MPIMG

SPON2 and its implication in epithelial-mesenchymal transition

Dissertation zur Erlangung des akademischen Grades des
Doktors der Naturwissenschaften (Dr.rer.nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von

Artur Muradyan
aus Yerevan, Armenien

September 2009

Diese Dissertation wurde in der Zeit vom Januar 2005 bis Mai 2009 am Max-Planck-Institut für molekulare Genetik in Berlin, in der Abteilung von Herrn Prof. Dr. Hans-Hilger Ropers, in der Arbeitsgruppe von Herr Dr. Reinhard Ullmann angefertigt.

The work presented in this thesis has been conducted from January 2005 till May 2009 at Max Planck Institute for Molecular Genetics (Berlin), in the department of Prof. Dr. Hans-Hilger Ropers under supervision of Dr. Reinhard Ullmann.

1. Gutachter: Prof. Dr. Hans-Hilger Ropers
Max-Planck-Institut für molekulare Genetik

2. Gutachter: Prof. Dr. Constance Scharff
Freie Universität Berlin

Tag der Disputation: **04.02.2010**
Day of the Disputation: **04.02.2010**

Hiermit erkläre ich, die vorliegende Arbeit selbständig und ohne unerlaubte Hilfe angefertigt zu haben und alle verwendeten Hilfsmittel und Inhalte aus anderen Quellen als solche kenntlich gemacht zu haben.

Des Weiteren versichere ich, dass die vorliegende Arbeit nie in dieser oder anderer Form Gegenstand eines früheren Promotionsverfahrens war.

I hereby declare that the work presented in this thesis has been conducted independently and without any inappropriate support and that all sources of information, be it experimental or intellectual, are aptly referenced.

I hereby declare that this thesis has not been submitted, either in the same or a different form, to this or any other university for a degree.

Artur Muradyan

Berlin, September 2009

To my parents, brother and
in loving memory of my sister Irina

ACKNOWLEDGMENT

I would like to thank Prof. Dr. Hans.-H. Ropers for establishing a scientific framework, in which I pursued my PhD, and for giving me an opportunity to carry out my thesis in his department. I am indebted to Prof. Dr. Constance Scharff for reviewing and evaluating my thesis.

I want to highlight with sincere gratitude my group leader Dr. Reinhard Ullmann, who was persistently supporting and guiding over the years I spent in his group. I highly appreciate him for shaping my self-motivation and independence in science.

I acknowledge Max Planck Society for financial support.

A special chapter of compliments deserved a brilliant and kind person Dr. Sybille Krauss for her generous and immense help in many experiments and fruitful discussions. I express my gratitude to Dr. Chandan Goswami for valuable discussions and technical tips.

I am thankful to Prof. H.Popper, Prof. R.Pfragner, Dr. I.Halbwedl and Dr. H.Kothmeier for samples and SCLC cell line.

Special thanks goes to my dear Ines Mueller and Hannelore Madle, who assisted in some experiments and made my life in laboratory easier, to Fikret, Gennadiy, Murat, Masoud, Corinna and Lars for being always kind and helpful. I thank also my group members Ralph, Marei, Jan, Katrin, Stephanie, Grit, Anne, Vivien and Alischo.

Ein herzliches Dankeschön meinem besten Freund und Kollegen Georg Wieczorek, mit dem ich viele gute und unvergessliche Zeiten in Deutschland verbracht habe, und seiner Frau Christiana, die immer so nett und gastfreundlich war. Georg ist derjenige, der mir die Schönheiten Deutschlands gezeigt hat.

During my entire project I benefited from the expertise of the whole department, therefore I express my sincere gratitude and sympathy. Particularly, I want to point out the healthy atmosphere, being in tune with harmony and productive collaboration. I am thankful to secretariat of our department, namely Gabriele Eder and Hannelore Markert for their continuous support in all kinds of paperwork and legality.

I am thankful to my friend and colleague Dr. Stephan Heymann not only for scientific cooperation, but also for kind attitude and hospitality. I owe my appreciation to his wife Margarita as well.

I am grateful to Dr. Peter Stosiek and Arpik Nshdejan for their constant motivation and care.

I am thankful to my friends Meline, Bedros, Karine, Andranik, Hendrik, Vigen, Borik, Mariam, Nazeli, Masis and many others from Armenian Community in Berlin. I do love them all !

Looking back at the years I spent in Germany, I want to thank Berlin, one of my lovely cities in the world that gave me so much hospitality, positive impressions and memories.

I owe a special gratitude to my lovely and beautiful girlfriend Katerina Kraft for her love, kindness, support, care and charm.

Finally, I owe my deepest gratitude and appreciation to my parents **Karine Gasparyan** and **Grant Muradyan**, to my brother **Armen** and his wife **Lusine** for their permanent encouragement, everlasting support, attention, care and love. I thank my delightful nephew **Gor** and sweet niece **Natalie**, who with their endearing words were passing me supreme emotions and were making my life full of joy and love. I remember with gratitude my grandmother **Eugenia Sahakyan**, who was a big part of my education. And the last but not least, a very special merit belongs to my lovely, unforgettable and missed sister **Irina Muradyan**, who had an immense impact on my education and life.

Խորին երախտագիտությունս եմ հայտնում ծնողներիս՝ Կարինե Գասպարյանին և Հրանտ Մուրադյանին, եղբորս Արմենին և նրա տիկնոջը Լուսինեին, մշտական խրախուսանքի, օժանդակության, ուշադրության, խնամքի և սիրո համար: Շնորհակալություն եմ հայտնում նաև եղբորս հրապուրիչ և քաղցր բալիկներին՝ Գոռին և Նատալիին, կյանքս լուսավոր և ուրախ դարձնելու համար: Երախտագիտությամբ եմ հիշում տատիկիս՝ Եվգենիա Սահակյանին, կրթությանս մեջ մեծ ավանդ ունենալու համար: Հատուկ շնորհակալությամբ ուզում եմ հիշել կյանքից անժամանակ հեռացած քրոջս՝ Իրինա Մուրադյանին, իր բարության, արդարության, քնքշանքի, գեղեցկության և սիրո համար:
Աստված Ձեզ պահապան:

TABLE OF CONTENTS

ABBREVIATIONS

Ab	antibody
AC	adenocarcinoma
aCGH	array CGH
Amp	ampicillin
BAC	bacterial artificial chromosome
BSP	bisulfite sequencing PCR
CAF	cancer associated fibroblast
CAM	cell adhesion molecule
CGH	comparative genomic hybridization
CIA	chloroform-isoamyl alcohol
co-IP	co-immunoprecipitation
ECM	extracellular matrix
EMT	epithelial-mesenchymal transition
EtOH	ethanol
GEO	gene expression omnibus
GFP	green fluorescent protein
GO	gene ontology
HPRD	human protein reference database
IF	immunofluorescence
IP	immunoprecipitation
Kan	kanamycin
kb	kilobase
kDa	kilodalton
LB	Luria-Bertani broth
LSM	laser-scanning microscope
MCA	methylated CpG island amplification
MeDIP	methylated DNA immunoprecipitation
MET	mesenchymal-epithelial transition
MTOC	microtubule-organising centre
PCR	polymerase chain reaction
qRT-PCR	quantitative real-time PCR
RNS	reactive nitrogen species
ROS	reactive oxygen species
RT	room temperature
SCLC	small cell carcinoma
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SQCC	squamous cell carcinoma
WB	Western blot
WGA	whole genome amplification
w/o	without

SUMMARY

Epithelial-mesenchymal transition (EMT) describes a process of differentiation of epithelial cells into cells with mesenchymal properties. EMT is indispensable for normal development, including formation of mesoderm and neural crest, and the same mechanism is also abused by tumour cells to promote metastasis and generation of tumour derived supportive stroma.

Analysis of a human small cell lung carcinoma cell line that exhibited simultaneous adherent and suspensive growth demonstrated a mesenchymal gene expression signature in the adherent fraction, which was absent from the suspensive subpopulation. In xenograft transplantation experiments it turned out that only the mesenchymal subpopulation was able to invade the host tissue. *SPON2* was among the few differentially expressed genes in the two subpopulations, suggesting this gene as a cause or early consequence of EMT in this cell line.

The few reports published till now describe SPON2 as an extra cellular matrix protein involved in axon guidance, immune response and adhesion. More recently it has also been reported as a prognostic marker for ovarian and prostate cancer.

A fetal lung fibroblast cell line characterized by high SPON2 expression was chosen to broaden the understanding of the biological function of this protein. In interphase cells SPON2 was observed as fibrous meshwork in the cytoplasm and as distinct speckles within the nucleus. During mitosis and cytokinesis SPON2 expression was limited to the spindle apparatus and the midbody, respectively. The molecular consequences of transient siRNA knockdown were monitored at the RNA level. Genes deregulated by reduced levels of SPON2 were significantly enriched for Gene Ontology terms such as EMT, adhesion and motility. Pathway analysis indicated that these changes were likely to be mediated by triggering the TGF-beta signalling pathway. *In silico* prediction of potential disease association highlighted fibrosis, rheumatoid arthritis and tumour progression, all of which are known to be related to EMT. Co-immunoprecipitation followed by mass spectrometry significantly linked SPON2 to migration and the cytoskeleton, which nicely corroborated the results obtained by immunofluorescence. Both gene expression and co-immunoprecipitation data were subjected to network analysis. The vast majority of potential SPON2 interaction partners, as highlighted by these two independent experiments, were integrated in one comprehensive network indicating the plausibility of the results.

In conclusion, data generated in this thesis suggest that SPON2 has an intracellular function, and that deregulation of this protein in tumour cells may enhance their metastatic capacity via EMT. Furthermore congenital or somatic mutations of SPON2 may be responsible for EMT associated late onset disorders including rheumatoid arthritis and fibrosis.

ZUSAMMENFASSUNG

Der Prozess in dem Epithelzellen in Zellen mit mesenchymalen Eigenschaften übergehen wird allgemein als Epitheliale-Mesenchymale Transition (EMT) bezeichnet. EMT ist unerlässlich für normales Wachstum, einschließlich der Bildung des Mesoderms und der Neuralleiste, allerdings wird derselbe Mechanismus von Tumorzellen bei der Metastasenbildung und zur Entstehung von tumorösem Stroma missbraucht.

Die Untersuchung einer humanen kleinzelligen Lungenkarzinomzelllinie, die gleichzeitig adhärentes und suspensives Wachstum zeigt, hat für die adhärierende Fraktion ein mesenchymales Genexpressionsmuster ergeben, dass sich bei der suspensiven Subpopulation nicht beobachten lässt. In Xenograftexperimenten zeigte sich, dass nur die mesenchymale Subpopulation zur Invasion des Wirtsgewebes fähig war. Genexpressionsanalysen beider Subpopulationen identifizierten SPON2 als eines der wenigen Gene mit unterschiedlicher Expressionshöhe, was darauf hinweist das SPON2 die Ursache oder eine frühe Folge von EMT für diese Zelllinie darstellt.

Die wenigen bisher erschienenen Arbeiten beschreiben SPON2 als ein Protein der extrazellulären Matrix, das in die Axon Leitung/Zielfindung, die Immunantwort und die Adhäsion involviert ist. Jüngsten Berichten zufolge ist es auch prognostischer Marker für Ovarial- und Prostatakrebs.

Eine fötale Lungenfibroblastenzelllinie, die sich durch hohe SPON2-Expression auszeichnet, wurde für das vertiefte Studium der biologischen Funktionen dieses Proteins ausgewählt. In der Interphase ließ sich SPON2 als fibröses Netzwerk im Zytoplasma und als wohlabgegrenzte Flecken im Zellkern beobachten. Während der Mitose und Zytokinese war die Lokalisation von SPON2 auf den Spindelapparat bzw. die Zellteilungsfurche beschränkt. Die molekularen Konsequenzen eines transienten siRNA-Knockdown wurden auf RNA-Ebene beobachtet. Gene die durch den Knockdown von SPON2 dereguliert wurden wiesen eine signifikante Anreicherung der „Gene Ontology“ Begriffen wie EMT, Adhäsion und Motilität auf. Signalwegsanalysen ließen vermuten, dass diese Veränderungen durch das Auslösen des TGF-beta Signalwegs vermittelt werden. *In silico* Vorhersagen möglicher Krankheitsassoziationen verwiesen auf Fibrose, rheumatoide Arthritis und Tumorprogression, von denen alle bekanntermaßen mit EMT zusammenhängen. Koimmunopräzipitation, gefolgt von Massenspektrometrie, stellte die Verbindung zwischen SPON2 zu Migration und Zytoskelett her, was die Ergebnisse der Immunofluoreszenzuntersuchungen noch zusätzlich bekräftigten. Die Datenkonvolute aus den Expressionsstudien und der Koimmunopräzipitation, wurden unabhängigen Netzwerkanalysen unterzogen durch welche potentielle Interaktionspartner von SPON2 identifiziert wurden. Die überwiegende Mehrheit der hierdurch identifizierten Interaktionspartner konnte im Anschluss in ein umfassendes Netzwerk integriert werden, welches die Plausibilität der Ergebnisse unterstreicht.

Zusammenfassend deuten die Ergebnisse dieser Arbeit auf eine intrazelluläre Funktion von SPON2 hin, dessen Deregulation in Tumorzellen, deren metastatische Fähigkeit über EMT steigern kann. Darüberhinaus könnten erbliche oder somatische Mutationen von SPON2 für EMT-assoziierte, spät manifestierende Krankheiten wie rheumatoide Arthritis und Fibrosen verantwortlich sein.

Introduction

I INTRODUCTION

I-1 Cancer biology

The first description about cancer comes from ancient Egyptian papyrus dated ~1600 B.C. The document describes eight cancer cases of breast treated by cauterization.

The word "cancer" was introduced by Hippocrates (460-370 B.C.), who is known as the "Father of Medicine". He used *carcinos* and *carcinoma* to describe ulcer-forming and non ulcer-forming tumours [1, 2].

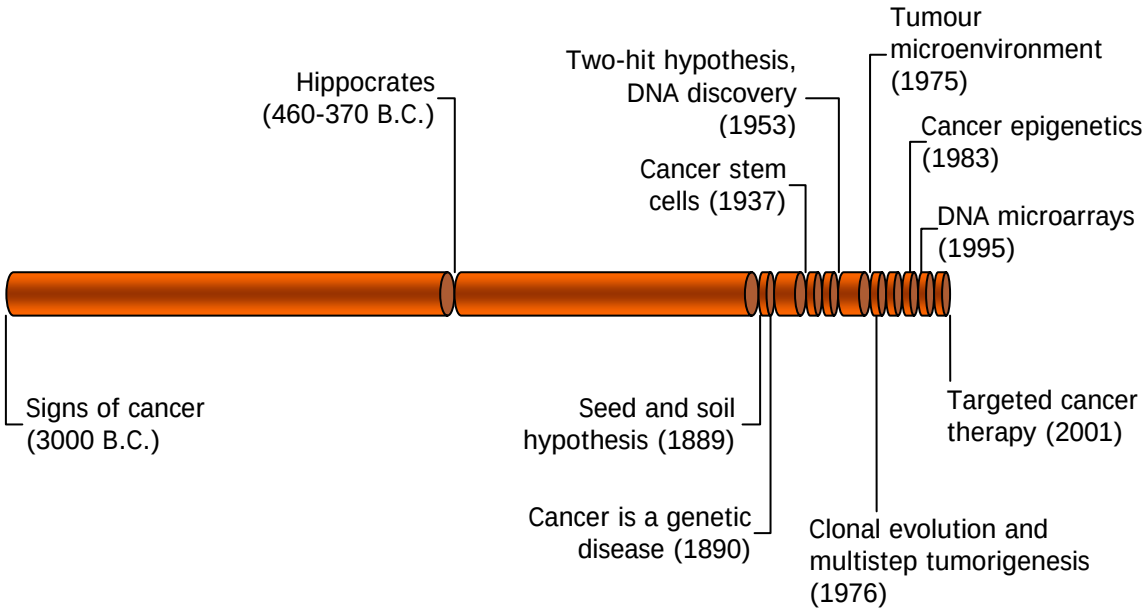


Figure I-1. Key events in cancer history (adapted from Nature Milestones Cancer, April 2006).

Cancer is the leading cause of human death, comprising around 13% of all the deaths (WHO Fact sheet No.297). It is not a simple disease, but rather a heterogeneous neo-formation or an abnormal "organ", built up by atypical cells and normal cells of other tissues (blood vessels, cells of immune response and fibroblasts). It is a multistep cellular evolution ending up in genomic instability, escape from apoptosis, invasion and self-reproduction of malignant cells. These cells may have various acquired abnormalities, including aneuploidy, chromosomal rearrangements, amplifications, deletions, gene rearrangements, and loss- or gain-of-function mutations.

Cancers are usually classified according to their tissue and organ of origin, histopathology and cytogenetics (especially in case of leukemias). Tumours arising from

epithelial cells are known as carcinomas, from glandular cells – adenomas, from embryonic mesodermal layer – sarcomas, leukemias and lymphomas arise from hematopoietic sells etc. The determination of cancer cell origin becomes even more complicated in cases as adenocarcinoma (carcinoma originating in glandular tissue) and squamous cell carcinoma (squamous differentiation), angiosarcoma (vascular channels) or chondrosarcoma (cartilaginous differentiation) [3]. However, exact molecular classification of cancers still remains complicated and debatable, whereas each tumour is individual and has its specific genetic and epigenetic signature. About 3% of all cancer patients suffer from cancers of unknown primary origin [4].

I-2 Cancer evolution

I-2.1 Knudson's two-hit hypothesis and field cancerisation

Cancerogenesis can be explained by several theories. One of them is Knudson's two-hit hypothesis and field cancerisation, which describe cancer cells arising after having accumulated two and more "hits". Knudson's two-hit hypothesis was proposed first in 1953 by Carl Nordling as multi-mutation cancer theory [5], and later was formulated by Alfred Knudson in 1971 [6]. This hypothesis was based on the observation that inherited forms of retinoblastomas occur at younger age when compared to sporadic forms of this disease. Moreover, in these familial cases tumours frequently developed in both eyes. This led Knudson to the conclusion that in sporadic retinoblastomas a given cell has to acquire two mutations in order to inactivate both alleles, while the familial form needs only one further hit as all cells already bear a mutation in Rb gene (Figure I-2).

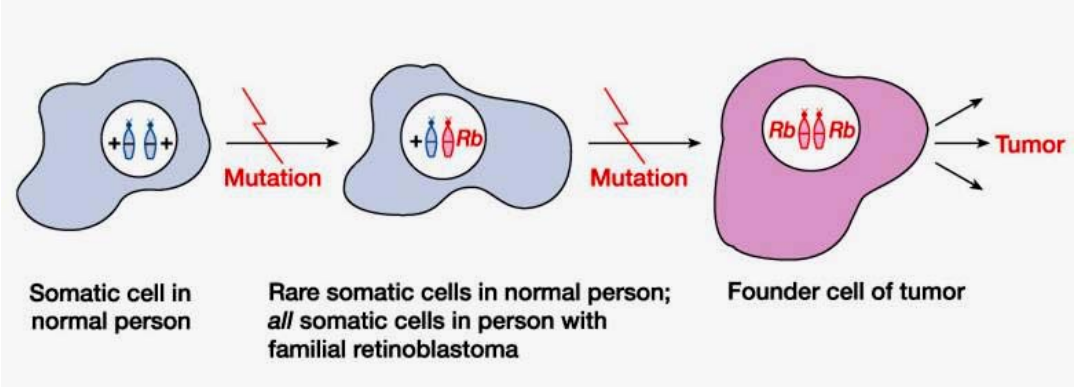


Figure I-2: Knudson's two hit hypothesis (Strachan and Read; Human Molecular Genetics 3)

Field cancerization is an extended form of Knudson's two-hit hypothesis [7]. Evolutionary our cells are endowed with sophisticated mechanisms to prevent sporadic malignization by DNA repair mechanisms and apoptosis. Single mutation does not suffice to circumvent these defence mechanisms and to provoke tumour development. According to field cancerisation theory several successive mutations are necessary for stepwise malignization, giving to affected cell growth advantage (Figure I-3).

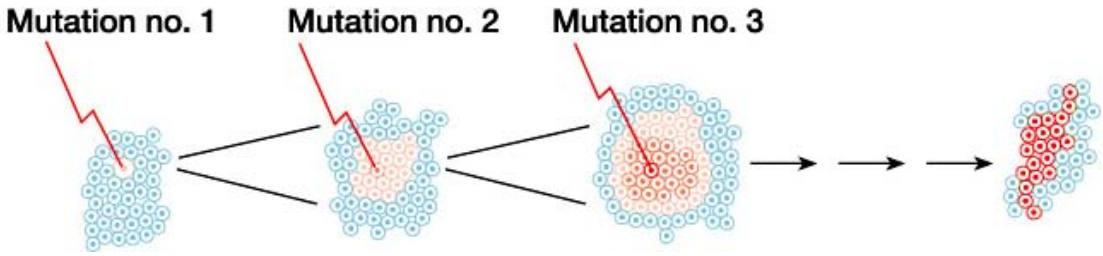


Figure I-3: Field cancerization

After getting a mutation cell acquires growth advantage and forms progeny with similar quality. When one of these cells receives the second mutation, it becomes more actively growing and is expanded further. Next successive mutations push a single cell to malignization (Image from Strachan and Read; Human Molecular Genetics 3)

Transformation of a normal cell into cancer one possibly requires six specific mutations in individual cell [8]. Proceeding on a mutation rate of 10^{-7} per gene per cell generation, it is extremely unlikely that a single cell suffers so many hits, if mutation rate is 10^{-7} per gene per cell generation. The probability that a cell (out of 10^{13} cells/person) acquires six mutations leading to tumour formation is 1 in 10^{29} [8]. Nevertheless, cancer happens due to a combination of at least two reasons:

- mutations promote cell proliferation leading to expanded populations of affected cells, which later on can receive the next hit, e.g. activating oncogenes or damaging tumour suppressor genes (Figure I-3);
- some mutations lead to genomic instability and impaired DNA repair, which results in increased mutation rate.

A summarizing theory was raised by Hanahan and Weinberg [9], proposing that the normal cell necessitates acquisition of at least six specific qualities in order to be transformed. These include: *(a)* independence from external growth factors, *(b)* insensitivity to anti-growth signals, *(c)* escaping apoptosis, *(d)* indefinite replication, *(e)* sustained angiogenesis and *(f)* ability to metastasize and invade.

I-2.2 Stem cell theory

Another suggestion for carcinogenesis is the cancer stem cell hypothesis. In 1858 Rudolf Virchow proposed that cancer can arise from embryonic-like cells [10]. The hypothesis was then extended by Cohnheim and Durante. Based on histological similarities of cancer and embryonic tissues they suggested that after activation the dormant rudiments of embryonic cells can become cancerous [11]. The term "cancer stem cell" is associated with cancer cell, which possess normal stem cell features (pluripotency and self renewal capacity). Dividing asymmetrically it produces two non-identical daughter cells: one gives rise to identical cancer stem cell and the second proliferates to a phenotypically different tumour cell population (Figure I-4).

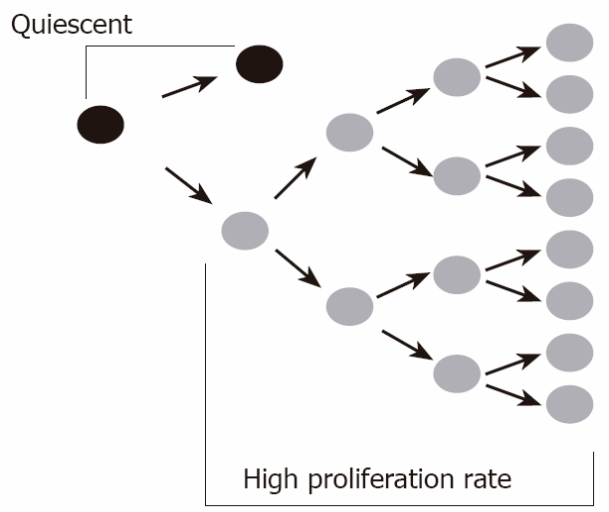


Figure I-4. Cancer stem cell model

The cancer stem cell (black) divides asymmetrically, giving rise identical to itself stem cell and a cell with high proliferative capacity (grey). (Image from [12]).

For a stem cell it should be enough to acquire only two of six proposed by Hanahan and Weinberg capabilities in order to drive it to cancer formation: independence from external growth factors and insensitivity to anti-growth signals [13].

Cancers often exist as a heterogeneous population of cells with different proliferative capacity. On the other hand, the most traditional cancer therapy agents target the fast growing and dividing cancer cells, in most of cases resulting in temporary remissions. Most likely it happens because the fast growing and dividing cells are eliminated. After therapy the slow dividing cancer stem cells survive and give rise to a new population of proliferating cancers cells [12, 14] (Figure I-5).

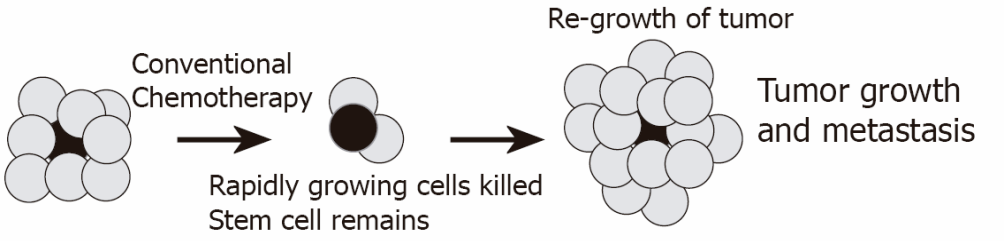


Figure I-5. Conventional chemotherapy principle.

Conventional chemotherapy kills rapidly growing cancer cells (grey). The putative cancer stem cell (black) survives and sustain the tumour mass. (Image from [12]).

The first evidence of cancer stem cell was documented by Lapidot and colleagues [15], when they transplanted human acute myeloid leukaemia (AML) cells in severe combined immunodeficiency (SCID) mice. The putative leukemic stem cells were isolated and determined with a very rare frequency 1:250.000 cells. Later, the cancer stem cells were also reported in solid tumours [16]. The involvement of microenvironment has also strong impact on stem cell fate. Many studies demonstrate that tissue-specific stem cells are regulated by microenvironment. For instance, GFP labelled neural stem cells, co-cultured with myoblasts, were differentiated into muscle cells bearing GFP protein [17]. However, it is still unclear, whether cancer stem cells are derived from normal stem cells, or derived from differentiated cells undergoing de-differentiation [18]. An additional supporting prerequisite of stem cell hypothesis can be cell-cell fusion and horizontal gene transfer. It is involved in various biological processes, as fertilization, muscle development (formation of multinuclear muscle fibres from mononuclear myoblasts), bones (formation of osteoclasts by macrophage fusion) and immune response (giant cell formation by macrophage fusion) (reviewed by [19]). Tumour cells also have considerable cell-cell fusion capacity. Fusion of cancer cells with somatic ones may result in even more aggressiveness, than parental cells [20]. Alternatively, fusion between mutated stem cells and somatic cells can generate cancer stem cells. This evidences indicate that inappropriate cell-cell or stem cell-cell fusions might be involved in cancer development [14, 18].

The horizontal gene transfer theory supposes that genetic material is able to be transported from an apoptotic cell to another via phagocytosis or endocytosis [21]. Both, cell-cell fusion and horizontal gene transfer theories are well reviewed by Bjerkvig *et al.* [18].

I-3 Epithelial tissue organization and carcinomas

Epithelium is a tissue composed of cells lining internal and external surfaces of our organs and body. It performs protective, secretory, transport, sensation and absorption functions. Direct exposure to exogenous influences makes epithelium often vulnerable to a wide range of damages. In order to survive these damages, epithelium has precursor cell compartments and regenerative capacity. Due to this feature the epithelial cells could be easy targets for transformation. Malignisation risk is even enhanced by combination of proliferative capacity and carcinogens, while epithelial tissues are often endangered to potential carcinogenic factors, including physical, chemical and biological agents. It is not surprising that the vast majority of all cancers are carcinomas (according to American Cancer Society).

A number of gene products are implicated in epithelial tissue organization and can be considered as epithelial markers, for example E-cadherin, different cytokeratins and cell junction proteins. A distinctive peculiarity of epithelial cells is their apico-basal polarity and regular organization, which is mainly realized by microtubules, actin cytoskeleton and spectrins [22] (Figure I-6). Healthy epithelial cells laterally adhere to each other via different junctions (see below) and to the basal lamina, which separates the epithelium from the underlying stroma. The stroma is a connective tissue that is populated by fibroblasts, cells of immune response, adipocytes and endothelial cells. An evident feature of the stroma is the considerable proportion of extracellular matrix, which is produced by fibroblast.

In contrast to regular epithelium, carcinoma cells are chaotically distributed over the tumour mass and can penetrate through the basal lamina into underlying stroma.

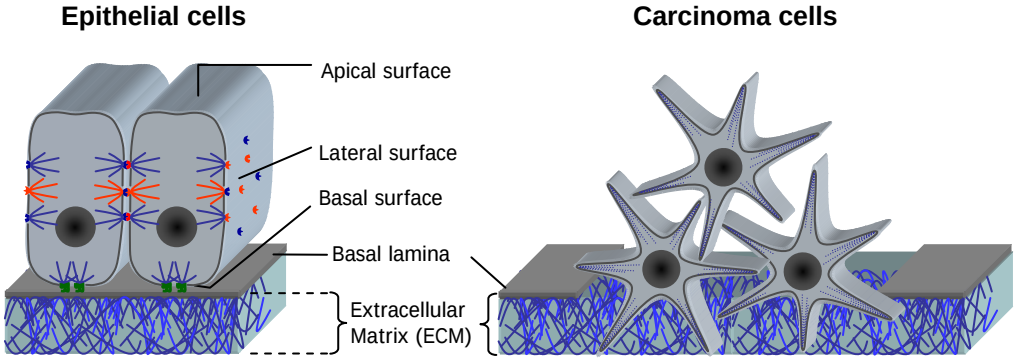


Figure I-6. Normal epithelial and carcinoma cells

Normal epithelial cells are polarized, whereas carcinoma cells loose their polarity and chaotically distributed in the tumour mass. The cancer cells have attenuated adhesion capacity and produce a set of enzymes that modulate and destruct ECM and the basal lamina.

I-4 Cell adhesion

Adhesion is an indispensable pre-requisite for the integration of cells into tissues. Cell adhesion molecules (CAMs) are highly conserved during evolution of metazoans. Even simple metazoans such as *Trichoplax adhaerens*, having no more than four cell types, exhibits great overlap of cell adhesion genes and synteny with humans [23]. The CAMs can be divided in four major classes: cadherins, integrins, selectins and immunoglobulin family proteins. Depending on the CAM types the adhesion can be homophilic (when connection is formed between two identical molecules) and heterophilic (when connection is formed between different CAMs). CAMs are composed of an extracellular domain, which interacts with other CAM and ECM, a transmembrane domain, and finally an intracellular domain, which interacts with the cytoskeleton.

In vertebrates the cell junctions can be categorised in two main classes. The first class is cell-cell adhesion, which is represented by tight junctions, gap junctions, adherens junctions and desmosomes. The second class, called cell-matrix adhesion, establishes adhesive contacts between cells and ECM proteins, and includes hemidesmosomes and focal contacts. Figure I-7a depicts the various types of adhesions necessary to maintain polarity and structure of epithelium.

Tight junctions (TJ) (Figure I-7b), also known as *zonula occludens*, have a pivotal role in integrating the cells together into tissue and is specific only to vertebrates [24]. In epithelial tissue TJ seal the cytoplasmatic membrane of neighbouring cells beneath the apical surface along the entire perimeter and prevent the inflow of water and ions between the cells. The protein complex, involved in tight junction generation mostly involves occludins, claudins, ZO-1, catenins and JAM [25].

Adherens junction (Figure I-7c), or *zonula adherens*, is an actin filament-based complex that is usually positioned underneath the tight junctions. This type of junction tailors the cytoskeleton of neighbouring cells with each other using vinculin, catenin and E-cadherin. The rim of spanning cytoskeleton filaments, tailored with these proteins, engirds the cell and supports its shape [26]. Adherens junctions are Ca^{++} dependant.

Gap junctions (Figure I-7d) are specialized intercellular connections composed by connexins. The connexins are transmembrane proteins that form cylindrical channels, capable to transport small molecules and ions from the cytoplasm of one cell to another. Gap junctions in some neurons function for transmission of electrical impulses [27].

Desmosomes (Figure I-7e) are randomly distributed button-like adhesion structures. In epithelial cells they are spreaded over lateral sides. Desmosomes are also found in muscle

cells. Desmosomes connect intracellular keratin filaments of neighbouring cells via a complex of proteins. Schematically it can be divided into inner and outer dense plaques, including desmoplakin, plakophilin and plakoglobin, which are localized in the cell, and transmembrane proteins, such as desmocollin, desmoglein and desmosomal cadherin [28].

Cell-extracellular matrix adhesion plays an important role in many biological processes including cell motility, proliferation and differentiation. In animals it is realized by hemidesmosomes and focal contacts. The hemidesmosomes are anchoring junctions, allowing epithelial cells to adhere on basal lamina; it is realized by interaction of integrins to cytoskeleton intermediate microfilaments. While desmosomes link two cells with each other, the hemidesmosomes establish adhesive contacts with ECM. In epithelial cells they are dispersed on basal surface and allow cells to anchor on basal lamina.

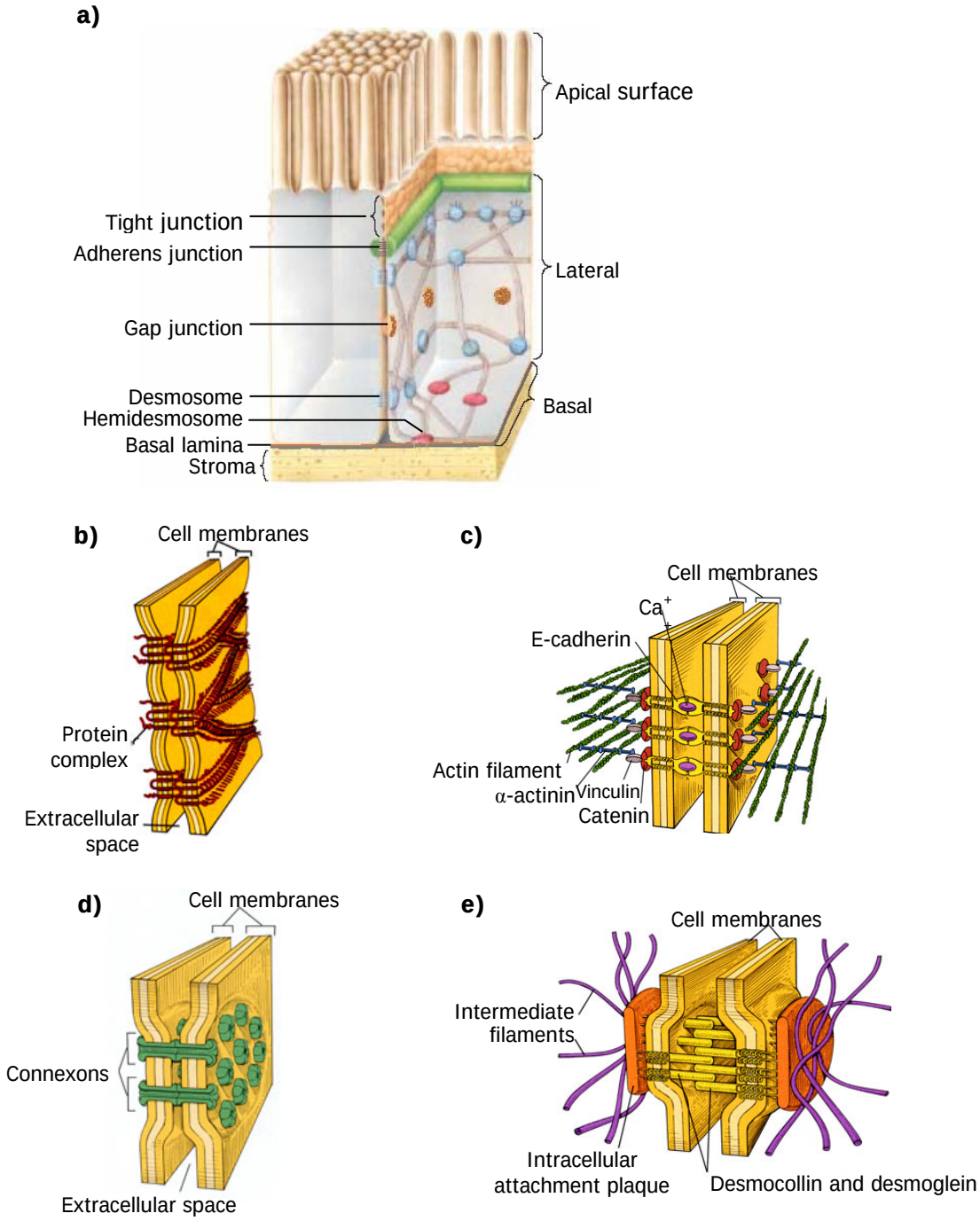


Figure I-7. Schematic overview of adhesions on example of epithelial cell (adapted from Lodish, Chapter 6 and http://anatomy.iupui.edu/courses/histo_D502/D502f04/lecture.f04/cell.f04/)

Epithelial cell *(a)* is a good example of cell, which sets its cell-cell adhesion via tight junctions *(b)*, adherens junctions *(c)*, gap junctions *(d)* and desmosomes *(e)*. The hemidesmosomes provide cell adhesion to ECM.

I-5 Tumour microenvironment and tumour-stroma interaction

Tumour microenvironment plays a pivotal role in tumour initiation and progression. A number of studies indicate that tumour growth is not only determined by cancel cells themselves, but also affected by the surrounding stroma [29]. The tumour stroma includes a specific ECM, as well as cellular components such fibroblasts, blood vessel cells, infiltrating immune cells and fat cells. This aberrant stroma produces a broad spectrum of microenvironmental cues, such as diverse growth factors, cytokines, proteolitic enzymes and angiogenic factors, and modulates tumour expansion and metastasis. Tumour-stroma interaction was demonstrated experimentally by different studies. Billingham and colleagues transplanted untreated skin epithelial cells on a stroma treated with mutagens and showed that this aberrant stroma was able to mutate epithelial cells [30]. Similarly, Barcellos-Hoff and Ravani induced neoplastic progression of healthy mammary epithelial cells by injecting it into irradiated mammary fat-pad [31]. Stroma normalization is another phenomenon that highlights the importance of tumour-stroma crosstalk. Thus, Bissell and colleagues succeeded to revert pre-malignant epithelial cells from the breast by reconstituting the basement membrane [32]. An interesting fact was demonstrated by Moinfar and colleagues, who studied genetic alterations in cancer-associated stroma of mammary carcinomas and found distinct chromosomal rearrangements, which were not present before in the carcinoma cells [33]. These facts indicate that aberrations in stroma cells may both commence and promote carcinoma development.

Cancer cells themselves can modify their adjacent stroma, making the environment supportive and favourable for tumour progression. This cancer prone stroma, observed in dysplasias, also called reactive [29]. Cancer cells produce a set of stroma-modulating factors, such as bFGF, VEGF, PDGF, EGFR, TGFβ, interleukins and others. These factors change stroma in a way that it resembles a healing wound, provoke inflammation and induce angiogenesis [34].

I-6 Inflammation and cancer

Inflammation is a complex biological response to injurious stimuli such as pathogens, physical damage and irritants. It is a protective system of the organism directed to restore tissue integrity. In the absence of inflammation a wound would not heal. Usually inflammation is self-limited and disappears after restoration of tissue integrity. In certain disease states the inflammation process becomes uncontrolled and acquires a persistent character. The pathways behind chronic inflammation, leading to diseases such as cancer, arthritis, type 2 diabetes and cardiovascular diseases are still unknown [35].

Cancers are thought to be "wounds that do not heal" [36]. It is assumed that chronic inflammation acts as a trigger of cancerogenesis in about 15-20% of all human cancers [37]. One of the best studied examples is chronic inflammation of the stomach induced by *Helicobacter pylori*, which can lead to gastritic cancer [38]. During inflammation the immune competent cells migrate to the site of injury and activate a cascade of events leading to the recognition of and fight with pathogens or irritant agents. This process is accompanied by secretion of a wide spectrum of immune factors, such as growth factors, cytokines, chemokines, and by release of reactive oxygen (ROS) and nitrogen (RNS) species by immune cells. In chronic inflammation intensive exposure to these factors can lead to proliferation of cells and microenvironment reorganization [39]. ROS and RNS unselectively target not only microbial infection, but also host cells, which results in generation of genomic mutations and rearrangements [40]. Another fact, highlighting the involvement of inflammation in cancerogenesis, is that growth factors and other wound resolving signals in direct or indirect way participate in malignization process and support tumour outgrowth [41]. Convincing evidence of such a phenomenon is tumour promotion by scar effect at surgical resection sites.

Despite the importance of tumour microenvironment in malignization, our understanding of tumour-stroma interaction still remains limited.

I-7 Role of fibroblasts in cancer

Fibroblasts are central components of the stroma. They are responsible for the synthesis and maintenance of the ECM. Fibroblasts are identified by their spindle-like shape and ability to adhere to plastic. They are highly heterogeneous, which is also defined by their anatomical topology. It was demonstrated experimentally by genome-wide expression analysis of fibroblasts from 43 unique anatomic sites [42].

Induced by different stimuli the fibroblasts can switch to an activated phenotype. This implicates fibroblasts not only in normal processes as tissue injury, inflammation and wound healing, but also in cancers. Once activated, fibroblasts acquire proliferative capacity and differ from their default counterparts. Such activated fibroblasts in the stroma that underlies tumours are known as carcinoma-associated fibroblasts (CAFs) or tumour associated fibroblasts. CAFs produce alpha-smooth muscle actin (α-SMA), which allows them to migrate into areas of damage. CAFs also secrete MMPs, which allows spreading of cells, as well as different growth and chemotactic factors that orchestrate inflammation and angiogenesis [43]. Cancer associated fibroblasts (CAF) share several features with activated fibroblasts, which participate in wound healing (see Figure I-8). The main difference between them is that activated fibroblasts are able to become inactive by attenuation of activating cues, whereas CAFs retain their activity continuously [44]. This continuing activity can result in tumour associated fibrosis, which finally can lead to organ destruction through excessive production and deposition of ECM by CAFs.

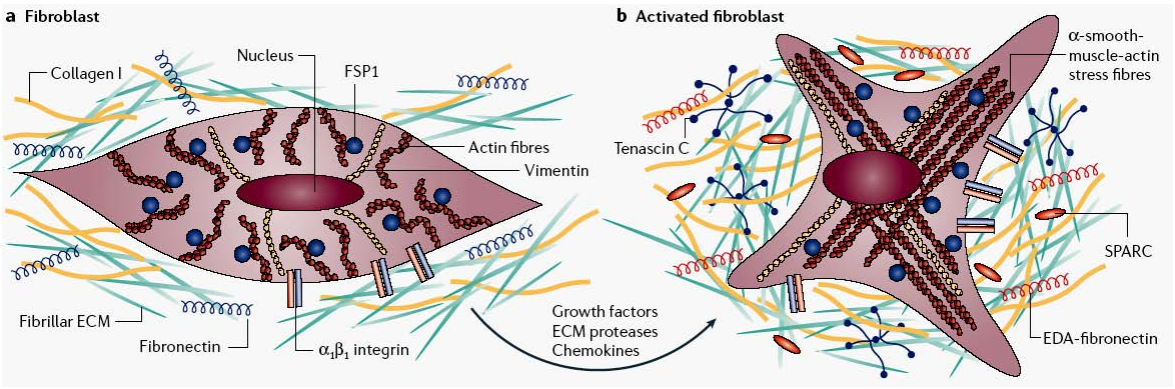


Figure I-8 Comparison of non active fibroblast with activated fibroblast.

In contrast to activated fibroblasts *(b)*, inactive fibroblast *(a)* has smaller endoplasmatic reticulum and flattened heterochromatic nucleus. Fibroblast is activated by different stimuli, arising after tissue injury. Such stimuli can be growth factors, ECM proteases and chemokines. Active fibroblast has large oval nucleus and rough endoplasmatic reticulum [45]. Activated fibroblast has proliferative activity and enhanced ECM protein production, e.g. type I collage, tenascin C, extra-domain a fibronectin and SPARC. Image from [44].

I-7.1 Seed and soil theory and metastasis

Metastases, rather than the primary tumour itself, are responsible for most (90%) cancer deaths [46]. The term metastasis comes from Greek (*meta*=next and *stasis*=placement) and describes a process of displacement. In tumour biology, metastasis is a multi-stage process, in which cancer cells leave the tumour site and migrate to distant organs and tissues via blood and lymphatic vessels to initiate a secondary tumour.

Metastases have organ-specific patterns. For example, colon carcinomas frequently metastasize to lung and liver, but seldom to skin, bone or brain and nearly never to intestine, kidneys or muscles [47]. Tumour dissemination process is inefficient and only few of many migrating cells successfully developed into distinct secondary tumour. During migration the metastasising cell is faced to many different stresses, such as loss of adhesion, haemodynamic sharing, amorphosis, anoikis, hypoxia, nutrient depletion, etc. For successful accomplishment of invasion metastasising cell must overcome these stresses [48].

Carcinoma cells need a set of proteins that can digest the ECM and basement membrane to enable the penetration into the stroma. Those cells, which survive anoikis and amorphosis are now capable to migrate and intravasate (a process, when the cell penetrates to blood and lymphatic vessels). After intravasation cells will be stressed mechanically by bloodstream, arrest in capillary, immune attack and other factors. In order to survive circulation, cancer cells have to adapt their genetic and physiological homeostasis for circumventing both apoptosis and recognition by immune surveillance [49]. Several studies showed that carcinoma cells injected into the bloodstream die either directly in the blood vessel, or shortly after extravasation [50, 51]. Furthermore, metastasising cells need to survive in the secondary site, recapitulate adhesivity, provoke angiogenesis and resist to cell death signals. For successful metastatic colonization cancer cells must either be adapted to, or modulate the microenvironment in site of metastasis [52]. Already in 1889 the English surgeon Stephen Paget (1855-1926) stated that metastases do not occur by chance. Instead, a metastatic cell (seed) needs an appropriate environment that provides growth advantage (soil) in order to propagate successfully [53]. He postulated “When a plant goes to seed, its seeds are carried in all directions but they can only live and grow if they fall on congenial soil”.

I-8 Cancer cell and motility

Cell migration is a key event in diverse biological processes such as development, trafficking of immune cells and wound healing [54]. A good example of cell migration in development is the gastrulation, where a collective of cells migrates as sheets inside the blastocyst and forms three embryonic germ layers: ectoderm, endoderm and mesoderm.

In immune surveillance the circulating immune cells migrate from blood vessels to the injured area, where they can resolve inflammation and participate in wound healing. Cell migration is also involved in numerous diseases, e.g. cancer metastasis, mental retardation, vascular disease and arthritis [55]. Active migration of immune cells is also observed in arthritis and other chronic inflammatory diseases. Impaired motility in neuronal cells at early stages of embryo development leads to congenital brain malformation and mental retardation.

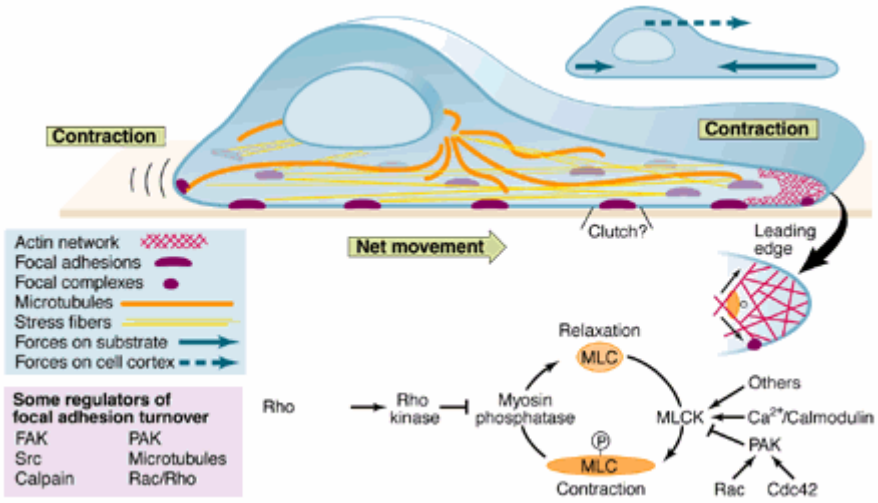


Figure I-9. Migrating cell [56].
The migrating cell modifies its cytoskeleton and adhesivity, changes the shape, and acquires or reorganizes polarization. To migrate, the cell must fulfill at least five different steps: *(a)* protrusion of the leading edge, *(b)* formation of focal contacts with ECM, *(c)* recruitment of surface proteases, *(d)* cell contraction by actin-myosin complexes, and *(e)* loss of adhesion at the rear side of migrating cell.

Migrating cells need several pivotal events, switching them from stationary to motile form (Figure I-9). At first, the cell polarizes and elongates, then it forms adhering to ECM protrusions. Protrusions are realized by actin filaments, which collaborate with other proteins and push the membrane outward. The establishment of cell-matrix interaction and formation of focal contacts commences with binding of integrins to the ECM. The

cytoplasmic tale of integrin interacts both directly and indirectly with various proteins including talin, α-actinin, focal adhesion kinase, tensin, vinculin and paxillin, and forms focal contacts [54]. Different in number and size focal contacts involve different types of integrins, which depends on cell types and environmental conditions. Different proteases participate in degradation of ECM proteins to generate space for the moving cell. Movement is fuelled by a complex of filamentous actin and myosin. This complex, called actomyosin, generates traction force and allows cell to move [57]. The last process in cell migration is adhesion release at the rear edge of the cell that is accompanied by disassembly of actin filaments and detachment of integrins from the ECM. The process of adhesion assembly and disassembly occurs continuously throughout the whole movement process.

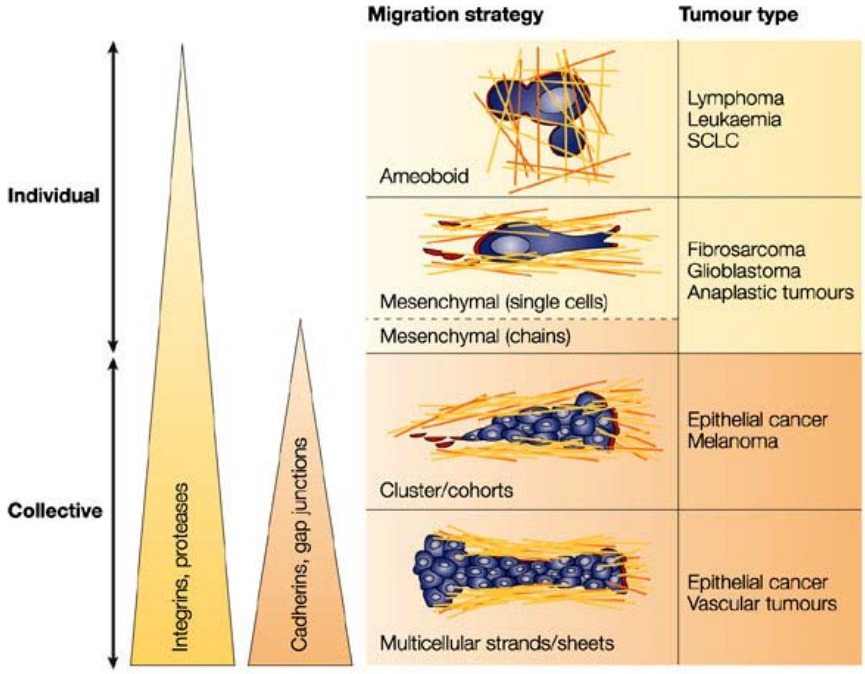


Figure I-10. Different mechanisms of tumour cell migration [54].
Individual or collective movement occurs by the use of different pathways, involving integrins, proteases, cadherins and gap junctions (triangles). Cadherins and gap junction proteins, being key players of cell-cell adhesion are involved in collective migration, as well as in mesenchymal chain migration. Integrins and proteases are more responsible for cell-matrix interactions; therefore they promote single cell migration.

Cancers exhibit different cell migration mechanisms and patterns. Depending on cancer type the cells may migrate as single cells (individual migration), or as cell groups and patches (collective migration) (Figure I-10) [54]. Individual migration strategy can be classified in two groups: amoeboid and mesenchymal. Amoeboid migration is typical to haematopoietic neoplasias (leukemias and lymphomas) and small cell lung carcinoma

(SCLC). By contrast, mesenchymal migration is mainly observed in tumours originating from connective tissues, including fibrosarcomas [58], gliomas [59], as well as in progressively de-differentiated carcinomas [60, 61]. In contrast to mesenchymal migration strategy, the amoeboid one is less adhesive (lacks mature focal contacts and stress fibres) and have higher migration velocity (comprehensively reviewed by Friedl and Wolf [54]).

I-9 Epithelial - mesenchymal (EMT) and mesenchymal - epithelial transition (MET)

Epithelia are sheets or tubes of tightly adhering cells that can originate from any germ layer. Mesenchyme refers to a tissue, which is loosely packed and descends from the mesoderm or neural crest [62]. Epithelial cells functionally and morphologically differ from mesenchymal cells. They have apico-basal and lateral polarity and build strong cell-cell (tight junctions, adherens junctions, gap junctions and desmosomes) and cell-matrix (focal adhesion) contacts. Normal epithelial cells usually do not exhibit trans-tissue migration and do not detach from basal lamina. The mesenchymal cells, indeed, are limited in adhesion capacity. They have front-to-back polarity and use filopodias for movement. A number of studies propose different markers used to discriminate epithelial and mesenchymal cells (Table I-1).

Table I-1. List of epithelial and mesenchymal markers [63].

Epithelial markers	Mesenchymal markers
E-cadherin (CDH1, or CD324)	Vimentin (VIM)
Claudins (CLDN)	Alpha-smooth muscle actin (α-SMA)
Occludin (OCLN)	Fibronectin (FN1)
Desmoplakin (DSP)	Vitronectin (VTN)
Different cytokeratins	S100A4 (FSP1)
Mucin 1 (MUC1)	Fibroblast growth factor receptor 2 (FGFR2, or CD332)

Epithelial cells can be transformed to fibroblast-like mesenchymal cells. First report of this transformation comes from 1960s [64]. Later, in 1982, it was demonstrated by Greenburg and Hay, who observed migration of individual mesenchymal-like cells by culturing embryonal and adult lens cells in 3D collagen gel [65, 66]. This phenomenon, called epithelial-mesenchymal transition (EMT), is orchestrated by a panel of intra- and extracellular cues. During EMT, epithelial cells undergo a stepwise transition to

mesenchymal cells characterized by cytoskeleton reorganisation, loss of apico-basal polarity and cell-cell adhesion (Figure I-11). These cells change theirs shape, detach from the basal membrane and acquire motile capacity. EMT is often accompanied by up-regulation of proteolitic enzymes such as MMPs. Different mechanisms and pathways, underlying EMT process, show overlap with biological processes such as cell adhesion, migration, survival and differentiation (reviewed by [63] and [67]). TGF-β, Notch and Wnt (both canonical and non-canonical) signalling pathways were proposed to play a crucial role in regulation of EMT in development and disease. Transition of epithelial to mesenchymal differentiation is often characterised by an increased expression of vimentin, N-cadherin, Snail1 (Snail), Snail2 (Slug), Twist, and a decrease of E-cadherin, desmoplakin, occludin and different cytokeratins. Another important hallmark of EMT is β-catenin dislocation from the cell membrane, where it serves as an adherens junction protein, into the nucleus, thus functioning as a transcription factor [68]. A description of proteins involved in EMT is presented in Table I-2.

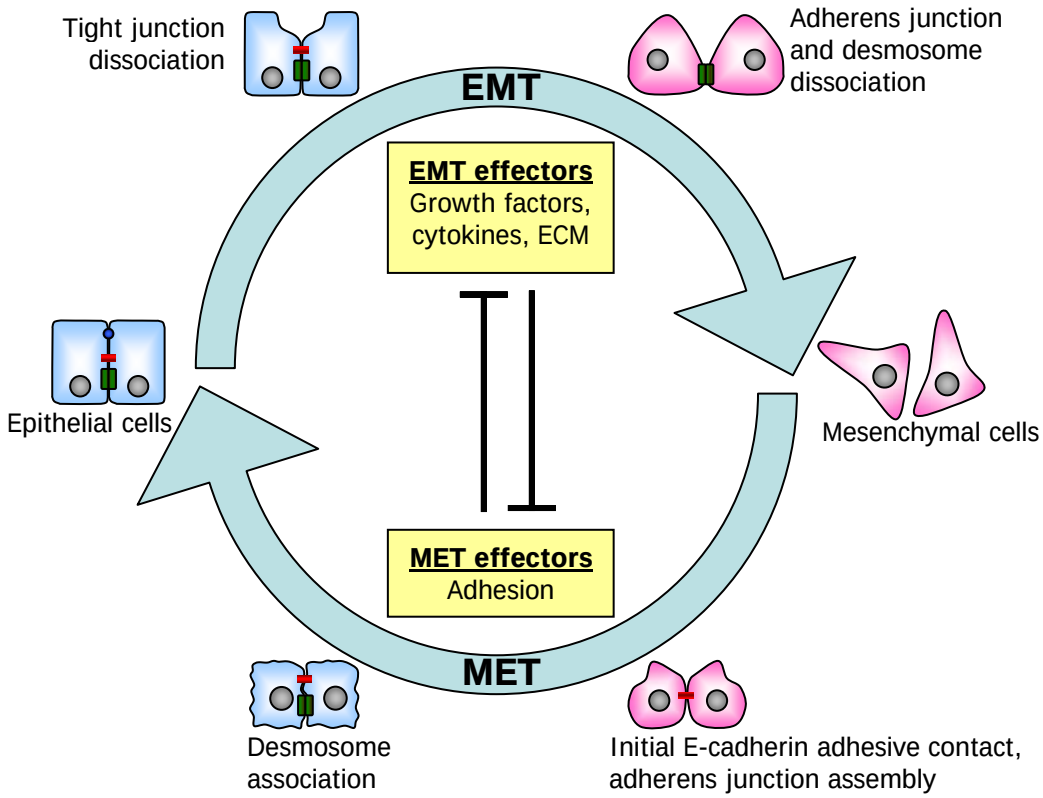


Figure I-11. The cycle of EMT and MET (adapted from [63])

Table I-2. Summary of EMT markers (adapted from [69])

Increased proteins	†Description
Vimentin	Along with the microfilaments (actins) and microtubules (tubulins), the intermediate filaments represent a third class of well-characterized cytoskeletal elements. Vimentins are found in various non-epithelial cells, especially mesenchymal cells.
N-cadherin	Neuronal cadherin is a calcium dependent cell-cell adhesion glycoprotein. Functions during gastrulation and is required for establishment of left-right asymmetry.
Fibronectin	Involved in cell adhesion and migration processes including embryogenesis, wound healing, blood coagulation, host defence, and metastasis
Snai1 (Snail)	A zinc finger transcriptional repressor which downregulates the expression of ectodermal genes within the mesoderm. The nuclear protein encoded by this gene is also thought to be critical for mesoderm formation in the developing embryo.
Snai2 (Slug)	Involved in epithelial-mesenchymal transitions and has antiapoptotic activity. Mutations in this gene may be associated with sporadic cases of neural tube defects.
Twist	Transcription factor, which seems to be involved in the negative regulation of cellular determination and in the differentiation of several lineages including myogenesis, osteogenesis, and neurogenesis. Also represses expression of proinflammatory cytokines such as TNFA and IL1B.
Goosecoid	Regulates chordin (CHRD). May play a role in spatial programming within discrete embryonic fields or lineage compartments during organogenesis.
FOXC2	Belongs to the forkhead family of transcription factors. May play a role in the development of mesenchymal tissues
Sox10	Encodes a member of the SOX (SRY-related HMG-box) family of transcription factors, involved in the regulation of embryonic development and in the determination of the cell fate. Acts as a nucleo-cytoplasmic shuttle protein and is important for neural crest and peripheral nervous system development.
MMP2	Matrix metalloproteinase-2, or Gelatinase A, involved in the breakdown of ECM in normal physiological processes, such as embryonic development, reproduction, and tissue remodelling, as well as in disease processes, such as arthritis and metastasis. Degrades type IV collagen, the major structural component of basement membranes.
MMP3	Matrix metalloproteinase-3, or Transin-1, degrades fibronectin, laminin, collagens III, IV, IX, and X, and cartilage proteoglycans. Thought to be involved in wound repair, progression of atherosclerosis, and tumour initiation.
MMP9	Matrix metalloproteinase-9, or Gelatinase B, degrades type IV and V collagens. Studies in rhesus monkeys suggest that the enzyme is involved in IL-8-induced mobilization of hematopoietic progenitor cells from bone marrow, and murine studies suggest a role in tumour-associated tissue remodelling.
Integrin $\alpha 5\beta 6$	Integrin $\alpha 5$ is receptor for vitronectin, cytotactin, fibronectin, fibrinogen, laminin, matrix metalloproteinase-2, osteopontin, osteomodulin, prothrombin, thrombospondin and von Willebrand factor. Integrin $\beta 6$ associates with integrin αV.

Decreased proteins	Description (RefSeq)
E-cadherin	Epithelial cadherin, a calcium dependent cell-cell adhesion glycoprotein. Loss of function is thought to contribute to progression in cancer by increasing proliferation, invasion, and/or metastasis.
Desmoplakin	An obligate component of functional desmosomes that anchors intermediate filaments to desmosomal plaques.
Cytokeratin	
Occludin	Located at tight junctions and may be involved in its formation and maintenance.

Accumulated in the nucleus proteins	Description
β-catenin	Adherens junction protein. Adherens junctions (AJ) are critical for the establishment and maintenance of epithelial layers. AJs regulate normal cell growth and behaviour at several stages of embryogenesis, wound healing, and tumour cell metastasis, cells form and leave epithelia. Disruption and reestablishment of epithelial cell-cell contacts may be regulated by the disassembly and assembly of AJs.
Smad-2	Mediates the signal of TGF-β, and thus regulates multiple cellular processes, such as cell proliferation, apoptosis, and differentiation.
NF-κβ	NF-κβ is a transcription regulator that is activated by various intra- and extra-cellular stimuli such as cytokines, oxidant-free radicals etc. Activated NF-κβ translocates into the nucleus and stimulates the expression of genes involved in a wide variety of biological functions.
Snai1 (Snail)	see above
Snai1 (Slug)	see above
Twist	see above

† Description obtained from RefSeq

EMT is an important process in embryonic development. It is implicated in gastrulation, where cells of the primitive ectoderm undergo morphogenesis, actively migrate inward in the blastula and in this way generate the third layer called mesoderm [70]. Later on the mesoderm gives rise to muscle and connective tissues [67]. Another example of EMT in development is formation of the neural crest. The neural crest is a population of precursor cells, which is formed in vertebrate embryos and is positioned between neural plate and epidermal ectoderm [71]. Here, EMT enables the cells to migrate long distances and to initiate formation of different tissues and cells, including bones, connective tissue, smooth muscles, melanocytes, neurons and glia of peripheral nervous system. EMT is also necessary for developmental event in cardiac valve and secondary plate formation [69].

EMT is not an irreversible process. For instance the reverse process called mesenchymal-epithelial transition (MET) is known to be crucial in nephron epithelium formation [72-74].

EMT and MET are not limited to development, but also relevant to diseases such as fibrosis [75-79], proliferative vitreoretinopathy [80-83], arthritis [84, 85] and, importantly, in cancer [67, 86-88]. For instance kidney fibrosis in chronic renal disease, which can lead to organ dysfunction, is caused by abundant secretion of ECM proteins by myofibroblasts. The myofibroblasts are thought to origin from tubular epithelial cells that have undergone EMT [89]. High levels of involved in EMT proteins such as TGFβ, FGF2 and EGF were observed in fibrosis. Moreover, TGFβ-induced EMT of epithelial cells of the lens led to cataracts and ocular fibrotic disease [90]. Additionally, EMT can result in cell migration and retina

detachment in proliferative vitreoretinopathies [83]. An EMT-like process also orchestrates transition of synovial tissue in rheumatoid arthritis [84, 91].

Occurrence of EMT in cancer metastasis is discussed by a number of scientists, insisting on high impact of EMT in carcinoma progression and metastasis [69]. According to them the carcinoma, being a malignant population of epithelial cells, undergoes transition to its abnormal and atypical mesenchymal counterparts and drastically facilitates tumour metastasis. The malignant cell, which underwent EMT and dislocated into distant site, may recapitulate secondary tumour via MET. Based on a fact that E-cadherin was expressed in metastasising breast carcinoma cells, Tarin and colleagues suggested that these cells migrate without undergoing EMT [92, 93]. Also, there are certain cases, where the carcinoma cells exhibit collective movement, and the metastasis is developed from invasion of those cells [94].

Materials & Methods

II Materials and Methods

II-1 Materials

II-1.1 Chemicals and chemical compounds

1.1.a General chemicals and compounds

Table I-3. Standard chemicals and reagents

Reagent Name	Supplier
Acetic acid	Merck, Germany
AdvaSon	Advalytix, Germany
Agarose, Electrophoresis Grade	Invitrogen, USA
Aqua ad iniectabilia	Baxter, Germany
Bromphenol blue	Serva, Germany
Cy3™ -dUTP	Amersham, USA
Chloroform	Sigma, USA
Cy5™ -dUTP	Amersham, USA
DAPI (4,6-diamino-2-phenylindole-2HCl·H_2O)	Serva, Germany
dATP	Roche, Germany
dCTP	Roche, Germany
Dextran Sulphate	USB, USA
dGTP	Roche, Germany
DIG-11-dUTP	Roche, Germany
DMSO	Sigma, USA
dUTP	Roche, Germany
EDTA	Merck, Germany
Ethanol	Merck, Germany
Ethidium Bromid	Sigma, USA
Formaldehyde – 37.0% (v/v)	Merck, Germany
Formamide	Merck, Germany
Glycerin	Merck, Sigma
Hydrochloric acid	Merck, Germany
Imperial protein stain	Pierce, USA
Isopropanol	Merck, Germany
Magnesium chloride	Merck, Germany
Methanol	Merck, Germany
Mineral oil	Sigma, USA
NP40	Merck, Germany
Paraformaldehyde	Sigma, USA
Phenol	Roth, Germany
Phenol:Chloroform:Isoamyl Alcohol 25:24:1 saturated with TE	Sigma, USA
S-adenosylmethionine (SAM)	New England Biolabs, UK
SDS, 10%	Merck, Germany
Sodium perchlorate	Merck, Germany
Sodium acetate	Merck, Germany
Sodium azide (NaN_3)	Sigma, USA

Sodium hydroxide	Merck, Germany
Sodium perchlorat	Merck, Germany
TE Buffer	Millipore, USA
Tris	Merck, Germany
Tris·HCl	Merck, Germany
Triton X-100	Sigma, USA
Tween-20	Sigma, USA
UltraPure Glycerol	Invitrogen, USA
β-mercaptoethanol	Merck, Germany

1.1.b Buffers and solutions

The buffer stock solutions were prepared with ddH_2O. These buffers were then diluted with Millipore water. All buffers used for IP, co-IP, methylated DNA immunoprecipitation (MeDIP), methylated CpG island amplification (MCA) and microarray experiments were completely prepared with Millipore water.

Table I-4. Buffers and solutions

Buffer Name	Composition/Supplier
10x dNTP mix	0.5mM dATP; 0.5mM dGTP; 0.5mM dCTP
10x Nick Translation buffer	0.5M Tris·HCl (pH7.9); 50mM $MgCl_2$, 0.5mg/ml BSA
ImmunoPure Gentle Ab Elution Buffer	Pierce, USA
10xPBS	1.37 M NaCl; 27.0 mM KCl; 100 mM Na2HPO4; 20.0 mM KH2PO4, pH 7.4
10xPEM	1M PIPES; 50mM EGTA; 20mM $MgCl_2$, pH7.0
10xTBE	900 mM Tris; 900mM boric acid; 20mM EDTA, pH 8.0
1x PBST	1xPBS; 1:1000 Tween-20
20xSSC	3.0M NaCl; 300mM sodium citrate, ddH_2O, pH 7.0
2x Magic Mix	48% Urea (BioRad); 15mM Tris·HCl pH 7,5; 8,7% Glycerin; 1%SDS; 0,004% Bromphenol blue; 143mM β-mercaptoethanol (add fresh)
50xTAE	2.0M Tris; 1.0M acetic acid; 50mM EDTA, pH 8.0
Blocking solution	0.1% SDS; 4xSSC; 0.5% BSA
Ethidium Bromide solution	10mg/ml EtBr in ddH_2O
FDST	0.8g dextran sulfate, 5ml formamide, 1ml 20xSSC, 7ml ddH_2O
Humidifying buffer	4% FDST
co-IP buffer	50mM Tris (pH7.5), 100mM NaCl, 5mM $MgCl_2$, 1mM DTT, Complete Protease Inhibitor (Roche, Germany), 0.5% NP40
5x SDS-PAGE buffer	1.5ml β-mercaptoethanol, 1.5g SDS, 0.15g bromphenol blue, 5.75ml 87% glycerin, buffer 2 (0.5M Tris, 0.4% SDS, pH6.8) up to 10ml
IP buffer (for MeDIP)	10mM sodium phosphate, 140mM NaCl, 0.05% Triton X-100, pH7.0
Laemmli buffer	25mM TRIS p.a, 190mM Glycin, 0.1% SDS in ddH_2O
Mounting medium	90% glycerol; 0.1M Tris-HCl pH 8.0; 2.3 % DABCO
PN buffer	100mM sodium phosphate (pH 8,0); 0,05% NP40
Prehybridization buffer	25% formamide; 4xSSC; 0,1%SDS

TE buffer	10mM Tris·HCl; pH 7.5/8.0; 1.0mM EDTA, pH 8.0
Tris·HCl	1.0M Tris·HCl, pH 7.0/7.5/8.0
Washing buffer	50% formamide; 2xSSC; 0,1%SDS
WB blocking buffer	5% milk in PBST
10x DIG-11-dUTP/dUTP	125µl 1mM DIG-11-dUTP; 750µl 0.5mM dUTP; 125µl TE (pH7.9)

1.1.c Transfection reagents

Table I-5. Transfection reagents

Reagent name	Supplier
FuGENE HD	Roche, Germany
Oligofectamine reagent	Invitrogen, USA

II-1.2 Biological reagents

1.2.a RNA/DNA

Table I-6. RNA/DNA

Name	Supplier
Herring sperm DNA	Invitrogen, USA
Yeast tRNA (Lyophilized)	Invitrogen, USA
DNA size marker Hyperladder I, 200-10000bp	Bioline, Germany
DNA size marker Hyperladder II, 50-2000bp	Bioline, Germany
Human Cot-1 DNA	Roche, Germany
SPON2 pCMV6 Entry, NM_012445.1 (SPON2)	OriGene, USA

1.2.b Primers

All PCR primers were ordered from MWG (Germany) and listed in Table I-7:

Table I-7. List of primers

Name	Aim	Forward sequence	Reverse sequence
Chr1-1-2	BSP	5'-ttgaggtaggagaatagttgga-3'	5'-atcccctctaaacctaaaaacc-3'
Chr1-2-3	BSP	5'-taaaaataatcgtttaaagtttgtagt-3'	5'-acaaaaacaaaacccacc-3'
Chr9-1	BSP	5'-gggtgaggttttttaggtaat-3'	5'-tacctacaacaaatcaaccaaa-3'
Chr11-1	BSP	5'-ggggagttttaggagggt-3'	5'-attctccatcatattaacctc-3'
Chr19-2	BSP	5'-ggggtagttggttttgttttt-3'	5'-aaaaaacctattccccaacct-3'
Chr3-3	BSP	5'-attatggtattgagggatgagg-3'	5'-aaccctaaaacaacaacaatcc-3'
Primer-12	Different	5´-taactagcatgc-3´	-
Primer-21	Different	-	5'-agtgggattccgcatgctagt-3'
RMCA	MCA	5'-ccgggcagaaag-3'	5'-ccaccgccatccgagcctttctgc-3'
Exon2	MeDIP	5'-gcccctctcctgaacttacc-3'	5'-ggtcactgtcactaccaaggc-3'
Exon3	MeDIP	5'-ctctgaatccagggccg-3'	5'-tgcaggtgcgtgtagttgag-3'
SPON2	qRT-PC	5'-ccagcagggacaatgagat-3'	5'-acgtagcgagtcctgctct-3'
HPRT	qRT-PCR	5'-ttgttgtaggatatgcccttg-3'	5'-ccagatgtttccaaactcaact-3'
MMP1	qRT-PCR	5'-aacaaatactggaggtatgatgaa-3'	5'-gtcaaaattctcttcgttttagca-3'
SPON2-orf	Sequencing	5'-gacgttgtatacgactcc-3'	5'-tgctgccagatcctcttc-3'

1.2.c siRNA

Table I-8. List of siRNAs

Name	Target	Target sequence	Supplier
Klaus	GFP	5'-gcaagctgaccctgaagttca-3'	Qiagen, Germany
M2	MMP1, NM_002421	5'-cagattctacatgcgcacaaa-3'	Qiagen, Germany
SP1	SPON2, NM_012445	5'-cagcatcaccttcacgggcaa-3'	Qiagen, Germany

1.2.d Proteins

Table I-9. Proteins

Protein name	Supplier
BSA, cryst. lyophil.	Sigma, USA
Human MMP1 recombinant protein	R&D Systems, USA
Human SPON2 recombinant protein	Abnova, Taiwan
Kaleidoscope Precision Protein Standards	Bio-Rad, USA

1.2.e Enzymes

Table I-10. List of used enzymes

Enzyme	Supplier
Alu I	Promega, USA
AmpliTaq® DNA Polymerase	Applied Biosystems, USA
BstU I	New England Biolabs, UK
DNA Polymerase I	Roche, Germany
DNase I (RNase free)	Roche, Germany
Hha I	New England Biolabs, UK
Hpa II	New England Biolabs, UK
M.Sss I (CpG Methylase)	New England Biolabs, UK
McrBC	New England Biolabs, UK
Pepsin	Serva, Germany
Proteinase K	Roche, Germany
RNase A	Roche, Germany
Rsa I	Promega, USA
Sma I	New England Biolabs, UK
T4 DNA Ligase	Roche, Germany
TITANIUM™ Taq DNA Polymerase	BD Clontech, USA
Xma I	New England Biolabs, UK

1.2.f Antibodies

Table I-11. Primary antibodies

Antigen	Conjugate	Clonality / Isotype	Supplier
5-methylcytidine	-	Monoclonal / mouse IgG_1	Eurogentec, Belgium
Mindin (N-13)	-	Polyclonal / goat IgG	Santa Cruz Biotechnology, USA
Mindin (W-18)	-	Polyclonal / goat IgG	Santa Cruz Biotechnology, USA
MMP-1 (3B6)	-	Monoclonal / mouse IgG1	Santa Cruz Biotechnology, USA
SPON2	-	Polyclonal / mouse	custom
Vimentin	Cy3	Monoclonal / mouse IgG_1	Sigma, USA
Tubulin	-	Monoclonal / rat IgG_{2a}	Serotec, UK
Lamin A/C	-	Monoclonal / mouse IgG_{2a}	Santa Cruz Biotechnology, USA
Tubulin YL1/2	-	Monoclonal / rat IgG_{2a}	Abcam, UK

Table I-12. Secondary antibodies

Antigen	Conjugate	Supplier
FLAG	FITC	Sigma, USA
FLAG	Cy3	Sigma, USA
Goat IgG	HRP	Sigma, USA
Mouse IgG	Cy3	Sigma, USA
Mouse IgG	FITC	Sigma, USA
Mouse IgG	HRP	
Rabbit IgG	HRP	
Rat IgG	Alexa 597	Molecular Probes

II-1.3 Cell Lines

Table I-13. Cell lines

Name	Description	Notes*	Medium
4981	Human, EBV-transformed 46,XY B-lymphocytes	T	10.0% (v/v) FBS; 100 U/ml penicillin; 68.6 µM streptomycin; 2mM L-Glu *ad volumina* RPMI-1640
HeLa	Human, epitheloid cervix carcinoma	T, A	10.0% (v/v) FBS; 100 U/ml penicillin; 68.6 µM streptomycin; 2.00 mM L-Glu, *ad volumina* DMEM
Hey	Human, ovarian cancer	T, A	
IMR-90	Human, fetal lung fibroblasts	P, A	10.0% (v/v) FBS; 100 U/ml penicillin; 68.6 µM streptomycin; 2mM L-Glu *ad volumina* EMEM
Jurkat	Human, T-lymphocyte	T, S	
K562	Human, Erythroleukemia cell line	T, S	
MCF-7	Human, breast adenocarcinoma	T, A	
Mel2A	Human, skin melanoma	T, A	
MeWo	Human, skin melanoma	T, A	
SCLC A	Human, small cell lung carcinoma	T, M	
SH-SY5Y	Human, Neuroblastoma SK-N-SH subtype from metastatic bone tumor	T, M	15.0% (v/v) FBS; 100 U/ml penicillin; 68.6 µM streptomycin; 2.00 mM L-Glu, *ad volumina* DMEM
SW13	Human, adrenal cortex adenocarcinoma	T, A	
SW480	Human, colon adenocarcinoma	T, A	
U373MG	Human, glioblastoma astrocytoma	T, A	10.0% (v/v) FBS; 100 U/ml penicillin; 68.6 µM streptomycin; 2.00 mM L-Glu; 1.00% (v/v) NEAA; 1.00 mM sodium pyruvate, *ad volumina* EMEM

Table I-14. Cell culture reagents and media

Cell culture reagents and media	Supplier
DMEM	Lonza, Switzerland
DMSO	Sigma, USA
DPBS	Lonza, Switzerland
EMEM	Lonza, Switzerland
FBS	Sigma, USA
L-Glutamine, 200mM	Lonza, Switzerland
OptiMEM	
Penicillin, 10 000U/ml	Lonza, Switzerland
Quantum 263 with L-Glu.	PAA, Austria
Sodium pyruvate, 100mM	Lonza, Switzerland
Streptomycin, 10 000U/ml	Lonza, Switzerland
Trypsin-EDTA	Lonza, Switzerland

* P=primary, T=transformed, A=adherent, S=suspension, M=mixed

II-1.4 Kits and Microarrays

Table I-15. Kits

Kit Name	Supplier
BigDye Terminator Mix V3.1	Applied Biosystems, USA
BioPrime® Array CGH Genomic Labeling System	Invitrogen, USA
BioPrime® Array CGH purification module	Invitrogen, USA
Bio-X-Act Long Mix	Bioline, Germany
Dynabeads® M-280 Sheep anti-Mouse IgG	Invitrogen, USA
EZ DNA Methylation-Gold Kit™	Zymo Research, USA
Gene Expression Hybridization Kit	Agilent, USA
Gene Expression Wash Buffer Kit	Agilent, USA
GenomePlex™ WGA Kit	Sigma, USA
Genomic DNA Labeling Kit PLUS	Agilent, USA
Hybridization Backing Kit, 1 x 244K format	Agilent, USA
Hybridization Backing Kit, 4 x 44K format	Agilent, USA
Microcon® Centrifugation Filter Devices YM-30	Millipore, USA
NoEndo Endotoxin-Free JETstar Plasmid Kit	Genomed, Germany
Oligo aCGH/ChIP-on-chip Hybridization Kit	Agilent, USA
Oligo aCGH/ChIP-on-chip Wash Buffer Kit	Agilent, USA
One-Color RNA Spike-In Kit	Agilent, USA
Qiaquick® PCR Purification Kit	Qiagen, Germany
QIAshredder	Qiagen, Germany
Quick Amp Labeling Kit, one-color	Agilent, USA
RNeasy® Mini-Kit	Qiagen, Germany
SilverQuest™ Silver Staining Kit	Invitrogen, USA
SulfoLink Coupling Gel	Perce, USA
SYBR® Green PCR Master Mix	Applied Biosystems, USA
TaqMan® Reverse Transcription Reagents	Applied Biosystems, USA
Western Lightning® Detection Kit	Perkin Elmer, USA

Table I-16. Microarrays

Microarray	Supplier
Human CpG Island Microarray, 1x244K	Agilent, USA
Custom High-Definition CGH Microarray, 4x44K	Agilent, USA
Whole Human Genome Oligo Microarray, 4x44K	Agilent, USA
Human Genome CGH Microarray 244A, 1x244K	Agilent, USA

II-1.5 Materials, plastic-/glassware and disposables

Table I-17. Materials, plastic-/glassware and disposables

Name	Supplier
0,025 µm filters	Millipore, USA
Cover Glasses, 18x18mm	Menzel-Glaser, Germany
Cover Glasses, 20x20mm	Menzel-Glaser, Germany
Cover Glasses, 24x24mm	Menzel-Glaser, Germany
Cover Glasses, 24x60mm	Engelbrecht, Germany
Cover Glasses, Ø 10mm	Menzel-Glaser, Germany
Dynal MPC-S magnetic rack	Invitrogen, USA
Eppendorf® Safe-Lock® Tubes, 1.5ml, 2.0ml	Eppendorf, Germany
Fuji Medical X-Ray film	Fuji, USA
MicroAmp™ Optical 96-Well Reaction Plate	Amersham, USA
MicroAmp™ Optical Adhesive Film	Amersham, USA
Microscope slides, SuperFrost	Roth, Germany
Nunc slides	Nunc, USA
PAP Pen, immunostaining	Kisker, Switzerland
Parafilm	Pechiney Plastic Packaging, USA
Pasteur pipette	Roth, Germany
Pipett boy	Integra Biosciences, Switzerland
Pipettes	Gilson and Eppendorf
RNaseZAP Wipes	Ambion, USA
Rubber cement Fixogum	Marabu, Germany

II-1.6 Equipment

Table I-18. Equipment

Equipment	Supplier
5415D, 5417R and 5810R centrifuges	Eppendorf, Germany
ABI 3730 capillary sequencer	Applied Biosystems
ABI Prism 9700	Applied Biosystems
AF30-WS Self Contained Ice Machine	Scotsman
Agilent hybridization oven	Agilent, USA
Agilent microarray scanner	Agilent, USA
Analytical balance BP61, d=0.1mg	Sartorius, Germany
Array Boosters SB400 and SB850	Advalytix, Germany
AxioCam MRm CCD camera	Zeiss, Germany
Axioplan 2 imaging microscope	Zeiss, Germany
Branson Sonifier 450A	Danbury, USA
BTR10 roller	Ratek, Australia
CA/REV 6 Cleanbench laminar flow hood	Clean Air, Netherlands
Carousel, 48-position	Agilent, USA
Cell counter Casy 1	Schärfe Systems, Germany
Controlled Environment Incubator Shaker	New Brunswick Scientific, USA
Curix 60	Agfa, Belgium
Cycler Primus HT	MWG Biotech, Germany

DCS-S75 digital camera	Sony, Japan
E.A.S.Y. 440K Gel Documentation System	Herolab, Germany
Electronic pipettes and dispensers	Eppendorf, Germany
Electrophoresis Power Supply Consort 425	Fröbel Labortechnik, Germany
Eppendorf Concentrator 5301	Eppendorf, Germany
GenePix 4000B laser scanner	Axon Instruments, USA
Heating Plate Digsi Therm 3434	Harry Gestigkeit GmbH, Germany
Horizon 11.14 Horizontal Gel Electrophoresis System	Life Technologies
Horizontal shaker KL-2	Edmund Bühler, Germany
Hybridization Chamber - SureHyb	Agilent, USA
Hybridization Chamber Rotator Rack	Agilent, USA
IQAir GC MultiGas air filtering and ozone depleting	Incen, Switzerland
Leica DMIRE2 inverted microscope	Leica, Germany
LSM 510meta confocal microscope	Zeiss, Germany
Magnetic mixer IKAMAG Rec G	Janke und Kunkel, Germany
Magnetic mixer IKAMAG Reo	Janke und Kunkel, Germany
Milli-Q Advantage	Millipore, USA
NanoDrop® ND-1000 UV-Vis spectrophotometer	NanoDrop Technologies, USA
Nikon Eclipse TS100	Nikon, Japan
Precision balance BP2100 S, d=0.01g	Sartorius, Germany
Rotanta 46 R cooled swing-out centrifuge	Hettich, Germany
Rotina 48 R swing-out centrifuge	Hettich, Germany
Steri-cycle CO2 incubator 371	Thermo Scientific, USA
Thermomixer 5320	Eppendorf, Germany
Thermomixer 5436	Eppendorf, Germany
Thermomixer comfort	Eppendorf, Germany
Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell	Bio-Rad, USA

II-1.7 Bioinformatics tools and software

Table I-19. Bioinformatics tools and software

Software name	Provider
ABI Prism SDS	Applied Biosystems, USA
CGH Analytics 3.4.27	Agilent, USA
CGHPro 3.1.1	in house, MPIMG, Berlin
CodonCode aligner 1.6.0 beta5	CodoneCode, USA
Cytoscape v2.6.0	Cytoscape.org
eArray	Agilent, USA
EndNote 9.0.1	Thomson ISI Research soft
Feature Extraction	Agilent, USA
GenePix Pro5.0	Axon
GeneSpring GX7.3.1	Agilent, USA
GenomeCat	in house, MPIMG, Berlin
Isis imaging	MetaSystems, Germany
LSM Image Examiner 4.0.0.241	Zeiss, Germany
Ingenuity Pathway Analysis (IPA)	Ingenuity Systems, USA
Methyl Primer Express v.1.0	Applied Biosystems , USA

II-1.8 Online Resources and databases

In silico sequence analysis, primer design, conformation checking and other analyses were performed using databases, alignment tools and other software available online. They are listed in alphabetical order in Table I-20.

Table I-20. Online resources and databases

Resource / database name	URL
ArrayExpress	http://www.ebi.ac.uk/microarray-as/ae/
BaCelLo Protein Subcellular Localization Predictor	http://gpcr.biocomp.unibo.it/bacello/
BLAST at NCBI	http://www.ncbi.nlm.nih.gov/BLAST/
Ensembl	http://www.ensembl.org/
GenBank	http://www.ncbi.nlm.nih.gov/Genbank/
GeneCards	http://www.genecards.org/
GEO – Gene Expression Omnibus	http://www.ncbi.nlm.nih.gov/geo/
HPRD – Human Protein Reference database	http://www.hprd.org/
IntAct	http://www.ebi.ac.uk/intact/
InterPro	http://www.ebi.ac.uk/interpro/
IPI-International Protein Index	http://www.ebi.ac.uk/IPI/
KEGG Pathway Database	http://www.genome.jp/kegg/
Mfold	http://www.bioinfo.rpi.edu/applications/mfold/
Online Mendelian Inheritance in Man (OMIM)	http://www.ncbi.nlm.nih.gov/Omim
Pfam, [95]	http://www.sanger.ac.uk/Software/Pfam/
Primer3 (v 0.4.0)	http://frodo.wi.mit.edu/
Prosite	http://www.expasy.org/prosite/
PubMed	http://www.pubmed.org
Reactome	http://www.reactome.org/
Sanger Institute	http://www.sanger.ac.uk/
SMART - Simple Modular Architecture Research Tool	http://smart.embl-heidelberg.de/
UCSC Genome Browser	http://genome.ucsc.edu/
Unigene	http://www.ncbi.nlm.nih.gov/UniGene
UniProtKB	http://www.uniprot.org

II-2 Methods

II-2.1 General methods

2.1.a Laser Captured Microdissection

Tissue sections, stained with Harris hematoxylin, were laser-captured microdissected by an experienced pathologist at the Institute of Pathology (Graz, Austria).

2.1.b DNA isolation

For DNA isolation from cell lines, the cell pellets were resuspended in 20ml lysis buffer (0.4M TRIS-HCl pH8.0, 0.06M Na-EDTA, 0.15M NaCl), then 32µl RNase A (10µg/ml) was added into the lysate and incubated at 37°C for 1h. Afterwards, 5ml sodium perchlorate was added to tubes and mixed. Then 20ml ice-cold chloroform was added, again mixed overhead 10-15 times and centrifuged 15min at 3500xg. The upper phase was transferred to a fresh tube and combined with 1vol ice-cold absolute EtOH. DNA was spooled with a Shepherd's crook and washed in 70% EtOH. After air-drying DNA was detached from crook in H_2O by incubating the tubes at 65°C for 30min and finally resuspended for overnight incubation at room temperature.

DNA from laser-captured microdissected cells was extracted at the Institute of Pathology, Graz, Austria. Shortly, cells were digested in 10µl TE, pH9.0, and Proteinase K (20mg/ml) for 18h at 55°C.

2.1.c RNA isolation

RNA from cell cultures was isolated by the use of RNeasy® RNA extraction kit in conjunction with QIAshredder® column according manufacturer's protocol. In order to protect RNA from digestion with endogenous RNases 10µl of β-mercaptoethanol was added to 1ml RLT buffer before lysing the cells.

2.1.d **DNA Sonication**

Necessary amount of DNA was diluted to 200µl in water. DNA was sonicated in 1.5µl Eppendorf tubes for 4sec by Branson Sonifier 450A (Danbury, USA). The settings on sonifier were the following: Timer - Hold, Duty Cycle - Constant, Output Control - 4. After sonication samples were immediately chilled on ice and purified by Qiaquick® PCR purification kit (Qiagen, Germany) according to the protocol, recommended by manufacturer. Column purification step was done in order to remove DNA sequences shorter than 100bp and bigger than 10kb.

2.1.e **DNA Purification**

e-1 Phenol:Chloroform:Isoamyl Alcohol purification

Phenol:Chloroform:Isoamyl alcohol DNA purification was done according to standard protocol (Molecular Cloning, 3rd edition, Sambrook and Russel, ISBN:978-087969577-4).

e-2 Ethanol Precipitation

For ethanol precipitation of DNA, 2.5vol of ice-cold absolute ethanol and 0.1vol of sodium acetate (pH 5.2) were added to the samples. Samples were incubated at –20°C for overnight and then were centrifuged for 20 min at 16 000×g at 4°C. Afterwards, the pellet was washed with 200µl of 70% ethanol, centrifuged for 10 min at 16000×g at 4°C, air dried and resuspended in TE buffer (pH 7.5) or H_2O (Baxter, Germany).

2.1.f **Whole Genome Amplification (WGA)**

Whole genome amplification of genomic DNA was performed with the use of GenomePlex Whole Genome Amplification (WGA) Kit (Sigma, USA) according to the manufacturer's recommendations. The technique is based on random chemical fragmentation and conversion of genomic DNA into a library of DNA molecules flanked by universal priming sites. DNA fragments were amplified by standard PCR using universal oligonucleotide primers, included in the kit. For amplification I used 0.75µl of TITANIUM™ Taq DNA Polymerase (BD Clontech, Mountain View, USA). The cycling program is presented in Table I-21.

Table I-21. PCR program for WGA

Temperature	Time	Cycles
95°C	3 min	1x
94°C 65°C	15 sec 2 min	14x
4°C	forever	

2.1.g DNA Sequencing

Samples were directly sequenced in both directions using the ABI 377 DNA sequencer (Applied Biosystems, USA).

For target amplification 2ng (7µl) DNA per 100bp DNA length for PCR products, and 100ng (7µl) of plasmid DNA were mixed with 10µl H_2O, 1µl sequencing primer (10pmol) and 2µl BigDye Terminator Mix V3.1 (Applied Biosystems, USA). Thermal profiles for the amplification and labelling are indicated in Table I-22 and the primers used for amplification/sequencing are indicated in Table I-23.

Table I-22. Thermal cycle for sequencing reactions

Temperature	Time	Cycles
96°C	1 min	1x
96°C Primer specific (Table I-23)	10 sec 5 sec	25x
60°C	4 min	
4°C	forever	

Table I-23. List of sequencing primers

Name	Genomic region/Gene†	Tm‡	Forward/Reverse primers
Chr11-1	chr11:122806412-122806737	56.0°C 54.0°C	5'-ggggagttttaggagggt-3' 5'-attctccatcatattaacctc-3'
Chr1-1-2	chr1:9065634-9065971, SLC2A5	56.5°C 56.5°C	5'-ttgaggtaggagaatagttgga-3' 5'-atcccctctaaacctaaaaacc-3'
Chr1-2-3	chr1:9065752-9066193, SLC2A5	52.8°C 49.1°C	5'-taaaaataatcgtttaaagtttgtagt-3' 5'-acaaaaacaaaacccacc-3'
Chr19-2	chr19:7218417-7218711, INSR	54.0°C 54.0°C	5'-ggggtagttggttttgttttt-3' 5'-aaaaaacctattccccaacct-3'
Chr3-3	chr3:53832663-53833025, CHDH	56.5°C 54.7°C	5'-attatggtattgagggatgagg-3' 5'-aaccctaaaacaacaacaatcc-3'
Chr9-1	chr9:34390693-34391129, C9orf25	54.7°C 52.8°C	5'-gggtgaggttttttttaggtaat-3' 5'-tacctacaacaaatcaaccaaa-3'
SPON2-orf	NA	53.7°C 56.0°C	5'-gacgttgtatacgactcc-3' 5'-tgctgccagatcctcttc-3'

† According to Hg.17 (May 2004)

‡ For each primer pair the lowest Tm was used. For sequencing reaction see I-**Error! Main Document Only.**.

For DNA precipitation and purifications 1µl of 2%SDS was added to samples and incubated at 98°C for 10 second. Then 25µl of 100% EtOH was added to each reaction, tubes were mixed thoroughly by inverting and centrifuged at 4000xg 60min at 4°C. The supernatant was carefully discarded. 150µl of 70% EtOH was added on samples and the tubes were inverted several times without disturbing the pellets, centrifuged at 4000xg for additional 30 min at 4°C and the supernatant was carefully discarded. This washing step was repeated twice and then the pellets were air-dried.

Sequence data were assembled and analysed using CodonCode aligner 1.6.0 beta5 software (CodonCode, USA).

2.1.h **Cell culturing**

All mammalian cell lines were cultured under a 5.0% CO_2 atmosphere at 37°C. Adherent cells were seeded at a density of 5.2 – 35.0×10^3 cells/cm^2. Suspensive cells were maintained at a density of 1×10^6 cells/ml. Table I-13 lists the culturing conditions for those cell lines relevant to this study.

For seeding, the cell were trypsinised and counted with CASY1 (Schärfe System, Germany). Trypsinization was done at 37°C for 5 min using Trypsin-EDTA (Lonza, Switzerland).

II-2.2 Sample preparation for microarray hybridization

2.2.a **DNA Labeling**

a-1 Random Primed Labeling

DNA samples were sonicated (see section 2.1.d) to generate fragments from 200 to 1000bp in size. Test and reference DNA were labeled by random priming employing the BioPrime DNA Labeling System (Invitrogen, California) with Cy3-dUTP (for test DNA) and Cy5-dUTP (for reference DNA) (Amersham Biosciences, USA) according Invitrogen's protocol.

a-2 Labeling by Agilent's protocol

Genomic DNA labeling for hybridization on Agilent 244K oligo array CGH was followed by Agilent's recommendation (Manual Part Number G4410-90010, version 5.0, Agilent, USA).

2.2.b **Methylated DNA Immunoprecipitation (MeDIP)**

Five microgram of genomic DNA was sonicated for 4sec in order to get sequences with average length of 500bp (see section 2.1.d). Two microgram of sonicated DNA was denatured for 10min at 95^{o}C and quickly cooled on ice. The volume of DNA was adjusted to 245µl by ice-cold IP buffer (10mM sodium phosphate, 140mM NaCl, 0.05% Triton X-100, pH7.0), and then mixed with 5µl (1µg/µl) of monoclonal antibody against 5-methylcytidine (Eurogentec, Belgium). Tubes were incubated for 3h at +4^{o}C with gentle rotation. U-bottom shaped 2ml Safe-Lock Eppendorf tubes (Eppendorf, Germany) were used in order to optimize the mixing effect and to avoid clumping at the bottom of the tubes. Then, 15µl of washed Dynabeads M-280 Sheep anti-Mouse IgG (Invitrogen, USA) were added on the reaction and incubated for overnight at +4^{o}C with gentle rotation. Dynabeads washing was done according manufacturer's recommendation. Afterwards, tubes were placed on a Dynal MPC-S magnetic rack (20µl-2ml) (Invitrogen, USA) for 3min, and the supernatant was removed. 700µl of fresh IP buffer was mixed with magnetic beads, carefully re-pipetted (to avoid making foam) and put back on the magnetic rack. This washing procedure was repeated 3 times in order to remove not-methylated DNA. Finally, magnetic beads were dissolved in 15µl of water (Baxter, Germany). After performing immunoprecipitation, DNA bound to magnetic beads (5µl) underwent whole genomic amplification by GenomePlex WGA kit (Sigma, USA) according to the manufacturer's protocol (see section 2.1.f), excepting DNA fragmentation step. After amplification DNA was checked on an agarose gel, purified by Qiaquick® PCR purification kit (Qiagen, Germany), according manufacturer's recommendations, and the concentration was measured by a NanoDrop® ND-1000 spectrophotometer (NanoDrop Tecnologies, USA). Afterwards, DNA samples were labeled by BioPrime DNA Labeling System (Invitrogen, California) with Cy3-dUTP and Cy5-dUTP (Amersham Biosciences, USA) following Invitrogen's protocol (see above).

2.2.c **Methylated CpG island amplification (MCA)**

Five microgram of genomic DNA was digested first with 5µl (100U) of Sma I (NEB, UK) with 5µl NEBuffer 4 (NEB, UK) at 25°C for overnight. Then 2µl (20U) of Xma I (NEB, UK) and 0.5µl of BSA (NEB, UK) were added to the reaction and incubated at 37°C for 6 h. After incubation the DNA was purified by Phenol:Chloroform:Isoamyl alcohol (see section II-2.1e-1). As the volume of the restriction digestion reaction is too small for PCI purification, the reaction volume was adjusted to 200µl by TE buffer. After ethanol precipitation (see section II-2.1e-2) DNA was diluted in 20µl of TE buffer, pH7.5 (Upstate, USA) and was labeled by BioPrime DNA Labeling System (Invitrogen, Carlsbad, California) with Cy3-dUTP and Cy5-dUTP (Amersham Biosciences, USA) according Invitrogen's protocol.

2.2.d **cRNA preparation for Gene Expression microarrays**

The whole genomic RNA was extracted from cell lines as described in section 2.1.c. cRNA preparation was done by Agilent Low RNA Input Linear Amplification Kit PLUS protocol, Version 4.0, January 2006 (Agilent, USA). The method uses T7 RNA polymerase, which simultaneously amplifies target material and incorporates Cy3- CTP. One Color Spike-In kit (Agilent, USA) was used in order to provide positive controls for monitoring microarray workflow from sample amplification and labeling to microarray processing. It consists of a set of 10 *in vitro* synthesized, polyadenylated transcripts derived from the Adenovirus E1A gene, optimized to anneal to complementary probes on the microarray with minimal self-hybridization or cross-hybridization. The amplification rate of Spike-In probes is independent from the quality of the starting RNA sample.

The workflow of cRNA preparation is the following (see also Figure I-12):

- Total template RNA is combined with One Color Spike-In kit (Agilent, USA), then the cDNA is synthesized using MMLV-RT and oligo dT-Promoter primers;
- cRNA is synthesized from cDNA and labeled by Cy3-CTP using T7 RNA polymerase;
- cRNA is purified using Qiagen's RNeasy mini spin columns (Qiagen, Germany);
- the quality (specific activity) and yield of synthesized cRNA is measured before proceeding microarray hybridization.

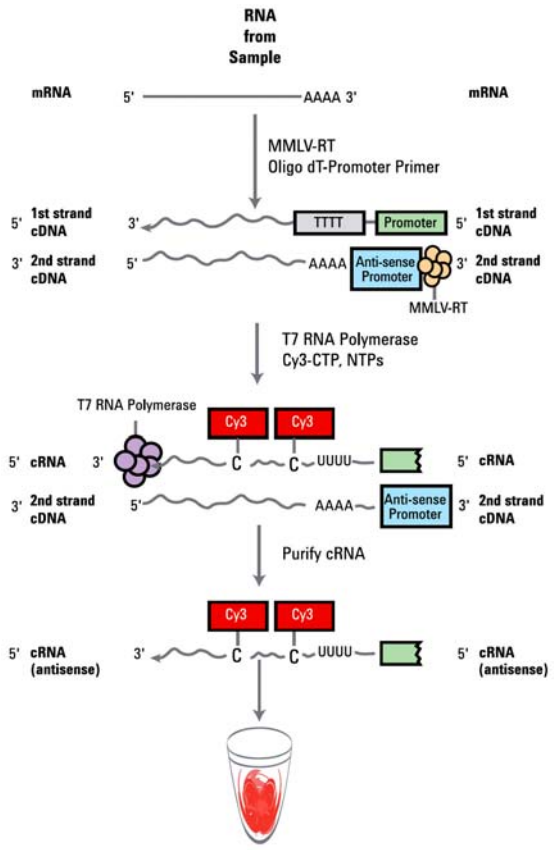


Figure I-12. cRNA preparation workflow for gene expression arrays (from Agilent's manual)

II-2.3 Microarray platforms

2.3.a In-house BAC array CGH

a-1 Array CGH platform

For the analysis of DNA copy number changes in-house 36K sub-megabase "tiling-path" BAC arrays were used, containing the 1Mb set (Welcome Trust Sanger Center, courtesy of Nigel Carter), the human 32K BAC Re-Array set (http://bacpac.chori.org/pHumanMinSet.htm, kindly provided by Pieter de Jong) ([96-98]) and a set of 390 subtelomeric clones (assembled by members of the COST B19 initiative: Molecular Cytogenetics of Solid Tumors).

a-2 Linker Adapter PCR of BAC DNA

Linker adapter PCR includes three steps: restriction enzyme digestion of BAC DNA, ligation of adaptor primers and ligation mediated PCR. First, DNA was digested with MseI and BfaI restriction enzymes. The reaction of restriction was prepared as follows:

Table I-24. Restriction reaction for Linker Adapter PCR

Restriction reaction	Vol.
10x NEBuffer 1	0.750 µl
100x BSA	0.075 µl
Mse I (50U/µl)	0.012 µl
Bfa I (5U/µl)	0.150 µl
H_2O	1.513 µl
Exonuclease digested DNA	5.000 µl
Total volume	7.5 µl

The reaction was placed in a PCR cycler for 3h at 37°C and then inactivated for 20min at 80°C. The length of restriction fragments was checked on 1% agarose gel (in the range of 100-1500 bp). Then the following reaction was set up (Table I-25):

Table I-25. Ligation Reaction for Linker Adapter PCR

Ligation reaction	Vol.
Primer-21 (100pmol/µl)	0.5 µl
Primer-12 (100pmol/µl)	0.5 µl
10x Ligase Buffer (Roche)	0.8 µl
H_2O (Baxter, Germany)	5.2 µl
Digested BAC DNA	1.0 µl
Total volume	8.0 µl

The reaction was carried out in a PCR cycler with heated lid. The program on cycler is showed in Table I-26.

Table I-26. Thermal program of ligation reaction

Step	Temperature	Time
Primer denaturation	65°C	2 min
Primer annealing	65°C→15°C with 1°C/min ramp	50 min

At 15°C 0.2µl of Ligase Buffer (Roche, Germany), 0.2µl of 5U/µl T4 Ligase (Roche, Germany) and 1.6µl of H_2O were added. Then the reaction was incubated at 15°C for 18-20h. After incubation the DNA went to the Ligation Mediated PCR:

Table I-27. Ligation Mediated PCR reaction

Ligation mediated PCR reaction	Vol.
Ligation product	1.0 µl
10x PE Buffer	5.0 µl
dNTP mix (1mM each)	10.0 µl
Primer-21 (100pmol/µl)	0.5 µl
H_2O	33.5 µl
Total volume	50.0 µl

30µl of Mineral oil (Sigma, USA) was added on the samples to prevent sample evaporation and placed in the thermal cycler at 68°C for 4min in order to denature Primer-12 from the BAC DNA. Then to each sample 1µl of DNA Taq Polymerase (10U/µl) was added and incubated in the cycler for additional 4min (extension step). Then the following program was loaded in the machine to perform DNA amplification:

Table I-28. Ligation Mediated PCR program

Temperature	Time	Cycles
95°C	3 min	1x
95°C 59°C	40 sec 30 sec (+2sec/sycle)	35x
72°C 4°C	7 min forever	1x

Several PCR products were randomly chosen and checked on 1% agarose gel. Size of PCR products should be in the range of 70-1500 bp, with the highest concentration at 200-800 bp.

a-3 Re-amplification of ligation amplified DNA

1.0µl of ligation mediated PCR product was used as a template for re-amplification in order to generate enough DNA for spotting. The amplification reaction and PCR program were the same as before (see Ligation Mediated PCR. Note: fill-in reaction is omitted).

a-4 Preparation of spotting solution

The re-amplified PCR products were precipitated overnight by adding 150µl ice-cold absolute ethanol and sodium acetate (pH 5.2) at -20°C. After pelleting by centrifugation it was dissolved in 3xSSC/1.5M Betaine.

a-5 Array production

The products were spotted onto epoxy coated glass slides (Nunc, Germany) using an in-house modified Qarray array printer (originally from Genetix, U.K.) and Pointech (Gibbon, MN) Tungston PTL 2500 slit pins). The microspotting technique was applied, where the spotting robot disposed the PCR products directly on the slides.

Epoxy slides were chosen due to several reasons. First, epoxy slides are especially suitable for covalent immobilization of oligonucleotides (10 to 80 bases), PCR products, as well as cDNA molecules. Second, additional amino-modifications of the nucleic acids are

not required. Third, their hydrophobic surface allows to get small spot diameters (100-130µm, depending on the type of pins and spotting buffer) to create high-density arrays. Finally their surface chemistry is very stable and remains active even during long spotting runs.

a-6 Slide Processing

Slides were blocked at 42°C for 1h in the 60ml of blocking solution (0,1% SDS, 4xSSC, 0,5% BSA), containing 200µl of Herring Sperm DNA (Invitrogen, USA). Afterwards, the slides were immediately rinsed 3 times with Millipore water and air-dried by centrifugation for 5 min at 150×g.

a-7 Hybridization of labeled genomic DNA

Principle of array CGH hybridization is shown on Figure I-13.

Genomic test and reference DNA labelled by Cy3 and Cy5, respectively were pooled together with 500µg (1µg/µl) of Human Cot-1 DNA (Roche, Germany), SpeedVac-ed to 200µl by Concentrator 5301 (Eppendorf, Germany). Then it was subjected to ethanol precipitation (see section II-2.1e-2). The precipitated DNA was dissolved in 6.8µl 10% SDS, 3.4µl Yeast tRNA (100µg/µl, Invitrogen), and 24µl hybridization mix (70% formamide, 2.8x SSC, 8% dextran sulphate), and denatured at 70°C for 15 min. After denaturation, the hybridization mix was incubated at 42°C for 2h to allow the Human Cot-1 DNA to anneal to the repetitive sequences. The labelled probes were then placed on the microarray under a coverslip (24x60mm, Menzel-Glaser, Germany). The arrays were incubated for 24 hours under humidified conditions using the Array Boosters SB400 and SB850 (Advalytix, Germany), respectively. After hybridization arrays were washed with 50% formamide 2xSSC, 0.1% SDS for 15min at 42°C, and followed by a 10 minute wash in PN buffer (0.2 M sodium phosphate with 0.001% NP40) at room temperature.

Arrays were incubated for 30sec in 1xPBS and then for 2-3sec in Millipore water. Finally the arrays were dried by centrifugation at 150×g for 5min.

a-8 Scanning

After hybridization the arrays were scanned at λ=532 nm for Cy3 and λ=635 nm for Cy5, using GenePix 4000B laser scanner (Axon Instruments, USA) and Agilent Microarray Scanner System (Agilent, USA), respectively, in order to read out the fluorescence signal

intensities in each channel. The resulting 16 bit TIFF images were analyzed by the use of GenePix Pro 5.0 software (Axon Instruments, USA).

Microarray hybridization, washing and scanning were done in an ozone depleted room. Ozone depletion and air filtering was done using multiple-filtering machine IQAir GC MultiGas (INCEN, Switzerland).

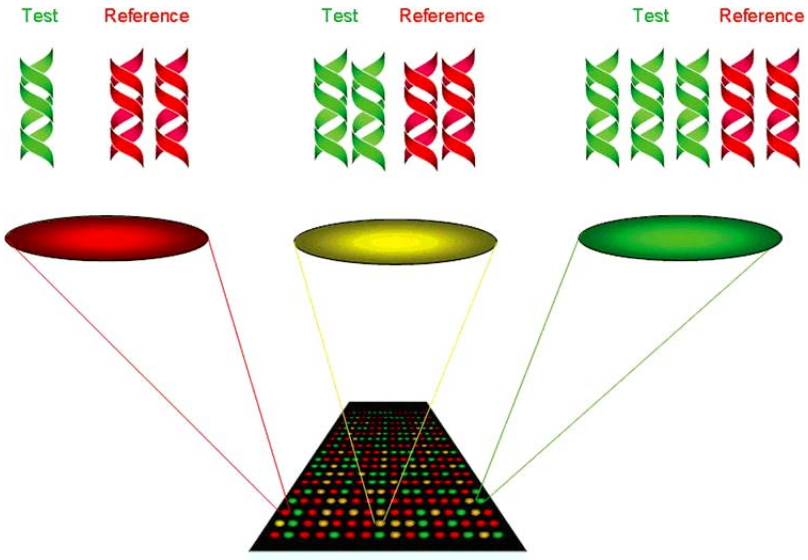


Figure I-13. Principle of array CGH hybridization.

Test and reference DNA are labeled with a green and red fluorochrome, respectively, and cohybridized on a DNA array. In case of a loss in the test DNA, the corresponding DNA spot on the array will appear red. A gain in the test DNA will result in a green spot. Note that with this technique, only changes in the test DNA, relative to the reference DNA, can be monitored.

2.3.b **Agilent microarrays**

The following Agilent arrays were used in this study:

- Human Genome CGH Microarray 244A (1x244K);
- Whole Human Genome Oligo Microarray (Gene Expression, 4x44K);
- Human CpG Island Microarray (1x244K);
- Custom High-Definition CGH Microarrays (4x44K).

b-1 Agilent Human Genome CGH Microarray 244A (1x244K)

Agilent's 60-mer array-based Comparative Genomic Hybridization (aCGH) application uses a "two-colour" approach to measure DNA copy number changes in an experimental sample relative to a reference sample. Microarray processing was done

following manufacturer's manual (Manual Part Number G4410-90010, version 5.0). Description of Agilent's Human Genome CGH Microarray 244A platform is presented in Table I-29.

Table I-29. Agilent Human Genome CGH Microarray 244A platform

Feature	Specification
Format	1x244K
Design ID	014693
Features/Microarray	243 504
Distinct Biological function	236 381
Replicated Biological Features (x3)	1 000
Internal QC features	5 045
Slide format	25mm x 75mm
Probe length	60-mer
Feature size	~65 µm
Sequence source	UCSC hg17 (NCBI Build 35, May 2004)

b-2 Agilent Custom High-Definition CGH Microarrays (4x44K)

Oligonucleotide probes were designed *in silico* by (Agilent, USA) eArray online software. Candidate probes representing unique genomic sequences were automatically selected by eArray program, scored, and filtered using bioinformatics prediction criteria for probe sensitivity, specificity, and responsiveness under appropriate conditions. Selection criteria were based on empirical testing of known genomic aberrations in comparative model systems, such as XX/XY hybridizations. For human, the resulting candidate probe collection consists of over 8.4 million predesigned aCGH probes with a median probe spacing of 192 bp. All CGH probes span coding and non-coding regions for comprehensive genome coverage. Microarray processing was done following manufacturer's manual (Manual Part Number G4410-90010, version 5.0). Agilent Custom High-Definition CGH Microarray platform is described in Table I-30.

Table I-30. Agilent Custom High-Definition CGH Microarray platform

Feature	Specification
Format	4x44K
Microarrays per slide	4
Slide format	25mm x 75mm
Probe length	60-mer
Feature size	~65 µm
Total features	45 220
Internal QC features	2 118
Sequence source	UCSC hg17 (NCBI Build 35, May 2004)

b-3 Agilent Human CpG Island Microarray (1x244K)

The Agilent's CpG Island microarrays are designed to analyze the genome wide measurement of methylation changes. Here the DNA samples obtained from MeDIP and methylated CpG island amplification (MCA) experiments (see sections 2.2.b and 2.2.c, respectively) were hybridized according manufacturer's protocol. Description of the Agilent's CpG Island microarrays is presented in Table I-31.

Table I-31. Agilent Human CpG Island Microarray platform

Feature	Specification
Format	1x244K
Design ID	014791
Features/Microarray	243 504
# CpG Islands	27 800 (covering 21Mb)
# CpG Probes	~195 000
Internal QC features	5 045
Slide format	25mm x 75mm
Probe length	60-mer
Feature size	~65 µm
Sequence source	UCSC hg18/NCBI build 36.1 (March 2006)

b-4 Agilent Whole Human Genome Oligo Microarray (Gene Expression, 4x44K)

Agilent's 4x44K Whole Human Genome Oligo Microarrays span conserved exons across the transcripts of the targeted full-length genes. These probes represent the human genome as we know it today, leveraging well-characterized, full-length and partial human genes from a number of major public sources. The sequence and annotation information are available through publicly-available databases such as RefSeq ID, HUGO/LocusLink gene symbol, GenBank ID, Ensembl and GoldenPath ID. The Whole Human Genome Oligo Microarray platform id described in Table I-32.

Table I-32. Agilent Whole Human Genome Oligo Microarray platform

Feature	Specification
Format	4x44K
Design ID	014850
Slide format	25mm x 75mm
Probe length	60-mer
Feature size	~65 µm
Dynamic Range	Over 3 orders of magnitude
Features/Microarrays	45 015
# Biological Features	43 376
# Unique Probes	41 000
Replicates of Biological Probes	264x10
# (+) Controls	1 468

Replicate # E1A Probes	10x32
# (-) Controls	153
Probe Orientation	Sense
Reproducibility	Median SD <0.013
Sequence Source	Goldenpath Ensembl Unigene Human Genome (Build 33)

II-2.4 Microarray Interpretation

2.4.a BAC aCGH Microarray Normalization and Interpretation

BAC arrays were normalized and visualized using CGHPRO 3.1.1 [99], which provides a variety of graphical data representation tools to visualize the data before and after normalization. The normalization is aimed to remove the systematic bias of fluorescence intensities, originating from different labelling efficiencies, scanning parameters, spatial effects, impact of ozone etc. All BAC arrays used in our study were normalized by subgrid LOWESS (locally weighted linear regression).

In CGHPRO all replicate spots are automatically identified by their common ID and their ratios are averaged.

Aberrations involving three or more neighbouring BAC clones were considered relevant, unless they coincided with a known polymorphism as listed in the Database of Genomic Variants (http://projects.tcag.ca/variation/). For estimation of segmental duplications (low copy repeats), BAC clones were classified and coloured into seven categories [99].

All the genomic coordinates provided in this thesis are reffering to Hg.17, May 2004 (NCBI Build 35).

2.4.b Agilent Microarray Normalization and Interpretation

Data, obtained with Agilent microarrays, were normalized by Feature Extraction software (Agilent, USA) using the default settings specified for each platform. For gene expression analysis data transformation and intra- and inter-experimental normalisation were done using GeneSpring GX7.3.1 software.

Human Genome CGH Microarrays 244A (1x244K) were visualized using either CGH Analytics 3.4.27, or CGHPRO 3.1.1. Agilent's 4x44K Whole Human Genome Oligo Microarrays were visualized using GeneSpring GX7.3.1 software.

II-2.5 Microarray Validation

2.5.a Validation of Methylation

a-1 MeDIP validation

In order to check the specificity of the monoclonal antibody used for MeDIP I designed the following experiment.

First I performed PCR amplification of a DNA, using two primer pairs:

- Exon2 – product size 232bp, contains 12 CpGs (chr11:133762596-133762827, hg.17),
- Exon3 – product size 384bp, contains 49 CpGs (chr11:133758971-133759354, hg.17).

After Qiaquick® column purification (Qiagen, Germany), half of each PCR products was methylated *in vitro* by M.Sss I (NEB, UK). The methylation state of both, methylated and not-methylated PCR products was checked by digestion with BstU I (NEB, UK).

Then, they were combined together so, that one tube contained methylated PCR product, containing 12 CpG-s and not-methylated one, containing 49 CpGs, and *vice versa*. These two tubes were treated as indicated in section 2.2.b and afterwards the Dynabeads were digested with Proteinase K (50µg/µl final concentration) for 3.5 hours at 50°C with 1min interval mixing for 10sec. After consequent CIA purification (see section II-2.1e-1) and ethanol precipitation (see section II-2.1e-2) they were visualized on 2% agarose gel.

a-2 Bisulfite Sequencing PCR (BSP)

Bisulfite conversion was done using EZ DNA Methylation-Gold Kit™ (Zymo Research, USA) according manufacturer's recommendation. The thermal program for conversion is indicated in Table I-33.

Table I-33. Thermal program of bisulfite conversion

Temperature	Time
98°C	10 min
64°C	2.5 hours
4°C	forever

The converted DNA was amplified by touch-down PCR using of Bio-X-Act Long Mix (Bioline, Germany). The PCR reaction was composed as follows: 12.5µl of Bio-X-Act Mix, 10.0µl water (Baxter, Germany), 0.75µl forward and 0.75µl reverse primers (100pmol/µl each) and 1.0µl bisulfite converted DNA. The PCR thermal program and primers with their genomic coordinates are listed in the Table I-34 and Table I-35, respectively.

Table I-34. PCR thermal program for BSP

Temperature	Time	Cycles
94°C	3 min	1x
94°C 58°C 72°C	30 sec 30 sec 30 sec	20x
94°C 48°C 72°C	30 sec 30 sec 30 sec	30x
72°C	7 min	1x
4°C	forever	1x

Table I-35. List of BSP primers and their amplification regions

Name	Genomic Position , hg.17	Forward primer / Reverse primer
Chr11-1	chr11:122806412-122806737	5'-ggggagttttaggagggt-3' 5'-attctccatcatattaacctc-3'
Chr19-2	chr19:7218417-7218711	5'-ggggtagttggttttgttttt-3' 5'-aaaaaacctattccccaacct-3'
Chr1-1-2	chr1:9065634-9065971	5'-ttgaggtaggagaatagttgga-3' 5'-atcccctctaaacctaaaaacc-3'
Chr1-2-3	chr1:9065752-9066193	5'-taaaaataatcgtttaaagtttgtagt-3' 5'-acaaaaacaaaacccacc-3'
Chr3-3	chr3:53832663-53833025	5'-attatggtattgagggatgagg-3' 5'-aaccctaaaacaacaacaatcc-3'
Chr9-1	chr9:34390693-34391129	5'-gggtgaggttttttaggtaat-3' 5'-tacctacaacaaatcaaccaaa-3'

PCR products were checked on a 2% agarose gel, cut from the gel and purified by QIAquick Gel Extraction kit (Qiagen, Germany) according manufacturer's protocol.

The sequencing of bisulfite converted DNA was done as described in Section 2.1.g. Obtained sequences were aligned using CodonCode aligner 1.6.0 beta5 (CodoneCode, USA) and BLASTN suite-2sequences algorithm [100].

2.5.b **Validation of Gene Expression data**

The validation of gene expression data was done by qRT-PCR. For the quantification of *MMP1* and *SPON2* expression levels I used the real time PCR system ABI 7900HT (Applied Biosystems, USA).

b-1 Reverse Transcription

Reverse transcription was performed by the use of TaqMan® Reverse Transcription Reagents (Applied Biosystems, USA) kit as follows:

Table I-36. Reverse Transcription reaction

Reverse Transcription reaction	Vol.
10x TaqMan RT Buffer	10.0 µl
$MgCl_2$ (25µM)	22.0 µl
dNTP Mixture	20.0 µl
Random Hexamer (50µM)	5.0 µl
RNase Inhibitor (20U/µl)	2.0 µl
Reverse Transcriptase (50U/µl)	2.5 µl
Genomic RNA (2µg)	X µl
RNAse free water	38.5-X µl
Total volume	100.0 µl

Table I-37. Thermal program for Reverse Transcription

Step	Temperature	Time
Incubation	25°C	10 min
Reverse Transcription	48°C	30 min
Inactivation	95°C	5 min

The incubation step is necessary to maximize primer-RNA template binding. After reverse transcription cDNA samples were stored at -20°C.

b-2 qRT-PCR

cDNA amplification was done in MicroAmp Optical 96-well reaction plates and MicroAmp Optical 384-well reaction plates (Applied Biosystems, USA) using SYBR Green PCR Master Mix (Applied Biosystems, USA).

Table I-38. qRT-PCR reaction

qRT-PCR reaction	Vol./96-well format	Vol./384-well format
SYBR Green PCR Master Mix	15.0 µl	10.0 µl
Primer Mix	5.0 µl	3.0 µl
cDNA	10.0 µl	7.0 µl
Total volume	30.0 µl	20.0 µl

Table I-39. Primer mix for qRT-PCR

Primer Mix	Vol.
For.primer (100pmol/µl)	5.0 µl
Rev.primer (100pmol/µl)	5.0 µl
H_2O	322.5 µl
Total volume	332.5 µl

Primers for *MMP1*, *SPON2* and *HPRT* were designed to yield PCR products from 100-150bp length (for primer sequences see Table I-7). Primer design was done by using Primer3 software [101]. Expected PCR products were addressed to Mfold software [102], which analyzed *in silico* the ability of expected PCR products on self-looping. Formation of loops within PCR product yields non-uniform amplification rate and distorts the results. Specificity of primers was checked by standard PCR and visualized on 2% agarose gel. qRT-PCR was done using ABI Prism 9700 (Applied Biosystems, USA). The thermal program is represented in Table I-40.

All samples and internal control were tested in three replicas. All primers were simultaneously tested in cDNA dilution series. Obtained data were normalized to internal control gene (*HPRT*) using SDS 2.1 software. Gene expression levels were determined by absolute and relative quantification.

Table I-40. Thermal program for qRT-PCR

Temperature	Time	Cycles
50°C	2 min	1x
95°C	10 min	1x
95°C 60°C	15 sec 60 sec	40x
95°C 60°C 95°C	15 sec 15 sec 15 sec	1x (dissociation stage)

I-1.1 siRNA knockdown

For siRNA knockdown of *MMP1* and *SPON2* I used Qiagen's predesigned siRNA (the sequences are listed in Table I-8).

Before transfection $1x10^5$ cells/well (IMR-90, passage 12) were seeded in 6-well plates (TPP, Switzerland) in 2ml medium. Transfection was done only after reaching the confluence to 40-50% (24 hours).

Table I-41. Composition of solutions A and B

Solution A	Solution B
5 µl Oligofectamin + 45 µl OptiMEM Incubate 5 min at RT	5 µl (20µM) siRNA + 145 µl OptiMEM Incubate 5 min at RT

After incubation 50 µl of solution A was combined with 150 µl of solution B, mixed well and incubated further for 20 min at RT. Then the medium was removed from cell culturing plates directly before siRNA administration and 800 µl of OptiMEM was added into each well. The transfection mix (Solution A+B, 200 µl) was slowly added into the wells, and the culturing flasks were incubated in Steri-cycle CO_2 incubator 371 (37°C, 5% CO_2 v/v).

II-2.7 SDS-PAGE

5 ml of separating gel and 1 ml of stacking gel were prepared per SDS-PAGE gel as showed in Table I-42.

Table I-42. Preparation of separating and stacking gels.

Components	Separating gel			Stacking gel
	10%	12%	15%	
ddH2O	1.9ml	1.6ml	1.2ml	0.68ml
Rotiphorese gel 30	1.7ml	2.0ml	2.5ml	0.17ml
Separating or stacking gel buffer	1.3ml (separating buffer)			0.13ml (stacking buffer)
Ammonium persulfate	0.05ml			0.01ml
TEMED	0.002ml			0.001ml

Separating gel was poured between short plate and spacer plate using the PROTEAN 3 system, covered with isopropanol. The gel was polymerized for 45 min at RT. Next, the isopropanol was removed and the gel was washed with ddH_2O. After the stacking gel was applied onto separating gel, the comb was set between two glass plates. After 30min the gel was ready for SDS-PAGE

II-2.8 Western blot (WB)

Protein samples were mixed with 2x Magic mix and denatured for 15 min at 95°C before loading. The separation of the samples was performed in 1x Laemmli buffer at 200V for 50-70 min using a Mini-PROTEAN 3 electrophoresis system (Bio-Rad, USA). After running, gels were equilibrated for 15 min in 1x blotting buffer.

To determine protein size the Kaleidoscope Precision Protein Standards (Bio-Rad, USA) were run in parallel with the samples.

The blotting PVDF membrane was equilibrated for 3 sec in methanol, 2 min in ddH_2O and 15 min in 1x blotting buffer. Then proteins were blotted onto the membrane by means of a Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell (Bio-Rad, USA) at 15 V for 30 min. To saturate the unbound regions on the blot, it was incubated in PBST/5% milk powder for 30 min at RT or overnight at 4°C. Then, it was incubated with the primary antibody diluted in PBST/1%BSA for 1 hour at RT or overnight at 4°C followed by three wash steps of 5 min each in PBST. The corresponding HRP conjugated secondary antibody was diluted in PBST and, subsequently, added to the blot for 30 min, followed by three washing steps as before. Protein signals were detected by incubating the blots with the Western Lightning Chemiluminescence Reagent Plus (PerkinElmer, USA) for 1 min and subsequent exposure to Fuji Medical X-Ray films in a hypercassette. Blots were developed in a Curix 60 automatic film processor (Agfa, Belgium).

II-2.9 Coomassie and silver staining

For Coomassie staining the gels were washed 3 times with Millipore water for 5min and stained with "Imperial protein stain" (Pierce, USA) solution for 48h on a shaking platform at 4^oC. Then the gels were destained by washing in Millipore water for 1 min.

Silver staining was done using the SilverQuest Silver staining kit (Invitrogen, USA) according manufacturer's protocol.

II-2.10 Immunofluorescence (IF)

$1x10^5$ cells/well were seeded on coverslips in 6-well plates (TPP, Switzerland) with 2ml medium. After reaching to sufficient confluence, the medium was removed, cells were washed with 1.2xPEM buffer and fixed by incubation in 3.7% Paraformaldehyde (diluted with 1.2xPEM buffer) for 10min at RT. After fixation, coverslips were shortly washed in 1xPBS and permeabilized at RT in 0,2% Triton X-100 (diluted in PBS) for 10 min. After washing 4 times with 1xPBS, the coverslips were blocked by 1% BSA (in 1xPBS) for 30 min at RT. After blocking, the primary antibody was diluted to appropriate concentration by 1% BSA, applied on coverslips, covered with Parafilm and incubated for 1h at RT. After incubation the directly labeled antibodies were ready for washing and DAPI staining, whereas unlabeled ones were incubated with fluorescently labeled secondary antibody. Shortly, the Parafilm was removed and coverslips were washed 4 times in 1xPBS at RT.

Then the secondary antibody was diluted in 1% BSA to the appropriate concentration, applied on coverslips, covered by Parafilm and incubated for 60 min. Then the coverslips were again washed 4 times in 1xPBS. Finally, 15 µl of 0.5µg/ml DAPI in mounting medium was applied on coverslips, inverted on clean slides, slightly pressed by Whatman paper in order to suck abundant DAPI/mounting medium.

Cells were visualized by epi-fluorescent Axioplan 2 (Zeiss, Germany) or LSM 510meta confocal (Zeiss, Germany) microscopes. The imaging on LSM 510meta confocal microscope was done with beam configuration settings indicated in Table I-43. Imaging was done using MultiTrack imaging with sequential scanning, which dramatically reduces the crosstalk between channels.

Table I-43. LSM 510meta beam configuration settings

Wavelength, color	Laser	Dichroic beam splitter (HFT)	Emission splitter (NFT)	Emission filters
488nm green	Argon 30mW	HFT405/488	NFT490	BP505-530
543nm red	Helium-Neon 1mW	HFT488/543	NFT545	LP610
405nm blue	Diode 50mW	HFT405/488	mirror	BP420-480

II-2.11 SPON2 overexpression

SPON2 overexpression vector (pCMV6Entry, Figure I-14) was obtained from OriGene (USA). SPON2 was tagged at C-terminus with Myc and DDK (FLAG™). DDK is an octapeptide, which can be specifically detected by anti-FLAG antibody.

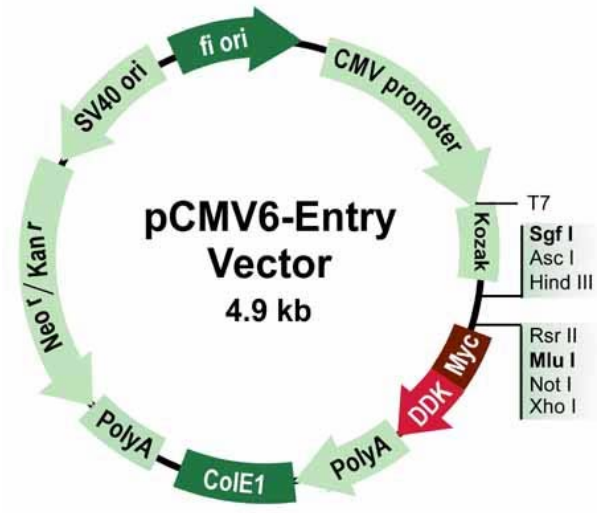


Figure I-14. pCMV6-Entry Vector (Myc-DDK Tagged)

2.11.a **Vector transformation into E.coli and validation**

pCMV6Entry vector was obtained as 10µg dried plasmid on a filter paper. It was diluted in 50µl TE buffer and transformed into *E.coli* TOP1D strain. Shortly, 10µl of plasmid DNA was added on 30µl of chemically competent bacteria, chilled on ice for 30min, moved in 42^{o}C heat block for 1.5 min (heat shock) and then shortly moved on ice. Afterwards, transformed bacterial culture was incubated in 200µl of LB medium (w/o antibiotic) by shaking for 30min at 37^{o}C. After short incubation, the cells were plated on LB+Kan for overnight at 37^{o}C. Next day, a single colony was picked and cultured in 200ml liquid LB+Kan medium for overnight at 37^{o}C. Afterwards, 5ml of bacterial culture underwent plasmid extraction according standard protocols and was subjected to sequencing (see section 2.1.g). For the sequencing SPON2-orf primers were used (see Table 1.2.b). Simultaneously, a control digestion of the vector was done by the use of XhoI and AscI in order to excise *SPON2* cDNA from the vector and visualized on a 1% agarose gel.

2.11.b **Vector extraction and purification**

After confirmation of the SPON2-Myc-DDK ORF, plasmid DNA the rest of bacterial culture was extracted using NoEndo Endotoxin-Free JETstar Plasmid Kit (Genomed, Germany) according manufacturer's recommendations.

2.11.c **Transfection in to mammal cells**

Before transfection 1x10^{5} cells/well were seeded in 6-well plates (TPP, Switzerland) in 2ml medium. Transfection was done only after reaching the confluence to 40-50% (24 hours). Shortly, 100µl of OptiMEM was combined with 5µl FuGENE HD (Roche, Germany) and 2µg of plasmid DNA, mixed well and incubated 15 min at RT. Directly before adding the transfection mixture to the cells, the medium was removed and 2ml of medium w/o antibiotic was added into each well. The transfection mix was slowly added into the wells and the culturing flasks were incubated in Steri-cycle CO_2 incubator 371 (37°C, 5% CO_2 v/v).

II-2.12 Co-Immunoprecipitation (co-IP)

SPON2 transfected cells (16 plates) were washed with PBS, scratched and pooled together. After centrifugation (3000×g, 10min) the pellet was diluted in 1ml co-IP buffer (50mMTris (pH7.5), 100mM NaCl, 5mM $MgCl_2$, 1mM DTT, complete protease inhibitor (Roche, Germany), 0.5% NP40). The first run of co-immunoprecipitation, designed for mass spectrometry of cut bands, was done in co-IP buffer without NP40. The runs for mass spectrometry analysis of whole lanes and of whole co-immunoprecipitates were done in the presence of 0.5% NP40 in IP Buffer. Then the cells were sonicated for 5 cycles per 10sec each (Timer - Hold, Duty Cycle - Constant, Output Control – Max). The samples were cleaned by centrifugation at 14000×g at 4°C for 10min and the supernatants were transferred in new tubes (this step was repeated 3 times). Afterwards 100µl of Mouse IgG-Agarose (Sigma, USA) was added to the supernatants and incubated for one hour at 4°C with overhead rotation. Tubes were centrifuged at 14000×g at 4°C for 10min and the supernatant was split in two new tubes. 1ml of co-IP buffer. 100µl of Anti-FLAG M2-Agarose (Sigma, USA) was added into the first tube, and 100µl of IgG-Agarose was added on second tube. Afterwards, the tubes were incubated at 4°C with overhead rotation for overnight. Next day the agarose beads were washed (10 times) as follows: tubes were centrifuged at 3000×g for 1 min and the pellets (beads) were resuspended in 1ml co-IP buffer. After the last washing step beads were resuspended in 20µl of 2x SDS-PAGE buffer. Before loading onto the gel, the samples were denatured at 95°C for 5min.

II-2.13 Mass Spectrometry

For Mass Spectrometry, co-IP from SPON2-FLAG transfected cells was sent to the Mass Spectrometry/Proteomics core facility of the Centre of Medical Research, Medical University, Graz, Austria. Primary data analysis was done by this core facility using Spectrum Mill (Agilent, USA) and Mascot database (Matrix Science, UK).

II-2.14 Custom antibody production against SPON2 protein

A custom polyclonal antibody against SPON2 protein was raised by BioGenes (Berlin, Germany) in rabbits. The predicted epitopes proposed by the company were C-AHSSDYSMWRKNQY-amide (76-89aa) and C-QDTVTEITSSSPSH-amide (207-220aa).

Additionally, I checked these epitopes by IEDB analysis Resource (http://tools.immuneepitope.org). For each of the two epitopes 2 rabbits were immunized (day 0) and boosted 5-6 times (days 7, 14, 28, 35, 77 and 105). Serum from control bleedings (days 35, 84) was routinely checked by dot-blot and compared with preimmune serum (day 0, negative control).

Serums from the final bleedings were purified by SulfoLink Coupling Gel (Pierce, USA) according manufacturer's recommendations with some modifications concerning the antibody elution step. After running the serum through the column, washing with 12ml of PBS was done in order to wash away non-bound components. Then the following buffers (8ml each) were used in consecutive order to perform elution of antibody from the column:

1. ImmunoPure Gentle Ab Elution Buffer (Pierce, USA);
2. 50mM Pipes pH5.8
3. 100mM Citrat pH5.0,
4. 50mM Glycin pH2.5
5. 100mM Triethylamin pH11.5

Each buffer was collected separately in 4 fractions (2ml each) and pH of all fractions was stabilized by adding 50µl 1M Tris pH8.0. The eluate fractions were conserved with NaN_3 (final conc. 0.5%).

The purified fractions (20 fractions for each animal) were checked by dot-blot, and Ab containing fractions were validated by WB, using the whole cell lysate (IMR-90 cell line) and recombinant SPON2 protein (Abnova, Taiwan). For negative control the Ab was pre-blocked with peptide. The Ab was also checked if it is suitable for IF.

II-3 Meta-analysis

Substantial technological developments during the last decade tremendously increased not only our knowledge in the field of biomedicine, but also produced a large amount of experimental data. Although generation of valuable results, presented by numerous genes, is a great advantage of high-throughput techniques, it necessitates *in silico* algorithms, enabling and facilitating interpretation of collected results.

II-3.1 Gene Ontology

Gene Ontology (GO) is an initiative, aimed in categorisation and annotation of genes in a hierarchical manner. In other words, GO project provides an ontology of keywords presenting pertinent features of gene products. The ontology contains three distinct categories: *biological process*, *cellular component* and *molecular function*. *Biological process* categorizes genes by their implication in specified processes (e.g. cell adhesion, morphogenesis, development etc.). The category *cellular component* indicates the localization of gene products in certain cellular components (e.g. tight junctions, nuclear inner membrane, extracellular matrix etc.). Finally, *molecular function* describes the role of gene products in specific molecular activity and biochemical reactions (enzyme activity, microtubule motor activity, lactase activity etc.) [103]. Each annotation in GO also includes an evidence code, indicating in what context the annotation to a particular term is supported (Figure I-15). These codes are grouped in the following categories: experimental, computational analysis, author statement and curator statement evidences. Apart from "inferred from electronic annotation" (IEA), all the codes are assigned by GO project curators.

The use of GO tremendously facilitates and speeds up the search for information about genes and processes. For instance, GO can easily provide a list of genes implicated in a certain process, function or cellular component without wasting time on collecting information from single sources. Although GO is an advantageous resource, it is not complete and lags behind the current research. Another limitation of GO is that individual curators may interpret data differently.

GO slims are simplified versions of the GO ontology, built up only by summarizing nodes, which give a broad overview without details of specific terms (daughter branches). GO slims are especially useful for annotation of large lists of genes, produced by microarrays.

To get a summarized signature of the expression pattern I also performed annotation enrichment using GOA and whole proteome slim [104-107]. The slim I used is a generic gene ontology annotation, which is combined with UniProtKB and IPI (International Protein Index) sets.

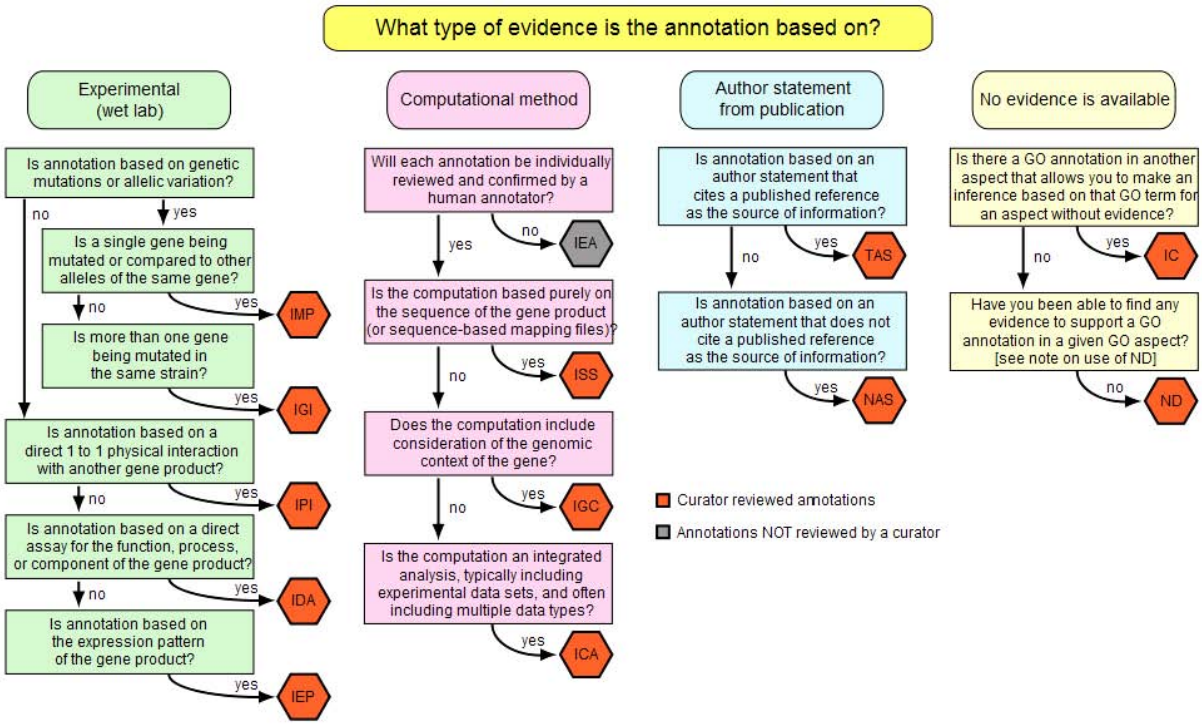


Figure I-15. GO evidence code decision tree (GO consortium).

II-3.2 Annotation and enrichment tools

Nowadays abundance of information, especially produced by high throughput platforms such as microarrays, needs appropriate tools for interpretation. For instance, functional annotation of differentially expressed genes is a compulsory step in the analysis of microarray data. Enrichment analysis can rapidly unravel its biological context and categorize large gene lists into separate functional groups (e.g. implication into biological functions, pathways and diseases, or enrichment with protein domains etc). Significance of enriched groups strongly depends on background (overall gene/protein set of experiment).

There are different commercial softwares and freeware, dedicated to this kind of analysis. In this study I was frequently using GeneSpring GX (Agilent, USA), Database for Annotation, Visualization and Integrated Discovery (DAVID, http://www.david.niaid.nih.gov) [108, 109] and Ingenuity Pathway Analysis (IPA, www.ingenuity.com).

GeneSpring GX is a program mainly oriented on the analysis of gene expression data. It involves a broad spectrum of statistical algorithms, gene annotation tools, as well as tools for KEGG pathway and GO annotation enrichment. It is especially useful for gene expression analysis, but is strongly limited to the given array platform. In contrast to

GeneSpring GX, the DAVID freeware provides more annotation flexibility and handles any type of gene lists [109]. Another important condition for successful analysis is quality of gene lists, analysed by enrichment algorithms.

Ingenuity Pathway Analysis (IPA) software does not only share many features with DAVID, but also enables modelling, analysis and understanding of complex biological and chemical systems with simultaneous calculation of context enrichment. It helps to find gene or protein oriented cellular and disease processes, as well as information about their implication in signalling and metabolic pathways. IPA also allows building of networks based on biological and chemical knowledge, and is especially useful for disease oriented annotation of results.

II-3.3 Network visualisation

GO study accompanied by enrichment of biological terms helps us to understand the biological context behind the results obtained by high-throughput approaches. Network visualisation, knitted by protein-protein interactions, enables another approach to data interpretation. Cytoscape is an open source bioinformatics platform for visualizing molecular interaction networks and integrating these interactions with gene expression profiles [110]. It allows not only the easy retrieval of interactome, but also acquires protein annotations and description. The main elements of Cytoscape are nodes and edges, which refer to proteins and interactions, respectively. Nodes (genes/proteins) are imported by Cytoscape from different web services (e.g. HPRD, BioGRID, IntAct, Reactome etc.), which assembles them with each other via available edges (interactions). Resulting network is not limited by direct interactions. Cytoscape retrieves also intermediate nodes that extend the interaction network.

Cytoscape's core analysis supports basic tools for data integration and visualization. This analysis can be divided into the following consecutive steps:

- data integration – transfer or upload of experimental data;
- transfer of annotation – retrieval of annotations and descriptions from web services;
- graphical layout – organisation of the network's structure using different algorithms (e.g. spring embedded, hierarchical, circular and other layouts). I used spring embedded layout, where nodes are treated like physical objects that repel each other. Here the layout algorithm sets the positions of the nodes in a way that minimizes the sum of forces in the network;

- attribute visual mapping (adjustment) – whereas layout determines the position of nodes and edges, the attribute visual mapping allows to control their appearance (e.g. node colour, size, shape, edge colour, thickness etc.);
- selection and filtering – reduces complexity of large networks using different criteria (e.g. filtering out nodes and edges that do not connect queried nodes);

Additional features (e.g. network and molecular profiling analyses, layouts, scripting, connections with databases) are accessible as plug-in components.

II-3.4 Gene Expression Omnibus (GEO)

Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) is a public repository for a wide range of high-throughput experimental data. These include microarray-based experiments (gene expression arrays, aCGH, ChIP-on-chip) and non-array techniques (e.g. SAGE, mass spectrometry peptide profiling, various types of quantitative sequence data). The National Center for Biotechnology Information (NCBI) initiated the GEO database in 2000 with the aim to support dissemination of gene expression data generated by high-throughput methodologies [111]. The GEO is composed of the following domains:

- platform (GPL) - defines the array template and contains sequence IDs;
- samples (GSM) - represents biological information, source and treatment protocol of applied on platform samples;
- series (GSE) - links together experimentally related samples;

Based on these publicly available datasets, GEO provides a number of meta-analysis tools:

- datasets (GDS) - experiment-centred data, based on biologically and statistically comparable samples. It includes clustering tools (hierarchical and K-means) and differential expression queries (t-test and fold difference);
- profiles - gene-centred data, correspond to expression profiles of single genes.

GEO is a great tool to share own data and compare with other publicly available data. Using GEO researchers can download expression data and extract task- or gene-specific information, which can be very useful avoiding redundant laboratory work.

Results

III RESULTS

III-1 Carcinoma and underlying stroma cells share common chromosomal aberrations

Aberrations, such as LOH and microsatellite instability, were repeatedly found in epithelium and stroma adjacent to carcinoma [33, 112-114]. In order to check if DNA copy number changes are also present in these adjacent components, we studied laser microdissected normal epithelium, stromal and tumour components derived from 8 adenocarcinoma (AC) and 9 squamous cell carcinoma (SQCC) of lungs using submegabase resolution array CGH [115]. In 4 cases (23.5%) array-CGH revealed that some chromosomal aberrations are also present in tumour associated stroma and adjacent epithelial cells (Figure I-16).

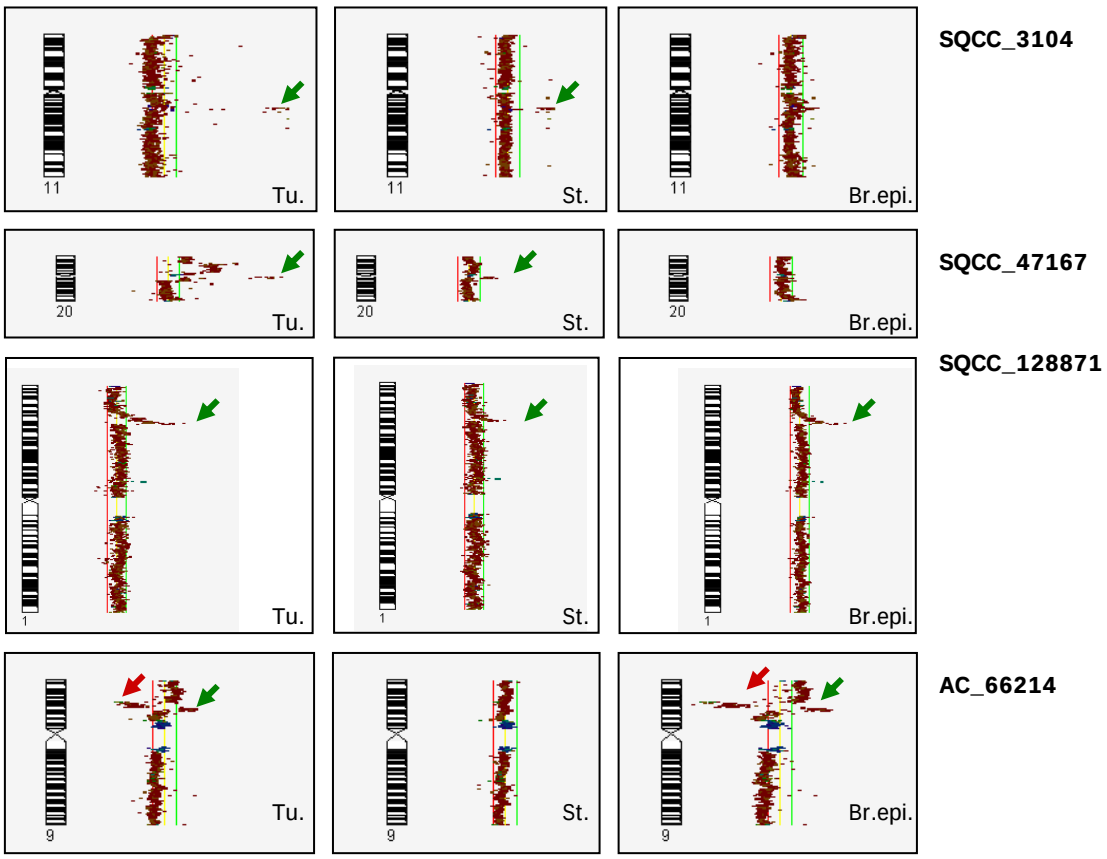


Figure I-16. DNA copy number changes, retained in stromal and normal epithelial components. Data analysis and visualization were performed by CGHPRO (Max Planck Institute for Molecular Genetics, Berlin, Germany). For each BAC clone Cy3/Cy5 signal intensity ratios are plotted alongside the chromosome idiograms. Losses and gains lead to a ratio shift to the left and right, respectively. Red and green lines correspond to log2 ratios -0.3 (loss) and 0.3 (gain), respectively. Tu - tumour, St – underlying stroma, and Br.epi – normal bronchial epithelium.

III-2 SCLC cell line as a possible model for EMT study

The majority of small cell carcinoma (SCLC) cell lines grow as floating aggregates [116]. The cell line I used was adapted from a patient diagnosed with small cell carcinoma of the lung (courtesy of Prof. Popper and Prof. Pfragner, Austria). It exhibits mixed growth properties (Figure I-17), where a major part of the cells was growing in suspension and only a very minor portion of the cells was adhering to the surface of the flasks. These adhesive cells showed spindle-shaped fibroblasts-like morphology (Figure I-17). Array CGH revealed that both fractions were characterised by the same pattern of chromosomal aberrations. This strongly argues against a putative contamination by fibroblasts (Figure I-20).

The suspensive fraction of SCLC cell line (SCLC-S) was carefully separated from its adhesive counterpart (SCLC-A) and cultured in a new flask. Although this procedure was repeated several times, the SCLC-S retained its ability to generating SCLC-A, and *vice versa*.

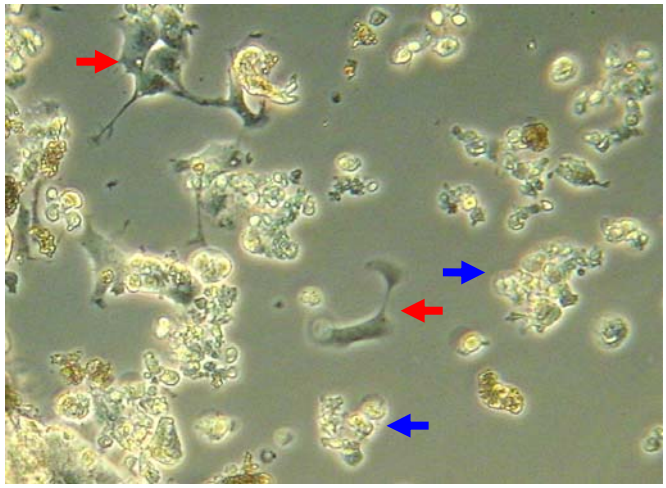
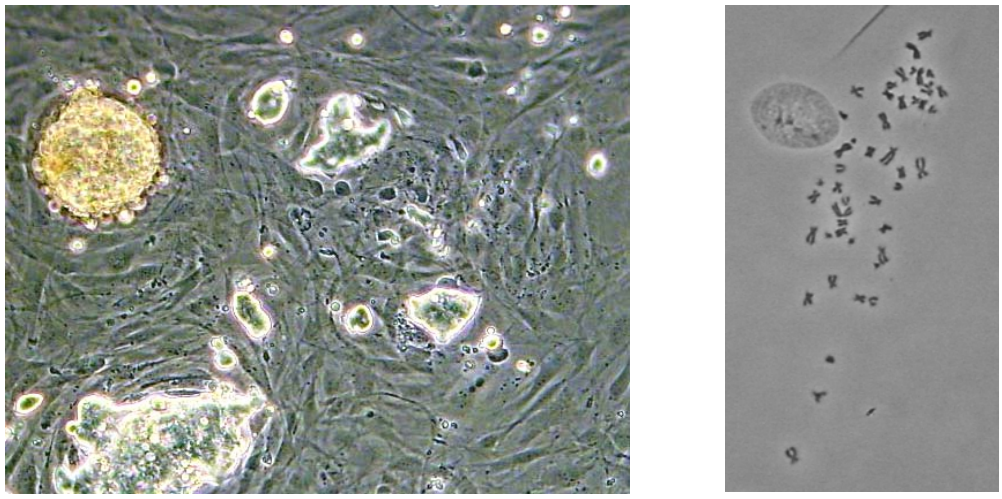


Figure I-17. Original SCLC cell line.
The red and blue arrows indicate adhesive and suspensive cells, respectively.

III-2.1 SCLC–A growth was accelerated after xenograft transplantation

SCLC cell line was transplanted into a nude mouse to accelerate growth of cancer cells (Laboratory Animal Science and Genetics, Medical University of Vienna, Himberg, Austria). SCLC cells re-seeded after xenograft transplantation displayed considerable changes of their growth properties. The great majority of cells grew adhesive with spindle shaped morphology and had increased proliferation rates (Figure I-18a). Karyotype analysis excluded potential contamination of transplanted cells by host (mouse) fibroblasts (Figure I-18b).



a) **b)**

Figure I-18. SCLC cell line after xenograft transplantation into a mouse (Passage 1).

(a) adhesive fibroblast-like cells dominated after xenograft transplantation. *(b)* Giemsa stained metaphase chromosomes of this cell line show no contamination by host (mouse) cells.

Additional verification was done by Agilent's 244k oligo aCGH analysis, comparing SCLC cell lines before and after transplantation. The patterns of DNA copy number changes were identical with only one exception, where the post-transplantation cell line showed a loss in chr3:71418986-71741160 region harbouring *FOXP1* (Figure I-19).

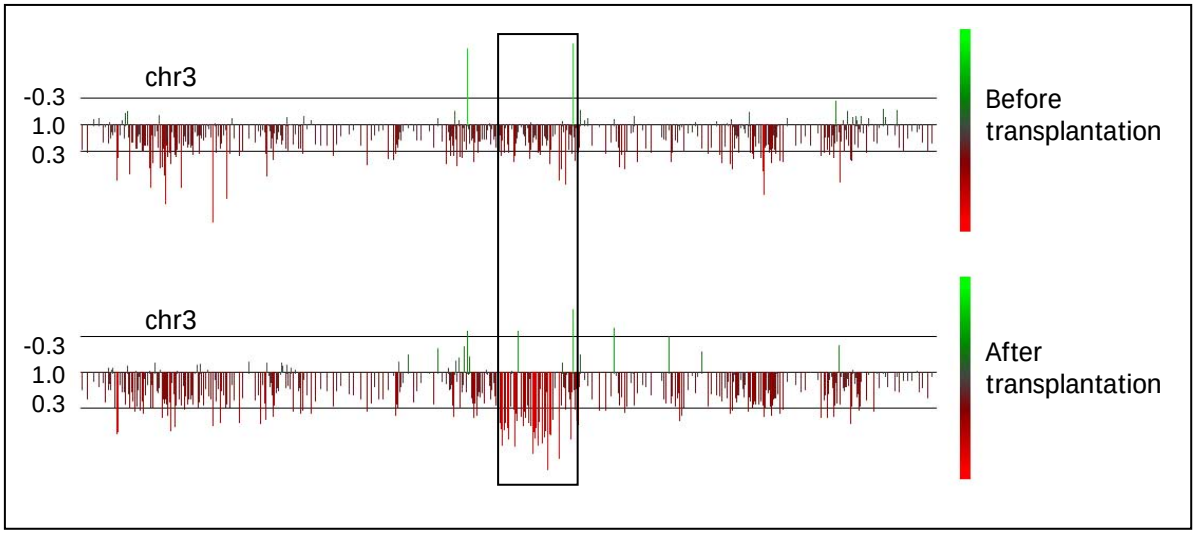


Figure I-19. aCGH analysis of SCLC cell line before and after xenograft transplantation.

The box indicates loss of chromosomal region chr3:71418986-71741160 in the cells after transplantation. Red and green bars represent losses and gains, respectively. The vertical lines show log2 ratio thresholds -0.3, 1.0 and 0.3. Visualisation was done with GenomeCAT 1.1.

III-3 aCGH analysis revealed no differences in DNA copy number between SCLC–A and –S

Although originating from the same tumour, SCLC–A and –S exhibit different morphological features. To figure out possible DNA copy number differences between SCLC–A and –S I performed concurrent hybridization of SCLC–A (Cy3) with reference DNA[§] (Cy5), SCLC–S (Cy3) with reference DNA (Cy5), and SCLC–A (Cy3) with SCLC–S (Cy5) on in-house 36K tiling path BAC arrays. Even if aCGH revealed various complex aberrations, I could not see any differences between SCLC–A and –S (Figure I-20). Same aberration pattern was revealed also when SCLC–A and –S were hybridized on 244K oligo-aCGH (Agilent), confirming our results.

[§] DNA from an EBV transformed 46, XY B lymphocytes. All references for array-CGH experiments were matched with their sex, either using DNA from 46, XY or 46, XX B lymphocytes.

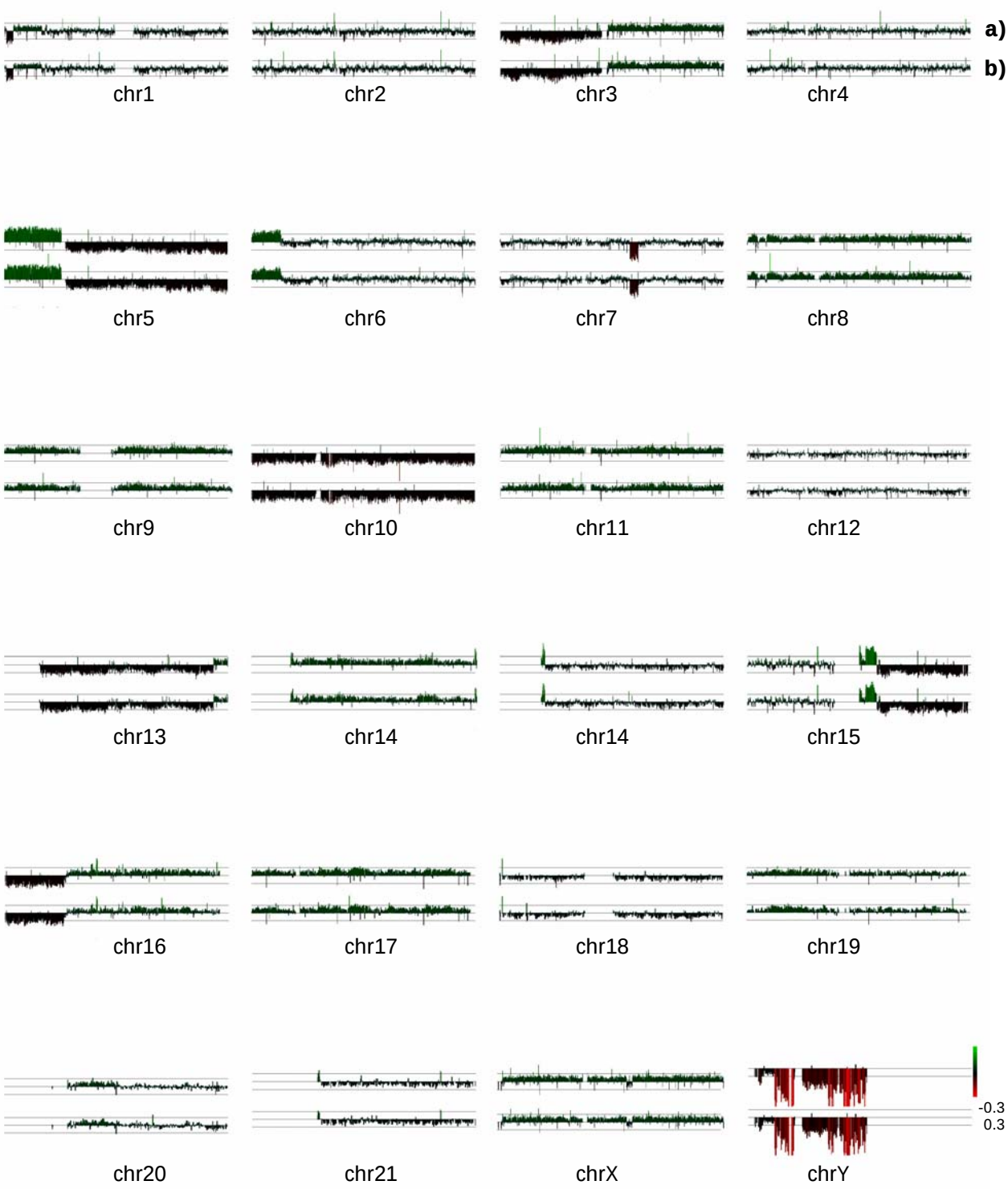


Figure I-20. Comparison of copy number changes in SCLC–A and –S by 36K BAC aCGH.

aCGH pattern of SCLC–A (a) and SCLC–S (b) shows no differences in DNA copy number. Analysed by GenomeCAT 1.1 software, the red and green (colour bar) represents loss and gain, respectively. The vertical lines show log2 ratio thresholds -0.3 and 0.3

III-4 BAC-array based MeDIP and MCA revealed no differences between SCLC–A and –S

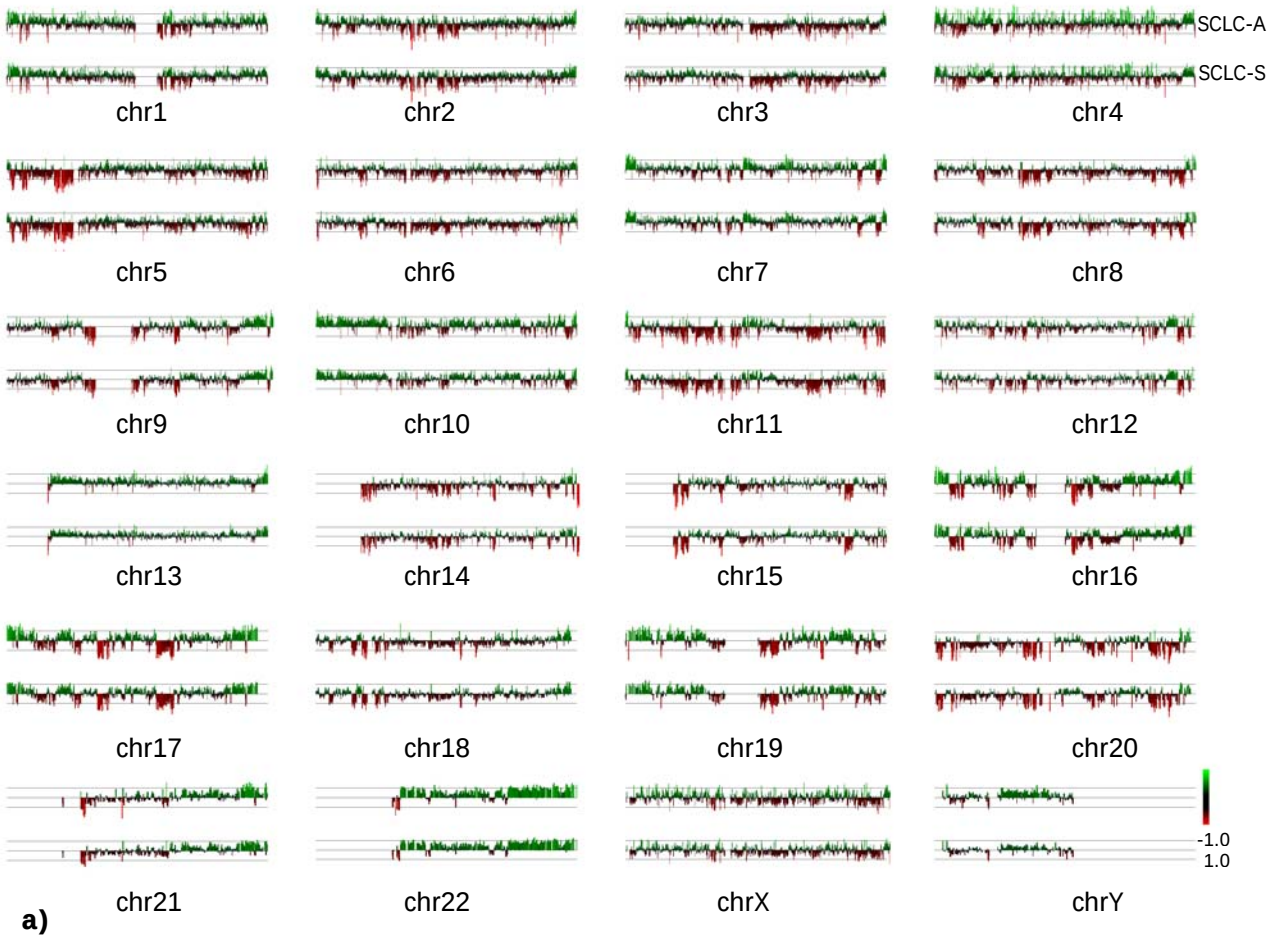
Based on fact that aCGH revealed no differences between SCLC–A and –S on genomic level, I supposed that adhesive and suspensive growth of SCLC cell line can be explained by differences in DNA methylation pattern. Therefore methylated DNA immunoprecipitation (MeDIP) and methylated CpG island amplification (MCA) techniques were established.

Figure I-21 and Figure I-22 show comparison of methylation patterns of SCLC–A and –S as analysed by MeDIP and MCA, respectively. No relevant differences were found in MeDIP and MCA on 36K BAC array. Moreover, I hybridized methylated DNA both, enriched by MeDIP and MCA, from SCLC-A versus SCLC–S on 36k BAC array and again revealed no relevant differences (Figure I-21b and Figure I-22b).

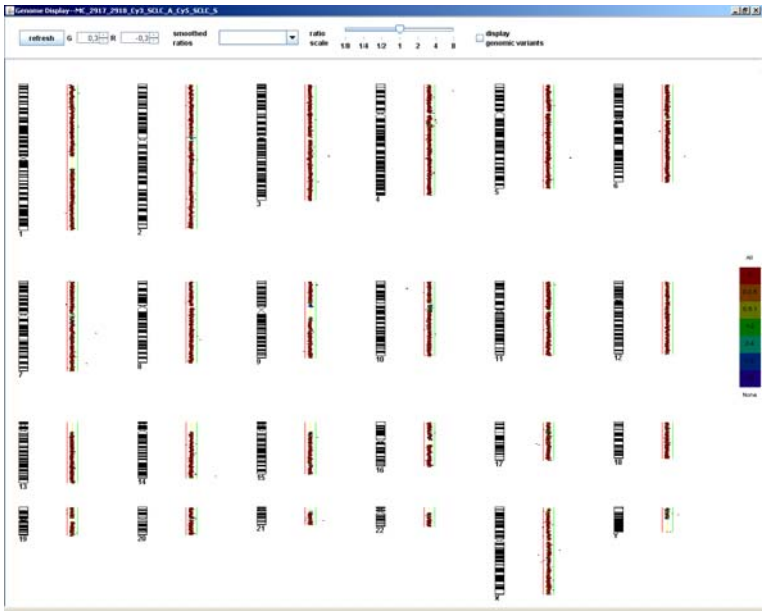
Hybridization of MeDIP and MCA products on Agilent's 244K CpG island microarrays gave unreliable results and consequently these results were not included in our study. DNA methylation status of several differentially methylated CpG islands was checked by bisulfite sequencing PCR (BSP). The last, accepted as a "gold standard" method for DNA methylation study, detected no differences between SCLC–A and –S. Unsuitability of the microarray platform was also corroborated by the manufacturer.

A comparison of MeDIP and MCA data revealed overlapping patterns of methylation (Figure I-23a). As expected, MCA depicts more refined methylation peaks when compared to MeDIP data.

Specificity of antibody against 5-methylcytidine was tested on two DNA sequences with 232bp and 384bp length, containing 12 and 49 CpGs, respectively. They were differentially methylated by M.Sss I enzyme, while methylation of not treated by this enzyme DNA was erased by PCR. The methylation state was checked by methylation sensitive restriction enzyme. These PCR products were then immunoprecipitated by Dynabeads and loaded on agarose gel (Figure I-23b).



a)

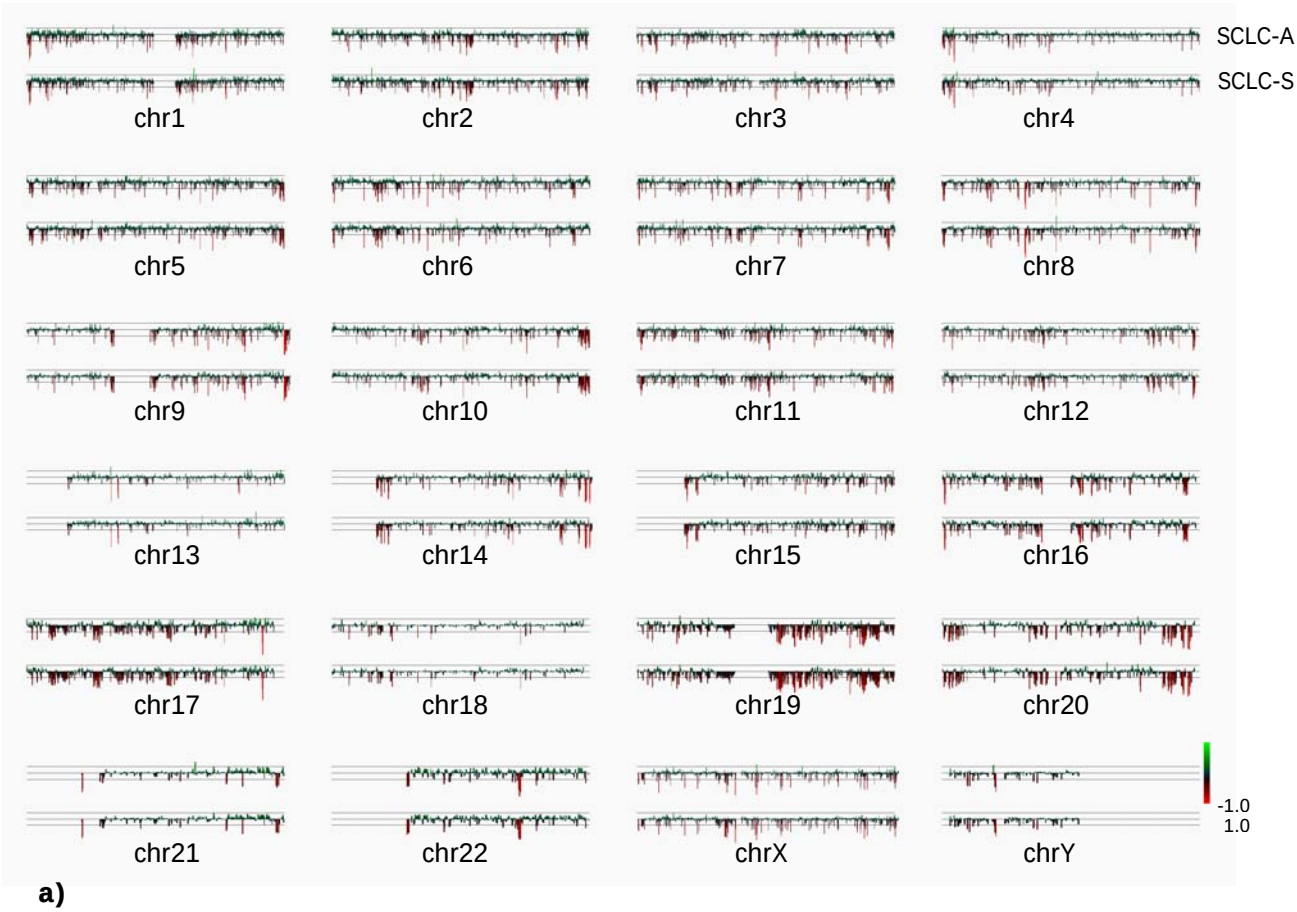


b)

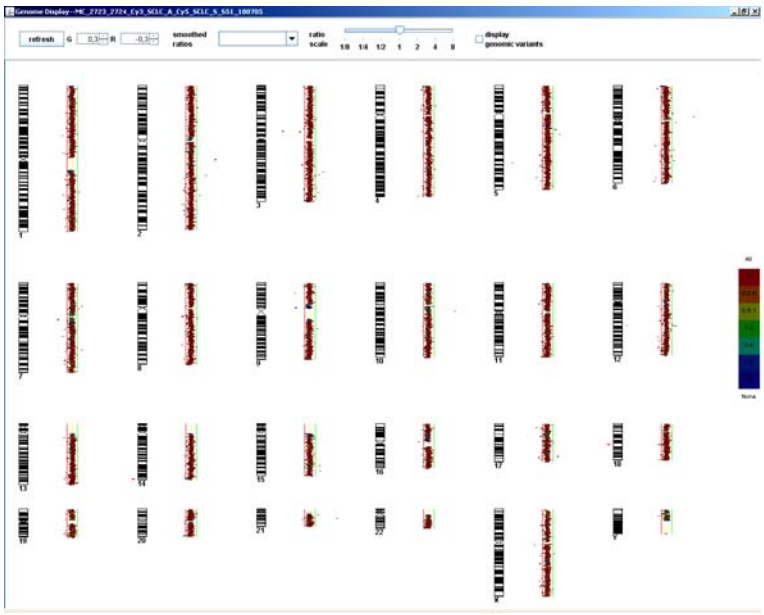
Figure I-21. Comparison of DNA methylation of SCLC–A and –S by MeDIP.

a - DNA methylation pattern of SCLC–A and SCLC–S show no relevant differences. Analysed by GenomeCAT 1.1 software, the red and green colour represents loss and gain, respectively. The vertical lines show log2 ratio thresholds -1.0 and 1.0.

b - Methylated DNA from SCLC-A enriched by MeDIP and labelled in Cy3 and SCLC-S labelled in Cy5 have been co-hybridised on BAC arrays. No relevant changes in methylation status were detected. Yellow, red and green lines show log2 ratios 1.0, -0.3 and 0.3, respectively.



a)



b)

Figure I-22. Comparison of DNA methylation of SCLC–A and –S by MCA.

a - DNA methylation pattern of SCLC–A and SCLC–S show no relevant differences. Analysed by GenomeCAT 1.1 software, the red and green colour represents loss and gain, respectively. The vertical lines show log2 ratio thresholds -1.0 and 1.0.

b - Methylated DNA from SCLC-A enriched by MCA and labelled in Cy3 and SCLC-S labelled in Cy5 have been co-hybridized on BAC arrays. No relevant changes in methylation status were detected. Yellow, red and green lines show log2 ratios 1.0, -0.3 and 0.3, respectively.

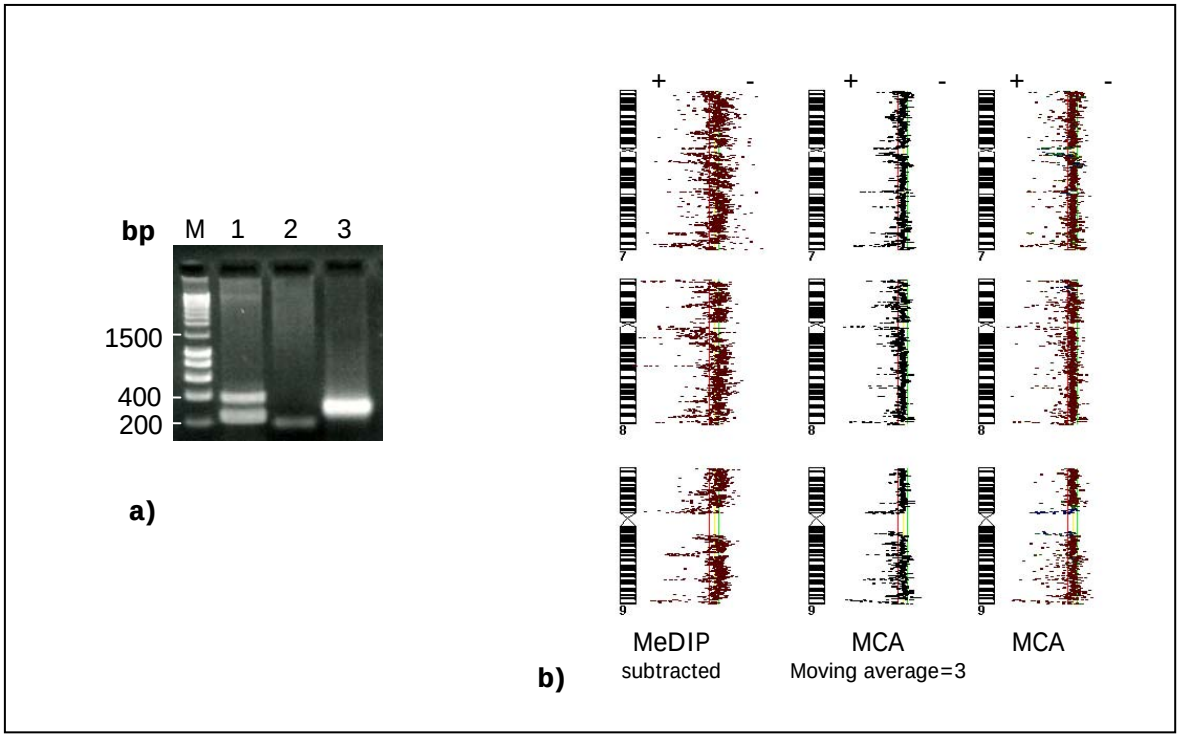


Figure I-23.

a - Validation of antibody against 5-methylcytidine (for MeDIP experiment). Lane 1 is a pull-down of methylated PCR products (both); lane 2 – pull-down of methylated 232bp and demethylated 384bp PCR products mixture; lane 3 – pull-down of demethylated 232bp and methylated 384bp PCR products mixture. The gel electrophoresis ensures that antibody specifically binds to methylated DNA.

b - Comparison MeDIP and MCA techniques on SCLC-A using 36k BAC platform. MeDIP subtracted was corrected, using pattern of non-cancerous cell line as reference (i.e. hyper- and hypomethylated peaks present in both, SCLC-A and reference DNA, were removed, depicting only SCLC-A specific methylation pattern). + and – represent hyper- and hypomethylation, respectively.

III-5 Gene expression analysis of SCLC cell line

Next I performed gene expression analysis of SCLC–A and –S. RNA was separately hybridized on Agilent's 4x44K Whole Human Genome Oligo Microarrays in three technical replicas. After intra- and inter-array normalizations and filtering by volcano plot (2-fold change, P<0.05, Figure I-24) I obtained a list of 60 differentially expressed genes; 22 upregulated in SCLC-A, and 38 in SCLC-S (supplementary CD, S1.txt and S2.txt, respectively). The most differentially expressed genes were *MMP1* and *SPON2*, which were overrepresented in SCLC-A. Overexpression of MMP1 and *SPON2* in SCLC-A was independently verified by qRT-PCR (Figure I-25).

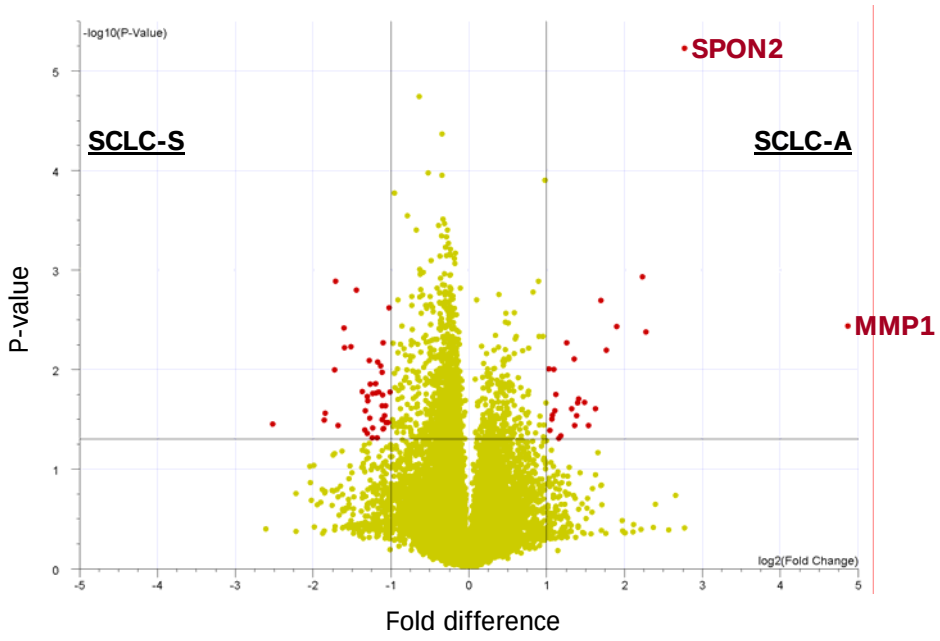


Figure I-24. Volcano plot of gene expression array of SCLC-A and –S.

Volcano plot shows significance and magnitude of differences in gene expression. The negative $\log_{10}$-transformed P-values are plotted against the log ratios ($\log_2$ fold change) in a two-sample experiment. Left and right sides of plot correspond to SCLC-S and –A, respectively. Yellow dots are genes that did not pass 2-fold change and P<0.05 thresholds (horizontal and vertical black lines). Red dots are differentially expressed genes (n=61) passing the thresholds. Analysed by GeneSpring 7.3.1 (Agilent, USA).

Differentially regulated genes as defined by 2.0-fold difference and P<0.05 filtering were subjected to GO annotation analysis by GeneSpring GX 7.3.1, genome update April 2009 (Table I-45 and Table I-45). GO "Biological Process" enrichment (P<0.01) by GOA and whole proteome slims, using differentially expressed genes in SCLC, revealed only two terms (Table I-46).

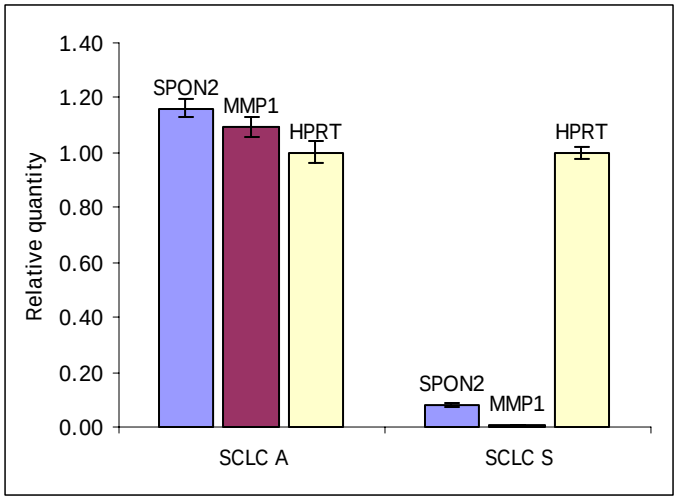


Figure I-25. Relative quantification of MMP1 and *SPON2* gene expression by qRT-PCR.

Relative quantity was analysed using Sequence Detection Software SDS 2.1 (Applied Biosystems, USA). Data were normalised by a series of dilutions (1:1, 1:2, 1:4, 1:8 and 1:16). *HPRT* gene served as internal control (standard detector). Specificity of all primers was checked on gel (single band for each, not shown) and additionally was validated by visual inspection of the dissociation curve. Relative quantity shows relative to undiluted HPRT quantity.

Table I-44. GO annotation enrichment analysis of upregulated genes in SCLC-A. Filtering criteria were 2.0-fold change and P<0.05.

GO Biological Process	Genes in Category	% of Genes in Category	Genes in List in Category	% of Genes in List in Category	p-Value
GO:6817: phosphate transport	144	0.70	4	25.00	3.94E-06
GO:7275: development	3595	17.50	11	68.75	8.51E-06
GO:15698: inorganic anion transport	227	1.11	4	25.00	2.38E-05
GO:48513: organ development	1356	6.60	7	43.75	3.62E-05
GO:6820: anion transport	268	1.30	4	25.00	4.55E-05
GO:30574: collagen catabolism	26	0.13	2	12.50	1.83E-04
GO:270: peptidoglycan metabolism	34	0.17	2	12.50	3.14E-04
GO:1955: blood vessel maturation	1	0.00	1	6.25	7.79E-04
GO:1501: skeletal development	304	1.48	3	18.75	1.56E-03
GO:35019: somatic stem cell maintenance	3	0.01	1	6.25	2.33E-03
GO:30512: negative regulation of transforming growth factor beta receptor signaling pathway	3	0.01	1	6.25	2.33E-03
GO:43062: extracellular structure organization and biogenesis	132	0.64	2	12.50	4.63E-03
GO:30198: ECM organization and biogenesis	132	0.64	2	12.50	4.63E-03
GO:6381: mRNA editing	6	0.03	1	6.25	4.66E-03
GO:16556: mRNA modification	6	0.03	1	6.25	4.66E-03
GO:19827: stem cell maintenance	8	0.04	1	6.25	6.21E-03
GO:320: re-entry into mitotic cell cycle	8	0.04	1	6.25	6.21E-03
GO:7155: cell adhesion	1053	5.13	4	25.00	7.61E-03
GO:6811: ion transport	1086	5.29	4	25.00	8.47E-03
GO:16547: RNA editing	11	0.05	1	6.25	8.53E-03
GO:35162: embryonic hemopoiesis	12	0.06	1	6.25	9.31E-03

Table I-45. GO annotation enrichment analysis of upregulated genes in SCLC-S. Filtering criteria were 2.0-fold change and P<0.05.

GO Biological Process	Genes in Category	% of Genes in Category	Genes in List in Category	% of Genes in List in Category	p-Value
GO:7618: mating	3	0.015	1	4.348	3.35E-03
GO:7620: copulation	2	0.010	1	4.348	2.24E-03
GO:43367: CD4-positive T cell differentiation	1	0.005	1	4.348	1.12E-03
GO:43374: CD8-positive T cell differentiation	1	0.005	1	4.348	1.12E-03
GO:19532: oxalate transport	4	0.020	1	4.348	4.47E-03
GO:15858: nucleoside transport	6	0.029	1	4.348	6.70E-03
GO:15860: purine nucleoside transport	2	0.010	1	4.348	2.24E-03

Table I-46. GOA and proteome slim enrichment analysis of differentially expressed genes in SCLC cell line (1.5-fold, P<0.05). Note that "cell motility" term in SCLC-S (marked with asterisk) is below the P=0.01 enrichment threshold.

GOA and proteome slim (Biological Process)	SCLC-A (p-value)	SCLC-S (p-value)
GO:69281: cell motility	4.14E-03	1.32E-01*
GO:7275: development	3.29E-05	3.86E-03

III-6 SPON2 and MMP1 siRNA knockdown in embryonal lung fibroblasts

Expression analysis of SCLC–A and –S revealed *SPON2* and *MMP1* as most differentially regulated genes. In order to find an appropriate cell line highly expressing *SPON2* and *MMP1* I used GEO database. A primary embryonal lung fibroblast primary cell line called IMR-90 was found to be the only one meeting our requirements (GEO dataset GDS395).

Whereas *MMP1* knockdown showed no morphological changes even after 168h, *SPON2* knockdown induced alterations in the cellular morphology starting from 72-96h after transfection (Figure I-26). Drastic changes in cell shape by *SPON2* knockdown suppose its implication in cell adhesivity and polarity. Genome-wide expression analysis of transfected cells was analysed using Agilent's 4x44K Whole Human Genome Oligo Microarrays (Figure I-27). The lists of differentially expressed genes are presented in supplementary CD (S3.txt, S4.txt, S5.txt, S6.txt, S7.txt and S8.txt files). Comparison of differentially expressed genes at 72h, 96h and 120h after *SPON2* siRNA treatment (4-fold changes, P<0.05) discovered 103 commonly deregulated genes (Figure I-27 b and c). The results of GO annotation enrichment based on genes deregulated by *SPON2* knockdown is given in Table I-47. Enrichment by GOA and proteome slims, picking up generalized terms, is shown on Table I-48 and Figure I-28.

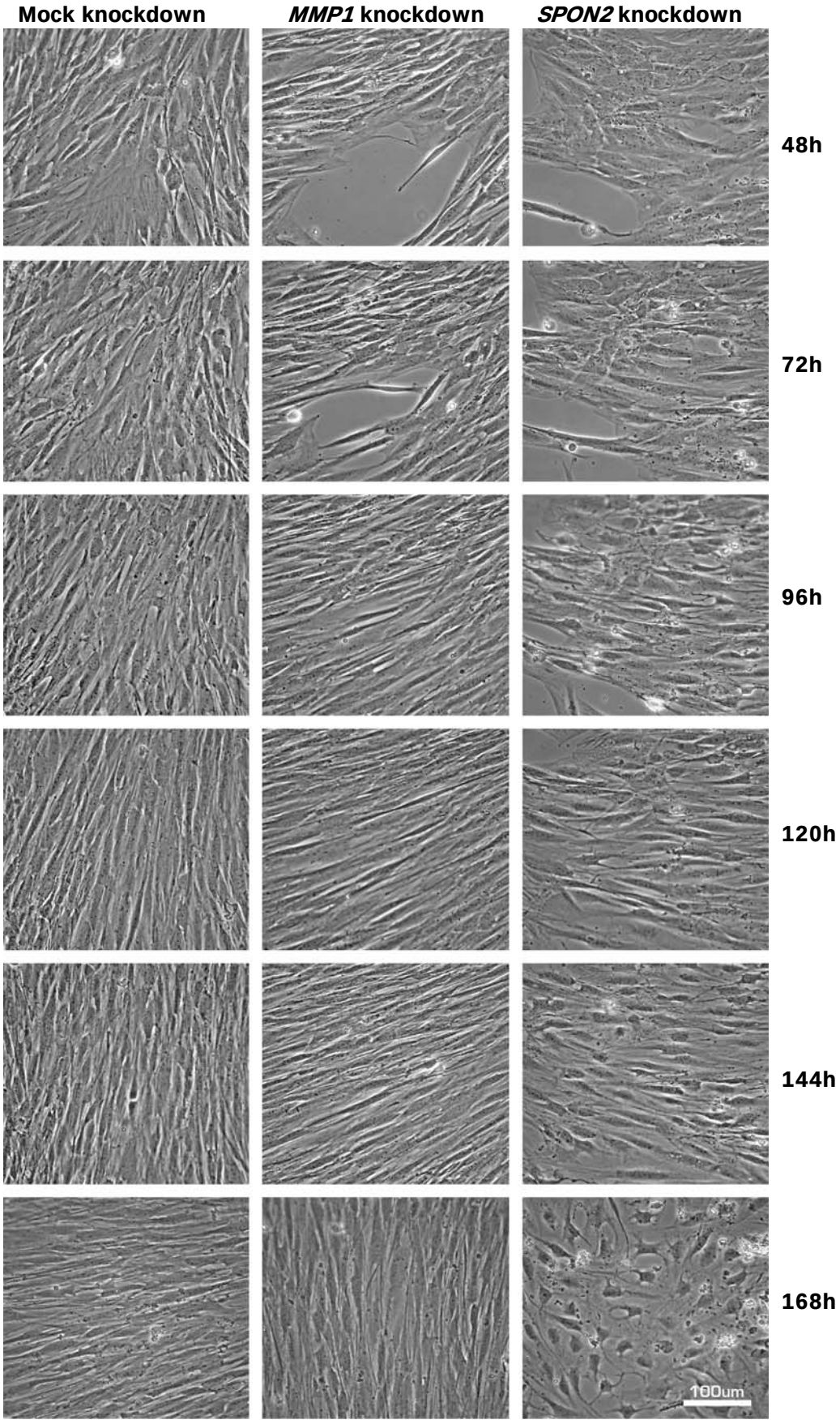


Figure I-26. Morphological changes in IMR-90 fibroblasts after *MMP1* and *SPON2* knockdown.

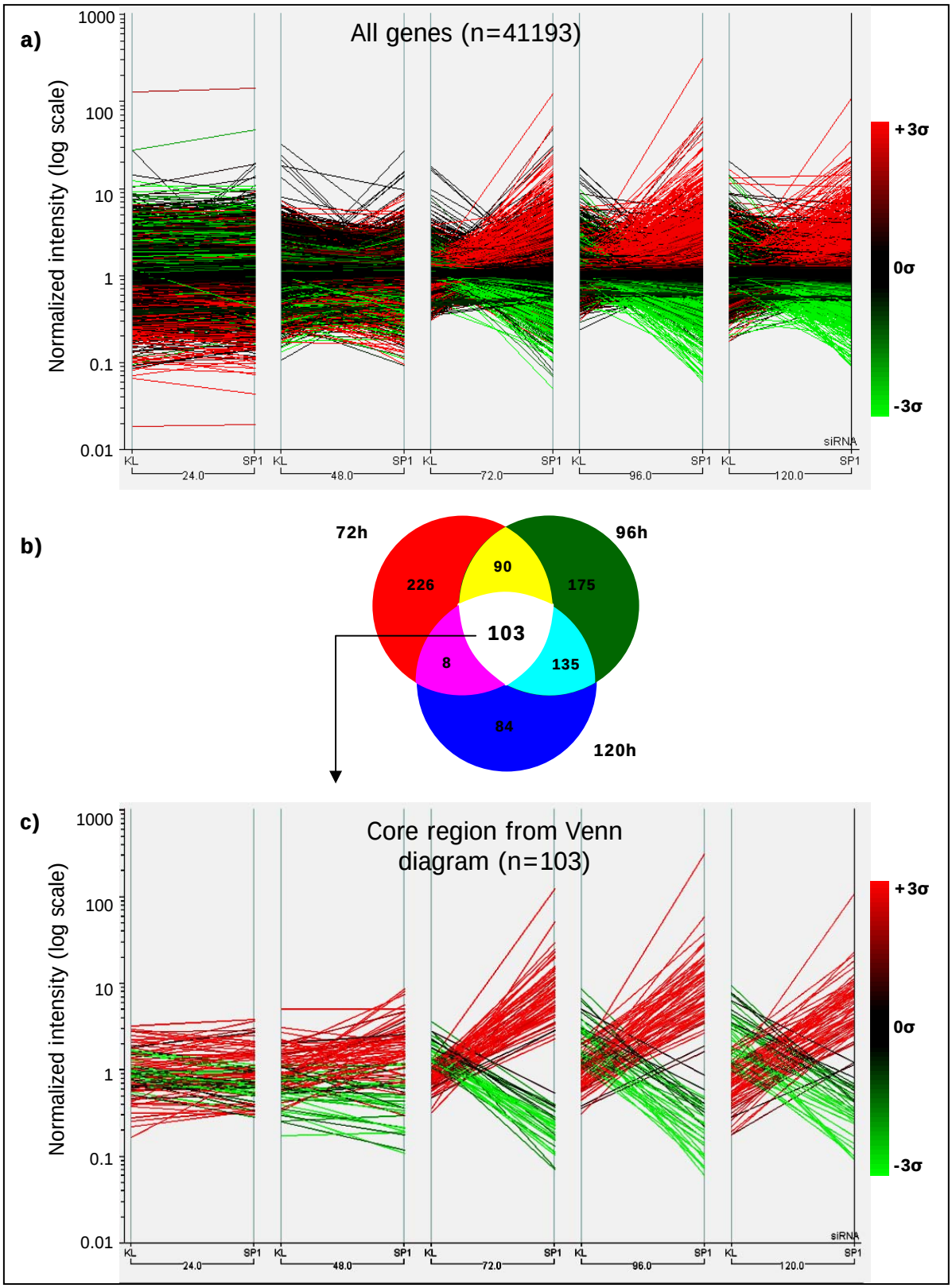


Figure I-27. Gene expression profile of IMR-90 cells at various time points after *SPON2* siRNA knockdown. Gene expression pattern before *(a)* and after *(c)* filtering. Differentially expressed genes were first filtered by volcano plot (4-fold changes, P-value 0.05), then the core region was extracted by Venn diagram *(b)*. List of filtered genes is given in supplementary CD, S9.txt file. The colour bar indicates significance; vertical scale is normalized intensity (log scale); 24, 48, 72, 96 and 120 – hours after transfection; KL – mock; SP1 – *SPON2* siRNA.

Table I-47. GO annotation enrichment based on genes deregulated by SPON2 knockdown (4-fold, P<0.05). Full table of annotation enrichment is listed in supplementary tables S-1, S-2 and S-3.

GO Biological Process	48h (P-value)	72h (P-value)	96h (P-value)	120h (P-value)
GO 6954: inflammatory response	0.693	1.47E-07	1.29E-09	3.47E-15
GO 42330: taxis	0.496	9.40E-09	1.39E-11	3.44E-15
GO 9611: response to wounding	0.165	2.27E-10	7.09E-10	7.21E-13
GO 6955: immune response	8.63E-03	3.61E-10	3.41E-11	2.57E-12
GO 42127: regulation of cell proliferation	4.51E-03	3.26E-08	9.41E-06	2.36E-05
GO 7275: multicellular organismal development	0.580	9.97E-07	5.97E-05	5.83E-05
GO 8285: negative regulation of cell proliferation	5.33E-04	4.49E-07	1.24E-04	7.91E-05
GO 8283: cell proliferation	1.40E-04	3.54E-06	8.04E-04	9.96E-04
GO 1837: epithelial-mesenchymal transition	1	0.0252	1.42E-04	3.05E-08
GO 30509: BMP signalling pathway	1	6.94E-03	8.83E-04	1.41E-05
GO 7155: cell adhesion	0.0667	0.0369	1.52E-04	6.65E-07
GO 30855: epithelial cell differentiation	1	0.0306	9.23E-04	1.71E-05
GO 6928: cell motion	0.578	5.08E-04	8.72E-05	3.23E-05
GO 51674: localization of cell	0.578	5.08E-04	8.72E-05	3.23E-05
GO 2009: morphogenesis of an epithelium	0.260	1.05E-03	1.77E-03	3.97E-05
GO 51094: positive regulation of developmental process	1	1.40E-03	8.36E-06	5.67E-05
GO 16477: cell migration	0.298	8.94E-05	7.34E-05	1.54E-04
GO 45785: positive regulation of cell adhesion	1	1	1	3.83E-04
GO 45123: cellular extravasation	1	1	1	2.95E-03
GO 1935: endothelial cell proliferation	2.98E-05	1.57E-03	1	1
GO 50673: epithelial cell proliferation	0.1	0.0619	1	1
GO Cellular Component				
GO 5576: extracellular region	1.42E-04	1.31E-16	1.56E-33	1.59E-31
GO 31012: extracellular matrix	8.95E-03	0.0725	5.79E-05	7.50E-05
GO 5610: laminin-5 complex	1	1	0.0439	1.14E-06
GO 5923: tight junction	2.82E-05	2.54E-03	0.0168	6.53E-04
GO 30054: cell junction	2.55E-04	6.76E-04	0.0791	4.89E-03
GO 5911: cell-cell junction	1.18E-04	2.03E-04	0.0430	2.06E-03
GO 5915: zonula adherens	1	1	1	0.0413
GO 45298: tubulin complex	1	0.0666	0.0721	0.0514

Table I-48. GOA and proteome slim enrichment analysis of deregulated genes (4-fold, P-value 0.05) in IMR-90 cells after *SPON2*-siRNA knockdown (48-120h)

GOA and proteome slim (Biological Process)	48h (P-value)	72h (P-value)	96h (P-value)	120h (P-value)
GO 7154: cell communication	0.0995	1.83E-10	9.41E-13	8.24E-08
GO 69281: cell motility	0.578	5.08E-04	8.72E-05	3.23E-05
GO 50896: response to stimulus	0.0426	6.42E-08	9.75E-06	4.32E-05
GO 30154: cell differentiation	0.504	5.31E-04	4.31E-05	4.51E-05
GO 7275: development	0.580	9.97E-07	5.97E-05	5.83E-05
GO 7610: behaviour	0.454	9.48E-06	2.41E-05	9.06E-05
GO 43062: extracellular structure organisation and biogenesis	1	7.91E-03	6.95E-04	0.0508
GO 6810: transport	0.329	5.02E-03	0.638	0.996

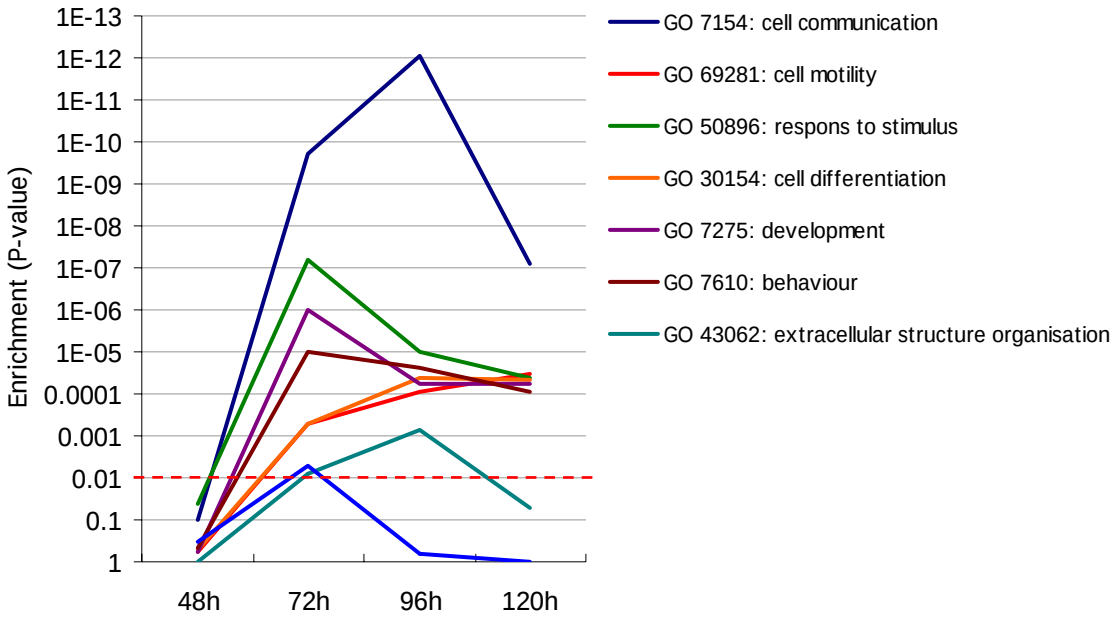


Figure I-28. GOA and proteome slim enrichment plot based on Table I-48.
The dotted line shows the P=0.01. GO terms below P=0.01 I considered as not relevant.

III-7 SPON2 knockdown specificity

The SPON2 protein contains a unique N-terminal Spondin_N domain, which has no detectable sequence homology to any known human protein, and one C-terminal TSP1 (thrombospondin type-1) domain (Figure I-29a). The latter can be identified in over 40 human proteins [117]. In order to prove the specificity of *SPON2* knockdown, all spondin family and TSP-1 domain containing proteins were analysed in GeneSpring software (Figure I-29b and c). *SPON2* knockdown did not directly affect members of the spondin family and TSP1 domain containing proteins, and shows relevant changes in expression level only in *SPON2* gene.

Transfection protocol and competence of cells were checked by Cy3 labelled GFP-targeting siRNA. The efficiency of knockdown was additionally validated by qRT-PCR and Western blot (Figure I-30). qRT-PCR analysis also confirmed that *SPON2* knockdown increases *MMP1* expression, while *MMP1* silencing does not affect *SPON2* expression.

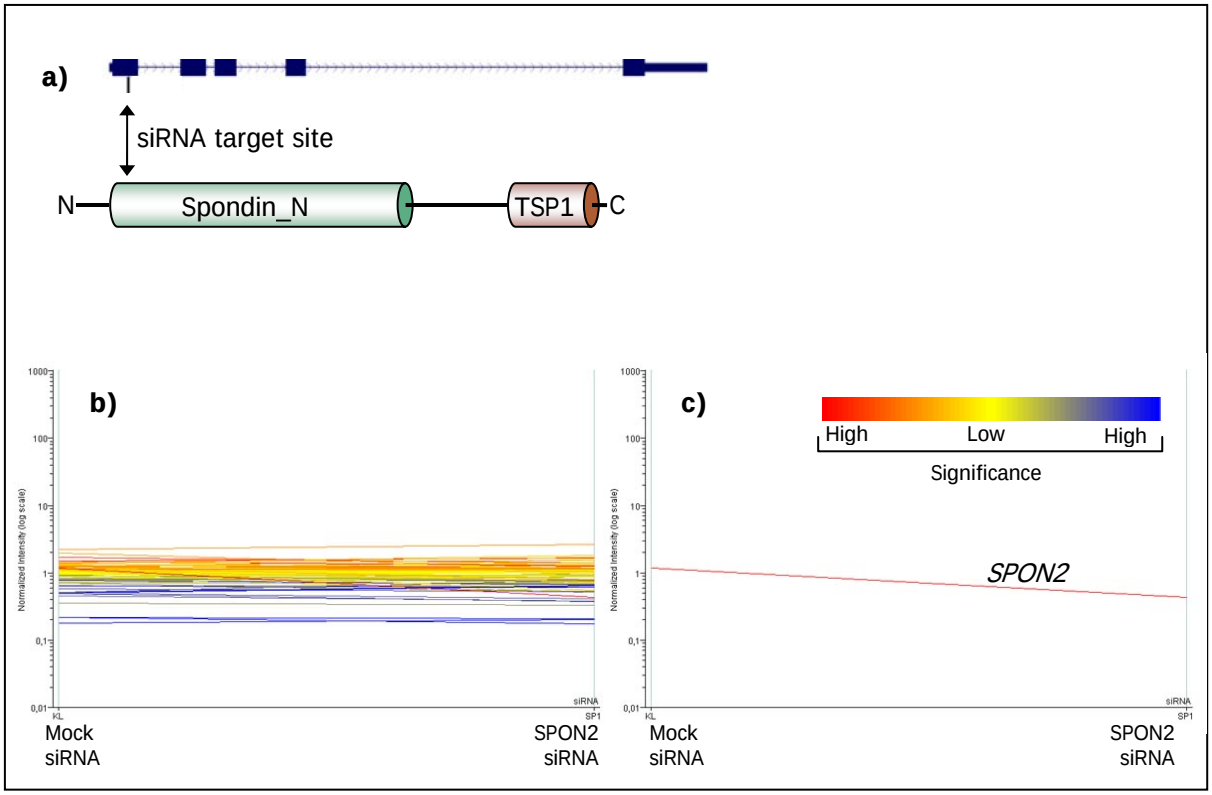


Figure I-29.

a – SPON2 protein contains two domains: Spondin_N and TSP1. Target site of siRNA is aligned on exon1 (Spondin_N domain).

b – Spondin family and TSP1 domain containing proteins before filtering by 2-fold and P<0.05 criteria. 24h after SPON2 silencing in IMR-90 cells.

c – Spondin family and TSP1 domain containing proteins after filtering by 2-fold and P<0.05 criteria. 24h after SPON2 silencing in IMR-90 cells. The single red line is *SPON2.* Colour bar indicates expression level by significance (red-high significance/high expression, blue-high significance/low expression, yellow-low significance).

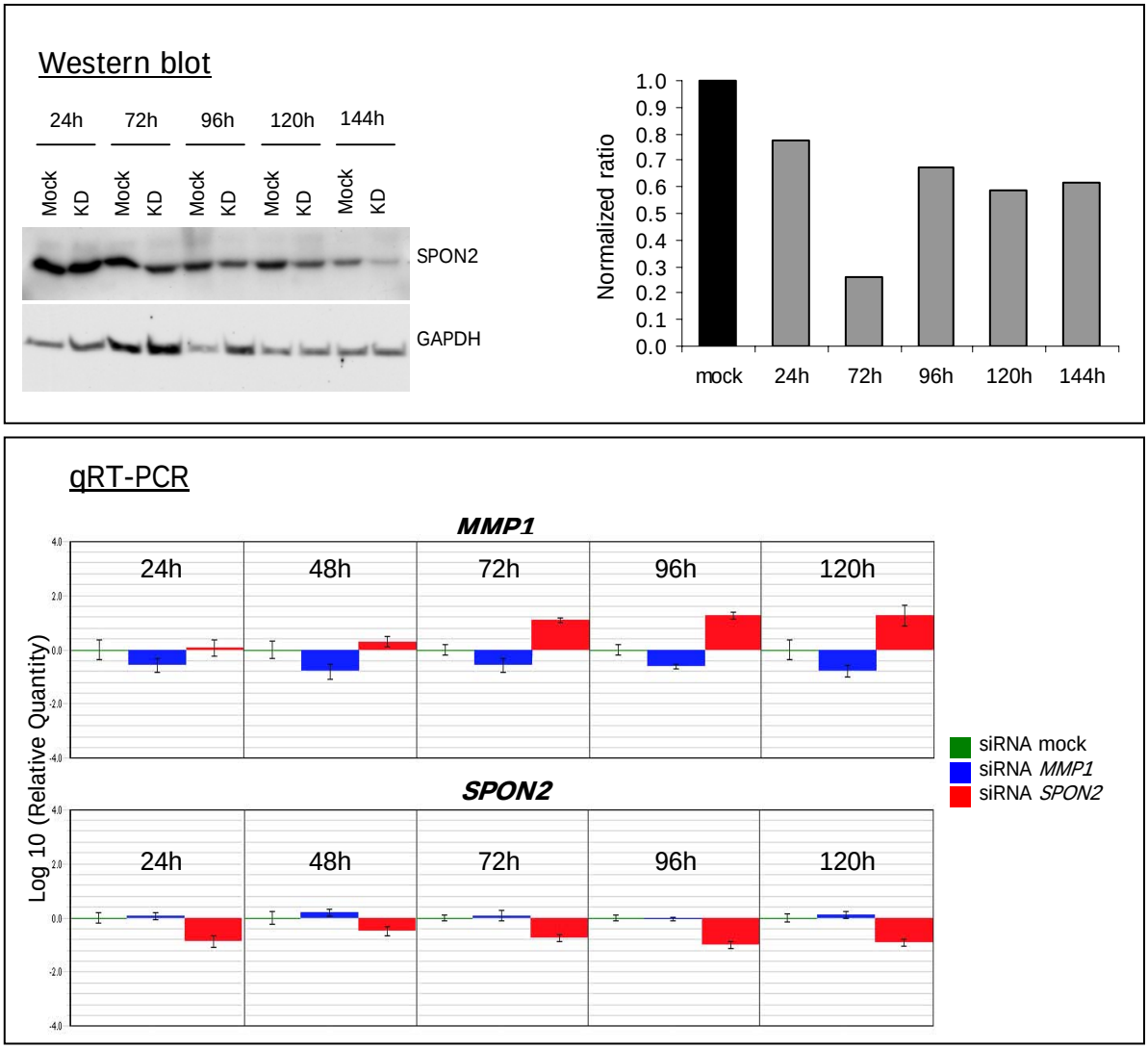


Figure I-30. Knockdown validation by Western blot (SPON2) and qRT-PCR (*SPON2* and *MMP1*).

Relative quantification was done using Sequence Detection Software SDS 2.1, where each time point was calibrated by its mock. *HPRT* - endogenous control. NTC (no template control) showes no amplification. RQ min/max confidence 99.9%. KD – *SPON2* knockdown

III-8 SPON2 protein study

III-8.1 SPON2 overexpression and generation of a SPON2-Ab

Commercially available antibodies were not working for immunofluorescent assays (IF) and co-immunoprecipitation (co-IP) in our hands. Therefore I used the following strategies to detect SPON2 in IF and in IP: *a)* transient transfection of N-terminally FLAG-tagged SPON2 and *b)* generation of a SPON2-antibody.

The overexpression procedure was validated by co-IP and following Western blot (Figure I-31) and mass spectrometry (see Chapter III-8.3).

For the detection of endogenous SPON2 four polyclonal antibodies were generated by immunisation of rabbits with peptides C-AHSSDYSMWRKNQY-amide (76-89aa) and C-QDTVTEITSSSPSH-amide (207-220aa). Immunisation, boosting and bleedings were done by BioGenes (Berlin, Germany). Before final bleeding of rabbits the antibody titre in animal blood serum was checked by an enzyme linked immunosorbent assay (Figure I-33a) and validated in parallel by dot blot analysis (Figure I-33b). After final bleeding, the antibody was isolated from the serum by immunopurification. For that purpose antigenic peptides were linked to a column via sulfhydryl chemistry using its N-terminal cysteine residue. The resulting 20 elution fractions were analysed by dot blot and those with the highest signal were assayed by Western blot (Figure I-33c) and IF (Figure I-32).

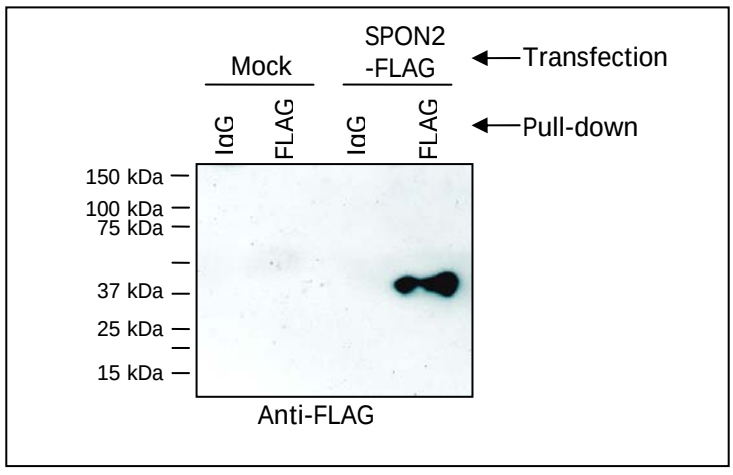


Figure I-31. Western blot of exogenous SPON2 in IMR-90 (48h after transfection) cells using anti-FLAG antibody.

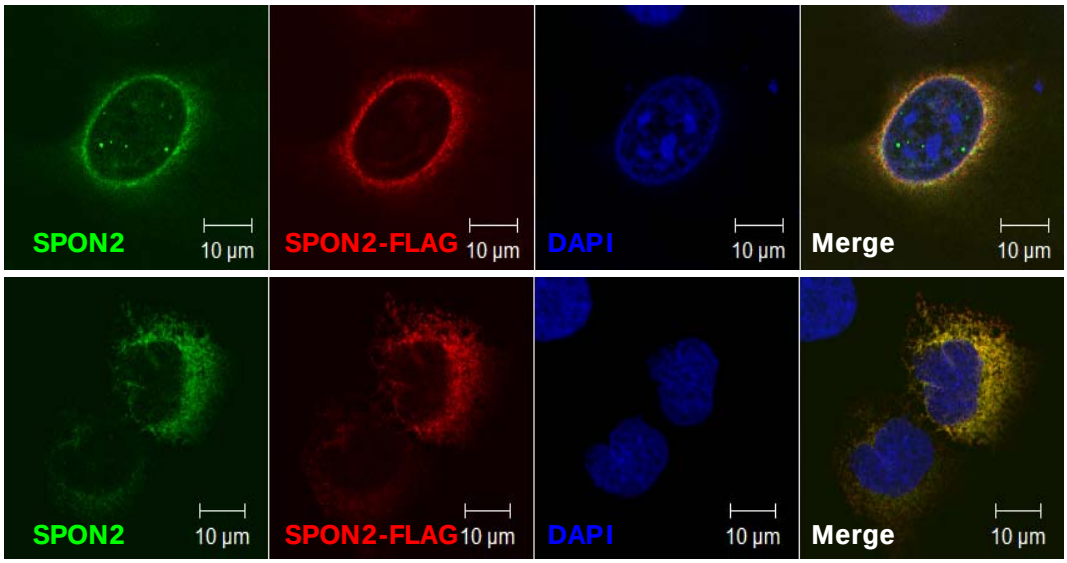


Figure I-32. Detection of SPON2 by custom-made rabbit antibody. SPON2 tagged with FLAG is transiently expressed (48h) in HaCaT cell line. The cells are immunostained by anti-FLAG antibody (red) and by custom-made SPON2 antibody (green), which shows strong co-localisation. Images captured by LSM-510 meta (Zeiss, Germany).

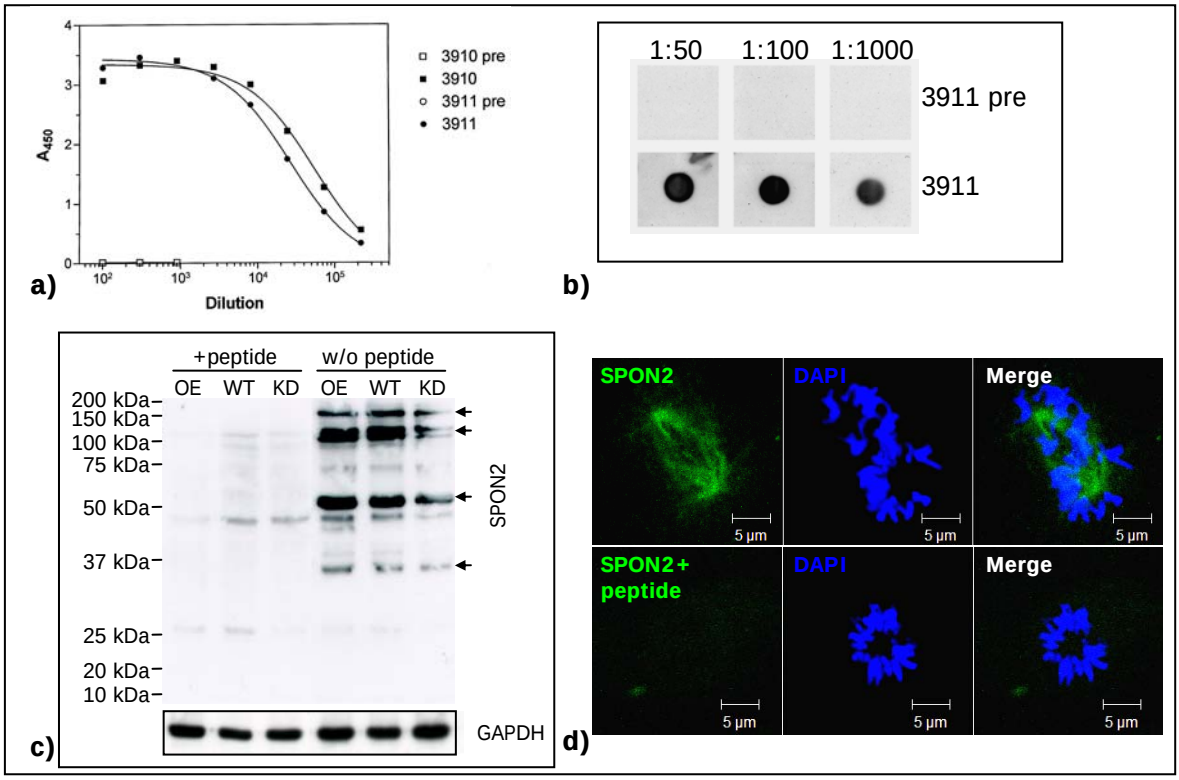


Figure I-33. Titration and validation of anti-SPON2 antibody (epitope C-QDTVTEITSSSPSH-amide).

a - ELISA, measuring absorbance at 450 nm (OD450). Shows titre difference in post-immune (3910 and 3911) and pre-immune sera (3910 pre and 3011 pre). Courtesy of BioGenes (Berlin, Germany).

b - Dot blot of the pre- (3911 pre) and post-immunization (3911) sera from immunized rabbit.

c - Western blot analysis by purified antibody on IMR-90 whole cell lysate. OE - overexpression, WT - wild type, KD - knockdown. GAPDH is a loading control.

d - IF of non-transfected IMR-90 cells, using custom SPON2 antibody (3911). Antibody, blocked with peptide, does not detect SPON2. Images captured by LSM-510 meta (Zeiss, Germany).

III-8.2 Subcellular localization of SPON2

8.2.a SPON2 is localized on nuclear envelope

In order to investigate the intracellular localization of exogenous SPON2 I performed immunostaining using anti-FLAG antibody. Confocal microscopy showed SPON2 localization on nuclear envelope (Figure I-34a). This localisation was validated by Western blot using separated cytosolic and nuclear fraction of untransfected IMR-90 cell line (Figure I-34c). A potential cross-contamination of nuclear and cytosolic fractions was excluded by incubation of the same blot with antibodies against lamin A/C (nuclear marker) and tubulin (cytosolic marker).

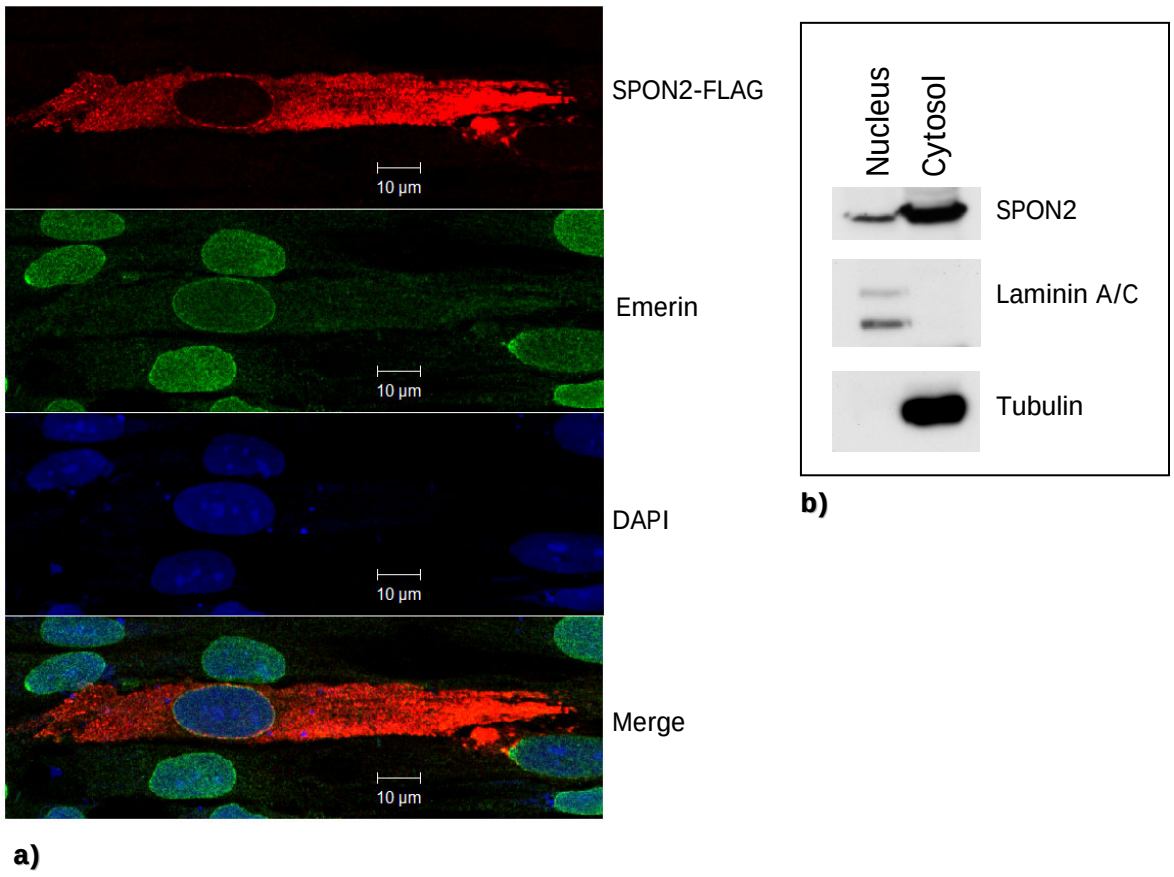


Figure I-34. Nuclear envelope localisation of exogenous SPON2 protein.

a - SPON2 expression vector tagged with FLAG (red) transiently expressed (96h) in IMR-90 cell line shows its nuclear localisation. Emerin was used as a nuclear membrane marker. Images captured by LSM-510 meta (Zeiss, Germany).

b - Nuclear localisation of SPON2 protein in IMR-90 detected by separation of nuclear and cytosolic fractions and subsequent Western blot analysis. Lamin A/C and Tubulin are nuclear and cytosolic markers, respectively. SPON2 was detected by commercially available antibody Mindin W-18 (Santa Cruz Biotechnology, USA).

8.2.b **SPON2 is localized on mitotic spindles, midbody and microtubules**

IF microscopy of HaCaT and IMR-90 cell lines, using custom antibody against endogenous SPON2, showed a clear localization on microtubules (Figure I-36), mitotic spindles and midbody (Figure I-35). Similar localisation pattern was observed using anti-FLAG antibody in overexpressed cells as well (Figure I-37).

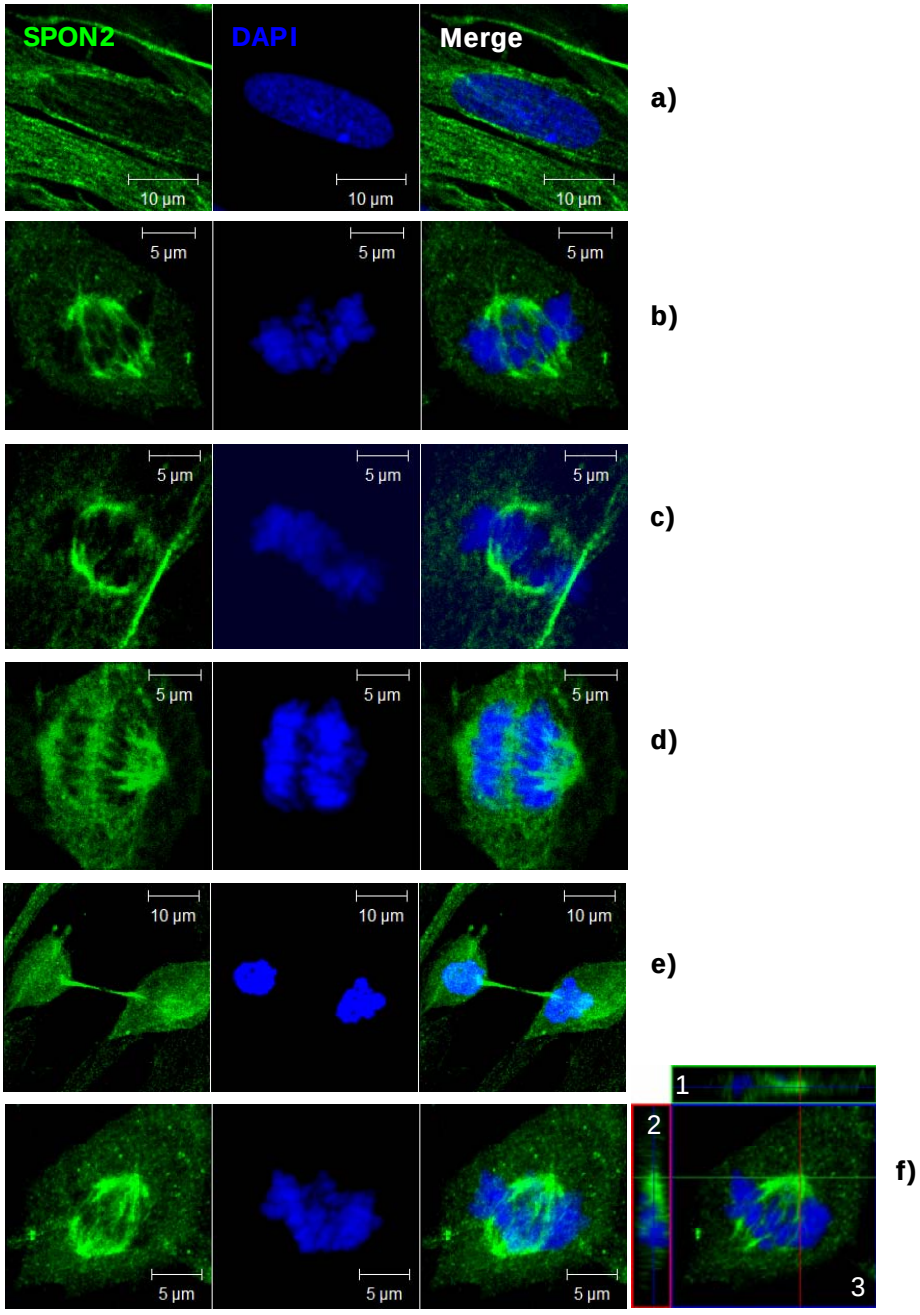


Figure I-35. SPON2 localisation on mitotic spindles in different stages of IMR-90 cell division.

(a), *(b)*, *(c)*, *(d)* and *(e)* are interphase, early metaphase, metaphase, early anaphase and cytokinesis, respectively. *(f)* is three dimensional orthogonal slice projection; x axis (green line), y axis (red line), xz slice (1), yz slice (2), blue line represents the z-stack position of the central panel (3). Images captured by LSM-510 meta (Zeiss, Germany).

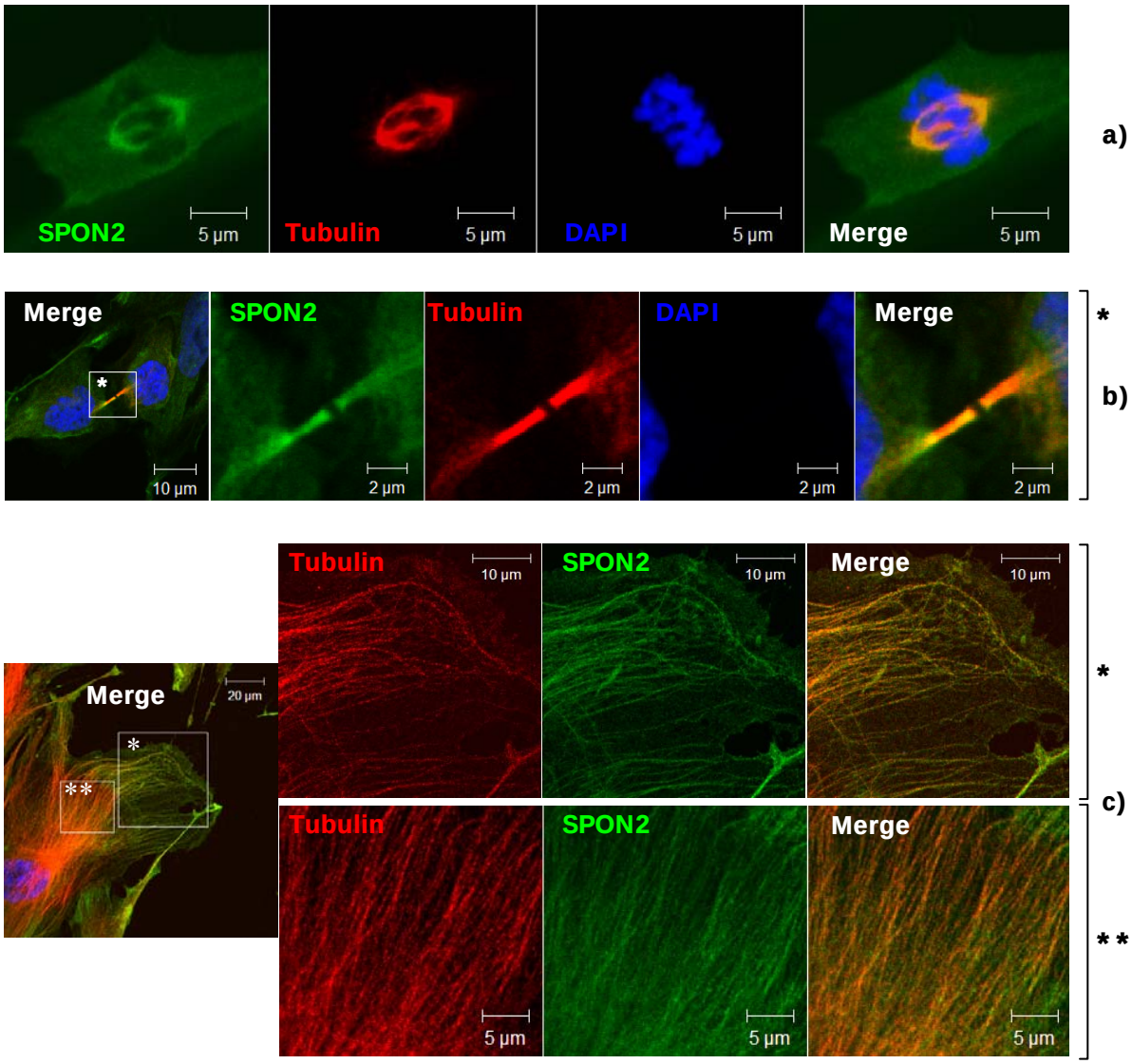


Figure I-36. SPON2 co-localisation with tubulin on mitotic spindles *(a)*, midbody *(b)* and cytoplasmatic microtubules (c) in IMR-90 fibroblasts. * and ** are zoomed areas. *(d)* – localisation of exogenous SPON2-FLAG (96h after transfection). Detected by anti-FLAG antibody. Images captured by LSM-510 meta (Zeiss, Germany).

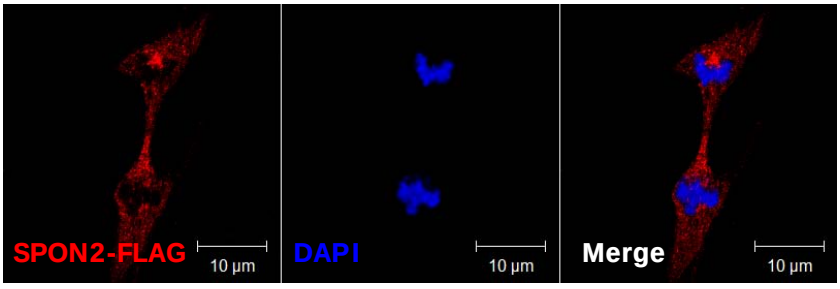


Figure I-37. Localisation of exogenous SPON2-FLAG in IMR-90 cells (96h after transfection). Detected by anti-FLAG antibody. Images captured by LSM-510 meta (Zeiss, Germany).

8.2.c **SPON2 is possibly localized on adhesion contacts**

Previous studies by Feinstein *et al.* [118] describe SPON2 as a protein promoting adhesion of hippocampal neurons. IF microscopy of F11 cell line using antibody against endogenous SPON2 revealed its localisation on growing cones and filopodias (Figure I-39a and b). The F11 cell line is a fusion product of mouse neuroblastoma cell line N18TG-2 with embryonic rat dorsal-root ganglion (DRG) neurons [119]. The IF microscopy of SPON2 overexpressing HaCaT cells showed some cells having SPON2 at cell protrusions (Figure I-38a). Overexpressed SPON2 localisation at putative cell-matrix adhesion sites was also detected using antibody against endogenous protein (Figure I-38b).

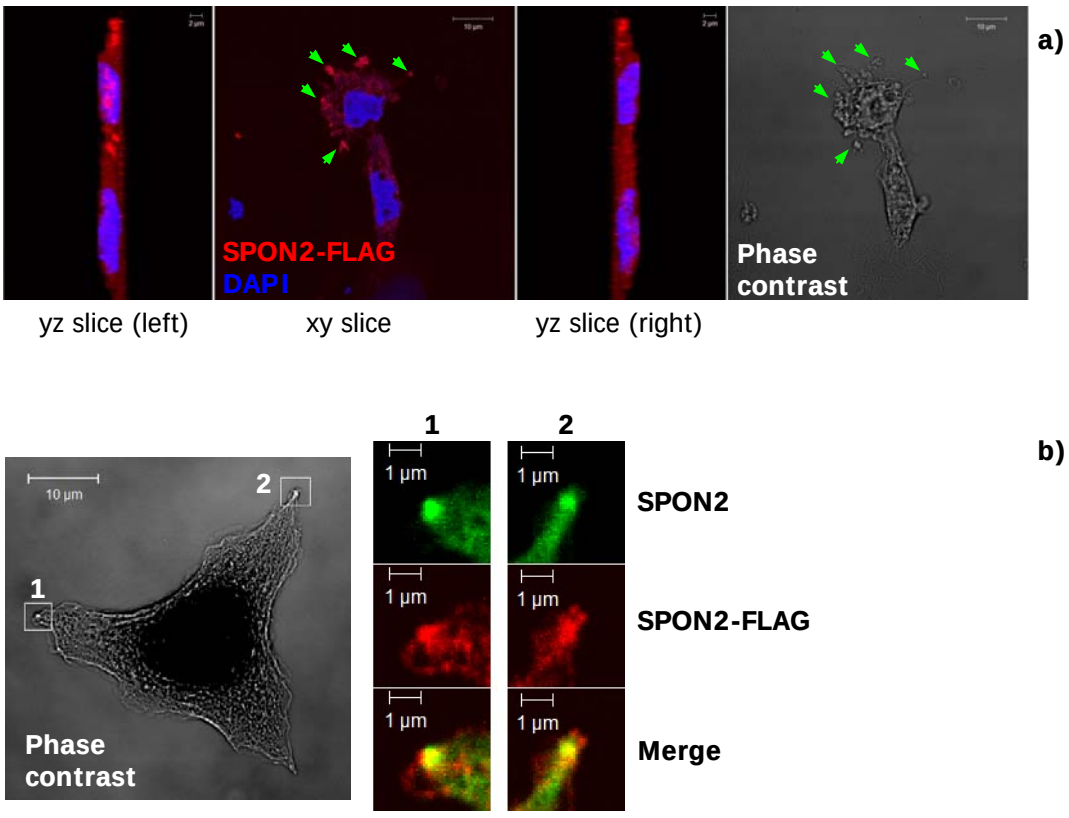


Figure I-38. Confocal microscopy of HaCaT cells.

a - Three dimensional projection of a migrating HaCaT cell, transiently overexpressed with SPON2 (120h). The green arrows show possible cell-matrix adhesion contacts. yz slices are generated by projection of 9 stacks; stack size - 4.8µm, stack scaling - 0.6µm.

b - HaCaT cell, transiently overexpressed with SPON2-FLAG (48h). 1 and 2 are zoomed areas. Images are captured by LSM-510 meta (Zeiss, Germany).

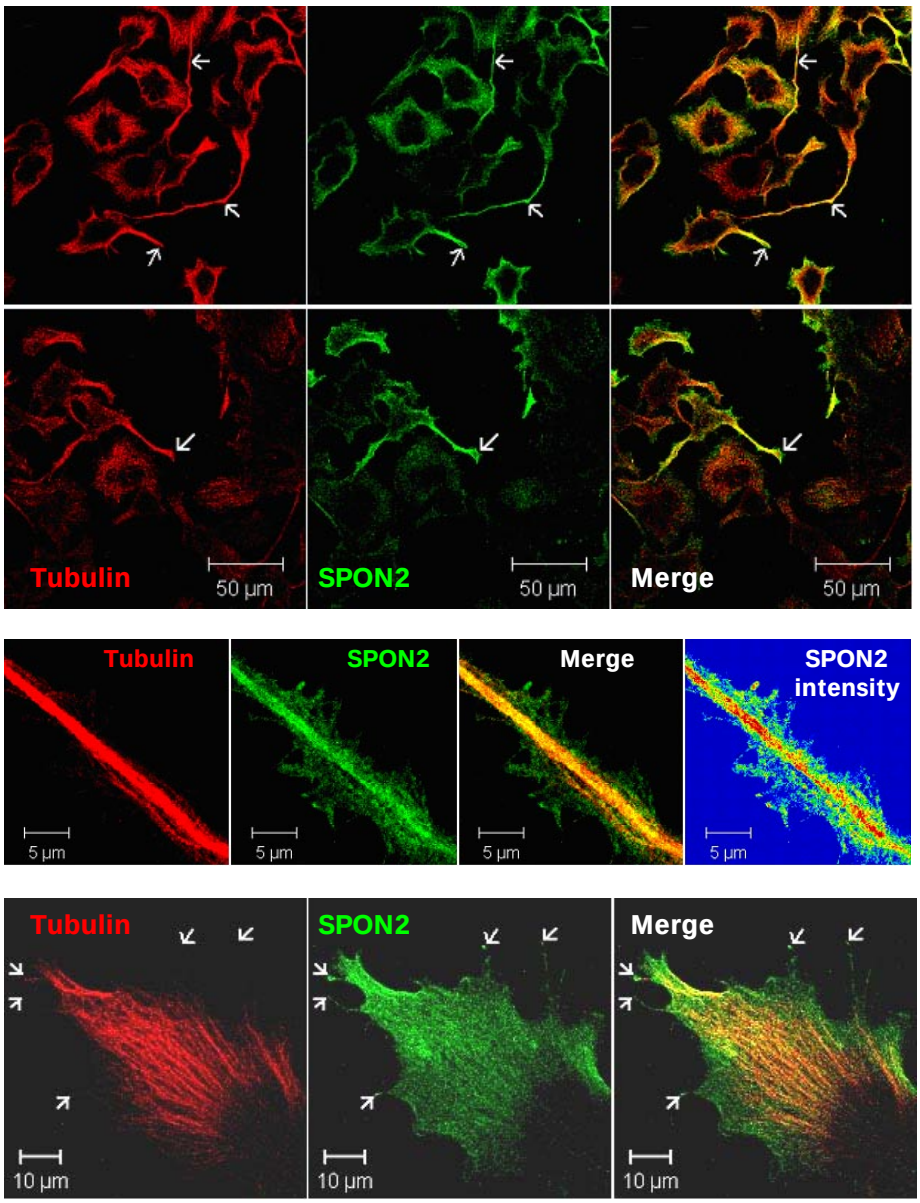


Figure I-39. Confocal microscopy of F11 cells (peripheral neuronal cells).

a - SPON2 is associated with neuronal growth cones (white arrows).

b - SPON2 localisation on filopodias. SPON2 intensity plot is obtained by rainbow palette (LSM Image Analyser). Images are captured by LSM-510 meta (Zeiss, Germany).

c - SPON2 is associated with neuronal growth cones (white arrows).

III-8.3 SPON2 co-IP and Mass Spectrometry

To identify SPON2-interacting proteins we performed co-IP using overexpressed SPON2 tagged by FLAG-tag. The co-immunoprecipitated proteins were separated on SDS-gels, visualized with Coomassie or silver staining, and identified by mass spectrometry. Three different strategies were done using: *a)* analysis of single pull-down specific bands from the Coomassie stained gels (Figure I-40a); *b)* analysis of the whole lane from Coomassie stained gel (Figure I-40b), which was done in two biological replicas; and *c)* analysis of the whole co-immunoprecipitate without separation on a SDS-gel.

The whole lane and whole co-immunoprecipitate mass spectrometry analysis were done by comparing SPON2-FLAG co-IP with the following controls:

- co-IP with IgG of SPON2-FLAG overexpressed IMR-90 cell lysate;
- co-IP with IgG of mock IMR-90 cell lysate;
- co-IP with anti-FLAG antibody of mock IMR-90 cell lysate.

Mass spectrometry was done in the Mass Spectrometry/Proteomics core facility, Centre of Medical Research, Medical University, Graz, Austria. A list of candidate proteins, recovered by mass spectrometry and their description are shown in Table I-49 and Table I-51, respectively. GO term enrichment, using DAVID annotation tool, is presented in Table I-50.

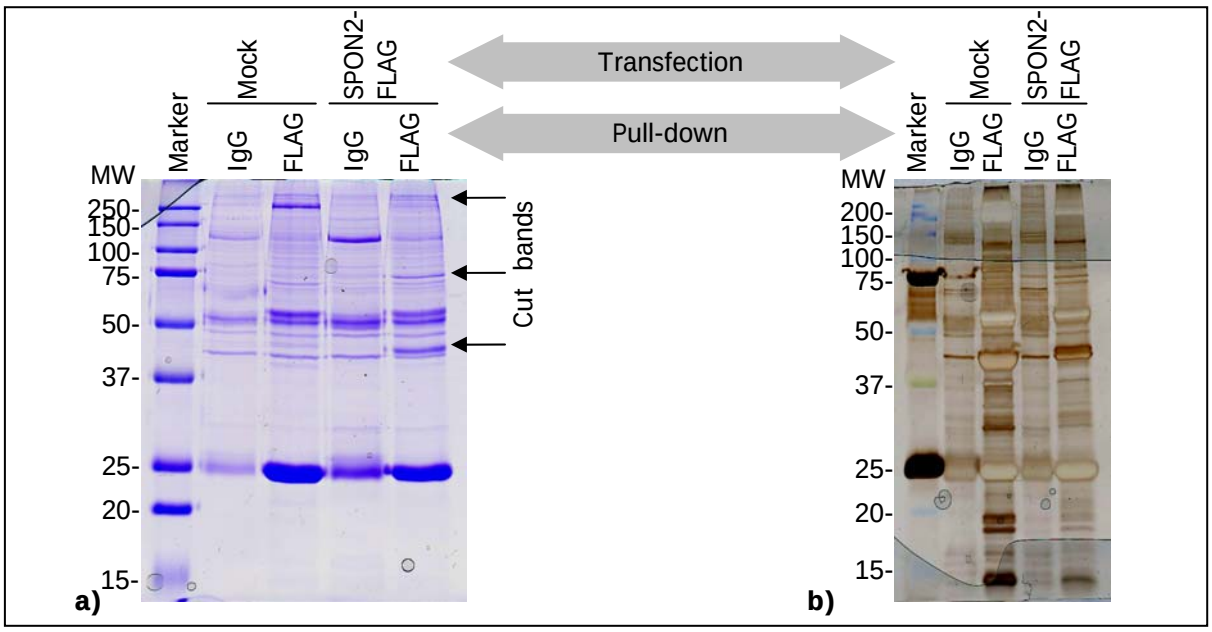


Figure I-40. Coomassie and Silver staining of co-IP-ed by SPON2-FLAG protein. MW (left margins), molecular weight, in kDa.

Table I-49. List of candidate SPON2 interaction proteins, revealed by mass spectrometry. Proteins, detected in two and more experiments are highlighted by blue colour.

Cut bands (n=3)	**Whole lane 1**	**Whole lane 2**	**Whole IP**
ACTA2/-1, ACTB, ACTG1/-2	ACTA2/ -1, ACTB, ACTG1/ -2	CALD1	HSPA5
HSPA5	CALD1	GSN	TPR
TPR	GSN	HSPA5	KIF11
CAPN2	HSPA5	TMOD2/ -3	PRMT5
COL6A3	TMOD2/ -3	TPM1	
DDX17/-4	TPM1	LIMA1	
DYNC1H1	TPM4		
EEF1A1	CAPZB		
EIF4A2			
FLNA			
HNRPM			
MIG2			
PFKL			
PLTN			
Tubulin alpha			
VIM			
DES			

Table I-50. GO annotation enrichment of candidate proteins, co-IP-ed with SPON2

GO Biological Process	**Benjamini-Hochberg P-value**	**P-Value**	**Proteins**
cytoskeleton organization and biogenesis (GO:0007010)	3.49E-10	6.64E-14	PLEC1, ACTA1, TUBB, FLNA, ACTB, TUBA1A, VIM, TUBB4, GSN, CAPZB, KIF11, TUBA4A, DYNC1H1, DES, LIMA1
organelle organization and biogenesis (GO:0006996)	7.95E-07	3.03E-10	PLEC1, ACTA1, TUBB, FLNA, PRMT5, ACTB, TUBA1A, VIM, TUBB4, GSN, KIF11, CAPZB, TUBA4A, DYNC1H1, DES, LIMA1
cellular component organization and biogenesis (GO:0016043)	8.26E-04	4.72E-07	PLEC1, ACTA1, TUBB, FLNA, PRMT5, EIF4A2, ACTB, TUBA1A, VIM, TUBB4, GSN, KIF11, CAPZB, TPR, TUBA4A, DYNC1H1, DES, LIMA1
microtubule-based movement (GO:0007018)	2.13E-03	1.62E-06	KIF11, TUBA4A, DYNC1H1, TUBB, TUBA1A, TUBB4
cytoskeleton-dependent intracellular transport (GO:0030705)	3.68E-03	3.51E-06	KIF11, TUBA4A, DYNC1H1, TUBB, TUBA1A, TUBB4
protein polymerization (GO:0051258)	5.14E-03	5.89E-06	TUBA4A, TUBB, TUBA1A, TUBB4, GSN
cell motility (GO:0006928)	5.95E-03	7.95E-06	CALD1, CAPZB, TPM4, TPM1, TUBB, FLNA, VIM, ACTB
localization of cell (GO:0051674)	5.95E-03	7.95E-06	CALD1, CAPZB, TPM4, TPM1, TUBB, FLNA, VIM, ACTB
actin cytoskeleton organization and biogenesis (GO:0030036)	2.05E-02	3.54E-05	CAPZB, ACTA1, FLNA, ACTB, LIMA1, GSN
microtubule-based process (GO:0007017)	2.25E-02	4.33E-05	KIF11, TUBA4A, DYNC1H1, TUBB, TUBA1A, TUBB4
actin filament-based process (GO:0030029)	2.32E-02	4.92E-05	CAPZB, ACTA1, FLNA, ACTB, LIMA1, GSN

Table I-51. Description of candidate proteins, co-IP-ed with SPON2.

Protein	Description
Actin	Major cytoskeleton protein
Caldesmon-1	Calmodulin- and actin-binding protein that plays an essential role in the regulation of muscle contraction. The conserved domain of this protein possesses binding capacity to Ca(2+)-calmodulin, actin, tropomyosin, myosin, and phospholipids.
Calpain-2	Calcium-regulated non-lysosomal thiol-protease which catalyze limited proteolysis of substrates involved in cytoskeleton remodelling and signal transduction
Cytoplasmic dynein 1 heavy chain 1	Acts as a motor for the intracellular retrograde motility of vesicles and organelles along microtubules.
Desmin	Class-III intermediate filaments found in muscle cells. Involved in cytoskeleton organisation [120] and muscle contraction [121]
Elongation factor 1-alpha 1	Found in a nuclear export complex. [122, 123]
F-actin-capping protein subunit beta	Binds in a Ca(2+)-independent manner to the fast growing ends of actin filaments (barbed end) thereby blocking the exchange of subunits at these ends. Unlike other capping proteins (such as gelsolin and severin), these proteins do not sever actin filaments
Filamin A	Actin-binding protein, crosslinks actin filaments into orthogonal networks in cortical cytoplasm and participates in the anchoring of membrane proteins for the actin cytoskeleton. Remodelling of the cytoskeleton is central to the modulation of cell shape and migration. Filamins are involved in EMT [124, 125].
Gelsolin	Calcium-regulated, actin-modulating protein that binds to the plus (or barbed) ends of actin monomers or filaments, preventing monomer exchange (end-blocking or capping). It can promote the assembly of monomers into filaments (nucleation) as well as sever filaments already formed
Heat shock 70 kDa protein 5	Binds to a many proteins. Probably plays a role in facilitating the assembly of multimeric protein complexes inside the ER [126]
Kindlin-2	Participates in the connection between ECM adhesion sites and the actin cytoskeleton and also in the orchestration of actin assembly and cell shape modulation [127]
Kinesin-like protein KIF11	Motor protein required for establishing a bipolar spindle. Blocking of KIF11 prevents centrosome migration and arrest cells in mitosis with monoastral microtubule arrays [128]
LIM domain and actin-binding protein 1	Binds to actin monomers and filaments. Increases the number and size of actin stress fibres and inhibits membrane ruffling. Inhibits actin filament depolymerisation. Bundles actin filaments, delays filament nucleation and reduces formation of branched filaments [129]
Nucleoprotein TPR	Component of the cytoplasmic fibrils of the nuclear pore complex implicated in nuclear protein import [130]
Plectin 1	Actin binding [131]
Shk1 kinase-binding protein 1	Regulates mitosis [132]
Tropomodulin-2 and -3	Involved in EMT [133]. Blocks the elongation and depolymerisation of the actin filaments at the pointed end. Binds tropomyosin [134].
Tropomyosin-1 and -4	Binds to actin filaments. Plays a central role in muscle contraction. Smooth muscle contraction is regulated by interaction with caldesmon. In non-muscle cells is implicated in stabilizing cytoskeleton actin filaments [135, 136]. Tropomyosins are involved in EMT [124]
Tubulin alpha	Major cytoskeleton protein
Vimentin	Major cytoskeleton protein

III-8.4 Meta-analysis

Aimed to understand the intrinsic interplay of obtained by *SPON2* knockdown deregulated genes, as well as to get a rationale overview of these data, I addressed them to interactome visualization using Cytoscape. Based on deregulated by *SPON2* knockdown genes, I retrieved protein-protein interactions from Pathway Commons (www.patwaycommons.org), restricting search results to HPRD database (release 01-Sep-07). Figure I-41 displays a network, built up with genes deregulated by *SPON2* knockdown (4-fold change, $P<0.05$, 120h after oligofection). The majority of proteins were integrated into a single network, interconnected either directly or via intermediate proteins (grey nodes). It is worth mentioning that both genes, up- and downregulated by *SPON2* knockdown, are evenly distributed within the network and do not form expression-based clusters. This suggests that GO enrichment analysis does not necessitate a separation of up- and downregulated genes.

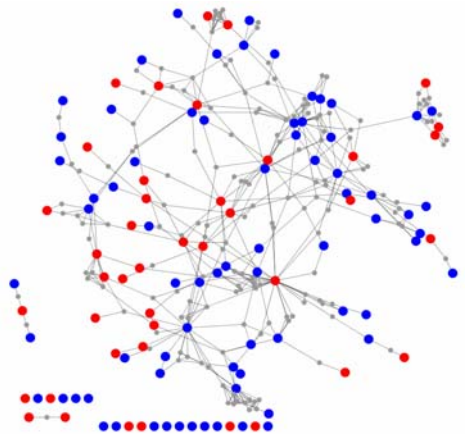


Figure I-41. Interaction network of deregulated by *SPON2* knockdown genes.

The blue and red nodes represent upregulated genes in *SPON2* knockdown and mock genes, respectfully. A query encompassing 245 genes returned interaction partners for 108 of them (HPRD, release 01-Sep-07). Analysis and visualisation by Cytoscape 2.6.0, using Pathway Commons Web Service Client and HPRD database.

The same kind of analysis was done also with proteins that were revealed in all co-IP experiments (Figure I-42a). As a control I built up a network using a random protein list (Figure I-42b) with the same parameters and query size. In contrast to the SPON2 co-IP interactome, the control network failed to build up a complex network and was represented by multiple discrete clusters (comparison of their network formation capacity is shown on Figure I-42c).

Merging the interaction networks, which is built up by *SPON2* knockdown (Figure I-43a) and SPON2 co-IP (Figure I-43b), revealed seven overlapped nodes (6.5% of deregulated by

knockdown genes and 20.0% of co-IP-ed proteins), coloured in green. The merged network did not show these seven genes linking two separate structures. Indeed, co-IP-ed proteins were completely (excluding CAPN2 and CAPZB) integrated into the network of deregulated by *SPON2* knockdown genes without cluster outlining (Figure I-43c).

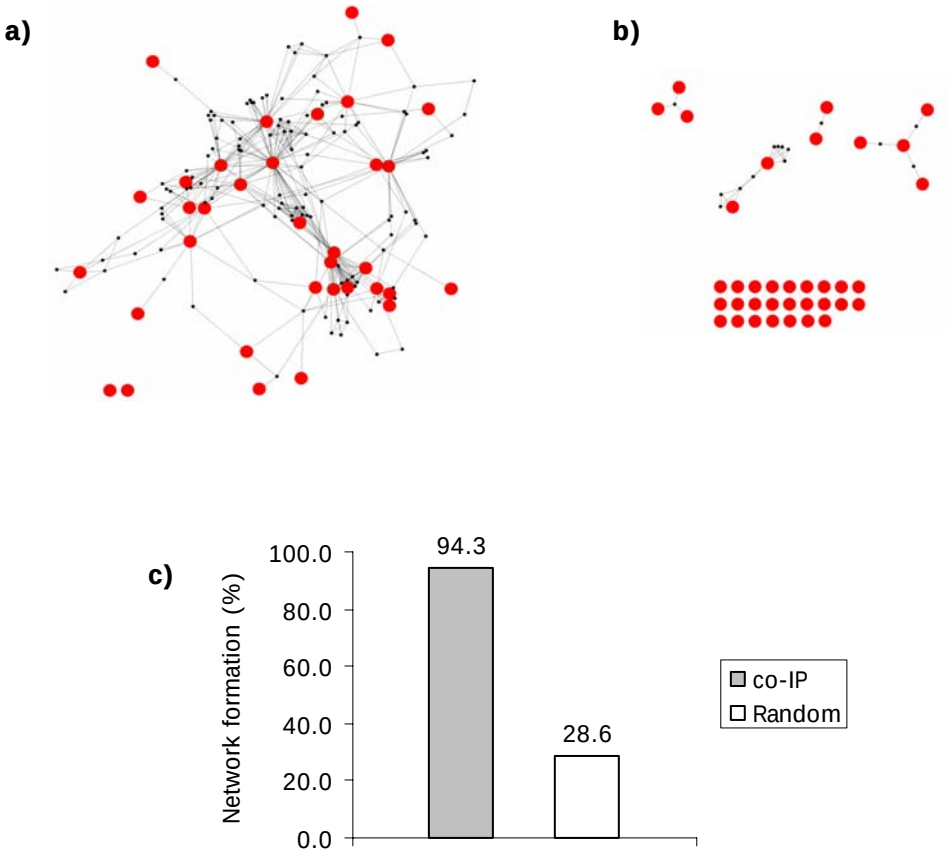


Figure I-42. Interaction networks of co-IP-ed by SPON2 proteins *(a)* and a random list *(b)*.

Interaction networks of co-IP-ed by SPON2 proteins and a random list are on supplementary CD (Network1.cys and Network2.cys, respectively). Both networks were analysed in the same way by Cytoscape 2.6.0 using Pathway Commons Web Service Client and HPRD database (release 01-Sep-07). Layout by Cytoscape Spring Embedded layout. *(c)* - Percentage of implicated in networks *(a)* and *(b)* nodes.

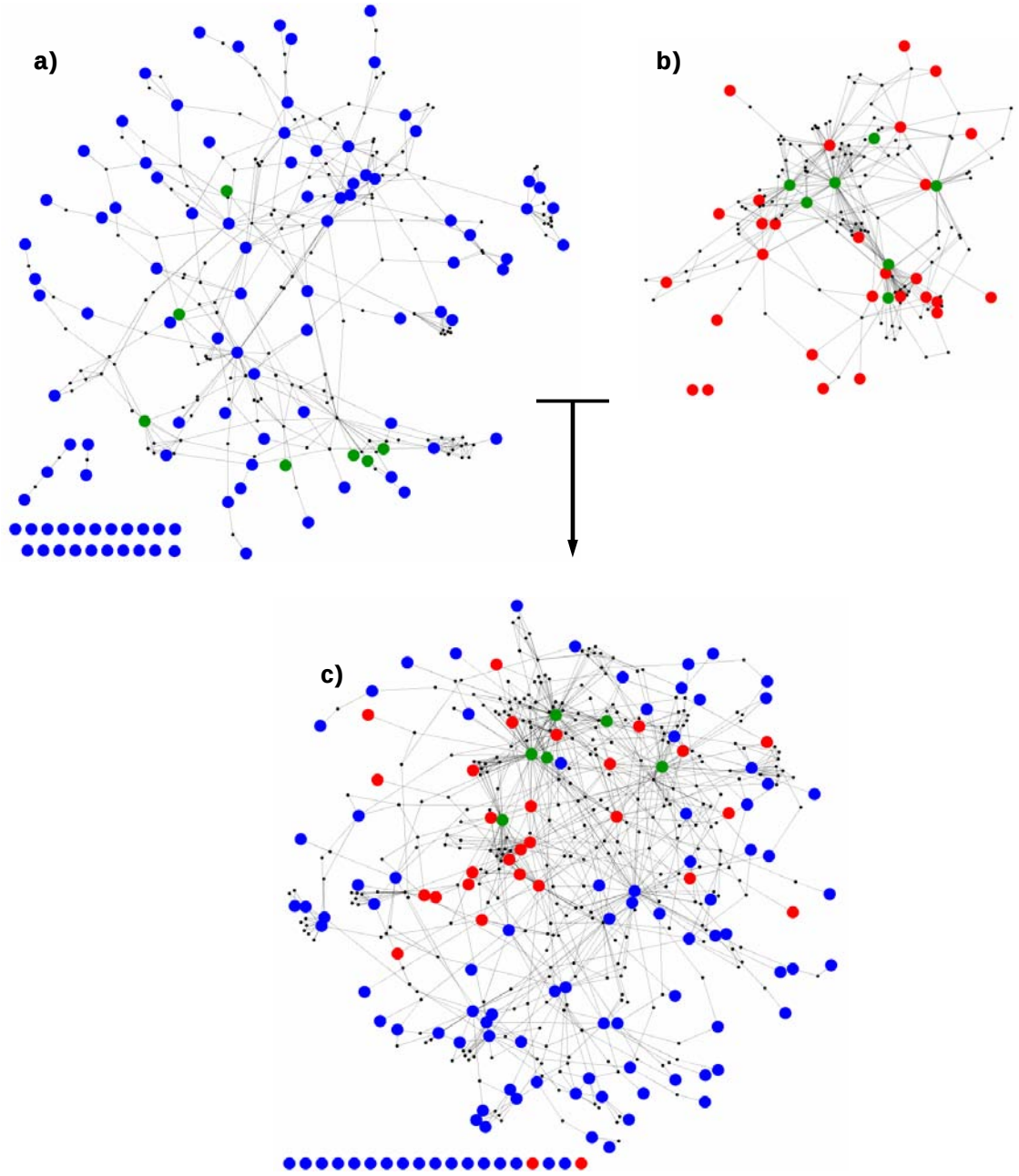


Figure I-43. HPRD based interaction visualization network (supplementary CD, Network1.cys).

a - Interaction network of deregulated by *SPON2* knockdown genes (120h, 4-fold, P<0.05)

b - Interaction network of SPON2 co-IP-ed proteins, listed in Table I-49.

c - Network generated by merging networks of genes deregulated by *SPON2* knockdown *(a)* and SPON2 interaction partners as indicated by co-IP *(b)*.

Analysed by Cytoscape 2.6.0, using Pathway Commons Web Service Client and HPRD database (release 01-Sep-07). Layout by Cytoscape Spring Embedded algorithm. Blue nodes – genes deregulated by knockdown, red nodes – co-IP-ed proteins, green nodes – genes and proteins highlighted by both experiments, small black nodes – intermediate interaction partners, lines – interactions.

Discussion

IV DISCUSSION

The tumour microenvironment has a great impact on cancer progression. In light of *in vitro* data, demonstrating that aberrant stroma can promote tumour growth, a number of researchers set out to investigate the presence of mutations in the stroma, which in part were also overlapping with those identified in the tumour [33, 112-114]. This prompted us to analyse DNA copy number changes in a series of squamous cell carcinoma (SQCC) and adenocarcinoma (AC) of lung, tumour associated stroma and neighbouring epithelial cells. I have detected chromosomal aberrations in the stroma in 23.5% of cases (Figure I-16). All of them were also present as high copy amplification in the tumour. Taking into account that all samples were microdissected by an experienced pathologist, who could not identify any cancer cell in tumour associated stroma and adjacent epithelium, I suspected presence of malignant cells in the stroma, which was concealed by fibroblast-like looking morphology. Thus, mimicking fibroblasts cancer cells raised our interest for epithelial-mesenchymal transition (EMT).

SCLC xenograft in nude mouse gave growth advantage to SCLC-A

The SCLC cell line was transplanted into a nude mouse with the aim to study plasticity of adhesive and suspensive cells. When these cells were recovered from the host and cultured *in vitro*, a dramatic preponderance of adhesive fibroblast-like cells was observed. Thus, transplantation revealed that SCLC-A has more plasticity than its suspensive counterpart. In comparison with pre-transplant SCLC cell line, array-CGH study of post-transplant cells revealed only an additional loss of chr3:71418986-71741160 region, involving *FOXP1* gene (Figure I-19). Appearance of *FOXP1* loss could be explained either as a *de novo* event or by selection process. Taking into account that cancers can be composed of heterogeneous subset of different clones, it is reasonable to assume that the *FOXP1* loss was already present before transplantation, and that this specific clone was adapted to the new microenvironment.

FOXP1 belongs to forkhead box (FOX) transcription factor family. This gene is reported as transcriptional suppressor with substantial impact on lung epithelium differentiation [137]. In cancers *FOXP1* can act as both, an oncogene [138] or a tumour suppressor gene [139].

SCLC A has mesenchymal pattern and may serve as an EMT model

SCLC are highly invasive cancers with occurrence of metastases in broad spectrum of tissues and organs [140]. The cell line that I used in this study exhibits simultaneously suspensive (SCLC-S) and adhesive (SCLC-A) growth. Separation of SCLC-S from –A and multiple passaging showed that the adhesive cells were continuously generated from its suspensive counterparts. Additionally, after separation adhesive cells turned to suspensive ones. This process, mimicking reverse of EMT, is known as mesenchymal-epithelial transition (MET). The possible contamination of SCLC cell line with fibroblasts was excluded array-CGH analysis of suspensive and adhesive SCLC cells, which revealed identical patterns of DNA copy number changes.

Whereas DNA copy number and methylation study of SCLC-A and –S showed no relevant differences, the gene expression study revealed an intriguing list of differentially expressed genes. After analysing these genes by GO "Biological Process" annotation enrichment ($P<0.01$), SCLC-A showed significant enrichment for terms such as development, cell adhesion, stem cell maintenance, ECM organisation, ECM biogenesis etc. (Table I-44). In contrast to SCLC-A, the overrepresented in SCLC-S genes showed low enrichment of irrelevant for our study GO terms. Generic GOA and Proteome Slim [104] based annotation enrichment of genes upregulated in SCLC-A (1.5-fold, $P<0.05$) significantly revealed cell motility term ($P=4.14E-03$). Such features, as cell motility and some properties of stem cells are distinct characteristics of EMT process [141].

Top 20 annotated genes ($P<0.05$), overrepresented in SCLC-A and –S, are listed in Table I-52 and Table I-53.

Table I-52. Top 10 annotated genes upregulated in SCLC-A

Genbank	Description	Fold Change
NM_002421	matrix metalloproteinase 1 (interstitial collagenase) (MMP1)	29,21
NM_012445	spondin 2, extracellular matrix protein (SPON2)	6,844
NM_001553	insulin-like growth factor binding protein 7 (IGFBP7)	4,836
NM_004370	collagen, type XII, alpha 1 (COL12A1), transcript variant long	4,691
NM_003118	secreted protein, acidic, cysteine-rich (osteonectin) (SPARC)	3,730
NM_000093	collagen, type V, alpha 1 (COL5A1)	3,405
NM_001849	collagen, type VI, alpha 2 (COL6A2), transcript variant 2C2	3,241
NM_053056	cyclin D1 (PRAD1: parathyroid adenomatosis 1) (CCND1)	3,096
NM_002775	protease, serine, 11 (IGF binding) (PRSS11)	2,798
NM_001233	caveolin 2 (CAV2), transcript variant 1	2,610

Table I-53. Top 10 annotated genes upregulated in SCLC-S

Genbank	Description	Fold Change
NM_005315	goosecoid-like (GSCL)	0,175
NM_002962	S100 calcium binding protein A5 (S100A5)	0,278
NM_145727	lipoprotein, Lp(a)-like 2 (LPAL2), transcript variant 2	0,306
NM_024318	leukocyte immunoglobulin-like receptor, member 6 (LILRA6)	0,313
NM_022901	leucine rich repeat containing 19 (LRRC19)	0,330
NM_016382	CD244 natural killer cell receptor 2B4 (CD244)	0,351
NM_001147	angiopoietin 2 (ANGPT2)	0,398
BC001868	zinc finger protein 44 (KOX 7)	0,406
NM_004633	interleukin 1 receptor, type II (IL1R2), transcript variant 1	0,406
NM_012147	double homeobox, 2 (DUX2)	0,414

Expression analysis of SCLC cell line depicted *SPON2* and *MMP1* as the most differentially regulated genes, which will be discussed in next chapters. Next gene – *IGFBP*, is known as tumour-derived adhesion factor and involved in cell adhesion [142, 143]. Another interesting gene is *SPARC* (osteonectin), coding a matrix-associated protein, is involved in migration by promoting cell motility [144]. *SPARC* is known as EMT-induced marker and found to be upregulated in invasive breast carcinomas [88]. *MMP1* (matrix metalloprotease-1) and *PRSS11* (serine protease, also known as *HTRA1*) genes encode proteases, degrading and/or remodelling the ECM, and paving a way for cell migration and expansion [54]. Inhibition of MMPs and serine proteases in different epithelial cancer models was repeatedly reported impairing migration of cancer cells *in vitro* [145-147]. *CAV2* (caveolin-2), a gene overrepresented in SCLC-A, is a major component of the inner surface of caveolae. Caveolae are small invaginations of the plasma membrane, playing an essential role in different cellular processes, including cell migration and cargo transport [148].

Additionally, I compared differentially expressed genes in SCLC cell line with data available in Gene Expression Omnibus (GEO). Interesting results were obtained, when top 20 differentially expressed annotated genes (Table I-52 and Table I-53) were compared with expression profiles of mesenchymal and epithelial compartments of the developing intestine (GDS 2699)[149]. Surprisingly, all the genes upregulated in SCLC-A were highly expressed in mesenchyme, whereas those of SCLC-S were expressed in epithelium (two genes) and the rest eight genes were either not present or showed no distinct epithelial or mesenchymal signature (ungrouped) (Figure I-44).

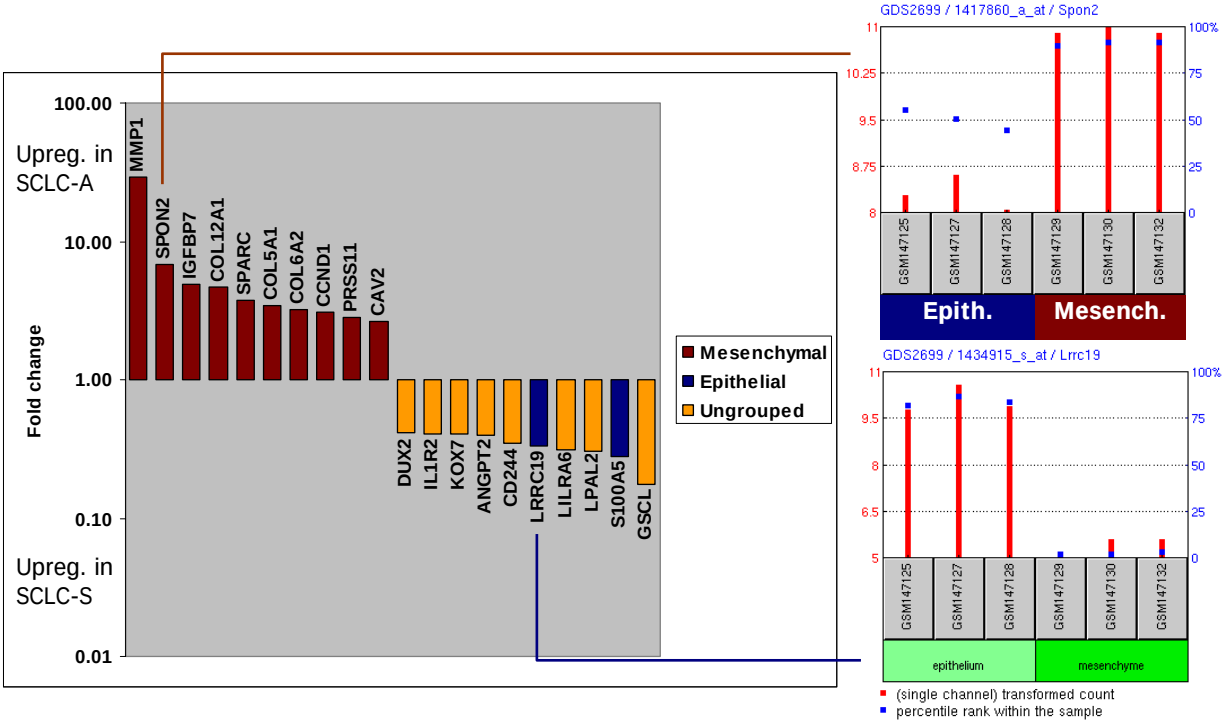


Figure I-44. Comparison of SCLC expression (top 10 annotated genes) with GDS2699 (GEO).

The left plot shows expression level of differentially expressed genes in SCLC-A and –S (Table I-52 and Table I-53, respectively). The bars are coloured by theirs expression level in GDS2699 (right plots). The right plots are expression profiles of *SPON2* and *LRRC19* genes obtained from GDS2699.

Another interesting dataset from GEO, involving expression study of foreskin, embryonic fibroblast like, undifferentiated H9 and embryonic carcinoma cells (GDS2423[150], supplementary figure S-1), highlighted all upregulated in SCLC-A genes as highly expressed in foreskin and embryonic fibroblast like, while upregulated in SCLC-S nine genes showed no expression and one was not present at all.

Collectively, abovementioned facts suggest our hypotheses that SCLC-A has mesenchymal pattern and that these adhesive cells originate from SCLC-S most likely via EMT process.

Characterisation of SPON2

SPON2 is the second most upregulated gene in SCLC-A. The gene is localized on chr4:1150663-1156427, and encodes a 331aa protein with a calculated theoretical weight of 35.8kDa. It is known as a highly conserved extracellular matrix protein [151, 152], which is abundantly expressed in the spleen, lymph nodes [153], peripheral nerves, prostate, uterus and cervix (GeneHub-GEPIS). SPON2 (or Mindin) is an axon guidance protein, promoting outgrowth and adhesion of embryonic hippocampal neurons [118]. SPON2 functions also as pattern recognition molecule for microbial pathogens [153]. It was found as an integrin ligand, which primes T-cells and recruits inflammatory cells [151, 154]. The integrin binding site is localized within the Spondin_N domain [152] (Figure I-45 and Figure I-46). Despite its expression in a number of different cell types, surprisingly, the phenotype of Spon2$^{-/-}$ mice is impaired to immunological problems. Bacterial and influenza clearance and response to a broad range of microbial pathogens have been found impaired. Moreover, problems with the recruitment of macrophages and mast cells to the sites of inflammation have been reported [153, 155]. Jia and colleagues demonstrated *in vitro* that direct binding of neutrophils to immobilized mindin promotes migration [151]. Spon2 also functions as an opsonin for macrophage phagocytosis of bacteria, as a trigger for initiation of innate immune response [153]. Additionally, expression of Rac1 and Rac2 is regulated by mindin via integrins on dendritic cells [154]. Rac1 actively collaborates with Rho guanosine triphosphatase (Rho GTPase) and by this orchestrates EMT [156, 157].

UniProtKB/Swiss-Prot database lists SPON2 as secreted protein localised in the extracellular space and extracellular matrix. The same localisation of SPON2 was also described in the literature [117, 151, 153, 158, 159]. Importantly, by immunofluorescent microscopy (IF) and Western blotting I could show its localization on nuclear envelope and mitotic spindles.

SPON2 consists of two domains: unique N-terminal Spondin_N domain, which has no detectable sequence homology to any known human protein, and one C-terminal TSP1 (thrombospondin type-1) domain, which was identified in over 40 human proteins [117] (Figure I-45). Additionally, based on amino acid sequence, BaCelLo (Balanced Cell Localization predictor [160]) *in silico* prediction tool revealed SPON2 nuclear localisation. The N-terminal part (1-26aa) of SPON2 has high hydrophobicity (Figure I-45). Additionally, Li and colleagues determined the Spondin_N domain structure at resolution of 1.8-Å, revealing eight-stranded antiparallel beta-sandwich motif resembling to membrane-targeting C2

domains (Figure I-46a and b) [152]. Li and colleagues addressed SPON2 structure to a homology search tool, which, in spite of absence of sequence homology, found high structural similarity with C2 domain (Figure I-46c). The top score in homology analysis was identified by C2 domain of phosphoinositide 3-kinase (PI3K). PI3K is a well know regulator of EMT [161-163].

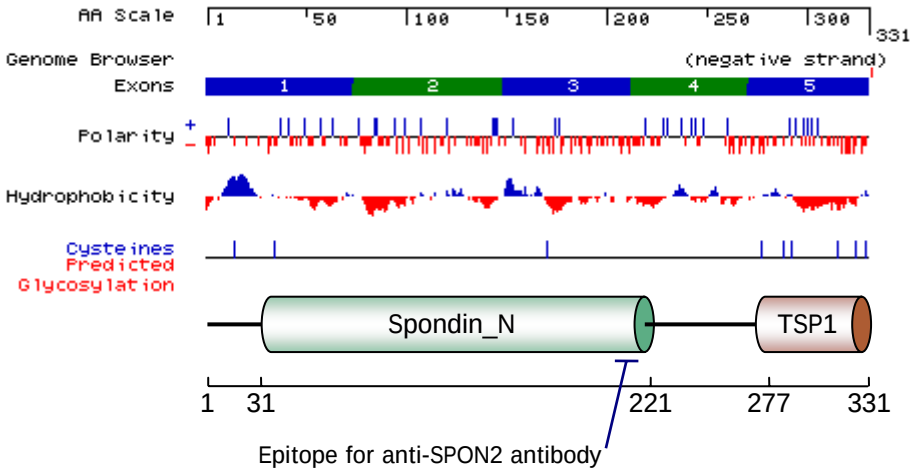


Figure I-45. Structure and features of SPON2 protein (adapted from UCSC Proteome Browser, [164]).

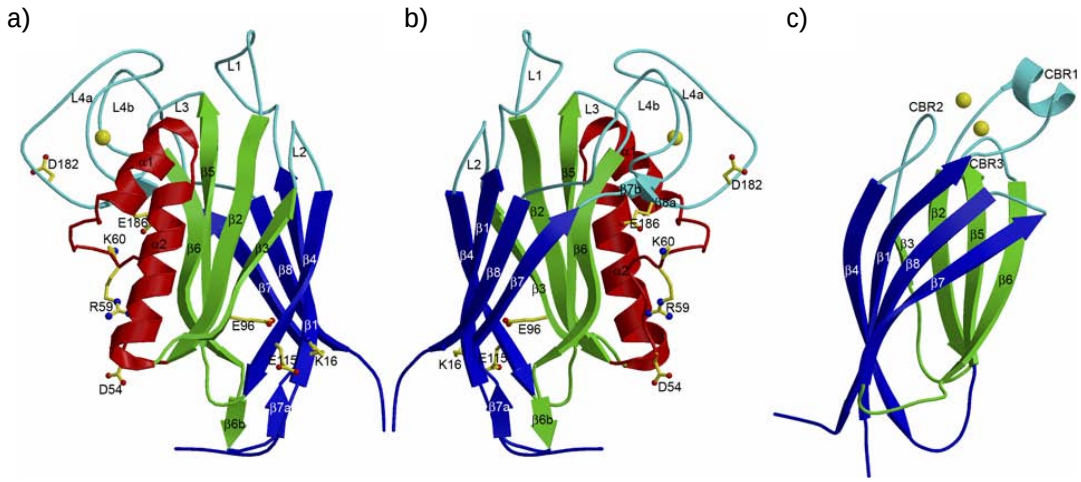


Figure I-46. Structure of the Spondin-N (or mindin-FS) domain of SPON2. Image from [152].

a - Side view of mindin-FS. The two four-stranded β-sheets, β6 β5 β2 β3 (front) and β4 β1 β8 β7 (back), are green and blue, respectively. α1 and α2 helices are red. The bound Ca^{2+} ion shown in yellow.

b - Mindin-FS is rotated around 180° about the vertical axis with respect to the view in *(a)*.

c - Side view of the C2 domain of cytosolic phospholipase A_2. The orientation is similar to that of mindin-FS in *(b)*. The two four-stranded β-sheets, β4 β1 β8 β7 (front) and β6 β5 β2 β3 (back), are blue and green, respectively. The bound Ca^{2+} ion is yellow.

SPON2 also has nine cysteine residues, which may form intra- and/or intermolecular disulfide bonds (Figure I-45). This explains the ability of SPON2 protein to form dimers and larger oligomers that I observed by Western blot (Figure I-33c). Similar di- and oligomer formation was also reported by He et al. [153].

Characterisation of MMP1

MMP1 is a matrix metalloprotease, implicated in embryonic development, inflammation and tissue remodelling processes, as well as in diseases, such as cancer [165, 166] and arthritis [167, 168]. Its overexpression was shown in tumours and it was suggested to be associated with tumour invasion and metastasis [169]. MMP1 (or fibroblast collagenase, interstitial collagenase) is actively involved in ECM degradation. Targets of this protein, however, are not limited to ECM proteins. MMP1 cleaves non-ECM molecules as well. For instance, MMP1, produced by stromal fibroblasts, cleaves and activates PAR1 of cancer cells. Thereafter, this receptor modification triggers migration and invasion mechanisms via a cascade of signalling events, involving G protein and Rho [170] (Figure I-47).

In order to learn more about the functional role of MMP1, SPON2 and their possible interaction partners, I performed temporal siRNA knockdown experiments.

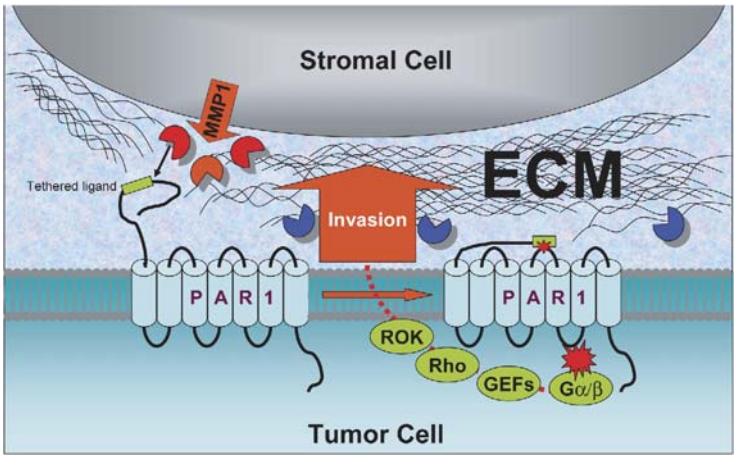


Figure I-47. PAR1 activation by MMP1 protein.

MMP1, secreted by stroma cell, is activated and cleaves type I collagen. It also cleaves the N-terminal extracellular part of PAR1. This receptor modification activates G protein, which via cascade of events leads to invasion phenotype. Figure from [170].

SPON2 and MMP1 knockdown

Taking advantage of the GEO database I found IMR-90 as the only cell line expressing both *SPON2* and *MMP1* in relatively high level. IMR-90 is a human primary embryonic fibroblast cell line derived from lung tissue of a 16 week old female Caucasian foetus. Transfection of these cells by *SPON2* siRNA led to dramatic morphological changes, whereas *MMP1* knockdown did not demonstrate visible morphological changes (Figure I-26).

When mRNA from treated cells was studied by expression microarrays and qRT-PCR, I found that *SPON2* knockdown leads to elevation of *MMP1* expression, but not *vice versa* (Figure I-30). Similar results were also obtained by Western blotting. This fact suggests that *MMP1* is a downstream target of *SPON2* gene.

SPON2 in EMT and cell motility

In contrast to SCLC-S, gene expression analysis of SCLC-A revealed a mesenchymal pattern. Accordingly, presence of "mesenchymal" cells in an epithelial cancer cell line gave us indications that morphological bifurcation of SCLC cell line might be mediated by EMT.

Furthermore, implication of SPON2 in EMT was suggested by temporal knockdown experiments. Whereas mock transfected cells exhibit clear spindle-shape morphology, *SPON2* silencing triggered non-uniform branching of fibroblasts, which was apparently caused by affected cell adhesion and polarity (Figure I-26). Alteration of these processes highlights *SPON2* role in EMT, where during transition epithelial cells weaken cell-cell adhesion, modulate their polarity, reorganize the cytoskeleton, and become isolated and motile [171-173]. Indeed, the temporal gene expression analysis of IMR-90 cells, treated with SPON2 siRNA, showed a significant enrichment for the GO term "epithelial-mesenchymal transition" (P=3.05E-08, Table I-47 and Figure I-48). This category encompasses 17 genes in total. Four of them, *TGFB3*, *BMP7*, *BMP2* and *S100A4* were influenced by SPON2 knockdown.

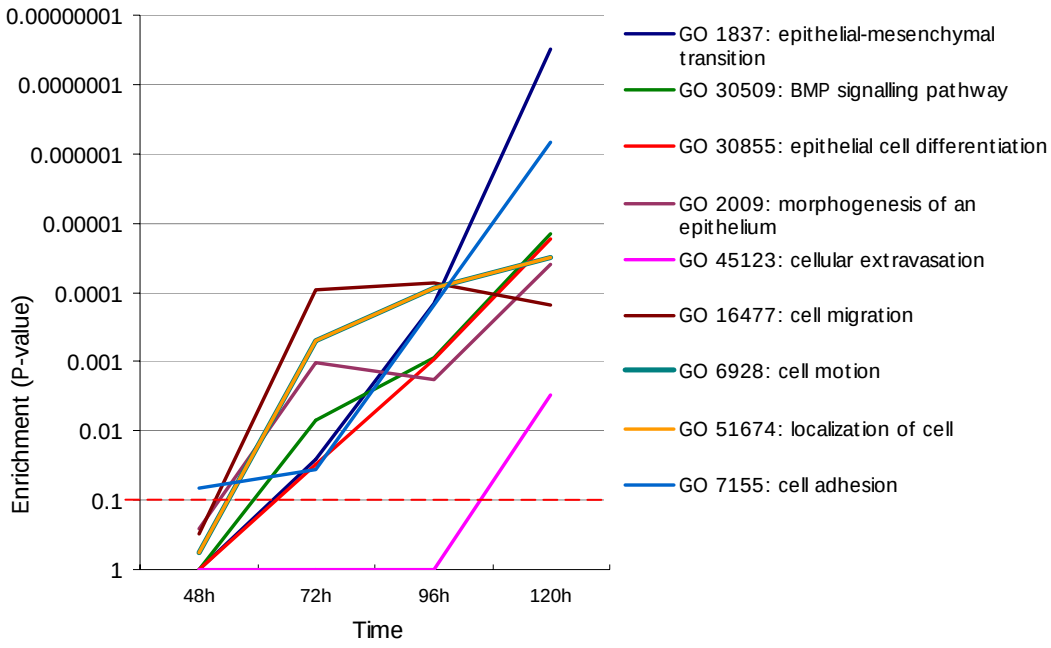


Figure I-48. GO annotation plot based of EMT and EMT related processes in *SPON2* siRNA treated IMR-90 cells. Differentially expressed genes were filtered by 4-fold change and P<0.05 criteria. Dotted red line shows the P=0.01.

TGFB3 (transforming growth factor, beta 3) is a gene well known to provoke EMT process [174, 175]. In comparison with mock siRNA transfection, it was downregulated in *SPON2* siRNA treated cells. In contrast to *TGFB3*, *BMP7* (bone morphogenic protein 7) counteracts EMT [176, 177], and markedly induces reverse to EMT process, known as mesenchymal-epithelial transition (MET) [178]. *BMP2* (bone morphogenic protein 2) expression is documented to induce EMT process in atrioventricular myocardium [179]. Both, BMP2 and BMP7 belong to transforming growth factor family and are involved in different cellular processes, including development, differentiation, proliferation and apoptosis [180]. The last of the four genes, suggesting a link of SPON2 to EMT, is *S100A4* (S100 calcium binding protein A4). S100A4 is a calcium-binding protein, which is expressed in epithelial cells during EMT [75]. The exact function of S100A4 is unclear, however, its interaction with components of the cytoskeleton proposes that it may be also involved in maintenance of cell shape and induction of motility [181].

Gene ontology is a very useful annotation tool for high throughput studies that tremendously facilitates collecting information about numerous genes or proteins. However, GO database *per se* is not complete and the annotation classifications do not describe fully all the data due to the following reasons: *(a)* knowledge is not static, it is updated constantly and often delayed; and *(b)* individual curators evaluate the data differently. For instance, it

was recently published that deficiency of heme oxygenase-1 (*HMOX1*), downregulated in *SPON2* siRNA treated cells, directly promotes EMT [182]. Yet, this gene was not flagged with "epithelial-mesenchymal transition" term. Another important gene, implicated in EMT but not included in the respective GO category, is *CLDN1*. *SPON2* knockdown led to significant upregulation of *CLDN1* (18.2-fold, 120h after siRNA administration). This protein is one of the most vital components of tight junctions [63]. Tight junctions provide cell-cell adhesion in epithelial and endothelial cells, functioning as a physical barrier that prevents water and solutes from passing freely through the intercellular space [183]. For successful EMT epithelial cells need to disassemble tight junctions prior to detachment from organized epithelial layer [63]. Furthermore, *CLDN1* influences cell polarity via interaction between tight junctions and F-actin [184]. Interestingly, F-actin was detected in SPON2 co-immunoprecipitate (co-IP) by mass spectrometry. Moreover, induction of EMT by transforming growth factor-beta (TGF-β) occurs by *CLDN1* downregulation [185]. *TAGLN* (transgelin) is another interesting gene, which was downregulated in *SPON2* siRNA treated IMR-90 cells. Transgelin is a transformation and shape-change sensitive actin binding protein and its downregulation may be an early and sensitive marker for the commencement of transformation [186 187]. TAGLN is a direct target of TGF-β/Smad-3 and is involved in epithelial cell migration in lung fibrosis [188].

Other genes, highlighting the impact of *SPON2* in EMT, are decorin (*DCN*) and alpha-2-macroglobulin (*A2M*). They attenuate EMT by hindering TGF-β binding to its receptors [189]. Intriguingly, both *DCN* and *A2M* were upregulated in SCLC-A.

The assumption that *SPON2* might be implicated into EMT process is also supported by recent manuscripts. Madar and colleagues studied gene expression in normal and cancer associated fibroblasts (CAFs) and revealed *SPON2* being 35.8-fold upregulated in CAFs [190]. Furthermore, among genes, upregulated in normal fibroblasts, were *SERPINB2*, *CLDN1* and *CXCL6*, which in our study were upregulated by *SPON2* knockdown. Essential role of EMT in the generation of CAFs (or myofibroblasts) was incontestably shown by a number of studies [171, 191-193]. It is believed that CAFs arise by EMT. Another evidence, supporting *SPON2* involvement in EMT, was presented by a gene expression study of extravillous trophoblasts (high invasiveness) and cytotrophoblasts (poor invasiveness) [194]. Trophoblasts are invasive and metastasizing cells of the placenta, acting similarly to tumour, e.g. by provoking angiogenesis, pathfinding behaviour and active crosstalk with surrounding microenvironment [195]. Importantly, EMT is reported to occur also during trophoblasts differentiation [196]. Here, the *SPON2*, along with *S100A4* and *FABP7*, was among the most

differentially upregulated genes in invasive extravillous trophoblasts. Interestingly, *HMOX1*, which deficiency directly mediates EMT [182], was downregulated in invasive cells and shown as a negative regulator of trophoblasts motility via up-regulation of PPARγ protein. Although these two studies observed *SPON2* among the highest deregulated genes in CAFs and extravillous trophoblasts, they do not discuss this in details, particularly in the context of EMT.

I already described before that Rac1 and Rho GTPase proteins, orchestrating EMT process, are regulated by Spon2 [154]. I also mentioned that SPON2 structurally strongly resembles a membrane-targeting C2 domain, where PI3K showed the highest similarity hit [152]. Thus, a putative mechanism by which SPON2 may regulate EMT could be via RAC1 and PI3K. RAC1 and PI3K are proteins playing an important role in EMT process [156, 157, 161-163].

The list of presumptive SPON2 interaction partners, analysed by DAVID annotation tool, exhibited significant enrichment of GO terms such as cytoskeleton organization and biogenesis, microtubule-based movement, cytoskeleton-dependent intracellular transport, cell motility and localization. (Figure I-49). Unexpectedly, it did not reveal any term directly referring to immune response, which most likely can be explained by fact that experiments were done on fibroblasts and not on immune cells.

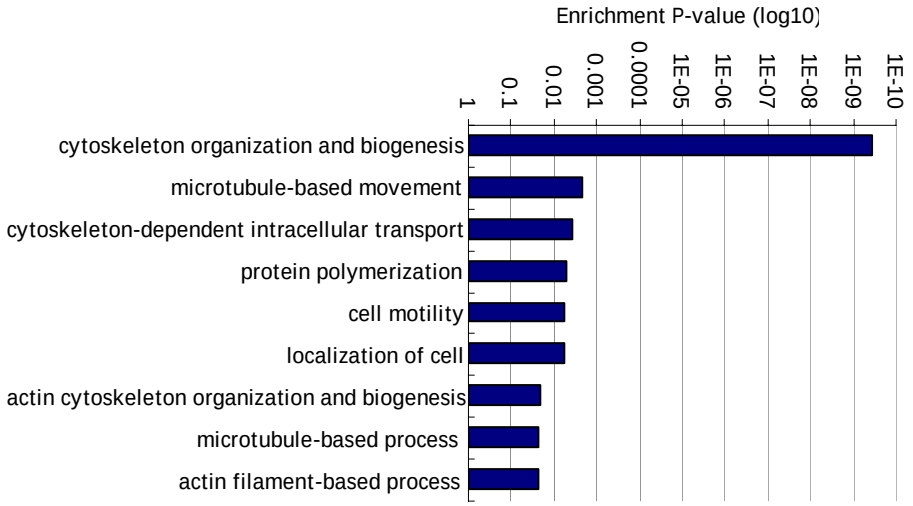


Figure I-49. GO annotation enrichment of co-IP-ed by SPON2 proteins.

Analyzed by DAVID annotation tool. Enrichment is based on GO Biological process; the human whole genome was used as background. * - Benjamini-Hochberg corrected P-value (P<0.05).

Modulation of cell adhesion is necessary, but insufficient for successful EMT. Additional application of developmental properties to gain migratory and invasive functions, involving extensive reorganization of the actin cytoskeleton, cell shape change and formation of membrane protrusions, is required for cancer invasion and EMT [197]. The fact of coincidence of GO enrichment terms in our results, which were obtained by different strategies (gene expression data from knockdown study and protein co-IP), mostly came on cell motility and localization.

Collectively, GO enrichment coincidence of two different experimental datasets (gene expression and co-IP) indicates the relevance of obtained data, highlights direct involvement of *SPON2* in cell motility and cytoskeleton reorganisation and, thus, strongly supports our hypothesis that *SPON2* is involved in EMT process.

SPON2 is possibly involved in amoeboid migration

Another process, where *SPON2* could be implicated is amoeboid migration. This kind of rapid locomotion is defined by crawling movement of single cells. In higher eukaryotes amoeboid movement was observed in leukocytes, neutrophils, macrophages, some cancer cells (lymphoma, SCLC, small cell prostate cancer etc.) and myofibroblasts (or CAFs) [54, 198-201]. Mesenchymal migration, which is typical to fibroblasts, smooth-muscle cells and different cancers (e.g. connective tissue cancers, some carcinomas etc.), needs proteolitic and ECM remodelling enzymes, depends on integrin mediated adhesion and presence of traction forces (reviewed by Friedl and Wolf [54]). In contrast to mesenchymal migration, the amoeboid one does not necessitate proteolitic processes, allowing cells rather to traverse than to degrade the ECM. Integrin-mediated adhesion is not essential for amoeboid movement and its velocity is up to 20 times higher when compared to mesenchymal movement [202, 203].

Taken together that SPON2 protein is needed for successful migration of macrophages and neutrophils [151, 158], and that both cell types move my amoeboid migration [54, 198-200] suggests that SPON2 protein can be also implicated in amoeboid migration.

SPON2 in cell-matrix adhesion

Implication of SPON2 in adhesion was described before in several manuscripts [151] [204]. Jia *et al*. demonstrated experimentally that SPON2 directly binds to integrins and by this

plays an essential role in recruitment of inflammatory cells. Indeed, gene expression study of *SPON2* silencing in IMR-90 fibroblasts (24-120h after knockdown) revealed an array of genes either directly or indirectly involved in cell adhesion. GO enrichment analysis of these genes showed significant enrichment (P-value up to 6.65E-07) with cell adhesion term (Table I-47). Interestingly, the number of genes flagged with adhesion term was drastically increased starting from 72h. This coincides with morphological changes observed by *SPON2* knockdown, where the treated fibroblasts exhibit apparent loss of adhesion starting from 72h and reaching the maximum at 168h (Figure I-26). Whereas GO annotation enrichment did not reveal if *SPON2* promotes or inhibits cell adhesion, direct observation of morphological changes in siRNA treated IMR-90 fibroblasts indicates it as adhesion promoting protein.

Immunofluorescence microscopy of HaCaT and F11 cell lines demonstrated SPON2 localization on the tips of protrusions, possible cell-matrix adhesion contacts and filopodias (Figure I-39 and Figure I-38). Cell motility strongly depends on adhesion. Thus, involvement of *SPON2* in cell-matrix adhesion can be strongly supported by fact, where *SPON2* null mice demonstrate impaired recruitment and migration of eosinophils [158], neutrophils and macrophages [151].

SPON2 in immune response

The role of SPON2 in immunity was already described in studies using Spon2 knockout mice [153, 155]. These mice showed impaired clearance of intranasal influenza viruses and have defective response to broad spectrum of bacterial pathogens. SPON2 interacts directly with integrins and promotes recruitment of inflammatory cells, and accordingly, contributes to immune response [151]. Li and colleagues succeeded blocking SPON2-mediated migration of neutrophils using antibodies against integrin–α4, –αM and –β2 [152]. Indeed, our knockdown experiments revealed a panel of deregulated cytokines (n=43-50, 72-120h, 4-fold changed, P=0.05), which was also demonstrated by GO enrichment analysis (Figure I-50).

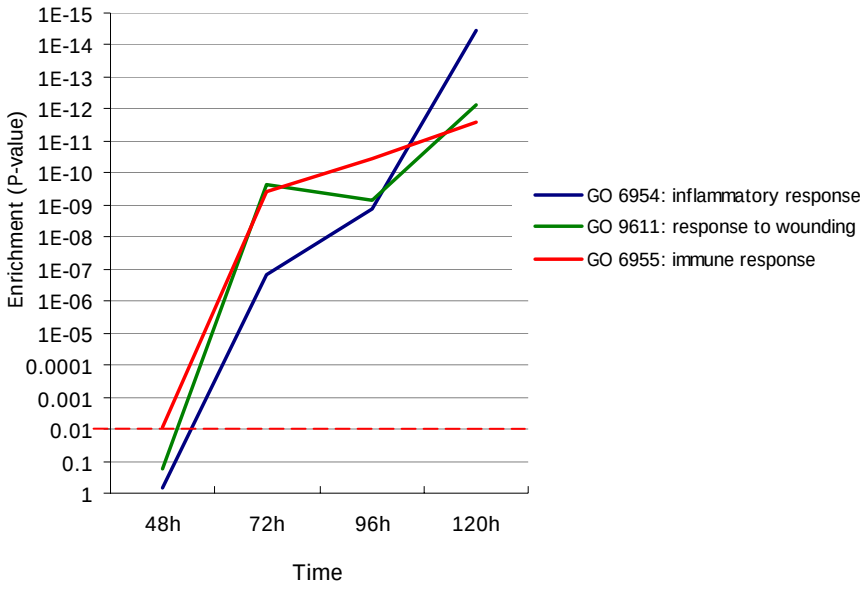


Figure I-50. GO annotation plot based on immunity related processes in SPON2 siRNA treated IMR-90 cells. Differentially expressed genes were filtered by 4-fold change and P<0.05 criteria. Dotted red line shows the P=0.01.

Interactome network visualisation highlights conformity of obtained results

For a multidimensional analysis obtained results were studied by Cytoscape network visualisation software. Cytoscape included 80.6% of genes deregulated in *SPON2* knockdown in a single network (Figure I-41). Even distribution of upregulated (blue) and downregulated (red) genes in *SPON2* knockdown hints that GO enrichment analysis did not necessitate separation of these genes by their up- and downregulation. Similar network visualisation was performed with the putative SPON2 interaction partner proteins as detected by co-IP (Table I-49). In contrast to a control set of random proteins, the co-IP based interactome showed impressive integration (94.3%) of candidate proteins in a single network (Figure I-42). The only orphans were calpain-2 (CAPN2) and F-actin capping protein subunit beta (CAPZB). CAPN2 is a Ca-regulated protease, which is involved in cytoskeleton remodelling and facilitates actin reorganization [205]. Calpain deficient mice show impaired migration and reduction of focal contacts in embryonic fibroblasts [206]. CAPZB is a protein that regulates actin filament growth by capping the barbed end [207], potentially involved in cell motion (GO Biological Process, traceable author statement referred to [208]). The reason why these two proteins nevertheless appeared as orphans could

be due to their insufficient annotation in current HPRD release. Taking into account that proteins integrated into the network are implicated in cytoskeleton organisation and remodelling, these two orphans (CAPN2 and CAPZB) have direct functional relation to generated structure.

The consistency of obtained results was visualised by merging separate networks built up by expression analysis of *SPON2* knockdown and by SPON2 co-IP (Figure I-43). Surprisingly, co-IP-ed proteins were completely (excluding two abovementioned orphans) integrated into another network without cluster outlining. The merging also revealed seven common nodes (6.5% of deregulated by knockdown genes and 20.0% of co-IP-ed proteins). These common nodes were not linking two separate structures, built up by gene expression and co-IP data, but were evenly scattered within unified network.

Collectively, the interaction network analysis indicates the conformity of our results, obtained by two different strategies.

Intracellular localization of SPON2

Until now SPON2 has not been described as a protein involved in maintenance of cytoskeleton architecture. Moreover, its localization was described only as secreted or extracellular matrix protein and was not even mentioned in intracellular context [155, 158]. Surprisingly, I observed this protein localized also on microtubules, nuclear envelope, mitotic spindles and midbody, as well as in filopodias and growth cones of neuronal cells (Figures I-32, I-33, I-34, I-35 and I-36). Initially this localisation was first detected by IF against overexpressed SPON2-FLAG, but similar expression patterns were also detected by application of an antibody against endogenous SPON2. This association with microtubules and mitotic spindles was further supported by co-IP and following mass spectrometry analysis. The list of candidate proteins was significantly enriched with actin maintenance, cytoskeleton organisation and biogenesis, and cell movement (Figure I-49). Such putative SPON2 interaction partners as kinesin like protein KIF11 and Shk1 kinase-binding protein are actively involved in cytokinesis and spindle dynamics [128, 132] (Table I-49). Similarly, the elongation factor 1-alpha 1 (EEF1A1) and nucleoprotein TPR proteins are known to be implicated in nuclear import and export [122, 123, 130].

Our results, combined with those of literature show contribution of SPON2 to the processes, listed in Table I-54.

Table I-54. Processes with possible contribution of SPON2

Process	Source
Axon guidance	gene expression of fibroblasts with *SPON2* knockdown and [118, 159, 209]
Cancer	gene expression of SCLC cell line and fibroblasts with *SPON2* knockdown, co-IP and [210-213]
Cell division	SPON2 IF and co-IP
Cell motility	gene expression of fibroblasts with SPON2 knockdown, SPON2 co-IP, IF and [158]
Cytoskeleton organisation and biogenesis	SPON2 co-IP
Cytoskeleton-dependent intracellular transport	SPON2 co-IP
EMT	*SPON2* knockdown in fibroblasts and its gene expression, SPON2 co-IP
Immune response and secretion	gene expression of fibroblasts with *SPON2* knockdown and [151-155, 209]

When converging the processes listed in Table I-54, and trying to find a common event, which underlies all these processes, I assumed that SPON2 could be implicated in vesicle trafficking.

Vesicle transport is essential for the communication of signals between distal neuronal parts and the nucleus in the context of axon guidance and synapse formation [214-216]. *SPON2* is highly expressed in peripheral nerves and is involved in axon guidance [118, 159, 209]. Additionally, SPON2 was found deregulated in mental retardation and neuronal diseases [217].

The vesicle trafficking in immune response is essential as well. It mediates secretion of cytotoxic elements, antigen presentation and endocytosis [153]. Intracellular logistic, orchestrating also cytoskeleton organisation and cell motility, benefits from motor proteins (e.g. myosin, kinesin, dynein), which actively moves cargos along the cytoskeleton. For instance, integrins repeatedly circulate from rear to leading edge in migrating cells by motor proteins. In complex with cargo, motor proteins move along cytoskeleton [218]. Importantly, whereas CAV2 (caveolin-2, a component of caveolae) and cytoplasmatic dynein were found by gene expression study of SCLC cell line, SPON2 co-IP detected cytoplasmatic dynein and kinesin-11. Tropomyosin-1 and -4, detected by SPON2 co-IP as well (Table I-49), are described as key determinants of actin cytoskeleton and mediate cell migration, intracellular vesicle transport and cytokinesis [218, 219]. Collaboration of motor proteins with vesicle transport plays a critical role in cytokinesis. For example, F-actin associated vesicle transport on central spindles delivers membrane components to the cleavage furrow [220].

Based on IF and co-IP results, I supposed that SPON2 might be of importance also in cell division by its implication in spindle organisation and nuclear envelop breakdown and reassembly. In animals during mitosis the nuclear envelop undergoes breakdown,

allowing penetration of spindle microtubules into the nuclear space and access to the kinetochores. The initial event in envelop disassembly is dispersion of nucleoporins (proteins of the nuclear pore complex) [221]. They are not evenly distributed within the cell, but concentrate in the vicinity of microtubule organizing centres (MTOC). After breakdown of nuclear envelop nucleoporins are packed into membranous vesicles, and later on reused for nuclear envelop reassembly [221]. The nucleoporins are also well known to regulate spindle assembly.

Kinesin like protein KIF11 is associated with spindle microtubules and promotes separation of spindle poles [221, 222]. Dynein is a minus-end directed motor protein that is implicated in spindle orientation, and the spindles are crosslinked by KIF11. These two proteins were co-immunoprecipitated with SPON2 in our experiments. Additionally, dynein was also shown to be overexpressed in SCLC-A. Most likely SPON2, together with its interaction partners revealed by co-IP, orchestrates actin and tubulin organisation, including mechanical movement of chromosomes.

Cell pathfinding, motility and cytoskeleton reorganization are key events in cancer progression and EMT. Moreover, cancer cell actively uses immune and secretory functions especially for its invasion and microenvironment mounting [223].

SPON2 in cancer and other diseases

Until now I was discussing the role of *SPON2* in different biological processes, such as EMT, cell motility, immune response etc. However, it would be reasonable to combine information available in literature and different databases with our results in the context of diseases.

SPON2 was suggested as a diagnostic marker for ovarian and prostate cancers [211, 213]. *SPON2*, along with many other genes, is also mentioned in several patents. These gene lists are mentioned in patents as: deregulated in colon (WIPO patent WO/2007/112330) and gastritic (US Patent 20070184439) cancers, markers for prostate cancer (US Patent 20060068425), ovarian, colon, breast and stomach cancer risk assessment protein (US Patent 7294704), extracellular tumour endothelial marker (US Patent 20060210975). The latter also claims that *SPON2* is involved in rheumatoid arthritis and diabetic retinopathy. Strikingly, EMT is known to be involved in abovementioned diseases [69, 83, 84, 224]. *SPON2* implication in cancer progression, and specifically in metastasis, is also supported by its

upregulation in CAFs when compared to normal fibroblasts [190]. *SPON2* was also differentially expressed in autism [217].

Additionally, I performed meta-analysis of differentially expressed genes in IMR-90 fibroblasts after *SPON2* knockdown (72, 96 and 120h), using Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems, www.ingenuity.com). This software enables the analysis of complex biological and chemical systems, using a broad range of databases and sources. In respect to disease association, our data showed significant ($p<0.0001$) enrichment of genes, implicated in tumorigenesis, rheumatic disease, arthritis and fibrosis (Figure I-51).

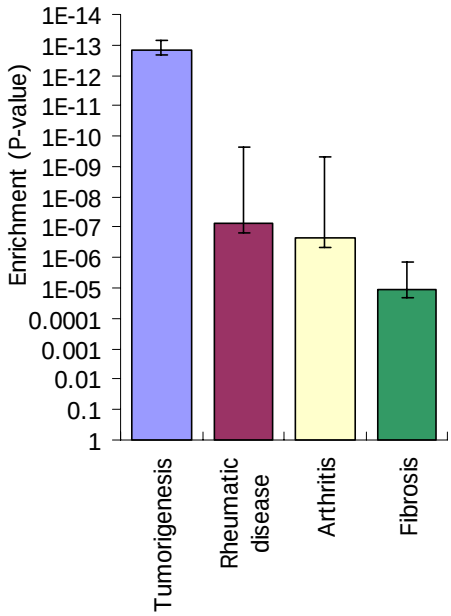


Figure I-51. Disease association of genes deregulated by *SPON2* knockdown.

Enrichment analysis was done by Ingenuity Analysis Tool 7.1 (IPA), based on genes deregulated after 72, 96 and 120h of SPON2 knockdown.

In summary this study proposes *SPON2* gene as a good candidate for multiple biological processes such as EMT, vesicle transport and cytoskeleton organization, which fits well to abovementioned diseases.

SUPPLEMENTARY DATA

Table S-1. GO annotation enrichment of deregulated genes (4-fold, P<0.05) in IMR-90 cells by SPON2 knockdown (72h)

GO Category	Genes in Category	% of Genes in Category	Genes in List in Category	% of Genes in List in Category	p-Value
GO:50874: organismal physiological process	2995	14.58	99	36.8	9.37E-20
GO:7267: cell-cell signaling	823	4.005	41	15.24	1.66E-13
GO:7154: cell communication	5765	28.06	124	46.1	1.83E-10
GO:9611: response to wounding	626	3.047	31	11.52	2.27E-10
GO:6955: immune response	1232	5.996	45	16.73	3.61E-10
GO:8015: circulation	203	0.988	17	6.32	1.59E-09
GO:9605: response to external stimulus	1090	5.305	40	14.87	3.47E-09
GO:6952: defense response	1352	6.58	45	16.73	6.55E-09
GO:42330: taxis	200	0.973	16	5.948	9.40E-09
GO:6935: chemotaxis	200	0.973	16	5.948	9.40E-09
GO:9607: response to biotic stimulus	1407	6.848	45	16.73	2.17E-08
GO:42127: regulation of cell proliferation	606	2.949	27	10.04	3.26E-08
GO:50896: response to stimulus	3247	15.8	77	28.62	6.42E-08
GO:7166: cell surface receptor linked signal transduction	1994	9.705	55	20.45	7.49E-08
GO:6954: inflammatory response	343	1.669	19	7.063	1.47E-07
GO:8285: negative regulation of cell proliferation	298	1.45	17	6.32	4.49E-07
GO:9613: response to pest, pathogen or parasite	669	3.256	26	9.665	8.27E-07
GO:7275: development	3595	17.5	79	29.37	9.97E-07
GO:42221: response to chemical stimulus	644	3.134	25	9.294	1.39E-06
GO:51216: cartilage development	44	0.214	7	2.602	1.54E-06
GO:48513: organ development	1356	6.6	39	14.5	3.00E-06
GO:7165: signal transduction	4610	22.44	93	34.57	3.27E-06
GO:8283: cell proliferation	1009	4.911	32	11.9	3.54E-06
GO:43207: response to external biotic stimulus	731	3.558	26	9.665	4.19E-06
GO:35295: tube development	123	0.599	10	3.717	5.12E-06
GO:8284: positive regulation of cell proliferation	268	1.304	14	5.204	1.29E-05
GO:9628: response to abiotic stimulus	803	3.908	26	9.665	2.17E-05
GO:51239: regulation of organismal physiological process	320	1.557	15	5.576	2.28E-05
GO:9888: tissue development	404	1.966	17	6.32	2.61E-05
GO:51234: establishment of localization	4721	22.98	91	33.83	2.97E-05
GO:42592: homeostasis	454	2.21	18	6.691	3.31E-05
GO:45073: regulation of chemokine biosynthesis	6	0.0292	3	1.115	4.31E-05
GO:48661: positive regulation of smooth muscle cell proliferation	6	0.0292	3	1.115	4.31E-05
GO:51179: localization	4772	23.22	91	33.83	4.54E-05
GO:9653: morphogenesis	1479	7.198	38	14.13	5.06E-05
GO:6950: response to stress	1600	7.787	40	14.87	5.75E-05
GO:30324: lung development	54	0.263	6	2.23	7.26E-05
GO:42033: chemokine biosynthesis	7	0.0341	3	1.115	7.47E-05
GO:50755: chemokine metabolism	7	0.0341	3	1.115	7.47E-05
GO:48251: elastic fiber assembly	7	0.0341	3	1.115	7.47E-05
GO:30323: respiratory tube development	55	0.268	6	2.23	8.06E-05

GO:6875: metal ion homeostasis	204	0.993	11	4.089	8.27E-05
GO:16477: cell migration	320	1.557	14	5.204	8.94E-05
GO:19226: transmission of nerve impulse	405	1.971	16	5.948	9.46E-05
GO:7610: behavior	728	3.543	23	8.55	9.48E-05
GO:48660: regulation of smooth muscle cell proliferation	8	0.0389	3	1.115	1.18E-04
GO:43118: negative regulation of physiological process	1429	6.955	36	13.38	1.21E-04
GO:15718: monocarboxylic acid transport	21	0.102	4	1.487	1.44E-04
GO:7626: locomotory behavior	607	2.954	20	7.435	1.57E-04
GO:30538: embryonic genitalia morphogenesis	2	0.00973	2	0.743	1.71E-04
GO:30241: muscle thick filament assembly	2	0.00973	2	0.743	1.71E-04
GO:48745: smooth muscle development	2	0.00973	2	0.743	1.71E-04
GO:48746: smooth muscle fiber development	2	0.00973	2	0.743	1.71E-04
GO:19371: cyclooxygenase pathway	2	0.00973	2	0.743	1.71E-04
GO:31034: myosin filament assembly	2	0.00973	2	0.743	1.71E-04
GO:1660: fever	2	0.00973	2	0.743	1.71E-04
GO:51549: positive regulation of keratinocyte migration	2	0.00973	2	0.743	1.71E-04
GO:51547: regulation of keratinocyte migration	2	0.00973	2	0.743	1.71E-04
GO:31033: myosin filament assembly or disassembly	2	0.00973	2	0.743	1.71E-04
GO:30278: regulation of ossification	22	0.107	4	1.487	1.75E-04
GO:1937: negative regulation of endothelial cell proliferation	9	0.0438	3	1.115	1.76E-04
GO:48659: smooth muscle cell proliferation	9	0.0438	3	1.115	1.76E-04
GO:9887: organ morphogenesis	661	3.217	21	7.807	1.78E-04
GO:30003: cation homeostasis	224	1.09	11	4.089	1.89E-04
GO:50801: ion homeostasis	263	1.28	12	4.461	1.93E-04
GO:30005: di-, tri-valent inorganic cation homeostasis	191	0.93	10	3.717	2.20E-04
GO:51243: negative regulation of cellular physiological process	1378	6.707	34	12.64	2.78E-04
GO:46850: regulation of bone remodeling	25	0.122	4	1.487	2.93E-04
GO:6873: cell ion homeostasis	237	1.153	11	4.089	3.06E-04
GO:50920: regulation of chemotaxis	11	0.0535	3	1.115	3.39E-04
GO:7268: synaptic transmission	366	1.781	14	5.204	3.57E-04
GO:42060: wound healing	167	0.813	9	3.346	3.66E-04
GO:7399: nervous system development	967	4.706	26	9.665	4.29E-04
GO:48468: cell development	420	2.044	15	5.576	4.57E-04
GO:6874: calcium ion homeostasis	138	0.672	8	2.974	4.76E-04
GO:6811: ion transport	1086	5.285	28	10.41	5.01E-04
GO:45079: negative regulation of chemokine biosynthesis	3	0.0146	2	0.743	5.08E-04
GO:51546: keratinocyte migration	3	0.0146	2	0.743	5.08E-04
GO:6928: cell motility	566	2.755	18	6.691	5.08E-04
GO:51674: localization of cell	566	2.755	18	6.691	5.08E-04
GO:30154: cell differentiation	1371	6.673	33	12.27	5.31E-04
GO:40011: locomotion	569	2.769	18	6.691	5.41E-04
GO:50817: coagulation	144	0.701	8	2.974	6.32E-04
GO:48731: system development	997	4.852	26	9.665	6.75E-04
GO:31214: biomineral formation	112	0.545	7	2.602	6.86E-04
GO:1503: ossification	112	0.545	7	2.602	6.86E-04
GO:7599: hemostasis	146	0.711	8	2.974	6.92E-04
GO:7186: G-protein coupled receptor protein signaling pathway	1004	4.886	26	9.665	7.47E-04
GO:48523: negative regulation of cellular process	1518	7.388	35	13.01	7.75E-04
GO:48519: negative regulation of biological process	1638	7.972	37	13.75	7.94E-04
GO:1936: regulation of endothelial cell proliferation	15	0.073	3	1.115	8.99E-04
GO:46849: bone remodeling	119	0.579	7	2.602	9.82E-04
GO:46942: carboxylic acid transport	119	0.579	7	2.602	9.82E-04
GO:45778: positive regulation of ossification	4	0.0195	2	0.743	1.01E-03

GO:35313: wound healing, spreading of epidermal cells	4	0.0195	2	0.743	1.01E-03
GO:15849: organic acid transport	120	0.584	7	2.602	1.03E-03
GO:2009: morphogenesis of an epithelium	88	0.428	6	2.23	1.05E-03
GO:50900: immune cell migration	36	0.175	4	1.487	1.22E-03
GO:50877: neurophysiological process	1042	5.071	26	9.665	1.27E-03
GO:44241: lipid digestion	17	0.0827	3	1.115	1.32E-03
GO:30299: cholesterol absorption	17	0.0827	3	1.115	1.32E-03
GO:48518: positive regulation of biological process	1511	7.354	34	12.64	1.40E-03
GO:51094: positive regulation of development	93	0.453	6	2.23	1.40E-03
GO:50878: regulation of body fluids	163	0.793	8	2.974	1.41E-03
GO:48522: positive regulation of cellular process	1281	6.234	30	11.15	1.47E-03
GO:30199: collagen fibril organization	18	0.0876	3	1.115	1.57E-03
GO:1935: endothelial cell proliferation	18	0.0876	3	1.115	1.57E-03
GO:7605: sensory perception of sound	205	0.998	9	3.346	1.57E-03
GO:45597: positive regulation of cell differentiation	65	0.316	5	1.859	1.61E-03
GO:1505: regulation of neurotransmitter levels	96	0.467	6	2.23	1.65E-03
GO:45669: positive regulation of osteoblast differentiation	5	0.0243	2	0.743	1.66E-03
GO:8211: glucocorticoid metabolism	5	0.0243	2	0.743	1.66E-03
GO:31649: heat generation	5	0.0243	2	0.743	1.66E-03
GO:46852: positive regulation of bone remodeling	5	0.0243	2	0.743	1.66E-03
GO:50954: sensory perception of mechanical stimulus	207	1.007	9	3.346	1.68E-03
GO:6805: xenobiotic metabolism	41	0.2	4	1.487	1.99E-03
GO:6807: nitrogen compound metabolism	588	2.862	17	6.32	2.02E-03
GO:7596: blood coagulation	139	0.676	7	2.602	2.40E-03
GO:43066: negative regulation of apoptosis	306	1.489	11	4.089	2.45E-03
GO:15721: bile acid transport	6	0.0292	2	0.743	2.47E-03
GO:46627: negative regulation of insulin receptor signaling pathway	6	0.0292	2	0.743	2.47E-03
GO:50892: intestinal absorption	21	0.102	3	1.115	2.48E-03
GO:9308: amine metabolism	549	2.672	16	5.948	2.50E-03
GO:6836: neurotransmitter transport	72	0.35	5	1.859	2.54E-03
GO:9410: response to xenobiotic stimulus	44	0.214	4	1.487	2.58E-03
GO:43069: negative regulation of programmed cell death	309	1.504	11	4.089	2.64E-03
GO:7188: G-protein signaling, coupled to cAMP nucleotide second messenger	110	0.535	6	2.23	3.28E-03
GO:42640: anagen	7	0.0341	2	0.743	3.43E-03
GO:6562: proline catabolism	7	0.0341	2	0.743	3.43E-03
GO:46626: regulation of insulin receptor signaling pathway	7	0.0341	2	0.743	3.43E-03
GO:30282: bone mineralization	24	0.117	3	1.115	3.66E-03
GO:30001: metal ion transport	572	2.784	16	5.948	3.73E-03
GO:19933: cAMP-mediated signaling	115	0.56	6	2.23	4.08E-03
GO:51260: protein homooligomerization	50	0.243	4	1.487	4.12E-03
GO:8217: regulation of blood pressure	51	0.248	4	1.487	4.43E-03
GO:1933: negative regulation of protein amino acid phosphorylation	8	0.0389	2	0.743	4.54E-03
GO:50732: negative regulation of peptidyl-tyrosine phosphorylation	8	0.0389	2	0.743	4.54E-03
GO:42532: negative regulation of tyrosine phosphorylation of STAT protein	8	0.0389	2	0.743	4.54E-03
GO:42518: negative regulation of tyrosine phosphorylation of Stat3 protein	8	0.0389	2	0.743	4.54E-03
GO:45763: negative regulation of amino acid metabolism	8	0.0389	2	0.743	4.54E-03
GO:1659: thermoregulation	8	0.0389	2	0.743	4.54E-03
GO:46426: negative regulation of JAK-STAT cascade	8	0.0389	2	0.743	4.54E-03
GO:50727: regulation of inflammatory response	26	0.127	3	1.115	4.62E-03
GO:7223: frizzled-2 signaling pathway	26	0.127	3	1.115	4.62E-03

GO:6810: transport	4138	20.14	72	26.77	5.02E-03
GO:19725: cell homeostasis	340	1.655	11	4.089	5.40E-03
GO:1943: hair follicle maturation	9	0.0438	2	0.743	5.79E-03
GO:6568: tryptophan metabolism	9	0.0438	2	0.743	5.79E-03
GO:7638: mechanosensory behavior	9	0.0438	2	0.743	5.79E-03
GO:50918: positive chemotaxis	9	0.0438	2	0.743	5.79E-03
GO:50926: regulation of positive chemotaxis	9	0.0438	2	0.743	5.79E-03
GO:50927: positive regulation of positive chemotaxis	9	0.0438	2	0.743	5.79E-03
GO:50930: induction of positive chemotaxis	9	0.0438	2	0.743	5.79E-03
GO:1932: regulation of protein amino acid phosphorylation	55	0.268	4	1.487	5.81E-03
GO:6519: amino acid and derivative metabolism	444	2.161	13	4.833	5.95E-03
GO:42475: odontogenesis (sensu Vertebrata)	29	0.141	3	1.115	6.30E-03
GO:30595: immune cell chemotaxis	29	0.141	3	1.115	6.30E-03
GO:6521: regulation of amino acid metabolism	57	0.277	4	1.487	6.59E-03
GO:19221: cytokine and chemokine mediated signaling pathway	57	0.277	4	1.487	6.59E-03
GO:42107: cytokine metabolism	91	0.443	5	1.859	6.91E-03
GO:30509: BMP signaling pathway	30	0.146	3	1.115	6.94E-03
GO:7167: enzyme linked receptor protein signaling pathway	402	1.956	12	4.461	6.98E-03
GO:1501: skeletal development	304	1.48	10	3.717	7.00E-03
GO:8016: regulation of heart contraction rate	58	0.282	4	1.487	7.01E-03
GO:45936: negative regulation of phosphate metabolism	10	0.0487	2	0.743	7.17E-03
GO:42326: negative regulation of phosphorylation	10	0.0487	2	0.743	7.17E-03
GO:19369: arachidonic acid metabolism	10	0.0487	2	0.743	7.17E-03
GO:42303: molting cycle	10	0.0487	2	0.743	7.17E-03
GO:42633: hair cycle	10	0.0487	2	0.743	7.17E-03
GO:50921: positive regulation of chemotaxis	10	0.0487	2	0.743	7.17E-03
GO:6916: anti-apoptosis	259	1.261	9	3.346	7.32E-03
GO:16338: calcium-independent cell-cell adhesion	31	0.151	3	1.115	7.61E-03
GO:6800: oxygen and reactive oxygen species metabolism	132	0.642	6	2.23	7.91E-03
GO:43062: extracellular structure organization and biogenesis	132	0.642	6	2.23	7.91E-03
GO:30198: extracellular matrix organization and biogenesis	132	0.642	6	2.23	7.91E-03
GO:48666: neuron development	265	1.29	9	3.346	8.44E-03
GO:50793: regulation of development	265	1.29	9	3.346	8.44E-03
GO:45667: regulation of osteoblast differentiation	11	0.0535	2	0.743	8.69E-03
GO:42402: biogenic amine catabolism	11	0.0535	2	0.743	8.69E-03
GO:43119: positive regulation of physiological process	1150	5.597	25	9.294	9.12E-03
GO:35239: tube morphogenesis	63	0.307	4	1.487	9.35E-03
GO:51240: positive regulation of organismal physiological process	137	0.667	6	2.23	9.41E-03
GO:6520: amino acid metabolism	368	1.791	11	4.089	9.51E-03
GO:7243: protein kinase cascade	579	2.818	15	5.576	9.65E-03

Table S-2. GO annotation enrichment of deregulated genes (4-fold, P<0.05) in IMR-90 cells by SPON2 knockdown (96h)

Category	Genes in Category	% of Genes in Category	Genes in List in Category	% of Genes in List in Category	p-Value
GO:50874: organismal physiological process	2995	14.58	99	33.22	2.92E-16
GO:7154: cell communication	5765	28.06	141	47.32	9.41E-13
GO:42330: taxis	200	0.973	20	6.711	1.39E-11
GO:6935: chemotaxis	200	0.973	20	6.711	1.39E-11
GO:6955: immune response	1232	5.996	50	16.78	3.41E-11
GO:9605: response to external stimulus	1090	5.305	45	15.1	2.19E-10
GO:9611: response to wounding	626	3.047	32	10.74	7.09E-10
GO:6952: defense response	1352	6.58	50	16.78	8.60E-10
GO:9613: response to pest, pathogen or parasite	669	3.256	33	11.07	9.31E-10
GO:6954: inflammatory response	343	1.669	23	7.718	1.29E-09
GO:48513: organ development	1356	6.6	49	16.44	2.80E-09
GO:9607: response to biotic stimulus	1407	6.848	50	16.78	3.27E-09
GO:43207: response to external biotic stimulus	731	3.558	33	11.07	8.35E-09
GO:51239: regulation of organismal physiological process	320	1.557	21	7.047	1.01E-08
GO:51216: cartilage development	44	0.214	9	3.02	1.14E-08
GO:7165: signal transduction	4610	22.44	107	35.91	7.23E-08
GO:42221: response to chemical stimulus	644	3.134	28	9.396	2.60E-07
GO:8015: circulation	203	0.988	15	5.034	2.98E-07
GO:7166: cell surface receptor linked signal transduction	1994	9.705	57	19.13	4.59E-07
GO:1501: skeletal development	304	1.48	18	6.04	5.35E-07
GO:7267: cell-cell signaling	823	4.005	31	10.4	1.28E-06
GO:9628: response to abiotic stimulus	803	3.908	30	10.07	2.27E-06
GO:45597: positive regulation of cell differentiation	65	0.316	8	2.685	4.39E-06
GO:9435: NAD biosynthesis	9	0.0438	4	1.342	5.16E-06
GO:19674: NAD metabolism	9	0.0438	4	1.342	5.16E-06
GO:48659: smooth muscle cell proliferation	9	0.0438	4	1.342	5.16E-06
GO:6950: response to stress	1600	7.787	46	15.44	6.17E-06
GO:51094: positive regulation of development	93	0.453	9	3.02	8.34E-06
GO:42127: regulation of cell proliferation	606	2.949	24	8.054	9.41E-06
GO:50896: response to stimulus	3247	15.8	76	25.5	9.75E-06
GO:6569: tryptophan catabolism	4	0.0195	3	1.007	1.20E-05
GO:46218: indolalkylamine catabolism	4	0.0195	3	1.007	1.20E-05
GO:42436: indole derivative catabolism	4	0.0195	3	1.007	1.20E-05
GO:6936: muscle contraction	241	1.173	14	4.698	1.23E-05
GO:7517: muscle development	243	1.183	14	4.698	1.35E-05
GO:7610: behavior	728	3.543	26	8.725	2.41E-05
GO:42692: muscle cell differentiation	60	0.292	7	2.349	2.52E-05
GO:7223: frizzled-2 signaling pathway	26	0.127	5	1.678	3.18E-05
GO:45445: myoblast differentiation	44	0.214	6	2.013	3.93E-05
GO:30154: cell differentiation	1371	6.673	39	13.09	4.31E-05
GO:45073: regulation of chemokine biosynthesis	6	0.0292	3	1.007	5.85E-05
GO:48661: positive regulation of smooth muscle cell proliferation	6	0.0292	3	1.007	5.85E-05
GO:7275: development	3595	17.5	79	26.51	5.97E-05
GO:16055: Wnt receptor signaling pathway	179	0.871	11	3.691	6.37E-05
GO:16477: cell migration	320	1.557	15	5.034	7.34E-05

GO:7626: locomotory behavior	607	2.954	22	7.383	8.33E-05
GO:6928: cell motility	566	2.755	21	7.047	8.72E-05
GO:51674: localization of cell	566	2.755	21	7.047	8.72E-05
GO:9888: tissue development	404	1.966	17	5.705	9.32E-05
GO:40011: locomotion	569	2.769	21	7.047	9.39E-05
GO:42033: chemokine biosynthesis	7	0.0341	3	1.007	1.01E-04
GO:50755: chemokine metabolism	7	0.0341	3	1.007	1.01E-04
GO:8285: negative regulation of cell proliferation	298	1.45	14	4.698	1.24E-04
GO:40007: growth	376	1.83	16	5.369	1.31E-04
GO:1837: epithelial to mesenchymal transition	19	0.0925	4	1.342	1.42E-04
GO:19363: pyridine nucleotide biosynthesis	19	0.0925	4	1.342	1.42E-04
GO:7155: cell adhesion	1053	5.125	31	10.4	1.52E-04
GO:48660: regulation of smooth muscle cell proliferation	8	0.0389	3	1.007	1.60E-04
GO:48628: myoblast maturation	20	0.0973	4	1.342	1.75E-04
GO:48741: skeletal muscle fiber development	58	0.282	6	2.013	1.90E-04
GO:48747: muscle fiber development	58	0.282	6	2.013	1.90E-04
GO:30538: embryonic genitalia morphogenesis	2	0.00973	2	0.671	2.10E-04
GO:30241: muscle thick filament assembly	2	0.00973	2	0.671	2.10E-04
GO:48745: smooth muscle development	2	0.00973	2	0.671	2.10E-04
GO:48746: smooth muscle fiber development	2	0.00973	2	0.671	2.10E-04
GO:19371: cyclooxygenase pathway	2	0.00973	2	0.671	2.10E-04
GO:31034: myosin filament assembly	2	0.00973	2	0.671	2.10E-04
GO:1660: fever	2	0.00973	2	0.671	2.10E-04
GO:51549: positive regulation of keratinocyte migration	2	0.00973	2	0.671	2.10E-04
GO:51547: regulation of keratinocyte migration	2	0.00973	2	0.671	2.10E-04
GO:31033: myosin filament assembly or disassembly	2	0.00973	2	0.671	2.10E-04
GO:1822: kidney development	38	0.185	5	1.678	2.10E-04
GO:48627: myoblast development	21	0.102	4	1.342	2.14E-04
GO:48637: skeletal muscle development	60	0.292	6	2.013	2.30E-04
GO:31214: biomineral formation	112	0.545	8	2.685	2.30E-04
GO:1503: ossification	112	0.545	8	2.685	2.30E-04
GO:6568: tryptophan metabolism	9	0.0438	3	1.007	2.38E-04
GO:1655: urogenital system development	40	0.195	5	1.678	2.69E-04
GO:46849: bone remodeling	119	0.579	8	2.685	3.48E-04
GO:42402: biogenic amine catabolism	11	0.0535	3	1.007	4.57E-04
GO:6941: striated muscle contraction	45	0.219	5	1.678	4.71E-04
GO:19722: calcium-mediated signaling	45	0.219	5	1.678	4.71E-04
GO:30574: collagen catabolism	26	0.127	4	1.342	5.04E-04
GO:50727: regulation of inflammatory response	26	0.127	4	1.342	5.04E-04
GO:50793: regulation of development	265	1.29	12	4.027	5.21E-04
GO:7178: transmembrane receptor protein serine/threonine kinase signaling pathway	98	0.477	7	2.349	5.65E-04
GO:6586: indolalkylamine metabolism	12	0.0584	3	1.007	6.03E-04
GO:42430: indole and derivative metabolism	12	0.0584	3	1.007	6.03E-04
GO:42434: indole derivative metabolism	12	0.0584	3	1.007	6.03E-04
GO:30335: positive regulation of cell migration	12	0.0584	3	1.007	6.03E-04
GO:7435: salivary gland morphogenesis	3	0.0146	2	0.671	6.23E-04
GO:45079: negative regulation of chemokine biosynthesis	3	0.0146	2	0.671	6.23E-04
GO:51546: keratinocyte migration	3	0.0146	2	0.671	6.23E-04
GO:6959: humoral immune response	131	0.638	8	2.685	6.61E-04
GO:43062: extracellular structure organization and biogenesis	132	0.642	8	2.685	6.95E-04
GO:30198: extracellular matrix organization and biogenesis	132	0.642	8	2.685	6.95E-04
GO:8283: cell proliferation	1009	4.911	28	9.396	8.04E-04
GO:30509: BMP signaling pathway	30	0.146	4	1.342	8.83E-04

GO:30855: epithelial cell differentiation	52	0.253	5	1.678	9.23E-04
GO:6937: regulation of muscle contraction	52	0.253	5	1.678	9.23E-04
GO:43067: regulation of programmed cell death	685	3.334	21	7.047	1.09E-03
GO:9187: cyclic nucleotide metabolism	54	0.263	5	1.678	1.10E-03
GO:42364: water-soluble vitamin biosynthesis	32	0.156	4	1.342	1.13E-03
GO:45663: positive regulation of myoblast differentiation	4	0.0195	2	0.671	1.23E-03
GO:7431: salivary gland development	4	0.0195	2	0.671	1.23E-03
GO:35313: wound healing, spreading of epidermal cells	4	0.0195	2	0.671	1.23E-03
GO:6769: nicotinamide metabolism	33	0.161	4	1.342	1.27E-03
GO:48523: negative regulation of cellular process	1518	7.388	37	12.42	1.33E-03
GO:43119: positive regulation of physiological process	1150	5.597	30	10.07	1.39E-03
GO:50776: regulation of immune response	182	0.886	9	3.02	1.40E-03
GO:9110: vitamin biosynthesis	34	0.165	4	1.342	1.43E-03
GO:270: peptidoglycan metabolism	34	0.165	4	1.342	1.43E-03
GO:30516: regulation of axon extension	16	0.0779	3	1.007	1.47E-03
GO:9074: aromatic amino acid family catabolism	16	0.0779	3	1.007	1.47E-03
GO:46457: prostanoid biosynthesis	16	0.0779	3	1.007	1.47E-03
GO:1516: prostaglandin biosynthesis	16	0.0779	3	1.007	1.47E-03
GO:50900: immune cell migration	36	0.175	4	1.342	1.77E-03
GO:2009: morphogenesis of an epithelium	88	0.428	6	2.013	1.77E-03
GO:43066: negative regulation of apoptosis	306	1.489	12	4.027	1.80E-03
GO:8284: positive regulation of cell proliferation	268	1.304	11	3.691	1.94E-03
GO:43069: negative regulation of programmed cell death	309	1.504	12	4.027	1.95E-03
GO:6787: porphyrin catabolism	5	0.0243	2	0.671	2.04E-03
GO:8211: glucocorticoid metabolism	5	0.0243	2	0.671	2.04E-03
GO:31649: heat generation	5	0.0243	2	0.671	2.04E-03
GO:42219: amino acid derivative catabolism	18	0.0876	3	1.007	2.10E-03
GO:42036: negative regulation of cytokine biosynthesis	18	0.0876	3	1.007	2.10E-03
GO:30199: collagen fibril organization	18	0.0876	3	1.007	2.10E-03
GO:1818: negative regulation of cytokine production	18	0.0876	3	1.007	2.10E-03
GO:42981: regulation of apoptosis	676	3.29	20	6.711	2.16E-03
GO:6939: smooth muscle contraction	38	0.185	4	1.342	2.17E-03
GO:7167: enzyme linked receptor protein signaling pathway	402	1.956	14	4.698	2.34E-03
GO:7565: pregnancy	93	0.453	6	2.013	2.35E-03
GO:51272: positive regulation of cell motility	19	0.0925	3	1.007	2.46E-03
GO:40017: positive regulation of locomotion	19	0.0925	3	1.007	2.46E-03
GO:48520: positive regulation of behavior	19	0.0925	3	1.007	2.46E-03
GO:19362: pyridine nucleotide metabolism	40	0.195	4	1.342	2.63E-03
GO:48675: axon extension	20	0.0973	3	1.007	2.87E-03
GO:7520: myoblast fusion	6	0.0292	2	0.671	3.03E-03
GO:7263: nitric oxide mediated signal transduction	6	0.0292	2	0.671	3.03E-03
GO:50777: negative regulation of immune response	43	0.209	4	1.342	3.43E-03
GO:48468: cell development	420	2.044	14	4.698	3.47E-03
GO:42640: anagen	7	0.0341	2	0.671	4.20E-03
GO:48251: elastic fiber assembly	7	0.0341	2	0.671	4.20E-03
GO:48609: reproductive organismal physiological process	139	0.676	7	2.349	4.22E-03
GO:19439: aromatic compound catabolism	23	0.112	3	1.007	4.31E-03
GO:302: response to reactive oxygen species	23	0.112	3	1.007	4.31E-03
GO:50876: reproductive physiological process	140	0.681	7	2.349	4.39E-03
GO:6916: anti-apoptosis	259	1.261	10	3.356	4.71E-03
GO:48519: negative regulation of biological process	1638	7.972	37	12.42	4.79E-03
GO:7411: axon guidance	109	0.53	6	2.013	5.15E-03
GO:51242: positive regulation of cellular physiological process	1099	5.349	27	9.06	5.39E-03
GO:6692: prostanoid metabolism	25	0.122	3	1.007	5.48E-03

GO:6693: prostaglandin metabolism	25	0.122	3	1.007	5.48E-03
GO:45661: regulation of myoblast differentiation	8	0.0389	2	0.671	5.54E-03
GO:1659: thermoregulation	8	0.0389	2	0.671	5.54E-03
GO:6807: nitrogen compound metabolism	588	2.862	17	5.705	5.68E-03
GO:6733: oxidoreduction coenzyme metabolism	50	0.243	4	1.342	5.91E-03
GO:8217: regulation of blood pressure	51	0.248	4	1.342	6.35E-03
GO:48522: positive regulation of cellular process	1281	6.234	30	10.07	6.67E-03
GO:45595: regulation of cell differentiation	190	0.925	8	2.685	6.69E-03
GO:9891: positive regulation of biosynthesis	82	0.399	5	1.678	6.82E-03
GO:42035: regulation of cytokine biosynthesis	82	0.399	5	1.678	6.82E-03
GO:1943: hair follicle maturation	9	0.0438	2	0.671	7.06E-03
GO:1658: ureteric bud branching	9	0.0438	2	0.671	7.06E-03
GO:35272: exocrine system development	9	0.0438	2	0.671	7.06E-03
GO:46890: regulation of lipid biosynthesis	9	0.0438	2	0.671	7.06E-03
GO:7519: striated muscle development	118	0.574	6	2.013	7.53E-03
GO:1816: cytokine production	118	0.574	6	2.013	7.53E-03
GO:46942: carboxylic acid transport	119	0.579	6	2.013	7.83E-03
GO:15849: organic acid transport	120	0.584	6	2.013	8.15E-03
GO:50770: regulation of axonogenesis	29	0.141	3	1.007	8.35E-03
GO:42475: odontogenesis (sensu Vertebrata)	29	0.141	3	1.007	8.35E-03
GO:1656: metanephros development	29	0.141	3	1.007	8.35E-03
GO:30595: immune cell chemotaxis	29	0.141	3	1.007	8.35E-03
GO:12501: programmed cell death	1030	5.013	25	8.389	8.44E-03
GO:6811: ion transport	1086	5.285	26	8.725	8.63E-03
GO:19369: arachidonic acid metabolism	10	0.0487	2	0.671	8.74E-03
GO:42303: molting cycle	10	0.0487	2	0.671	8.74E-03
GO:42633: hair cycle	10	0.0487	2	0.671	8.74E-03
GO:7274: neuromuscular synaptic transmission	10	0.0487	2	0.671	8.74E-03
GO:35295: tube development	123	0.599	6	2.013	9.15E-03
GO:17148: negative regulation of protein biosynthesis	30	0.146	3	1.007	9.18E-03
GO:42089: cytokine biosynthesis	89	0.433	5	1.678	9.57E-03
GO:43118: negative regulation of physiological process	1429	6.955	32	10.74	9.65E-03
GO:8016: regulation of heart contraction rate	58	0.282	4	1.342	9.96E-03

Table S-3. GO annotation enrichment of deregulated (4-fold, P<0.05) genes in IMR-90 cells by SPON2 knockdown (120h)

Category	Genes in Category	% of Genes in Category	Genes in List in Category	% of Genes in List in Category	p-Value
GO:6954: inflammatory response	343	1.669	26	12.04	3.47E-15
GO:9613: response to pest, pathogen or parasite	669	3.256	34	15.74	2.32E-14
GO:42330: taxis	200	0.973	20	9.259	3.44E-14
GO:6935: chemotaxis	200	0.973	20	9.259	3.44E-14
GO:43207: response to external biotic stimulus	731	3.558	34	15.74	2.93E-13
GO:9611: response to wounding	626	3.047	31	14.35	7.21E-13
GO:6955: immune response	1232	5.996	43	19.91	2.57E-12
GO:50874: organismal physiological process	2995	14.58	71	32.87	8.14E-12
GO:9605: response to external stimulus	1090	5.305	39	18.06	1.52E-11
GO:6952: defense response	1352	6.58	43	19.91	5.24E-11
GO:9607: response to biotic stimulus	1407	6.848	43	19.91	1.85E-10
GO:51216: cartilage development	44	0.214	8	3.704	1.68E-08
GO:42221: response to chemical stimulus	644	3.134	25	11.57	2.03E-08
GO:1837: epithelial to mesenchymal transition	19	0.0925	6	2.778	3.05E-08
GO:6950: response to stress	1600	7.787	41	18.98	8.12E-08
GO:7154: cell communication	5765	28.06	97	44.91	8.24E-08
GO:7267: cell-cell signaling	823	4.005	27	12.5	1.67E-07
GO:50727: regulation of inflammatory response	26	0.127	6	2.778	2.43E-07
GO:48513: organ development	1356	6.6	36	16.67	2.46E-07
GO:9888: tissue development	404	1.966	18	8.333	2.99E-07
GO:30595: immune cell chemotaxis	29	0.141	6	2.778	4.89E-07
GO:7155: cell adhesion	1053	5.125	30	13.89	6.65E-07
GO:1501: skeletal development	304	1.48	15	6.944	8.47E-07
GO:9628: response to abiotic stimulus	803	3.908	25	11.57	1.29E-06
GO:51239: regulation of organismal physiological process	320	1.557	15	6.944	1.61E-06
GO:50900: immune cell migration	36	0.175	6	2.778	1.88E-06
GO:31214: biomineral formation	112	0.545	9	4.167	2.86E-06
GO:1503: ossification	112	0.545	9	4.167	2.86E-06
GO:6569: tryptophan catabolism	4	0.0195	3	1.389	4.55E-06
GO:46218: indolalkylamine catabolism	4	0.0195	3	1.389	4.55E-06
GO:42436: indole derivative catabolism	4	0.0195	3	1.389	4.55E-06
GO:46849: bone remodeling	119	0.579	9	4.167	4.72E-06
GO:7178: transmembrane receptor protein serine/threonine kinase signaling pathway	98	0.477	8	3.704	9.16E-06
GO:30509: BMP signaling pathway	30	0.146	5	2.315	1.41E-05
GO:30855: epithelial cell differentiation	52	0.253	6	2.778	1.71E-05
GO:42127: regulation of cell proliferation	606	2.949	19	8.796	2.36E-05
GO:6928: cell motility	566	2.755	18	8.333	3.23E-05
GO:51674: localization of cell	566	2.755	18	8.333	3.23E-05
GO:40011: locomotion	569	2.769	18	8.333	3.46E-05
GO:42033: chemokine biosynthesis	7	0.0341	3	1.389	3.89E-05
GO:50755: chemokine metabolism	7	0.0341	3	1.389	3.89E-05
GO:2009: morphogenesis of an epithelium	88	0.428	7	3.241	3.97E-05
GO:50896: response to stimulus	3247	15.8	57	26.39	4.32E-05
GO:30154: cell differentiation	1371	6.673	31	14.35	4.51E-05

GO:6953: acute-phase response	38	0.185	5	2.315	4.64E-05
GO:7166: cell surface receptor linked signal transduction	1994	9.705	40	18.52	4.85E-05
GO:51094: positive regulation of development	93	0.453	7	3.241	5.67E-05
GO:7275: development	3595	17.5	61	28.24	5.83E-05
GO:45597: positive regulation of cell differentiation	65	0.316	6	2.778	6.20E-05
GO:8285: negative regulation of cell proliferation	298	1.45	12	5.556	7.91E-05
GO:7626: locomotory behavior	607	2.954	18	8.333	7.94E-05
GO:7610: behavior	728	3.543	20	9.259	9.06E-05
GO:6568: tryptophan metabolism	9	0.0438	3	1.389	9.19E-05
GO:30241: muscle thick filament assembly	2	0.00973	2	0.926	1.10E-04
GO:48745: smooth muscle development	2	0.00973	2	0.926	1.10E-04
GO:48746: smooth muscle fiber development	2	0.00973	2	0.926	1.10E-04
GO:19371: cyclooxygenase pathway	2	0.00973	2	0.926	1.10E-04
GO:31034: myosin filament assembly	2	0.00973	2	0.926	1.10E-04
GO:1660: fever	2	0.00973	2	0.926	1.10E-04
GO:31033: myosin filament assembly or disassembly	2	0.00973	2	0.926	1.10E-04
GO:50776: regulation of immune response	182	0.886	9	4.167	1.36E-04
GO:16477: cell migration	320	1.557	12	5.556	1.54E-04
GO:42402: biogenic amine catabolism	11	0.0535	3	1.389	1.78E-04
GO:6586: indolalkylamine metabolism	12	0.0584	3	1.389	2.35E-04
GO:42430: indole and derivative metabolism	12	0.0584	3	1.389	2.35E-04
GO:42434: indole derivative metabolism	12	0.0584	3	1.389	2.35E-04
GO:16338: calcium-independent cell-cell adhesion	31	0.151	4	1.852	2.99E-04
GO:8015: circulation	203	0.988	9	4.167	3.08E-04
GO:7435: salivary gland morphogenesis	3	0.0146	2	0.926	3.28E-04
GO:50902: leukocyte adhesive activation	3	0.0146	2	0.926	3.28E-04
GO:30431: sleep	3	0.0146	2	0.926	3.28E-04
GO:51040: regulation of calcium-independent cell-cell adhesion	3	0.0146	2	0.926	3.28E-04
GO:51041: positive regulation of calcium-independent cell-cell adhesion	3	0.0146	2	0.926	3.28E-04
GO:45785: positive regulation of cell adhesion	14	0.0681	3	1.389	3.83E-04
GO:7165: signal transduction	4610	22.44	70	32.41	4.53E-04
GO:9072: aromatic amino acid family metabolism	36	0.175	4	1.852	5.37E-04
GO:30516: regulation of axon extension	16	0.0779	3	1.389	5.80E-04
GO:9074: aromatic amino acid family catabolism	16	0.0779	3	1.389	5.80E-04
GO:46457: prostanoid biosynthesis	16	0.0779	3	1.389	5.80E-04
GO:1516: prostaglandin biosynthesis	16	0.0779	3	1.389	5.80E-04
GO:51240: positive regulation of organismal physiological process	137	0.667	7	3.241	6.24E-04
GO:7431: salivary gland development	4	0.0195	2	0.926	6.51E-04
GO:6030: chitin metabolism	17	0.0827	3	1.389	6.99E-04
GO:30593: neutrophil chemotaxis	17	0.0827	3	1.389	6.99E-04
GO:42219: amino acid derivative catabolism	18	0.0876	3	1.389	8.32E-04
GO:8283: cell proliferation	1009	4.911	22	10.19	9.96E-04
GO:8211: glucocorticoid metabolism	5	0.0243	2	0.926	1.08E-03
GO:31649: heat generation	5	0.0243	2	0.926	1.08E-03
GO:48675: axon extension	20	0.0973	3	1.389	1.14E-03
GO:48628: myoblast maturation	20	0.0973	3	1.389	1.14E-03
GO:48627: myoblast development	21	0.102	3	1.389	1.32E-03
GO:30278: regulation of ossification	22	0.107	3	1.389	1.52E-03
GO:45073: regulation of chemokine biosynthesis	6	0.0292	2	0.926	1.60E-03
GO:7263: nitric oxide mediated signal transduction	6	0.0292	2	0.926	1.60E-03
GO:19439: aromatic compound catabolism	23	0.112	3	1.389	1.74E-03
GO:302: response to reactive oxygen species	23	0.112	3	1.389	1.74E-03

GO:6690: icosanoid metabolism	50	0.243	4	1.852	1.87E-03
GO:50778: positive regulation of immune response	123	0.599	6	2.778	1.93E-03
GO:30155: regulation of cell adhesion	84	0.409	5	2.315	1.93E-03
GO:8544: epidermis development	167	0.813	7	3.241	1.97E-03
GO:50793: regulation of development	265	1.29	9	4.167	2.03E-03
GO:6692: prostanoid metabolism	25	0.122	3	1.389	2.22E-03
GO:6693: prostaglandin metabolism	25	0.122	3	1.389	2.22E-03
GO:46850: regulation of bone remodeling	25	0.122	3	1.389	2.22E-03
GO:42640: anagen	7	0.0341	2	0.926	2.23E-03
GO:48251: elastic fiber assembly	7	0.0341	2	0.926	2.23E-03
GO:30574: collagen catabolism	26	0.127	3	1.389	2.49E-03
GO:48523: negative regulation of cellular process	1518	7.388	28	12.96	2.61E-03
GO:6959: humoral immune response	131	0.638	6	2.778	2.65E-03
GO:1659: thermoregulation	8	0.0389	2	0.926	2.95E-03
GO:48247: lymphocyte chemotaxis	8	0.0389	2	0.926	2.95E-03
GO:45123: cellular extravasation	8	0.0389	2	0.926	2.95E-03
GO:19221: cytokine and chemokine mediated signaling pathway	57	0.277	4	1.852	3.04E-03
GO:50770: regulation of axonogenesis	29	0.141	3	1.389	3.42E-03
GO:42475: odontogenesis (sensu Vertebrata)	29	0.141	3	1.389	3.42E-03
GO:7398: ectoderm development	187	0.91	7	3.241	3.71E-03
GO:1943: hair follicle maturation	9	0.0438	2	0.926	3.77E-03
GO:35272: exocrine system development	9	0.0438	2	0.926	3.77E-03
GO:9435: NAD biosynthesis	9	0.0438	2	0.926	3.77E-03
GO:19674: NAD metabolism	9	0.0438	2	0.926	3.77E-03
GO:6807: nitrogen compound metabolism	588	2.862	14	6.481	3.87E-03
GO:48519: negative regulation of biological process	1638	7.972	29	13.43	3.94E-03
GO:7517: muscle development	243	1.183	8	3.704	4.29E-03
GO:19369: arachidonic acid metabolism	10	0.0487	2	0.926	4.68E-03
GO:42303: molting cycle	10	0.0487	2	0.926	4.68E-03
GO:42633: hair cycle	10	0.0487	2	0.926	4.68E-03
GO:43119: positive regulation of physiological process	1150	5.597	22	10.19	4.94E-03
GO:30216: keratinocyte differentiation	34	0.165	3	1.389	5.39E-03
GO:6044: N-acetylglucosamine metabolism	34	0.165	3	1.389	5.39E-03
GO:270: peptidoglycan metabolism	34	0.165	3	1.389	5.39E-03
GO:46456: icosanoid biosynthesis	35	0.17	3	1.389	5.85E-03
GO:7411: axon guidance	109	0.53	5	2.315	5.92E-03
GO:50767: regulation of neurogenesis	69	0.336	4	1.852	6.03E-03
GO:6041: glucosamine metabolism	36	0.175	3	1.389	6.34E-03
GO:48168: regulation of neuronal synaptic plasticity	12	0.0584	2	0.926	6.77E-03
GO:9792: embryonic development (sensu Metazoa)	159	0.774	6	2.778	6.79E-03
GO:42476: odontogenesis	37	0.18	3	1.389	6.84E-03
GO:1822: kidney development	38	0.185	3	1.389	7.37E-03
GO:6939: smooth muscle contraction	38	0.185	3	1.389	7.37E-03
GO:8284: positive regulation of cell proliferation	268	1.304	8	3.704	7.61E-03
GO:6182: cGMP biosynthesis	13	0.0633	2	0.926	7.95E-03
GO:46068: cGMP metabolism	13	0.0633	2	0.926	7.95E-03
GO:7271: synaptic transmission, cholinergic	13	0.0633	2	0.926	7.95E-03
GO:1655: urogenital system development	40	0.195	3	1.389	8.50E-03
GO:30239: myofibril assembly	14	0.0681	2	0.926	9.21E-03
GO:50729: positive regulation of inflammatory response	14	0.0681	2	0.926	9.21E-03
GO:30514: negative regulation of BMP signaling pathway	14	0.0681	2	0.926	9.21E-03

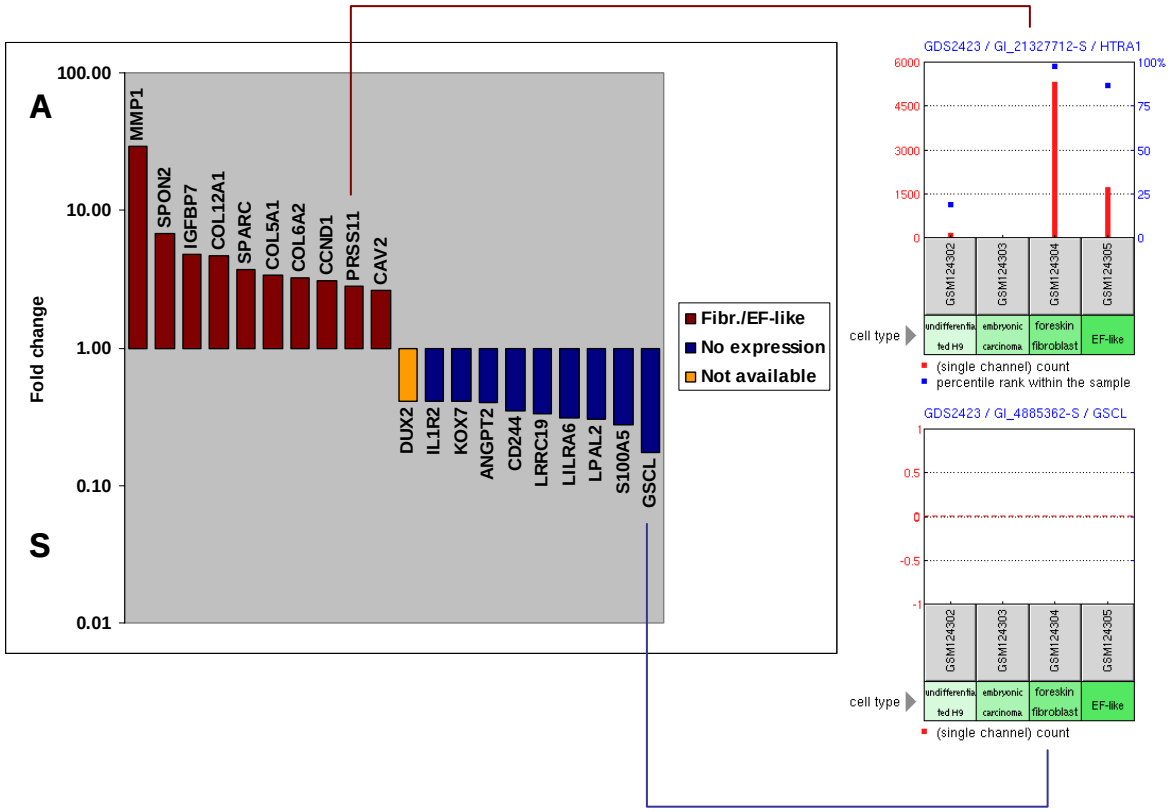


Figure S-1. Comparison of SCLC expression (top 10 annotated genes) with GDS2423 (GEO).

The left plot shows expression level of differentially expressed genes in SCLC-A and –S (Table I-52 and Table I-53, respectively). The bars are coloured by theirs expression level in GDS2423 (right plots). The right plots are expression profiles of *HTRA1* (or *PRSS11*) and *GSCL* genes obtained from GDS2423.

SUPPLEMENTARY CD

S1.txt	Genes upregulated in SCLC A. Filtered by volcano plot, 2-fold change, P<0.05.
S2.txt	Genes upregulated in SCLC S. Filtered by volcano plot, 2-fold change, P<0.05.
S3.txt	Genes upregulated in IMR-90 cells by SPON2 knockdown (72h). Filtered by volcano plot, 4-fold change, P<0.05.
S4.txt	Genes downregulated in IMR-90 cells by SPON2 knockdown (72h). Filtered by volcano plot, 4-fold change, P<0.05.
S5.txt	Genes upregulated in IMR-90 cells by SPON2 knockdown (96h). Filtered by volcano plot, 4-fold change, P<0.05.
S6.txt	Genes downregulated in IMR-90 cells by SPON2 knockdown (96h). Filtered by volcano plot, 4-fold change, P<0.05.
S7.txt	Genes upregulated in IMR-90 cells by SPON2 knockdown (120h). Filtered by volcano plot, 4-fold change, P<0.05.
S8.txt	Genes downregulated in IMR-90 cells by SPON2 knockdown (120h). Filtered by volcano plot, 4-fold change, P<0.05.
S9.txt	Commonly deregulated genes in IMR-90 cells by SPON2 knockdown (72h, 96h and 120h). Filtered by volcano plot, 4-fold change, P<0.05.
Network1.cys	Interaction networks of *(a)* co-immunoprecipitated proteins listed in Table I-49, *(b)* genes deregulated by SPON2 siRNA knockdown (120h), and *(c)* merged (a) and (b) networks.
Network2.cys	Interaction network using a random protein list.

REFERENCES

1. Kardinal, C. G. & Yarbro, J. W. A conceptual history of cancer. Semin Oncol 6, 396-408 (1979).
2. Gallucci, B. B. Selected concepts of cancer as a disease: from the Greeks to 1900. Oncol Nurs Forum 12, 67-71 (1985).
3. Alison, M. R. The Cancer Handbook (John Wiley and Sons Ltd, 2005).
4. Briasoulis, E. & Pavlidis, N. Cancer of Unknown Primary Origin. Oncologist 2, 142-152 (1997).
5. Nordling, C. O. A new theory on cancer-inducing mechanism. Br J Cancer 7, 68-72 (1953).
6. Knudson, A. G., Jr. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 68, 820-3 (1971).
7. Slaughter, D. P., Southwick, H. W. & Smejkal, W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer 6, 963-8 (1953).
8. Tom Strachan, A. P. R. Human Molecular Genetics (Garland Science, 2003).
9. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. Cell 100, 57-70 (2000).
10. RLK, V. Cellular pathology (Berlin, 1858).
11. J.Cohnheim. Ueber Entzundung und Eiterung. Path Anat Physiol Klin Med, 1-79 (1867).
12. Li, H. C., Stoicov, C., Rogers, A. B. & Houghton, J. Stem cells and cancer: evidence for bone marrow stem cells in epithelial cancers. World J Gastroenterol 12, 363-71 (2006).
13. Li, L. & Neaves, W. B. Normal stem cells and cancer stem cells: the niche matters. Cancer Res 66, 4553-7 (2006).
14. Miller, S. J., Lavker, R. M. & Sun, T. T. Interpreting epithelial cancer biology in the context of stem cells: tumor properties and therapeutic implications. Biochim Biophys Acta 1756, 25-52 (2005).
15. Lapidot, T. et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature 367, 645-8 (1994).
16. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 100, 3983-8 (2003).
17. Galli, R. et al. Skeletal myogenic potential of human and mouse neural stem cells. Nat Neurosci 3, 986-91 (2000).
18. Bjerkvig, R., Tysnes, B. B., Aboody, K. S., Najbauer, J. & Terzis, A. J. Opinion: the origin of the cancer stem cell: current controversies and new insights. Nat Rev Cancer 5, 899-904 (2005).
19. Ogle, B. M., Cascalho, M. & Platt, J. L. Biological implications of cell fusion. Nat Rev Mol Cell Biol 6, 567-75 (2005).
20. Duelli, D. & Lazebnik, Y. Cell fusion: a hidden enemy? Cancer Cell 3, 445-8 (2003).
21. Holmgren, L. et al. Horizontal transfer of DNA by the uptake of apoptotic bodies. Blood 93, 3956-63 (1999).
22. Rodriguez-Boulan, E. & Nelson, W. J. Morphogenesis of the polarized epithelial cell phenotype. Science 245, 718-25 (1989).
23. Srivastava, M. et al. The Trichoplax genome and the nature of placozoans. Nature 454, 955-60 (2008).

24. Furuse, M. & Tsukita, S. Claudins in occluding junctions of humans and flies. Trends Cell Biol 16, 181-8 (2006).
25. Schneeberger, E. E. & Lynch, R. D. The tight junction: a multifunctional complex. Am J Physiol Cell Physiol 286, C1213-28 (2004).
26. Lodish, H., Berk, A., Matsudaira, P. Molecular Cell Biology (Palgrave Macmillan, 2004).
27. Rela, L. & Szczupak, L. Gap junctions: their importance for the dynamics of neural circuits. Mol Neurobiol 30, 341-57 (2004).
28. Yin, T. & Green, K. J. Regulation of desmosome assembly and adhesion. Semin Cell Dev Biol 15, 665-77 (2004).
29. Mueller, M. M. & Fusenig, N. E. Friends or foes - bipolar effects of the tumour stroma in cancer. Nat Rev Cancer 4, 839-49 (2004).
30. Billingham, R. E., Orr, J. W. & Woodhouse, D. L. Transplantation of skin components during chemical carcinogenesis with 20-methylcholanthrene. Br J Cancer 5, 417-32 (1951).
31. Barcellos-Hoff, M. H. & Ravani, S. A. Irradiated mammary gland stroma promotes the expression of tumorigenic potential by unirradiated epithelial cells. Cancer Res 60, 1254-60 (2000).
32. Roskelley, C. D. & Bissell, M. J. The dominance of the microenvironment in breast and ovarian cancer. Semin Cancer Biol 12, 97-104 (2002).
33. Moinfar, F. et al. Concurrent and independent genetic alterations in the stromal and epithelial cells of mammary carcinoma: implications for tumorigenesis. Cancer Res 60, 2562-6 (2000).
34. Werner, S. & Grose, R. Regulation of wound healing by growth factors and cytokines. Physiol Rev 83, 835-70 (2003).
35. Medzhitov, R. Origin and physiological roles of inflammation. Nature 454, 428-35 (2008).
36. Dvorak, H. F. Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. N Engl J Med 315, 1650-9 (1986).
37. De Marzo, A. M. et al. Inflammation in prostate carcinogenesis. Nat Rev Cancer 7, 256-69 (2007).
38. Konturek, P. C., Konturek, S. J. & Brzozowski, T. Gastric cancer and Helicobacter pylori infection. J Physiol Pharmacol 57 Suppl 3, 51-65 (2006).
39. Azad, N., Rojanasakul, Y. & Vallyathan, V. Inflammation and lung cancer: roles of reactive oxygen/nitrogen species. J Toxicol Environ Health B Crit Rev 11, 1-15 (2008).
40. Coussens, L. M. & Werb, Z. Inflammation and cancer. Nature 420, 860-7 (2002).
41. Balkwill, F. Cancer and the chemokine network. Nat Rev Cancer 4, 540-50 (2004).
42. Rinn, J. L., Bondre, C., Gladstone, H. B., Brown, P. O. & Chang, H. Y. Anatomic demarcation by positional variation in fibroblast gene expression programs. PLoS Genet 2, e119 (2006).
43. Li, H., Fan, X. & Houghton, J. Tumor microenvironment: the role of the tumor stroma in cancer. J Cell Biochem 101, 805-15 (2007).
44. Kalluri, R. & Zeisberg, M. Fibroblasts in cancer. Nat Rev Cancer 6, 392-401 (2006).
45. Tarin, D. & Croft, C. B. Ultrastructural features of wound healing in mouse skin. J Anat 105, 189-90 (1969).
46. Nguyen, D. X. & Massague, J. Genetic determinants of cancer metastasis. Nat Rev Genet 8, 341-52 (2007).
47. Chambers, A. F., Groom, A. C. & MacDonald, I. C. Dissemination and growth of cancer cells in metastatic sites. Nat Rev Cancer 2, 563-72 (2002).

48. Mehlen, P. & Puisieux, A. Metastasis: a question of life or death. Nat Rev Cancer 6, 449-58 (2006).
49. Jakobisiak, M., Lasek, W. & Golab, J. Natural mechanisms protecting against cancer. Immunol Lett 90, 103-22 (2003).
50. Fidler, I. J. & Nicolson, G. L. Fate of recirculating B16 melanoma metastatic variant cells in parabiotic syngeneic recipients. J Natl Cancer Inst 58, 1867-72 (1977).
51. Fidler, I. J. Metastasis: guantitative analysis of distribution and fate of tumor embolilabeled with 125 I-5-iodo-2'-deoxyuridine. J Natl Cancer Inst 45, 773-82 (1970).
52. Sahai, E. Illuminating the metastatic process. Nat Rev Cancer 7, 737-49 (2007).
53. Paget, S. The distribution of secondary growths in cancer of the breast. 1889. Cancer Metastasis Rev 8, 98-101 (1989).
54. Friedl, P. & Wolf, K. Tumour-cell invasion and migration: diversity and escape mechanisms. Nat Rev Cancer 3, 362-74 (2003).
55. Ridley, A. J. et al. Cell migration: integrating signals from front to back. Science 302, 1704-9 (2003).
56. Horwitz, A. R. & Parsons, J. T. Cell migration--movin' on. Science 286, 1102-3 (1999).
57. Lauffenburger, D. A. & Horwitz, A. F. Cell migration: a physically integrated molecular process. Cell 84, 359-69 (1996).
58. Wolf, K. et al. Compensation mechanism in tumor cell migration: mesenchymal-amoeboid transition after blocking of pericellular proteolysis. J Cell Biol 160, 267-77 (2003).
59. Paulus, W., Baur, I., Beutler, A. S. & Reeves, S. A. Diffuse brain invasion of glioma cells requires beta 1 integrins. Lab Invest 75, 819-26 (1996).
60. Polette, M. et al. Association of fibroblastoid features with the invasive phenotype in human bronchial cancer cell lines. Clin Exp Metastasis 16, 105-12 (1998).
61. Tester, A. M., Ruangpanit, N., Anderson, R. L. & Thompson, E. W. MMP-9 secretion and MMP-2 activation distinguish invasive and metastatic sublines of a mouse mammary carcinoma system showing epithelial-mesenchymal transition traits. Clin Exp Metastasis 18, 553-60 (2000).
62. Gilbert, S. F., Singer, S. R., Tyler, M. S. & Kozlowski, R. N. Developmental biology (Sinauer Associates, Sunderland, Mass., 2000).
63. Thiery, J. P. & Sleeman, J. P. Complex networks orchestrate epithelial-mesenchymal transitions. Nat Rev Mol Cell Biol 7, 131-42 (2006).
64. Hay, E. D. Organization and fine structure of epithelium and mesenchyme in the developing chick embryo. In *Epithelial-Mesenchymal Interactions* (ed. RE, F. R. B.) (Williams & Wilkins, Baltimore, 1968).
65. Greenburg, G. & Hay, E. D. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. J Cell Biol 95, 333-9 (1982).
66. Greenburg, G. & Hay, E. D. Cytodifferentiation and tissue phenotype change during transformation of embryonic lens epithelium to mesenchyme-like cells in vitro. Dev Biol 115, 363-79 (1986).
67. Hugo, H. et al. Epithelial--mesenchymal and mesenchymal--epithelial transitions in carcinoma progression. J Cell Physiol 213, 374-83 (2007).
68. van Es, J. H., Barker, N. & Clevers, H. You Wnt some, you lose some: oncogenes in the Wnt signaling pathway. Curr Opin Genet Dev 13, 28-33 (2003).
69. Lee, J. M., Dedhar, S., Kalluri, R. & Thompson, E. W. The epithelial-mesenchymal transition: new insights in signaling, development, and disease. J Cell Biol 172, 973-81 (2006).

70. Nakaya, Y. & Sheng, G. Epithelial to mesenchymal transition during gastrulation: an embryological view. Dev Growth Differ 50, 755-66 (2008).
71. Yang, J. & Weinberg, R. A. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Dev Cell 14, 818-29 (2008).
72. Hay, E. D. & Zuk, A. Transformations between epithelium and mesenchyme: normal, pathological, and experimentally induced. Am J Kidney Dis 26, 678-90 (1995).
73. Davies, J. A. Mesenchyme to epithelium transition during development of the mammalian kidney tubule. Acta Anat (Basel) 156, 187-201 (1996).
74. Davies, J. A., Perera, A. D. & Walker, C. L. Mechanisms of epithelial development and neoplasia in the metanephric kidney. Int J Dev Biol 43, 473-8 (1999).
75. Kalluri, R. & Neilson, E. G. Epithelial-mesenchymal transition and its implications for fibrosis. J Clin Invest 112, 1776-84 (2003).
76. Rastaldi, M. P. Epithelial-mesenchymal transition and its implications for the development of renal tubulointerstitial fibrosis. J Nephrol 19, 407-12 (2006).
77. Li, M. X. & Liu, B. C. Epithelial to mesenchymal transition in the progression of tubulointerstitial fibrosis. Chin Med J (Engl) 120, 1925-30 (2007).
78. Saika, S. et al. Epithelial-mesenchymal transition as a therapeutic target for prevention of ocular tissue fibrosis. Endocr Metab Immune Disord Drug Targets 8, 69-76 (2008).
79. Bedi, S., Vidyasagar, A. & Djamali, A. Epithelial-to-mesenchymal transition and chronic allograft tubulointerstitial fibrosis. Transplant Rev (Orlando) 22, 1-5 (2008).
80. Casaroli-Marano, R. P., Pagan, R. & Vilaro, S. Epithelial-mesenchymal transition in proliferative vitreoretinopathy: intermediate filament protein expression in retinal pigment epithelial cells. Invest Ophthalmol Vis Sci 40, 2062-72 (1999).
81. Flanders, K. C. Smad3 as a mediator of the fibrotic response. Int J Exp Pathol 85, 47-64 (2004).
82. Saika, S. et al. Smad3 is required for dedifferentiation of retinal pigment epithelium following retinal detachment in mice. Lab Invest 84, 1245-58 (2004).
83. Lee, H., O'Meara, S. J., O'Brien, C. & Kane, R. The role of gremlin, a BMP antagonist, and epithelial-to-mesenchymal transition in proliferative vitreoretinopathy. Invest Ophthalmol Vis Sci 48, 4291-9 (2007).
84. Zvaifler, N. J. Relevance of the stroma and epithelial-mesenchymal transition (EMT) for the rheumatic diseases. Arthritis Res Ther 8, 210 (2006).
85. Lee, Y. H., Albig, A. R., Regner, M., Schiemann, B. J. & Schiemann, W. P. Fibulin-5 initiates epithelial-mesenchymal transition (EMT) and enhances EMT induced by TGF-beta in mammary epithelial cells via a MMP-dependent mechanism. Carcinogenesis 29, 2243-51 (2008).
86. Boyer, B., Tucker, G. C., Valles, A. M., Franke, W. W. & Thiery, J. P. Rearrangements of desmosomal and cytoskeletal proteins during the transition from epithelial to fibroblastoid organization in cultured rat bladder carcinoma cells. J Cell Biol 109, 1495-509 (1989).
87. Thiery, J. P. Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer 2, 442-54 (2002).
88. Polyak, K. & Weinberg, R. A. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat Rev Cancer 9, 265-73 (2009).
89. Liu, Y. Epithelial to mesenchymal transition in renal fibrogenesis: pathologic significance, molecular mechanism, and therapeutic intervention. J Am Soc Nephrol 15, 1-12 (2004).

90. de Iongh, R. U., Wederell, E., Lovicu, F. J. & McAvoy, J. W. Transforming growth factor-beta-induced epithelial-mesenchymal transition in the lens: a model for cataract formation. Cells Tissues Organs 179, 43-55 (2005).
91. Steenvoorden, M. M. et al. Transition of healthy to diseased synovial tissue in rheumatoid arthritis is associated with gain of mesenchymal/fibrotic characteristics. Arthritis Res Ther 8, R165 (2006).
92. Thompson, E. W., Newgreen, D. F. & Tarin, D. Carcinoma invasion and metastasis: a role for epithelial-mesenchymal transition? Cancer Res 65, 5991-5; discussion 5995 (2005).
93. Tarin, D., Thompson, E. W. & Newgreen, D. F. The fallacy of epithelial mesenchymal transition in neoplasia. Cancer Res 65, 5996-6000; discussion 6000-1 (2005).
94. Gaggioli, C. et al. Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. Nat Cell Biol 9, 1392-400 (2007).
95. Finn, R. D. et al. The Pfam protein families database. Nucleic Acids Res 36, D281-8 (2008).
96. Ishkanian, A. S. et al. A tiling resolution DNA microarray with complete coverage of the human genome. Nat Genet 36, 299-303 (2004).
97. Krzywinski, M. et al. A set of BAC clones spanning the human genome. Nucleic Acids Res 32, 3651-60 (2004).
98. Osoegawa, K., de Jong, P. J., Frengen, E. & Ioannou, P. A. Construction of bacterial artificial chromosome (BAC/PAC) libraries. Curr Protoc Hum Genet Chapter 5, Unit 5 15 (2001).
99. Chen, W., Erdogan, F., Ropers, H. H., Lenzner, S. & Ullmann, R. CGHPRO -- a comprehensive data analysis tool for array CGH. BMC Bioinformatics 6, 85 (2005).
100. Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. A greedy algorithm for aligning DNA sequences. J Comput Biol 7, 203-14 (2000).
101. Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132, 365-86 (2000).
102. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31, 3406-15 (2003).
103. The Gene Ontology project in 2008. Nucleic Acids Res 36, D440-4 (2008).
104. Biswas, M. et al. Applications of InterPro in protein annotation and genome analysis. Brief Bioinform 3, 285-95 (2002).
105. Camon, E. et al. The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. Nucleic Acids Res 32, D262-6 (2004).
106. Dimmer, E. C. et al. The Gene Ontology - Providing a Functional Role in Proteomic Studies. Proteomics (2008).
107. Lee, V., Camon, E., Dimmer, E., Barrell, D. & Apweiler, R. Who tangos with GOA?-Use of Gene Ontology Annotation (GOA) for biological interpretation of '-omics' data and for validation of automatic annotation tools. In Silico Biol 5, 5-8 (2005).
108. Dennis, G., Jr. et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 4, P3 (2003).
109. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44-57 (2009).
110. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498-504 (2003).

111. Barrett, T. et al. NCBI GEO: mining millions of expression profiles--database and tools. Nucleic Acids Res 33, D562-6 (2005).
112. Macintosh, C. A., Stower, M., Reid, N. & Maitland, N. J. Precise microdissection of human prostate cancers reveals genotypic heterogeneity. Cancer Res 58, 23-8 (1998).
113. Kurose, K. et al. Frequent somatic mutations in PTEN and TP53 are mutually exclusive in the stroma of breast carcinomas. Nat Genet 32, 355-7 (2002).
114. Matsumoto, N., Yoshida, T. & Okayasu, I. High epithelial and stromal genetic instability of chromosome 17 in ulcerative colitis-associated carcinogenesis. Cancer Res 63, 6158-61 (2003).
115. Erdogan, F. et al. Impact of low copy repeats on the generation of balanced and unbalanced chromosomal aberrations in mental retardation. Cytogenet Genome Res 115, 247-53 (2006).
116. Salge, U. et al. Transition from suspension to adherent growth is accompanied by tissue factor expression and matrix metalloproteinase secretion in a small cell lung cancer cell line. J Cancer Res Clin Oncol 127, 139-41 (2001).
117. Higashijima, S., Nose, A., Eguchi, G., Hotta, Y. & Okamoto, H. Mindin/F-spondin family: novel ECM proteins expressed in the zebrafish embryonic axis. Dev Biol 192, 211-27 (1997).
118. Feinstein, Y. et al. F-spondin and mindin: two structurally and functionally related genes expressed in the hippocampus that promote outgrowth of embryonic hippocampal neurons. Development 126, 3637-48 (1999).
119. Platika, D., Baizer, L. & Fishman, M. C. Sensory neurons "immortalized" by fusion with neuroblastoma cells. Trans Assoc Am Physicians 98, 301-4 (1985).
120. Munoz-Marmol, A. M. et al. A dysfunctional desmin mutation in a patient with severe generalized myopathy. Proc Natl Acad Sci U S A 95, 11312-7 (1998).
121. Goldfarb, L. G. et al. Missense mutations in desmin associated with familial cardiac and skeletal myopathy. Nat Genet 19, 402-3 (1998).
122. Bohnsack, M. T. et al. Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. Embo J 21, 6205-15 (2002).
123. Calado, A., Treichel, N., Muller, E. C., Otto, A. & Kutay, U. Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. Embo J 21, 6216-24 (2002).
124. Mathias, R. A. & Simpson, R. J. Towards understanding epithelial-mesenchymal transition: A proteomics perspective. Biochim Biophys Acta (2009).
125. Keshamouni, V. G. et al. Differential protein expression profiling by iTRAQ-2DLC-MS/MS of lung cancer cells undergoing epithelial-mesenchymal transition reveals a migratory/invasive phenotype. J Proteome Res 5, 1143-54 (2006).
126. Dana, R. C., Welch, W. J. & Deftos, L. J. Heat shock proteins bind calcitonin. Endocrinology 126, 672-4 (1990).
127. Tu, Y., Wu, S., Shi, X., Chen, K. & Wu, C. Migfilin and Mig-2 link focal adhesions to filamin and the actin cytoskeleton and function in cell shape modulation. Cell 113, 37-47 (2003).
128. Blangy, A. et al. Phosphorylation by p34cdc2 regulates spindle association of human Eg5, a kinesin-related motor essential for bipolar spindle formation in vivo. Cell 83, 1159-69 (1995).
129. Maul, R. S. et al. EPLIN regulates actin dynamics by cross-linking and stabilizing filaments. J Cell Biol 160, 399-407 (2003).

130. Byrd, D. A. et al. Tpr, a large coiled coil protein whose amino terminus is involved in activation of oncogenic kinases, is localized to the cytoplasmic surface of the nuclear pore complex. J Cell Biol 127, 1515-26 (1994).
131. Hijikata, T. et al. Plectin 1 links intermediate filaments to costameric sarcolemma through beta-synemin, alpha-dystrobrevin and actin. J Cell Sci 121, 2062-74 (2008).
132. Gilbreth, M. et al. Negative regulation of mitosis in fission yeast by the shk1 interacting protein skb1 and its human homolog, Skb1Hs. Proc Natl Acad Sci U S A 95, 14781-6 (1998).
133. Hill, J. J. et al. Glycoproteomic analysis of two mouse mammary cell lines during transforming growth factor (TGF)-beta induced epithelial to mesenchymal transition. Proteome Sci 7, 2 (2009).
134. Cox, P. R. & Zoghbi, H. Y. Sequencing, expression analysis, and mapping of three unique human tropomodulin genes and their mouse orthologs. Genomics 63, 97-107 (2000).
135. MacLeod, A. R. & Gooding, C. Human hTM alpha gene: expression in muscle and nonmuscle tissue. Mol Cell Biol 8, 433-40 (1988).
136. Thierfelder, L. et al. Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. Cell 77, 701-12 (1994).
137. Shu, W., Yang, H., Zhang, L., Lu, M. M. & Morrisey, E. E. Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors. J Biol Chem 276, 27488-97 (2001).
138. Barrans, S. L., Fenton, J. A., Banham, A., Owen, R. G. & Jack, A. S. Strong expression of FOXP1 identifies a distinct subset of diffuse large B-cell lymphoma (DLBCL) patients with poor outcome. Blood 104, 2933-5 (2004).
139. Koon, H. B., Ippolito, G. C., Banham, A. H. & Tucker, P. W. FOXP1: a potential therapeutic target in cancer. Expert Opin Ther Targets 11, 955-65 (2007).
140. Jafri, N. F., Ma, P. C., Maulik, G. & Salgia, R. Mechanisms of metastasis as related to receptor tyrosine kinases in small-cell lung cancer. J Environ Pathol Toxicol Oncol 22, 147-65 (2003).
141. Wu, Y. & Zhou, B. P. New insights of epithelial-mesenchymal transition in cancer metastasis. Acta Biochim Biophys Sin (Shanghai) 40, 643-50 (2008).
142. Akaogi, K. et al. Cell adhesion activity of a 30-kDa major secreted protein from human bladder carcinoma cells. Biochem Biophys Res Commun 198, 1046-53 (1994).
143. Oh, Y. et al. Synthesis and characterization of insulin-like growth factor-binding protein (IGFBP)-7. Recombinant human mac25 protein specifically binds IGF-I and -II. J Biol Chem 271, 30322-5 (1996).
144. Seno, T. et al. Downregulation of SPARC expression inhibits cell migration and invasion in malignant gliomas. Int J Oncol 34, 707-15 (2009).
145. Ntayi, C., Lorimier, S., Berthier-Vergnes, O., Hornebeck, W. & Bernard, P. Cumulative influence of matrix metalloproteinase-1 and -2 in the migration of melanoma cells within three-dimensional type I collagen lattices. Exp Cell Res 270, 110-8 (2001).
146. Hotary, K., Allen, E., Punturieri, A., Yana, I. & Weiss, S. J. Regulation of cell invasion and morphogenesis in a three-dimensional type I collagen matrix by membrane-type matrix metalloproteinases 1, 2, and 3. J Cell Biol 149, 1309-23 (2000).
147. Mignatti, P., Robbins, E. & Rifkin, D. B. Tumor invasion through the human amniotic membrane: requirement for a proteinase cascade. Cell 47, 487-98 (1986).

148. Navarro, A., Anand-Apte, B. & Parat, M. O. A role for caveolae in cell migration. Faseb J 18, 1801-11 (2004).
149. Li, X. et al. Deconvoluting the intestine: molecular evidence for a major role of the mesenchyme in the modulation of signaling cross talk. Physiol Genomics 29, 290-301 (2007).
150. Greber, B., Lehrach, H. & Adjaye, J. Fibroblast growth factor 2 modulates transforming growth factor beta signaling in mouse embryonic fibroblasts and human ESCs (hESCs) to support hESC self-renewal. Stem Cells 25, 455-64 (2007).
151. Jia, W., Li, H. & He, Y. W. The extracellular matrix protein mindin serves as an integrin ligand and is critical for inflammatory cell recruitment. Blood 106, 3854-9 (2005).
152. Li, Y. et al. Structure of the F-spondin domain of mindin, an integrin ligand and pattern recognition molecule. Embo J 28, 286-97 (2009).
153. He, Y. W. et al. The extracellular matrix protein mindin is a pattern-recognition molecule for microbial pathogens. Nat Immunol 5, 88-97 (2004).
154. Li, H., Oliver, T., Jia, W. & He, Y. W. Efficient dendritic cell priming of T lymphocytes depends on the extracellular matrix protein mindin. Embo J 25, 4097-107 (2006).
155. Jia, W., Li, H. & He, Y. W. Pattern recognition molecule mindin promotes intranasal clearance of influenza viruses. J Immunol 180, 6255-61 (2008).
156. Patel, S. et al. RhoGTPase activation is a key step in renal epithelial mesenchymal transdifferentiation. J Am Soc Nephrol 16, 1977-84 (2005).
157. Radisky, D. C. et al. Rac1b and reactive oxygen species mediate MMP-3-induced EMT and genomic instability. Nature 436, 123-7 (2005).
158. Li, Z. et al. The extracellular matrix protein mindin regulates trafficking of murine eosinophils into the airspace. J Leukoc Biol 85, 124-31 (2009).
159. Tzarfaty-Majar, V. et al. Plasmin-mediated release of the guidance molecule F-spondin from the extracellular matrix. J Biol Chem 276, 28233-41 (2001).
160. Pierleoni, A., Martelli, P. L., Fariselli, P. & Casadio, R. BaCelLo: a balanced subcellular localization predictor. Bioinformatics 22, e408-16 (2006).
161. Bakin, A. V., Tomlinson, A. K., Bhowmick, N. A., Moses, H. L. & Arteaga, C. L. Phosphatidylinositol 3-kinase function is required for transforming growth factor beta-mediated epithelial to mesenchymal transition and cell migration. J Biol Chem 275, 36803-10 (2000).
162. Larue, L. & Bellacosa, A. Epithelial-mesenchymal transition in development and cancer: role of phosphatidylinositol 3' kinase/AKT pathways. Oncogene 24, 7443-54 (2005).
163. Yan, W. et al. PI3 kinase/Akt signaling mediates epithelial-mesenchymal transition in hypoxic hepatocellular carcinoma cells. Biochem Biophys Res Commun 382, 631-6 (2009).
164. Hsu, F. et al. The UCSC Proteome Browser. Nucleic Acids Res 33, D454-8 (2005).
165. Pulukuri, S. M. & Rao, J. S. Matrix metalloproteinase-1 promotes prostate tumor growth and metastasis. Int J Oncol 32, 757-65 (2008).
166. Decock, J. et al. Plasma MMP1 and MMP8 expression in breast cancer: protective role of MMP8 against lymph node metastasis. BMC Cancer 8, 77 (2008).
167. Vincenti, M. P. & Brinckerhoff, C. E. Transcriptional regulation of collagenase (MMP-1, MMP-13) genes in arthritis: integration of complex signaling pathways for the recruitment of gene-specific transcription factors. Arthritis Res 4, 157-64 (2002).
168. Pay, S. et al. Comparison of synovial MMP-1 and TIMP-1 levels in patients with various inflammatory arthritides: is there any difference between rheumatoid

arthritis, Behcet's disease and familial Mediterranean fever? Clin Rheumatol 21, 511-5 (2002).
169. Sauter, W. et al. Matrix metalloproteinase 1 (MMP1) is associated with early-onset lung cancer. Cancer Epidemiol Biomarkers Prev 17, 1127-35 (2008).
170. Pei, D. Matrix metalloproteinases target protease-activated receptors on the tumor cell surface. Cancer Cell 7, 207-8 (2005).
171. Willis, B. C., duBois, R. M. & Borok, Z. Epithelial origin of myofibroblasts during fibrosis in the lung. Proc Am Thorac Soc 3, 377-82 (2006).
172. Moreno-Bueno, G., Portillo, F. & Cano, A. Transcriptional regulation of cell polarity in EMT and cancer. Oncogene 27, 6958-69 (2008).
173. Klymkowsky, M. W. & Savagner, P. Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe. Am J Pathol 174, 1588-93 (2009).
174. Kaartinen, V., Haataja, L., Nagy, A., Heisterkamp, N. & Groffen, J. TGFbeta3-induced activation of RhoA/Rho-kinase pathway is necessary but not sufficient for epithelio-mesenchymal transdifferentiation: implications for palatogenesis. Int J Mol Med 9, 563-70 (2002).
175. Nawshad, A. & Hay, E. D. TGFbeta3 signaling activates transcription of the LEF1 gene to induce epithelial mesenchymal transformation during mouse palate development. J Cell Biol 163, 1291-301 (2003).
176. Zeisberg, M. et al. BMP-7 counteracts TGF-beta1-induced epithelial-to-mesenchymal transition and reverses chronic renal injury. Nat Med 9, 964-8 (2003).
177. Zeisberg, M. et al. Fibroblasts derive from hepatocytes in liver fibrosis via epithelial to mesenchymal transition. J Biol Chem 282, 23337-47 (2007).
178. Zeisberg, M., Shah, A. A. & Kalluri, R. Bone morphogenic protein-7 induces mesenchymal to epithelial transition in adult renal fibroblasts and facilitates regeneration of injured kidney. J Biol Chem 280, 8094-100 (2005).
179. Ma, L., Lu, M. F., Schwartz, R. J. & Martin, J. F. Bmp2 is essential for cardiac cushion epithelial-mesenchymal transition and myocardial patterning. Development 132, 5601-11 (2005).
180. Zhang, J. & Li, L. BMP signaling and stem cell regulation. Dev Biol 284, 1-11 (2005).
181. Okada, H., Danoff, T. M., Kalluri, R. & Neilson, E. G. Early role of Fsp1 in epithelial-mesenchymal transformation. Am J Physiol 273, F563-74 (1997).
182. Kie, J. H., Kapturczak, M. H., Traylor, A., Agarwal, A. & Hill-Kapturczak, N. Heme oxygenase-1 deficiency promotes epithelial-mesenchymal transition and renal fibrosis. J Am Soc Nephrol 19, 1681-91 (2008).
183. Tsukita, S., Furuse, M. & Itoh, M. Multifunctional strands in tight junctions. Nat Rev Mol Cell Biol 2, 285-93 (2001).
184. Vinken, M., Vanhaecke, T. & Rogiers, V. Junctional structures and hepatocellular carcinoma: from the lab to the clinic? Liver Int 28, 432-4 (2008).
185. Medici, D., Hay, E. D. & Goodenough, D. A. Cooperation between snail and LEF-1 transcription factors is essential for TGF-beta1-induced epithelial-mesenchymal transition. Mol Biol Cell 17, 1871-9 (2006).
186. Shapland, C., Hsuan, J. J., Totty, N. F. & Lawson, D. Purification and properties of transgelin: a transformation and shape change sensitive actin-gelling protein. J Cell Biol 121, 1065-73 (1993).
187. Lawson, D., Harrison, M. & Shapland, C. Fibroblast transgelin and smooth muscle SM22alpha are the same protein, the expression of which is down-regulated in many cell lines. Cell Motil Cytoskeleton 38, 250-7 (1997).
188. Yu, H. et al. Transgelin is a direct target of TGF-beta/Smad3-dependent epithelial cell migration in lung fibrosis. Faseb J 22, 1778-89 (2008).

189. Neilson, E. G. Setting a trap for tissue fibrosis. Nat Med 11, 373-4 (2005).
190. Madar, S. et al. Modulated expression of WFDC1 during carcinogenesis and cellular senescence. Carcinogenesis 30, 20-7 (2009).
191. Zeisberg, M. & Kalluri, R. The role of epithelial-to-mesenchymal transition in renal fibrosis. J Mol Med 82, 175-81 (2004).
192. Potenta, S., Zeisberg, E. & Kalluri, R. The role of endothelial-to-mesenchymal transition in cancer progression. Br J Cancer 99, 1375-9 (2008).
193. Zavadil, J., Haley, J., Kalluri, R., Muthuswamy, S. K. & Thompson, E. Epithelial-mesenchymal transition. Cancer Res 68, 9574-7 (2008).
194. Bilban, M. et al. Identification of novel trophoblast invasion-related genes: heme oxygenase-1 controls motility via peroxisome proliferator-activated receptor gamma. Endocrinology 150, 1000-13 (2009).
195. Cross, J. C. et al. Trophoblast functions, angiogenesis and remodeling of the maternal vasculature in the placenta. Mol Cell Endocrinol 187, 207-12 (2002).
196. Vicovac, L. & Aplin, J. D. Epithelial-mesenchymal transition during trophoblast differentiation. Acta Anat (Basel) 156, 202-16 (1996).
197. Yilmaz, M. & Christofori, G. EMT, the cytoskeleton, and cancer cell invasion. Cancer Metastasis Rev 28, 15-33 (2009).
198. Araki, N., Hatae, T., Yamada, T. & Hirohashi, S. Actinin-4 is preferentially involved in circular ruffling and macropinocytosis in mouse macrophages: analysis by fluorescence ratio imaging. J Cell Sci 113 (Pt 18), 3329-40 (2000).
199. Alstergren, P. et al. Polarization and directed migration of murine neutrophils is dependent on cell surface expression of CD44. Cell Immunol 231, 146-57 (2004).
200. Graham, D. B. et al. ITAM signaling by Vav family Rho guanine nucleotide exchange factors regulates interstitial transit rates of neutrophils in vivo. PLoS ONE 4, e4652 (2009).
201. Shi, J., Badri, K. R., Choudhury, R. & Schuger, L. P311-induced myofibroblasts exhibit ameboid-like migration through RalA activation. Exp Cell Res 312, 3432-42 (2006).
202. Friedl, P., Zanker, K. S. & Brocker, E. B. Cell migration strategies in 3-D extracellular matrix: differences in morphology, cell matrix interactions, and integrin function. Microsc Res Tech 43, 369-78 (1998).
203. Friedl, P., Borgmann, S. & Brocker, E. B. Amoeboid leukocyte crawling through extracellular matrix: lessons from the Dictyostelium paradigm of cell movement. J Leukoc Biol 70, 491-509 (2001).
204. Paper, W. et al. Elevated amounts of myocilin in the aqueous humor of transgenic mice cause significant changes in ocular gene expression. Exp Eye Res 87, 257-67 (2008).
205. Turner, M. D. et al. Calpain facilitates actin reorganization during glucose-stimulated insulin secretion. Biochem Biophys Res Commun 352, 650-5 (2007).
206. Dourdin, N. et al. Reduced cell migration and disruption of the actin cytoskeleton in calpain-deficient embryonic fibroblasts. J Biol Chem 276, 48382-8 (2001).
207. Barron-Casella, E. A. et al. Sequence analysis and chromosomal localization of human Cap Z. Conserved residues within the actin-binding domain may link Cap Z to gelsolin/severin and profilin protein families. J Biol Chem 270, 21472-9 (1995).
208. Bruneel, A. et al. Proteomics of human umbilical vein endothelial cells applied to etoposide-induced apoptosis. Proteomics 5, 3876-84 (2005).
209. Feinstein, Y. & Klar, A. The neuronal class 2 TSR proteins F-spondin and Mindin: a small family with divergent biological activities. Int J Biochem Cell Biol 36, 975-80 (2004).

210. Manda, R. et al. Identification of genes (SPON2 and C20orf2) differentially expressed between cancerous and noncancerous lung cells by mRNA differential display. Genomics 61, 5-14 (1999).
211. Edwards, S. et al. Expression analysis onto microarrays of randomly selected cDNA clones highlights HOXB13 as a marker of human prostate cancer. Br J Cancer 92, 376-81 (2005).
212. Parry, R. et al. Identification of a novel prostate tumor target, mindin/RG-1, for antibody-based radiotherapy of prostate cancer. Cancer Res 65, 8397-405 (2005).
213. Simon, I. et al. Evaluation of the novel serum markers B7-H4, Spondin 2, and DcR3 for diagnosis and early detection of ovarian cancer. Gynecol Oncol 106, 112-8 (2007).
214. Rizzoli, S. O. & Betz, W. J. Synaptic vesicle pools. Nat Rev Neurosci 6, 57-69 (2005).
215. Otis, K. O., Thompson, K. R. & Martin, K. C. Importin-mediated nuclear transport in neurons. Curr Opin Neurobiol 16, 329-35 (2006).
216. Goldstein, A. Y., Wang, X. & Schwarz, T. L. Axonal transport and the delivery of pre-synaptic components. Curr Opin Neurobiol 18, 495-503 (2008).
217. Gregg, J. P. et al. Gene expression changes in children with autism. Genomics 91, 22-9 (2008).
218. Lanzetti, L. Actin in membrane trafficking. Curr Opin Cell Biol 19, 453-8 (2007).
219. Lin, J. J., Eppinga, R. D., Warren, K. S. & McCrae, K. R. Human tropomyosin isoforms in the regulation of cytoskeleton functions. Adv Exp Med Biol 644, 201-22 (2008).
220. Albertson, R., Cao, J., Hsieh, T. S. & Sullivan, W. Vesicles and actin are targeted to the cleavage furrow via furrow microtubules and the central spindle. J Cell Biol 181, 777-90 (2008).
221. Guttinger, S., Laurell, E. & Kutay, U. Orchestrating nuclear envelope disassembly and reassembly during mitosis. Nat Rev Mol Cell Biol 10, 178-91 (2009).
222. Valentine, M. T., Fordyce, P. M., Krzysiak, T. C., Gilbert, S. P. & Block, S. M. Individual dimers of the mitotic kinesin motor Eg5 step processively and support substantial loads in vitro. Nat Cell Biol 8, 470-6 (2006).
223. Pittet, M. J. Behavior of immune players in the tumor microenvironment. Curr Opin Oncol 21, 53-9 (2009).
224. Saika, S. et al. TGF beta in fibroproliferative diseases in the eye. Front Biosci (Schol Ed) 1, 376-90 (2009).