged to the

$EBV_{intermediate}$ group and seven to the $EBV_{high+low}$ (5 in EBV_{low} group and 2 in EBV_{high} group).

While the levels of EBV load inversely related to total CD3 and CD4 levels, as for DNs, such a relation was only detected during the first month after HSCT but not later. In general, CD8+ and DN cells specific for lytic EBV antigen were more frequent than those with specificity to the latent antigen. Two cases illustrating different patterns of EBV DNA load in relation to reconstitution of EBV specific CD8+ and CD4-CD8- cells were shown in detail in paper IV. Our data suggest that cell-bound EBV DNA load is an interesting reflection of the quality and balance of reestablishing the immune system. Moderate levels of EBV DNA load, reflects a balanced reconstitution, the parameters of which now should be better established.

In conclusion, in this study, we found a strong prognostic value of predefined levels of EBV DNA load in HSCT patients. Patients with very low or high levels of cell bound EBV-DNA in blood early after transplantation showed a poor prognosis, compared to patients with intermediate levels. When combined with two other risk factors, severe acute GVHD (aGVHD II-IV) and conditioning with high dose total body irradiation (TBI), none of these patients survived more than two years after transplantation. Data on EBV load adds a significant impact on survival and the patients with the three defined risk-factors had an extremely poor outcome. EBV load may be a surrogate marker for defining patients with dysfunctional immune reconstitution after HSCT. Patients groups were small and the study has to be extended to more patients, but if true the group with dramatically poor outcome deserves strong attention from a clinical handling point of view to improve this poor outcome of the intense treatment regime.

5 UNSOLVED PROBLEMS AND HYPOTHESES FOR FUTURE WORK

5.1 What does the stronger negative fluorescent signal in FCM mean?

In FCM, unstained cells and /or isotype control stained cells are applied as a negative control for the antibody labeled cells from the same culture. Rael cells stained with antibody anti-CD54 conjugated with Pacific blue (Figure 11a, in orange), with Pacific blue isotype control (Figure 11a, in blue) and unstained Rael parental control (Figure 11a, in black) were shown. The shadowed area represented the range of the unstained control and isotype control, and was defined as negative. The isotype control has a slightly stronger signal than unstained control but the range along the x axis of these two controls was almost the same, while anti-CD54 stained Rael was higher. The CD54+ Rael cells (those out of the range of the dark area, Figure 11a) were isolated, as shown in the blank box of figure 11b, part A. These cells reestablished the parental distribution within 7 days after isolation, with the majority (98.4%) of the isolated CD54+ cells becoming CD54- again, as we concluded in paper II. However, here we also notice that the CD54 negative cells did not distribute randomly in the negative area. The CD54 negative cells shift from right to the left along the x axis (the direction of CD54 high to low), until the parental distribution of CD54 was reestablished. So here is a question we have no answer: what does the reestablishment of parental negative distribution mean and how come these presumed background signals are detected by FCM? Is there any biological meaning of stronger or weaker negative signal in FCM? In figure 2E of paper II, to circumvent these questions, we present the data as dots instead of the histogram below (the same data, shown in different ways) (**Figure 11**).

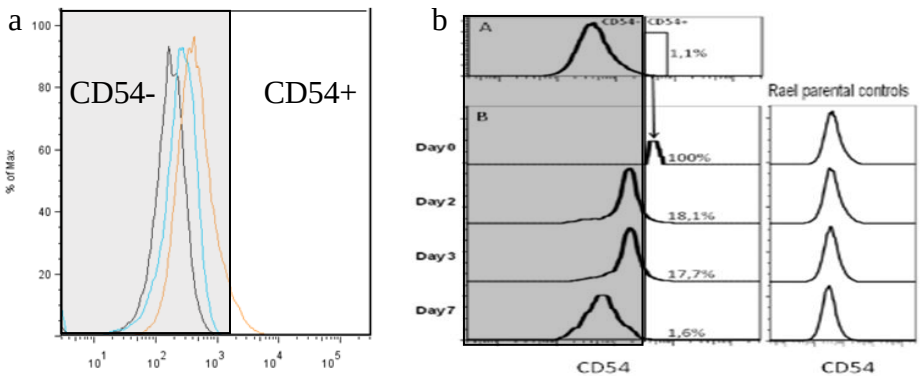

Figure 11. Stronger negative fluorescent signals.

5.2 Subpopulation suicide as a population survival strategy?

We noticed that the edge cells defined and isolated based on some parameter may transiently appear as two distinct subpopulations after isolation, one close to the basin of the attractor according to the defining parameter while the other is further away. The former one reestablishes the parental distribution while the latter gets fewer and fewer, until they disappear completely (Paper II, figure 3B and 3C). We have described this phenomenon briefly in the supplementary part in paper II (Suppl. 1.1.2). It appears to be not the result of experimental noise or erratic behavior since it was observed in two consecutive days and also in several repeated experiments, and both in cells with both strong and weak/negative surface marker expression, although the time point of its appearance varied. Here we show another observed example of this phenomenon (**Figure 12**): The isolated CD10 low edge cells showed two subpopulations during culture after isolation (Figure 12, in blue): one is of lower CD10 level, close to the Rael unstained negative control (Figure 12, in orange), which disappeared later, which the other is of higher CD10 level, close the Rael CD10 stained parental control, which reestablished the parental distribution(Figure 12, in green).

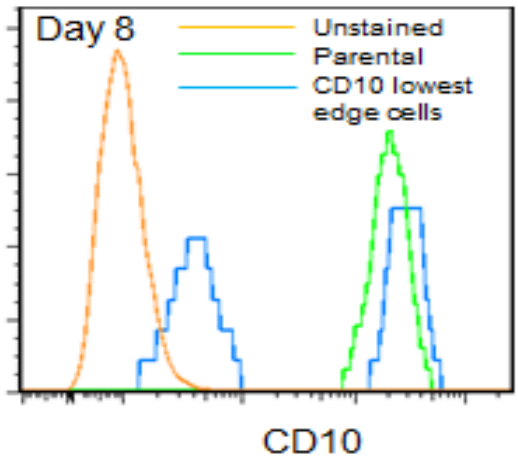

Figure 12. CD10 expression levels of Rael CD10lowest edge cells (in blue) cultured after isolation, compared with Rael unstained cells (in orange) and Rael CD10 labelled parental control (in green).

Hypotheses explaining this phenomenon:

5.2.1. Asymmetric cell division and self-suicide strategy

One cell obtains abnormal amounts of protein correlating to death while the other cell accumulates the proper amount of proteins to reach the range of the basin of attraction with higher chance of survival.

In this hypothesis, when cells are under extreme situations (in a status far away from its basin of attraction), e.g., extremely low or high expression level of certain proteins, beyond the range of its phenotype, which may be caused by external stimuli, or internal biological noise, the cells would experience asymmetric cell division. By doing so, when the cell population is short of biological resources (e.g., surface markers), one cell offers its own limited material to the other when dividing and commits suicide itself. (**Figure 13**).

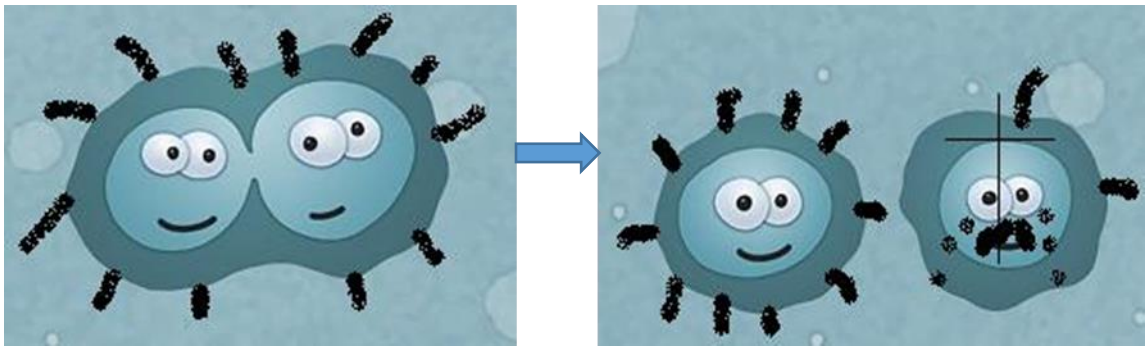

Figure 13. Asymmetric cell division--- subpopulation suicide strategy. (Modified from http://www.vladstudio.com/ipad-wallpapers/?kw=cell)

5.2.2 Delivered among cells or by cell-cell contact

Another hypothesis to explain this phenomenon would that there is biological material transfer between cells in such extreme situations, e.g., using exosomes. Or by transfer by cell-cell contact.

6 ACKNOWLEDGMENTS

I love the joyful land, and I love the lovely country. Sweden is amazing! Time flies, but it is worth. It gives me strength, to make choices following my heart. It gives me confidence, that I can fall down but I must be able to stand up. I am grateful to all those who have guided me, supported me, encouraged me and who love me during the long journey, especially:

My supervisor, **Prof. Ingemar Ernberg**, You opened the door to science for me and provided me excellent chances. I really appreciate and admire your wisdom, your broad interests, and your talent in communicate with different cultures. I could never forget how we and our nice collaborators worked the cell attractor project out, tough, exciting and full of fun! Thank you for being always interested in and respecting my ideas, for your full trust on me, for the tremendous support in my research and my life, for your love to Claire, for inviting and visiting my parents, for all the great moments together! I will keep the fundamental biology question 'what is life' in my mind, wherever I am.

My co-supervisor, **Dr. Jiezhi Zou**, thank you for guiding me from the very basic techniques, for your always available help, for your great patient for ordering reagents, for your excellent tips to solve many practical problems, for your love to Claire! My co-supervisor (though not written in paper), **Dr. Anders Wennborg**, thanks for all your comments for my papers and suggestions for my work. It is great to work with you. I benefit a lot from your strict training on data organization.

My mentor, **Prof. Markus Maeurer,** my honor to have such an excellent immunologist and nice person as my mentor. Thank you for your great help on my FACS learning, for your flowjo, for your always efficient support for my experiments, and your encouragement!

Prof. Maria Masucci. It is really my honor to have a wise and charming friend and nice mentor in my life like you. I am very impressed by your collection of china. Thank you for sharing your experience in science and in life with me, for your always available help, for your good suggestions, for your company in my hard time, for the joyful coffee time together, and for your love to Claire! The beautiful name is the best gift for her.

Present and previous members of Ingemar's group: **Fu Chen**(late), for your very patient guidance in lab when I just joint the group and kind help in life. **Lifu**, for your excellent suggestions and great support. Thanks for your comment on my thesis. **Aymeric**, very impressed by your great humor**! Anna Birgersdotter**, I like your smile! **Anna Aleman and Helena,** for your delicious cakes to Friday coffee! **Gösta,** very impressed by your talent in language learning **and Liudmila,** you are a perfect officemate, always helps to solve problems very efficiently and thanks for the dress! **Xiaoying,** thanks for all the great time we spent together, for the funny stories we share, and for your excellent tips for cooking! **Li-Sophie**, your Chinese is perfect, and also Taiji! Thanks for your always kind encouragement and nice words! **Elvira**, thanks a lot for sharing life experience with me, for your encouragement, for your great organization of the very nice trip in Novosibirsk and Altai! **Andrey**, for your guidance in bioinformatics! **Imran**, thanks for the interesting discussions about your culture, I learned a lot! And thanks for all your helps, including tips for this thesis via facebook! **Rozina**, you are such a sweet and warm hearted girl! Very impressed by your talent in organizing parties, clubs! Thanks for the great moments together! **Lotta,** you are so sweet and also so smart, keeping excellent balance in clinical work and research. I wish you a great future! **Sven ,** thanks for your great comment on my work from clinical point of view. It is great to have a very experienced clinician like you in the group! **Hoda,** you are always kind and helpful, great to know you before I leave. **Konrad**, for your great music! **Elisabeth, Tore, Maria Werner, Xue, Lydia, Gunnar Klein, Anna Friis, Johana, Yvonne, Ilya, Henrik,** for the great discussion in the group meeting, in Friday coffees, and Christmas parties!

Prof. Geogre Klein and **Prof. Eva Klein**'s group, **Geogre and Eva**, you are the most devoted scientists I have ever met. Thank you for your great comments and support to my work. **Barbro**, **Noemi**, **Liang**, **Benedek, Elena**, **Daniel**, **Eahsan**, **Harsha**, **Tatiana, Monika Adori**, **Suhas Darekar, Muhammad Mushtaq**, for the great discussion in the joint group meeting and all the nice talks and helps.

Other SPACE members: **Prof. Mats Remberger, Prof. Olle Ringdén, Prof. Jacek Winiarski, Nalini, Rajia, Lalit, Thomas, Isabelle, Brigitta, Åsa ,Katarina, Qingda, Karin,** for the nice collaboration in the HSCT project.

Prof. Sten Linnarsson, it was amazing how efficiently we started our collaboration! Thanks for the great guidance for the single cell experiment and also thank you for being in my half time seminar committee and giving valuable comments! **Anna**, thank you for your very detailed instruction for the experiments and for your always quick reply to me!

Prof. Rickard Sandberg, thank you for being interested in my ' cell attractor' work, for your help on the single cell RNAseq data analysis, and for your great advice for my postdoc position. **Mtakai**, it is really nice to collaborate with you. Thank you for your very efficient work, for your very patient and clear explanation on the data analysis to me! **Ilgar**, always nice to talk to you!

Prof. Laszlo Szekely, you are so creative and there are always good suggestions and new ideas from you! Thank you for the nice collaboration on single cell imaging! **Andrii,** for your great work for the imaging part!

Prof. Erik Aurell, thank you for the nice collaboration in FPE work, for your recommendation to the single cell analysis course in CSHL, and for introducing me to the stem cell discussion in Niels Bohr Institute. **Erez,** thank you for your great contribution to the FPE work. **Yuting,** I really appreciate your very efficient way of working! I wish you a great future!

Prof. Sui Huang at Institute for Systems Biology, Seattle. It was really a nice collaboration with you on the cell attractor project! I am very impressed by your critical thinking and elegant writing. Thanks for your comments in the project, your suggestions for career development and your great support!

Prof. Rolf Ohlsson and Dr. Anita Göndör, Dr. Satish Srinivas Kitambi and Dr. Jonas Fuxe in our corridor, for the nice discussion at Friday coffee! **Prof. Yihai Cao, Prof. Galina Selivanova, Prof. Lars-Gunnar Larsson, Prof. Marie Arsenian Henriksson, Prof. Pontus Aspenström, Prof. Roland Möllby, Dr. Velmurugesan Arulampalam, Prof. Sonia Lain, Prof. Anneka Ehrnst** (late)**, Prof. Eugene Zabarovesky** (late) **in MTC, Docent Teresa Frisan in CMB, Docent Dawei Xu in CMM, Dr. Ola Hermanson in Neuroscience,** thank you for all the interesting scientific discussions in my yearly follow ups, seminars, meetings and parties, thank you for your encouragement and support to my study!

My friends at MTC, **Åsa, Anita, Mia, Helena** in the administration for your great help and **Sándor**, **Magnus, Per, John and Torbjörn** (and many others) for your very good service. **Arnika, Agnieszka, Barbara, Carolina, Chengxi, Daeho, Emmanouil, Fan, Hai, Hamid, Hong,Honglei, Junwei, Marta, Noriyuki, Qinzi, Dr. Renhai Cao, Sharon, Xiaojuan, Xingqi, Yao, Yuan, Yunlong, Ziquan, and many others,** for the joyful time together!

Cold Spring Harbor Lab course group: all the fantastic classmates and teachers, especially **Prof. James Eberwine** at University of Pennsylvania, my supervisor in the course, thank you for the high quality course, for being interested in my research, and for your kind offer! Also thanks for your support for my postdoc position. **Anna,** my roommate, thanks for a great time together!

FACS core facility: **Barbro** in public health agency of Sweden, thank you for teaching me the FACS technique from the very beginning and always patient to answer my questions! **Birgitt** in MTC, I really appreciate your hard work. Thank you for staying after working time to help me to get enough cells! Without your support, the work could not be done!

Friends in Sweden, **Qianren and Xinjun's family,** for helping me for my housing when I just arrived in Sweden and for all the festivals we celebrated together! **Caihong and Jian's family,** it is great to meet you in Stockholm and thanks for your all your help, for all the happy time together! **Bei,** thank you for all the delicious food for me, for all the fun we had together. I miss you and looking forward to seeing you again! **Zhuochun,** thank you for all your help and for the great moment together. **Qing,** thank you for all the smart tips, for your help for my house moving and for the happy moments together! **Serena and Dome,** you are such a wonderful couple! Serena, so lucky to meet you. Thank you for all the delicious Italian (sometimes Indian :)) food, for all the parties, for taking care of the apartment after I left and for your love to Claire! **Fang,** thank you for introducing me to MTC! **Bin, Ci&Zheng, Xiaoda, Ximing, Yajuan, Jiaying, Wei, Ying, Yifeng, and many others friends at Karolinska,** thanks for the great time together! My corridor members in Lappis, all those lived in Amanusvagen 12, especially **Ola, Magali, Hui, Camila, Jiajia, Qiang and Qian**, for all the high-quality international corridor dinners and wonderful events, in our very young time! Friends in Master

program at Stockholm University: especially **Anusara** for sharing room with me and **Gishanti** for the help for practical things when I just arrived in Sweden.

Friends in US: **Zhenjun and Xiaojing, t**hank you for the pleasant trip to New York and New Jersey, for sharing practical tips and experience in baby care! **Jiangshui** for the nice trip in New York, and for your great help when my parents arrived in New York! My friends in Boston: **Sandro**, for your advice for my postdoc position! **Zhiyi,** you are such a considerate person and thank you for your very delicious food, for the nice talks! Warm regards to **Yuanyuan** and **Jifan**. **Zhonghua and Wen**, thank you for your help and also thank your mom for her excellent food! Yajing Bao and Prajna, amazing that we met in Boston and had baby almost in the same period. **Yajing**, thanks for all the helpful information regarding hospitals, baby cares, insurance, etc. Wonderful to meet you! **Prajna**, thanks for your encouragement and all the happy meals together in BWH! ! **Linxian,** thanks for all the encouragement, discussions and suggestions. **Yingxia,** thank you for all the pleasant lunch time together in BWH.

I thank all my friends in China who supported me during the last years, and special thanks to my high school classmates and teachers. Class 6 is amazing!

I thank my parents in law, sister and brother in law, and also my lovely little nephews, for their support for my study. I thank Uncle Du's family, for your support.

I thank all my family members for your great support over the years, my grandmas, grandpas, uncles, aunts and cousins, especially my sister Hong and Qing. **Hong, and Jialu**, you are a wonderful couple. Thank you for all the celebrations for birthdays and festivals with Mom and Dad, and for your love to Claire, when I am away. **Qing**, for all your company, from Lishui to Nanjing ,to Stockholm, for all your help, and for all your understanding and support!

爸爸妈妈，谨以此文，感谢你们多年悉心的付出。感谢你们对我的选择的支持和尊重，对我的爱和鼓励。 谢谢你们对 Claire 无尽的疼爱和悉心照料，没有你们这两年的全力帮助，我不可能完成这本论文，也希望 Claire 能记住这份深情。

Claire, thanks for all the happiness you bring to me. You are such a cute and sweet girl. Thanks for your excellent performance while I am away and I am so

proud of you! Life is beautiful with your company and I promise to be with you as long as I can, and as long as you need.

Ziming, thanks for your understanding, respect and support, for being nice when I am stressful and fixing practical issues when I am busy. Thank you for your love.

The work presented in this thesis was performed at the Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institutet, and supported by The Swedish Cancer Society, Cancerföreningen in Stockholm, the Maths O-Sundqvist family foundation

7 REFERENCES

1. Weiss L (2000) Metastasis of cancer: a conceptual history from antiquity to the 1990s. *Cancer metastasis reviews* 19(3-4):I-XI, 193-383.
2. Feinberg AP, Ohlsson R, & Henikoff S (2006) The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 7(1):21-33.
3. Hajdu SI (2011) A note from history: landmarks in history of cancer, part 1. *Cancer* 117(5):1097-1102.
4. Torre LA*, et al.* (2015) Global cancer statistics, 2012. *CA Cancer J Clin* 65(2):87-108.
5. Stewart BWW, Christopher P. (2014) *World Cancer Report 2014* (International Agency for Research on Cancer)p632.
6. Bernards R (2012) A missing link in genotype-directed cancer therapy. *Cell* 151(3):465-468.
7. Hanahan D & Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* 144(5):646-674.
8. Nowell PC (1976) The clonal evolution of tumor cell populations. *Science* 194(4260):23-28.
9. Hanahan D & Weinberg RA (2000) The hallmarks of cancer. *Cell* 100(1):57-70.
10. Loeb LA (2001) A mutator phenotype in cancer. *Cancer Res* 61(8):3230-3239.
11. Vogelstein B*, et al.* (1989) Allelotype of colorectal carcinomas. *Science* 244(4901):207-211.
12. Fearon ER & Vogelstein B (1990) A genetic model for colorectal tumorigenesis. *Cell* 61(5):759-767.
13. Vogelstein B*, et al.* (1988) Genetic alterations during colorectal-tumor development. *N Engl J Med* 319(9):525-532.
14. Weinberg RA (2006) The Biology of Cancer. (Garland Publishing Inc).
15. Spear BB, Heath-Chiozzi M, & Huff J (2001) Clinical application of pharmacogenetics. *Trends Mol Med* 7(5):201-204.
16. Moazed D (2011) Mechanisms for the inheritance of chromatin states. *Cell* 146(4):510-518.
17. Rutherford SL & Henikoff S (2003) Quantitative epigenetics. *Nat Genet* 33(1):6-8.
18. Rutherford SL & Lindquist S (1998) Hsp90 as a capacitor for morphological evolution. *Nature* 396(6709):336-342.
19. Sharma SV*, et al.* (2010) A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* 141(1):69-80.
20. Mack SC*, et al.* (2014) Epigenomic alterations define lethal CIMP-positive ependymomas of infancy. *Nature* 506(7489):445-450.
21. Versteeg R (2014) Cancer: Tumours outside the mutation box. *Nature* 506(7489):438-439.
22. Quintana E*, et al.* (2010) Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* 18(5):510-523.
23. Chaffer CL*, et al.* (2011) Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc Natl Acad Sci U S A* 108(19):7950-7955.
24. Janiszewska M & Polyak K (2015) Clonal evolution in cancer: a tale of twisted twines. *Cell Stem Cell* 16(1):11-12.
25. Stewart JM*, et al.* (2011) Phenotypic heterogeneity and instability of human ovarian tumor-initiating cells. *Proc Natl Acad Sci U S A* 108(16):6468-6473.
26. Hoek KS & Goding CR (2010) Cancer stem cells versus phenotype-switching in melanoma. *Pigment cell & melanoma research* 23(6):746-759.
27. Reya T, Morrison SJ, Clarke MF, & Weissman IL (2001) Stem cells, cancer, and cancer stem cells. *Nature* 414(6859):105-111.
28. Wicha MS, Liu S, & Dontu G (2006) Cancer stem cells: an old idea--a paradigm shift. *Cancer Res* 66(4):1883-1890; discussion 1895-1886.
29. Lapidot T*, et al.* (1994) A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 367(6464):645-648.
30. Bonnet D & Dick JE (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* 3(7):730-737.

31. Holzel M, Bovier A, & Tuting T (2013) Plasticity of tumour and immune cells: a source of heterogeneity and a cause for therapy resistance? *Nat Rev Cancer* 13(5):365-376.
32. Borst P (2012) Cancer drug pan-resistance: pumps, cancer stem cells, quiescence, epithelial to mesenchymal transition, blocked cell death pathways, persisters or what? *Open Biol* 2(5):120066.
33. Baylin SB (2011) Resistance, epigenetics and the cancer ecosystem. *Nat Med* 17(3):288-289.
34. Fidler IJ & Kripke ML (1977) Metastasis results from preexisting variant cells within a malignant tumor. *Science* 197(4306):893-895.
35. Gupta PB*, et al.* (2011) Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* 146(4):633-644.
36. Li Q*, et al.* (2016) Dynamics inside the cancer cell attractor reveal cell heterogeneity, limits of stability, and escape. *Proc Natl Acad Sci U S A* 113(10):2672-2677.
37. Aurell E & Sneppen K (2002) Epigenetics as a first exit problem. *Phys Rev Lett* 88(4):048101.
38. Eldar A & Elowitz MB (2010) Functional roles for noise in genetic circuits. *Nature* 467(7312):167-173.
39. Kholodenko BN (2006) Cell-signalling dynamics in time and space. *Nature reviews* 7(3):165-176.
40. Delbruck M (1949) Unités biologiques douées de continuité génétique Colloques Internationaux du Centre National de la Recherche Scientifique: CNRS.:33-35.
41. Baverstock K (2013) Life as physics and chemistry: A system view of biology. *Prog Biophys Mol Biol* 111(2-3):108-115.
42. Monod J & Jacob F (1961) Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* 26:389-401.
43. Huang S, Ernberg I, & Kauffman S (2009) Cancer attractors: A systems view of tumors from a gene network dynamics and developmental perspective. *Semin Cell Dev Biol* 20 (7.):869-876.
44. Kauffman S (1993) *Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press).
45. Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* 22(3):437-467.
46. Kauffman S (1969) Homeostasis and differentiation in random genetic control networks. *Nature* 224(5215):177-178.
47. Huang S & Kauffman S (2009) Complex gene regulatory networks - from structure to biological observables: cell fate determination. *Encyclopedia of Complexity and Systems Science*, ed Meyers RA (Springer), pp 1180-1213.
48. Baverstock K (2011) A comparison of two cell regulatory models entailing high dimensional attractors representing phenotype. *Prog Biophys Mol Biol* 106(2):443-449.
49. Huang S (2012) The molecular and mathematical basis of Waddington's epigenetic landscape: a framework for post-Darwinian biology? *BioEssays : news and reviews in molecular, cellular and developmental biology* 34(2):149-157.
50. Wang J, Xu L, Wang E, & Huang S (2010) The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. *Biophys J* 99(1):29-39.
51. Waddington CH (1957) the strategy of the genes. London: Allen and Unwin.
52. Schrödinger E (1944) *What is life* (Cambridge University Press, Cambridge).
53. Banerji CR*, et al.* (2013) Cellular network entropy as the energy potential in Waddington's differentiation landscape. *Sci Rep* 3:3039.
54. Kim J & Eberwine J (2010) RNA: state memory and mediator of cellular phenotype. *Trends Cell Biol* 20(6):311-318.
55. Takahashi K & Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126(4):663-676.
56. Yamanaka S & Takahashi K (2006) [Induction of pluripotent stem cells from mouse fibroblast cultures]. *Tanpakushitsu Kakusan Koso* 51(15):2346-2351.

57. Ohnishi K*, et al.* (2014) Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. *Cell* 156(4):663-677.
58. Ben-Porath I*, et al.* (2008) An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* 40(5):499-507.
59. Hong H*, et al.* (2009) Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature* 460(7259):1132-1135.
60. Kalmar T*, et al.* (2009) Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS Biol* 7(7):e1000149.
61. Munoz-Descalzo S, de Navascues J, & Arias AM (2012) Wnt-Notch signalling: an integrated mechanism regulating transitions between cell states. *BioEssays : news and reviews in molecular, cellular and developmental biology* 34(2):110-118.
62. Pisco AO & Huang S (2015) Non-genetic cancer cell plasticity and therapy-induced stemness in tumour relapse: 'What does not kill me strengthens me'. *Br J Cancer* 112(11):1725-1732.
63. Schadt EE, Friend SH, & Shaywitz DA (2009) A network view of disease and compound screening. *Nat Rev Drug Discov* 8(4):286-295.
64. Creixell P, Schoof EM, Erler JT, & Linding R (2012) Navigating cancer network attractors for tumor-specific therapy. *Nat Biotechnol* 30(9):842-848.
65. Pawson T & Linding R (2008) Network medicine. *FEBS Lett* 582(8):1266-1270.
66. Klein E, Kis LL, & Klein G (2007) Epstein-Barr virus infection in humans: from harmless to life endangering virus-lymphocyte interactions. *Oncogene* 26(9):1297-1305.
67. Dolcetti R, Dal Col J, Martorelli D, Carbone A, & Klein E (2013) Interplay among viral antigens, cellular pathways and tumor microenvironment in the pathogenesis of EBV-driven lymphomas. *Semin Cancer Biol* 23(6):441-456.
68. Steven NM, Leese AM, Annels NE, Lee SP, & Rickinson AB (1996) Epitope focusing in the primary cytotoxic T cell response to Epstein-Barr virus and its relationship to T cell memory. *J Exp Med* 184(5):1801-1813.
69. Yao QY*, et al.* (1996) Frequency of multiple Epstein-Barr virus infections in T-cell-immunocompromised individuals. *J Virol* 70(8):4884-4894.
70. Epstein MA, Achong BG, & Barr YM (1964) Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma. *Lancet* 1(7335):702-703.
71. Henle W, Diehl V, Kohn G, Zur Hausen H, & Henle G (1967) Herpes-type virus and chromosome marker in normal leukocytes after growth with irradiated Burkitt cells. *Science* 157(3792):1064-1065.
72. Kaneda A, Matsusaka K, Aburatani H, & Fukayama M (2012) Epstein-Barr virus infection as an epigenetic driver of tumorigenesis. *Cancer Res* 72(14):3445-3450.
73. Matsusaka K*, et al.* (2011) Classification of Epstein-Barr virus-positive gastric cancers by definition of DNA methylation epigenotypes. *Cancer Res* 71(23):7187-7197.
74. Saito M*, et al.* (2013) Role of DNA methylation in the development of Epstein-Barr virus-associated gastric carcinoma. *J Med Virol* 85(1):121-127.
75. Rowe M, Kelly GL, Bell AI, & Rickinson AB (2009) Burkitt's lymphoma: the Rosetta Stone deciphering Epstein-Barr virus biology. *Semin Cancer Biol* 19(6):377-388.
76. Dalla-Favera R*, et al.* (1982) Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc Natl Acad Sci U S A* 79(24):7824-7827.
77. Pajic A*, et al.* (2001) Antagonistic effects of c-myc and Epstein-Barr virus latent genes on the phenotype of human B cells. *Int J Cancer* 93(6):810-816.
78. Ichii M*, et al.* (2010) The density of CD10 corresponds to commitment and progression in the human B lymphoid lineage. *PLoS One* 5(9):e12954.
79. Maio M & Del Vecchio L (1992) Expression and functional role of CD54/Intercellular Adhesion Molecule-1 (ICAM-1) on human blood cells. *Leuk Lymphoma* 8(1-2):23-33.
80. Ringden O*, et al.* (1994) A randomized trial comparing busulfan with total body irradiation as conditioning in allogeneic marrow transplant recipients with leukemia: a report from the Nordic Bone Marrow Transplantation Group. *Blood* 83(9):2723-2730.

81. Li Q*, et al.* (2016) Both high and low levels of cellular Epstein-Barr virus DNA in blood identify failure after hematologic stem cell transplantation in conjunction with acute GVHD and type of conditioning. *Oncotarget*.
82. Dierickx D, Tousseyn T, & Gheysens O (2015) How we diagnose and treat posttransplant lymphoproliferative disorders. *Blood*.
83. Gustafsson A*, et al.* (2000) Epstein-Barr virus (EBV) load in bone marrow transplant recipients at risk to develop posttransplant lymphoproliferative disease: prophylactic infusion of EBV-specific cytotoxic T cells. *Blood* 95(3):807-814.
84. Thomas ED*, et al.* (1975) Bone-marrow transplantation (second of two parts). *N Engl J Med* 292(17):895-902.
85. Holtan SG, Pasquini M, & Weisdorf DJ (2014) Acute graft-versus-host disease: a bench-to-bedside update. *Blood* 124(3):363-373.
86. Aalto SM*, et al.* (2007) Epstein-Barr viral load and disease prediction in a large cohort of allogeneic stem cell transplant recipients. *Clin Infect Dis* 45(10):1305-1309.
87. Bakker NA*, et al.* (2007) Epstein-Barr virus-DNA load monitoring late after lung transplantation: a surrogate marker of the degree of immunosuppression and a safe guide to reduce immunosuppression. *Transplantation* 83(4):433-438.
88. Calabrese F*, et al.* (2010) Acute cellular rejection and Epstein-Barr virus-related post-transplant lymphoproliferative disorder in a pediatric lung transplant with low viral load. *Transpl Infect Dis* 12(4):342-346.
89. Hanto DW & Najarian JS (1985) Advances in the diagnosis and treatment of EBV-associated lymphoproliferative diseases in immunocompromised hosts. *J Surg Oncol* 30(4):215-220.
90. Jimenez S (2015) Epstein-Barr virus-associated post-transplantation lymphoproliferative disorder: potential treatments and implications for nursing practice. *Clin J Oncol Nurs* 19(1):94-98.
91. Liu Q*, et al.* (2013) Molecular monitoring and stepwise preemptive therapy for Epstein-Barr virus viremia after allogeneic stem cell transplantation. *Am J Hematol* 88(7):550-555.
92. Luskin MR*, et al.* (2015) The Impact of EBV Status on Characteristics and Outcomes of Posttransplantation Lymphoproliferative Disorder. *Am J Transplant* 15(10):2665-2673.
93. Ruf S & Wagner HJ (2013) Determining EBV load: current best practice and future requirements. *Expert Rev Clin Immunol* 9(2):139-151.
94. Stevens SJ, Pronk I, & Middeldorp JM (2001) Toward standardization of Epstein-Barr virus DNA load monitoring: unfractionated whole blood as preferred clinical specimen. *J Clin Microbiol* 39(4):1211-1216.
95. Uhlin M*, et al.* (2014) Risk factors for Epstein-Barr virus-related post-transplant lymphoproliferative disease after allogeneic hematopoietic stem cell transplantation. *Haematologica* 99(2):346-352.
96. Mautner J & Bornkamm GW (2012) The role of virus-specific CD4+ T cells in the control of Epstein-Barr virus infection. *Eur J Cell Biol* 91(1):31-35.
97. Nowakowska J*, et al.* (2015) T cells specific for different latent and lytic viral proteins efficiently control Epstein-Barr virus-transformed B cells. *Cytotherapy* 17(9):1280-1291.
98. Heslop HE & Rooney CM (1997) Adoptive cellular immunotherapy for EBV lymphoproliferative disease. *Immunol Rev* 157:217-222.
99. Ben-Bassat H*, et al.* (1977) Establishment in continuous culture of a new type of lymphocyte from a "Burkitt like" malignant lymphoma (line D.G.-75). *Int J Cancer* 19(1):27-33.
100. Klein G, Dombos L, & Gothoskar B (1972) Sensitivity of Epstein-Barr virus (EBV) producer and non-producer human lymphoblastoid cell lines to superinfection with EB-virus. *Int J Cancer* 10(1):44-57.
101. Ernberg I*, et al.* (1989) The role of methylation in the phenotype-dependent modulation of Epstein-Barr nuclear antigen 2 and latent membrane protein genes in cells latently infected with Epstein-Barr virus. *J Gen Virol* 70 (Pt 11):2989-3002.
102. Lyons AB & Parish CR (1994) Determination of lymphocyte division by flow cytometry. *J Immunol Methods* 171(1):131-137.

103. Islam S*, et al.* (2012) Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat Protoc* 7(5):813-828.
104. Dobin A*, et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1):15-21.
105. Chung NC & Storey JD (2015) Statistical significance of variables driving systematic variation in high-dimensional data. *Bioinformatics* 31(4):545-554.
106. Guldevall K*, et al.* (2010) Imaging immune surveillance of individual natural killer cells confined in microwell arrays. *PLoS One* 5(11):e15453.
107. Werner M, Ernberg I, Zou J, Almqvist J, & Aurell E (2007) Epstein-Barr virus latency switch in human B-cells: a physico-chemical model. *BMC Syst Biol* 1:40.
108. Roesch A*, et al.* (2010) A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* 141(4):583-594.
109. Pisco AO*, et al.* (2013) Non-Darwinian dynamics in therapy-induced cancer drug resistance. *Nat Commun* 4:2467.