UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Zentrum für Experimentelle Medizin, Institut für Tumorbiologie

Direktor: Prof. Dr. med. Klaus Pantel

SUMOylation-dependent regulation of the EMT-suppressor Grainyhead-like 2 (GRHL2) in breast cancer

Dissertation

zur Erlangung des Grades eines Doktors der Medizin /Zahnmedizin
an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:

Yang Xu

aus Heilongjiang (China)

Hamburg 2019

(wird von der Medizinischen Fakultät ausgefüllt)

Angenommen von der
Medizinischen Fakultät der Universität Hamburg am: 25.11.2019

Veröffentlicht mit Genehmigung der
Medizinischen Fakultät der Universität Hamburg.

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. Klaus Pantel

Prüfungsausschuss, zweite/r Gutachter/in: Prof. Dr. Christian Kubisch

Contents

Aim of the study 5

1. Introduction 6

1.1 Breast cancer metastasis 6

1.2 Epithelial-Mesenchymal Transition (EMT) 8

1.3 The transcription factor Grainyhead-like 2 (GRHL2) 13

1.3.1 The Grainyhead family of transcription factors 13

1.3.2 Structure of the GRHL2 gene and protein 13

1.3.3 Physiological significance of GRHL2 14

1.3.4 The role of GRHL2 in tumour development and progression 15

2. Material 19

3. Methods 19

3.1 Cell culture 19

3.1.1 Cultivation and passaging of cells 19

3.1.2 Freezing and thawing cells 20

3.1.3 Transfection of cells using polyethyleneimine (PEI) 20

3.1.4 Retroviral gene transfer 20

3.2 Molecular biology methods 21

3.2.1 Culturing *Escherichia coli (E. coli)* 21

3.2.2 Transformation of *E. Coli* 21

3.2.3 Small-scale DNA isolation 22

3.2.4 Large-scale DNA isolation 22

3.2.5 Determination of DNA concentration and purity 22

3.2.6 Restriction endonuclease digestion of DNA 22

3.2.7 Dephosphorylation of DNA 23

3.2.8 Agarose gel electrophoresis 23

3.2.9 Isolation of DNA from agarose gels 23

3.2.10 Ligation of DNA fragments 23

3.2.11 Polymerase Chain Reaction (PCR) for DNA amplification .. 24

3.2.12 Purification of PCR-amplification products 24

3.2.13 Isolation of total RNA from cultured cells 24

3.2.14 Reverse transcription of total RNA 25

3.2.15 Quantitative real-time PCR analysis (qRT-PCR) 25

3.2.16 DNA sequence analysis 25

3.2.17 Site-directed mutagenesis 26

3.3 Biochemical methods 26
3.3.1 Preparation of whole cell extracts (WCEs) 26
3.3.2 Measurement of protein concentration 27
3.3.3 SDS-PAGE 27
3.3.4 Semi-dry transfer 28
3.3.5 Immunoblot analysis 29
3.3.6 Indirect immunofluorescence analysis of cells 29
3.3.7 Pull-down assay 30
3.3.8 Luciferase Reporter Assays 30
3.4 Bioinformatic analyses 31
3.4.1 Computer-aided prediction of SUMOylation sites by in *silico*-analysis 31
4. Results 32
4.1 GRHL2 is post-translationally modified by SUMOylation 32
4.2 Identification of SUMOylation sites in GRHL2 using computational approaches 34
4.3 Identification of K159 as the major SUMOylation site in GRHL2 by site-directed mutagenesis 35
4.4 SUMOylation of GRHL2 is enhanced by PIAS proteins 38
4.5 The subnuclear localisation of GRHL2 is not influenced by SUMOylation 40
4.6 Regulation of the GRHL2 subnuclear localisation through association with PIAS proteins 41
4.7 SUMOylation positively regulates GRHL2 transcriptional activity 43
5. Discussion 47
6. Summary/Zusammenfassung 59
7. References 62
8. Appendix 76
9. Acknowledgement 86
10. Curriculum Vitae 87
11. Eidesstattliche Erklärung 88

Aim of the study

An increasing body of evidence suggests that the largely uncharacterized developmental transcription factor Grainyhead-like 2 (GRHL2) plays a crucial role in the regulation of the Epithelial-to-mesenchymal transition (EMT) during the metastatic spread of breast cancer. Thus, a better understanding of the regulation of GRHL2 activity and expression in cancer cells is pivotal for molecularly dissecting breast cancer metastasis.

This project is based on unpublished observations suggesting that GRHL2 is capable of physically interacting with various components of the SUMOylation pathway, implying that SUMOylation, which is the covalent attachment of SUMO proteins to substrates, could represent a novel posttranslational regulatory mechanism of the GRHL2 transcription factor. Major aim of this project is to start to investigate whether modification of GRHL2 by SUMO could affect its transcriptional activity and therefore also metastatic spread of breast cancer cells.

First, it was of interest to demonstrate that GRHL2 represents a novel substrate for SUMOylation. The lysine residue(s) serving as acceptor site(s) for the covalent attachment of SUMO-proteins should be identified using a wide range of biochemical assays in combination with computational approaches. As SUMOylation can be enhanced by SUMO E3 ligases, the role of members of the PIAS family of proteins in SUMOylation of GRHL2 should be investigated. Furthermore, extended indirect immunofluorescence studies should be conducted to illuminate the possible influence of SUMOylation and/or interaction with PIAS proteins on the subnuclear distribution of GRHL2. Finally, it was of uppermost interest to investigate whether SUMOylation either activates or suppresses GRHL2-dependent transcription of target genes.

The results of this study should help to elucidate possible SUMO-dependent regulatory mechanisms and to better understand the role of the GRHL2 transcription factor in EMT during breast cancer metastasis.

1. Introduction

Breast cancer is now the most commonly diagnosed cancer and the leading cause of cancer-related deaths among women worldwide. According to the global cancer statistics, over 2 million new breast cancer cases and 0.6 million breast cancer deaths were estimated for 2018, accounting for approximately 24 % of all cancers and 15 % of cancer-related deaths, respectively (Bray et al., 2018). Despite significant progress in screening and systemic treatment of primary tumours, strategies targeting metastatic breast cancer (MBC) are far less effective and therefore the metastatic spread of tumours still is the major cause of cancer-related mortality. A better understanding of the molecular mechanisms driving metastatic spread of breast cancers is crucial for combating MBC.

In the past decades, an enormous amount of research has indicated that epithelial-to-mesenchymal transition (EMT) plays a crucial role in tumour cell invasion and subsequent metastasis (Yeung and Yang, 2017). Several studies also demonstrated that the largely uncharacterized developmental transcription factor Grainyhead-like 2 (GRHL2) represents a potent suppressor of EMT during breast cancer metastasis (Cieply et al., 2013; Mooney et al., 2017; Werner et al., 2013). Aim of this work was to investigate the possibility of a SUMOylation-dependent regulation of the expression and activity of the metastasis suppressor GRHL2 in breast cancer.

1.1 Breast cancer metastasis

Metastatic breast cancer (also referred to as stage IV breast cancer) is characterized by the formation of secondary tumours in distant organs, predominantly in the skeleton, liver, lung and brain, with the most frequently site being the skeleton (Smid et al., 2008). At the time of diagnosis, 6 - 10 % of breast cancer patients present with detectable metastases and up to 30 % of patients without clinically secondary tumours are estimated to eventually develop metastatic lesions, irrespective of receiving systemic treatment (O'Shaughnessy, 2005; Pantel and Hayes, 2018). Treatment of breast cancer patients with metastatic disease still represents an extraordinary clinical challenge, as reflected by the extremely low 5-year survival rate.

Unfortunately, existing clinical strategies fall short in precisely classifying patients at high risk of developing metastases (Alix-Panabieres and Pantel, 2016). Treatment of MBC among many other parameters depends on the age, menopause statue, breast cancer subtype and location of the secondary tumours, and can include chemotherapy, radiation, surgery and hormonal therapy (Santa-Maria and Gradishar, 2015). Although a variety of options are available for treatment of MBC patients, what can be achieved in most of cases are optimizing quality of life by palliation of symptoms and prolongation of overall survival via delaying tumour progression (Senkus and Lacko, 2017). Currently, no matter how aggressive the treatment is, MBC remains incurable.

Metastasis represents a highly complex multistep process. To successfully develop a secondary site, tumour cells must complete all steps of a metastatic cascade comprising the following steps: local invasion, induction of blood vessel formation, intravasation, transport and survival in the circulation, extravasation, colonization of distant organs, and finally outgrowth to clinically detectable metastases (Fig. 1).

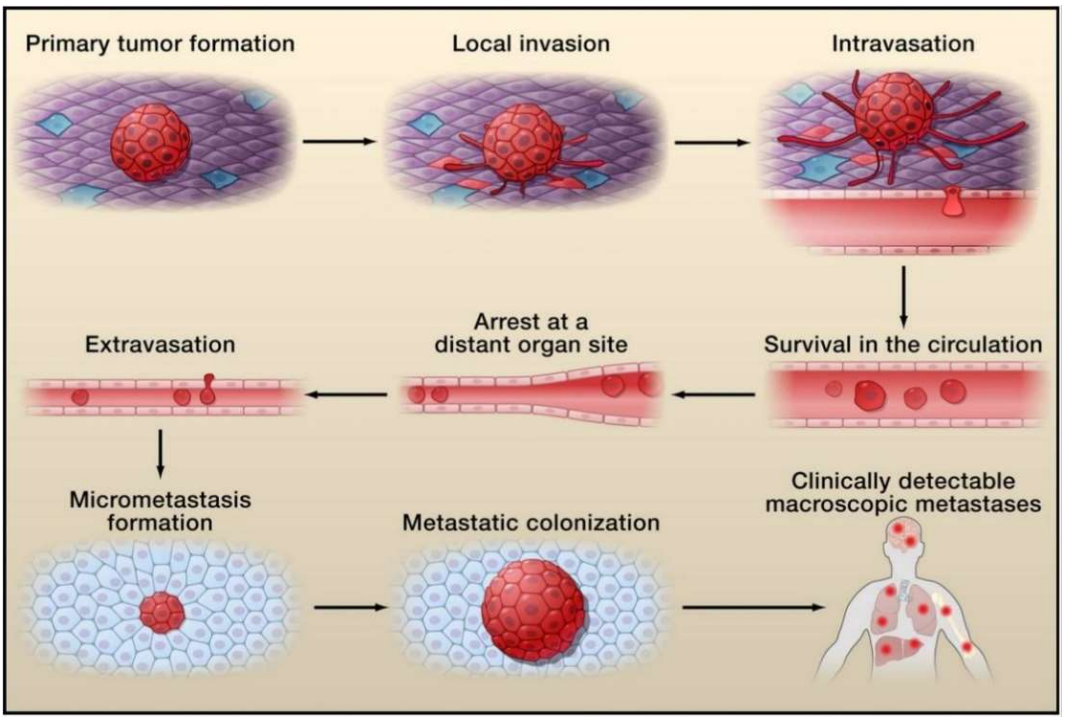

Fig. 1. The metastatic cascade. The figure illustrates sequential steps of the metastatic spread of tumour cells as described in the body text (Faltas, 2012).

Primary tumour cells acquire the ability to alter cell-to-cell adhesion and cell attachment to the extracellular matrix (ECM) allowing them to invade into the adjacent host tissue. To penetrate basement membranes and interstitial matrix, local degradation of ECMs is induced by tumour-associated proteases and tumour cell migration mostly driven by motility factors is initiated (Lamouille et al., 2014). Local invasion is complemented by angiogenesis, the development of new host blood vessels in the primary tumour. The process is induced at capillary or post-capillary venule levels by angiogenesis-promoting factors released by tumour or stromal cells (Bielenberg and Zetter, 2015).

Local invasion enables tumour cells to intravasate into blood vessels or lymphatic channels. Tumour-associated capillaries are notoriously defective, highly leaky and permeable, a property that favours the entry of tumour cells into the vasculature (Bielenberg and Zetter, 2015). Tumour cells discharged into the circulation can occur either as single cells or clusters of tumour cells. From rapidly growing tumours, thousands of aggressive tumour cells are believed to be shed into the circulation every day (Chambers et al., 2002).

Circulating tumour cells, which often are associated with platelets or fibrin, reach distant organs and through physical trapping as well as specific adhesion processes adhere to the endothelium of target organs. Tumour cells then induce retraction of endothelial cells and degradation of basement membranes, thus enabling them to extravasate into the organ parenchyma (Miles et al., 2008). Once settled in the organ parenchyma, tumour cell can grow to form clinically detectable metastases or may undergo cell cycle arrest and remain dormant for long periods of time, a phenomenon commonly described as tumour dormancy or minimal residual disease (MRD) (Werner and Pantel, 2017). Metastasis is a highly inefficient process in that only a small proportion of circulating tumour cells develop micrometastases and only minority of micrometastases eventually grow into macrometastases (Vanharanta and Massague, 2013).

1.2 Epithelial-Mesenchymal Transition (EMT)

Epithelial and mesenchymal cells are two of the basic cell types distinguished by their distinct morphology and cell physiology. They differ in several crucial characteristics such as cell junctions, the cell skeleton, cell polarity and gene

expression. Differentiated epithelial cells predominantly maintain structural integrity by establishing single layers or multilayer tissues via specialized intercellular junctions, including tight junctions, adherens junctions, gap junctions, and desmosomes, and are sited on basement membranes (Serrano-Gomez et al., 2016). Cell polarity is a fundamental trait of epithelial cells, featuring apical, lateral and basal membrane domains that function differently in sustaining cell-cell and cell-ECM contacts. By contrast, mesenchymal cells exist without defined cell polarity and are characterized by an elongated, fibroblast-like morphology. They normally are embedded inside the ECM, hardly establish direct contacts with adjacent cells, have a distinct cytoskeleton and attach to the ECM via distinct cell adhesion structures (Lee et al., 2006b; Thiery et al., 2009).

Epithelial-to-mesenchymal transition (EMT) describes a biological process that enables stationary epithelial cells to downregulate epithelial and to acquire a mesenchymal stem cell characteristics with enhanced migratory and invasive phenotype (Kalluri and Weinberg, 2009). The pioneering work of Elizabeth Hay first recognized EMT as a development program during embryogenesis in 1980s (Hay, 1995). However, subsequent studies demonstrated that the processes underlying EMT are reactivated during tissue repair and also tumour progression (Kalluri and Weinberg, 2009).

EMT is often classified into three subtypes based on distinct biological contexts: Type I EMT is associated with distinct embryogenic processes such as, for example, during implantation of the embryo into the endometrium, gastrulation when multipotent epithelial epiblast cells ingress to form mesoderm and endoderm, and neural crest formation when neuroectodermal epithelial cells transform to migratory neural crest cells (Bischof et al., 2006; Lim and Thiery, 2012).

Type II EMT occurs during tissue repair where resident epithelial cells transdifferentiate to fibroblasts and other related cells in order to synthesize ECM components and to narrow wound bed, enabling contract of the injured edge. In contrast to type I EMT, type II EMT represents an inflammation-associated event which ceases once inflammation is attenuated (Stone et al., 2016).

Type III EMT relates to tumour progression and metastasis and enables primary carcinoma cells to acquire invasiveness and subsequent metastatic abilities, thus promoting the formation of life-threatening metastatic lesions. Type III EMT represents a hall mark of cancer metastasis and has become a forefront area of

intensive research (Felipe Lima et al., 2016; Hanahan and Weinberg, 2011; Sato et al., 2016). While all three types of EMT represent distinct morphogenetic programs, a set of genetic and biological characteristics appears to be highly conserved. More research is still required to illuminate possible similarities and differences among the three types of EMT.

Hallmarks of EMT include dissolution of epithelial cell junctions, loss of apical-basal polarity and acquisition of front-rear polarity, rearrangement of cytoskeletal architecture and cell morphology, inactivation of epithelial-related genes and concomitant upregulation of an mesenchymal gene expression signature, increased cellular protrusions and enhanced motility, and, at least in some cases, expression of ECM-degrading enzymes (Lamouille et al., 2014). Furthermore, cells undergoing EMT acquire resistance to apoptosis and senescence (Thiery et al., 2009). Upon initiation of EMT, intercellular bridges consisting of adherens junctions, tight junctions, desmosomes and gap junctions are deconstructed, compromised by the degradation and/or re-localisation of core junctional proteins, such as epithelial cadherin (E-cadherin), claudin, occluding, zonula occludens 1 (ZO1) and connexin (Bax et al., 2011; Huang et al., 2012; Yilmaz and Christofori, 2009). Meanwhile, the disruption of junctional proteins further confers the loss of apical-basal polarity, which is maintained by polarity complexes, for instance partitioning-defective (PAR), Crumbs and Scribble complexes which physically and functionally associate with epithelial junctional architecture to define the apical and basolateral compartment (Huang et al., 2012; Navarro et al., 2005; Qin et al., 2005; St Johnston and Ahringer, 2010). Simultaneously, the expression of junctional complex proteins and polarity complex proteins are transcriptionally repressed to further destabilize the epithelial junctions and cellular polarity, among which the downregulation of E-cadherin is regarded as a hallmark of EMT progression (De Craene and Berx, 2013; Moreno-Bueno et al., 2008; Pecina-Slaus, 2003; Peinado et al., 2007).

Concomitantly, EMT triggers the upregulation of mesenchymal genes such as N-cadherin, vimentin and fibronectin, and other numerous non-epithelial cadherins and membrane proteins, thus allowing morphologic elongation and dynamic cell motility (Thiery and Sleeman, 2006) (Fig. 2). Cells undergoing EMT reorganize their cortical actin cytoskeleton into cytoplasmic and basal network of intermediate filaments and form new membrane projections, including lamellipodia, filopodia and actin-rich invadopodia, the latter one exerting a proteolytic function during ECM

degradation (McNiven, 2013; Yilmaz and Christofori, 2009, 2010). The transition to mesenchymal cytoskeletal architecture result in cell elongation, front-rear polarity formation, increase of cell contractility and cell motility which are essential for EMT during tumour cell invasion.

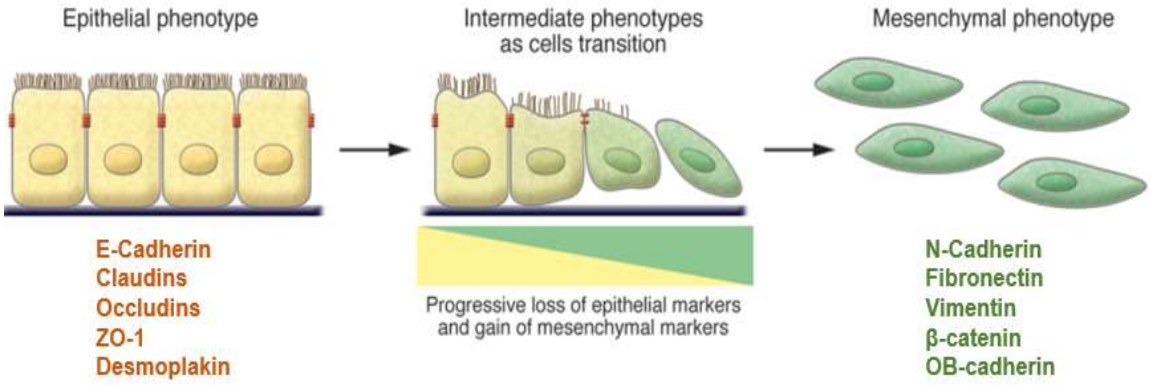

Fig. 2. Epithelial-to-Mesenchymal Transition. The figure represents the simplified EMT process. The commonly used epithelial and mesenchymal cellular markers are listed (Kalluri and Weinberg, 2009).

However, recent investigations suggest that, rather than undergoing a complete EMT process, cells may acquire mixed epithelial and mesenchymal properties, thus exhibiting a hybrid epithelial/mesenchymal (E/M) phenotype which enables cells to move collectively as clusters (Jolly et al., 2018). Following intravasation of these clusters into the bloodstream, they will give rise to clusters of circulating tumour cells (CTCs) - the chief “villains” of metastasis (Aceto et al., 2014; Cheung et al., 2016). Clustered cells better survive chemical gradients and migrate more efficiently under mechanical stress. Immune or stromal cells within clusters may help in evading immune surveillance during navigation. Also, cell-cell interactions mediated by epithelial adhesion molecules may provide pro-survival signals (Aceto et al., 2014; Al Habyan et al., 2018; Bardelli and Pantel, 2017; Giuliano et al., 2018; Jolly et al., 2017; Mohme et al., 2017). Furthermore, cells exhibiting a hybrid E/M phenotype often are resistant to apoptosis as compared to cells displaying a mere mesenchymal state and attain maximal stem cell character associated with higher tumour initiating potential (Pastushenko et al., 2018; Tan et al., 2014). Thus, promising *in vivo* result emphasize that hybrid E/M phenotypes could exist stably

and exhibit more aggressive properties in fostering cancer metastasis compared to a complete EMT phenotype (Jolly et al., 2015). Additionally, partial EMT could be more therapy-resistant compared with a wholly mesenchymal phenotype and more such hybrid phenotypes can be induced and/or selected under various therapies (Fustaino et al., 2017; Goldman et al., 2015; Jolly et al., 2017).

The EMT program is regulated by multiple pathways at different levels, including epigenetic modifications (e.g., DNA methylation) directly affecting gene expression, alterative splicing of mRNAs leading to distinct protein isoforms in mesenchymal cells, miRNA-mediated inhibition and/or degradation of gene transcripts, and post-translational modifications (e.g., phosphorylation, SUMOylation, methylation, etc.) governing the activity and stability of master transcription factors (Lamouille et al., 2014; Lee and Kong, 2016; Nieto et al., 2016). Additionally, EMT programs are also influenced by factors such as tissue hypoxia, mechanical stress or matrix stiffness (Diepenbruck and Christofori, 2016).

In response of numerous intracellular or extracellular stimuli, a substantial number of signalling pathways, for instance Transforming Growth Factor-β (TGF-β) signalling pathway, Wnt / β-catenin pathway, and Notch pathway, alone or through extensive crosstalk regulate the EMT program (Felipe Lima et al., 2016). Common to all these EMT pathways is the convergence on the regulation of core EMT-inducing transcriptional factors (EMT-TFs). Since the identification of Snail 1 as the first transcriptional repressor of E-cadherin, numerous EMT-TFs subsequently were discovered (Batlle et al., 2000). Based on their abilities to repress E-cadherin directly or indirectly, EMT-TFs can be classified into two groups. Direct regulators include SNAIL1, SNAIL2, ZEB1, and ZEB2, whereas transcription factors such as TWIST 1 and TWIST 2 are acting through indirect regulation of E-Cadherin expression (Lee and Kong, 2016). Apart from these EMT-inducing transcription factors, several transcription factors are also known to suppress EMT. One of these EMT-suppressor molecules represents the developmental transcription factor Grainyhead-like 2 (GRHL2).

1.3 The transcription factor Grainyhead-like 2 (GRHL2)

1.3.1 The Grainyhead family of transcription factors

The Grainyhead gene (dGrh) was first identified through a Drosophila mutant that exhibited neural tube closure defects leading to a “grainyhead” phenotype with discontinuous head skeleton which is the namesake of the family (Nusslein-Volhard et al., 1984). Subsequent studies identified Grainyhead as a crucial regulator of neuronal and cuticle-assembly target genes in Drosophila (Bray et al., 1989; Bray and Kafatos, 1991; Dynlacht et al., 1989). In mammals, several transcription factors related to Drosophila Grainyhead were described. Based on sequence homology, members of the LSF/GRH family gene family of transcription factors are grouped into two different branches (Wilanowski et al., 2002). The LSF/CP2 subfamily, which has Drosophila dCP2 as an ancestor gene, encompasses the transcription factors CP2, LBP-1a, and LBP-9d. The other phylogenetic branch, known as the GRH subfamily with dGrh as an ancestral gene, includes the GRHL1, GRHL2, and GRHL3 transcription factors which are considered as the three mammalian homologs of Drosophila dGrh (Ting et al., 2003; Wilanowski et al., 2002). The LSF/GRH transcription factors play critical roles in a variety of physiological and pathological contexts. For example, the three GRHL proteins participate in diverse physiological programs such as embryogenesis and wound healing, and more recently have been linked to diverse types of cancer (Auden et al., 2006; Frisch et al., 2017; Mlacki et al., 2015; Wang and Samakovlis, 2012). An increasing body of evidence suggests that especially the GRHL2 transcription factor plays a crucial role in various malignancies (Frisch et al., 2017; Mlacki et al., 2015).

1.3.2 Structure of the GRHL2 gene and protein

In humans, the GRHL2 gene is located on the long arm of chromosome 8 (8q22.3). The gene spans approximately 177 kb with 16 exons and translates into a 625 amino acid protein with a molecular weight of approximately 71 kDa (Wilanowski et al., 2002).

The GRHL2 protein consists of an N-terminal transactivation domain (TAD), a central DNA binding domain (DBD), and a C-terminal dimerization domain (DD). The transactivation domain encompasses residues 1-135 and can activate transcription

from a target gene promotor. The central DNA-binding domain, corresponding to residual 245-494, shares structural similarities with p53, and can attach to specific sequences of DNA. The C-terminal dimerization domain corresponds to residual 521-625 and contains ubiquitin-like folds. This domain participates in homo-dimerization and hetero-dimerization with GRHL1 and GRHL3 and possibly other yet unknown interaction partners (Fig. 3) (Kokoszynska et al., 2008; Wang and Samakovlis, 2012).

Fig. 3. The GRHL2 functional domains. The figure illustrates the location of transactivation domain, DNA binding domain and dimerization domain of GRHL2 as described in the body text.

1.3.3 Physiological significance of GRHL2

In adults, GRHL2 expression primarily occurs in epithelial cells, in organs such as epidermis, oral cavity, oesophagus, and kidney (Auden et al., 2006; Riethdorf et al., 2016). However, GRHL2 is not exclusively expressed in epithelial cells, the expression is also observed in non-epithelial cells, such as myocytes of left heart ventricle, smooth muscle cells of aorta media, neuronal cells of cerebrum, and dendritic cells within thymus (Riethdorf et al., 2016). The physiological significance of GRHL2 transcription factor has been intensively investigated using numerous engineered mouse models, and one fundamental role of GRHL2 relates to epithelial morphogenesis during embryonic development. Studies published by Werth et al showed that Grhl2-deficient mice exhibited neural tube closure defects and aberrant expression of apical junctional complex proteins in several epithelial tissues. With severe developmental defects, the Grhl2-null mutant embryos die around E11.5 and no embryos were viable and fertile at later stages (Werth et al., 2010). Likewise, in Grhl2-nonsense mutants induced by N-ethyl-N-nitrosourea embryonic lethality was induced at E12.5 due to a general defect in tissue closure and abnormalities in organ development (Pyrgaki et al., 2011). In addition to its fundamental role during embryogenesis, several groups also identified specific roles for the GRHL2 transcription factor in tubulogenesis of kidneys, morphogenesis of pulmonary

epithelium, and maintenance of cholangiocyte phenotype (Aue et al., 2015; Gao et al., 2013; Senga et al., 2012). Further analysis of underlying molecular mechanisms revealed that Grhl2 participate in epithelial differentiation both via direct regulation of expression of epithelial junctional molecules such as, for example, E-cadherin, Cldn3, and Cldn4, as well as via regulation of other epithelial-related transcription factors, including Nkx2-1, and OVOL2 (Aue et al., 2015; Senga et al., 2012; Varma et al., 2012; Werth et al., 2010). Although a substantial number of studies have established an indispensable role of GRHL2 in embryogenesis, the implication of GRHL2 in other physiological programs, for example wound healing and fibrosis, is still ill defined.

Two disease causing mutations in the human GRHL2 gene have been identified, providing the substantiation for GRHL2 being an autosomal dominant deafness gene (DFNA28). A frameshift mutation leading to a premature translation stop codon in exon 14 of GRHL2 was first linked to progressive non-syndromic sensorineural hearing loss (Peters et al., 2002). Subsequently, Vona et al. identified a heterozygous non-classical splice site mutation in exon 10, also resulting in a post-lingual hearing loss with a highly variable age of onset and progression (Vona et al., 2013). These evidences indicate that, apart from the physiological significance, GRHL2 also plays critical role in pathological conditions.

1.3.4 The role of GRHL2 in tumour development and progression

In recent years, the significance of GRHL2 transcription factor in cancer development and progression has attracted considerable interest. However, the role of GRHL2 in cancer is still poorly defined or even contradictory in some cases (Frisch et al., 2017; Mlacki et al., 2015). GRHL2 was identified as a tumour suppressor in many cancer types, including breast, ovarian, gastric and cervical cancer. In these tumour entities, the reduction in GRHL2 expression is associated with poor prognosis, aggressive subtype and advanced clinical stage (Chung et al., 2016; Cieply et al., 2013; Torres-Reyes et al., 2014; Werner et al., 2013; Xiang et al., 2013; Xiang et al., 2017). In line with this, an increased expression of GRHL2 correlated with improved overall survival in sarcoma patients (Somarelli et al., 2016). However, GRHL2 also was shown to promote tumour cell proliferation and to act as a proto-oncogene. In hepatocellular carcinoma and clear cell renal cell carcinoma,

increased GRHL2 expression correlated with higher risk of cancer relapse and poor prognosis (Butz et al., 2014; Tanaka et al., 2008). In addition, metastasis-promoting effects of GRHL2 in oral squamous cell carcinomas and murine breast cancer cell lines were also reported (Chen et al., 2016; Xiang et al., 2012). Taken together, depending on the cancer type, GRHL2 may exhibit both promoting and suppressing activities in cancer initiation and progression. The underlying molecular mechanisms for opposing GRHL2 functions in cancer cells are largely unclear and still await in-depth analysis. (Frisch et al., 2017; Mlacki et al., 2015).

Evidence is accumulating that GRHL2 could repress tumour progression, at least in context where the EMT program is a predominant oncogenic driver. A role of GRHL2 as a suppressor of EMT has directly been demonstrated by the identification of GRHL2 as a positive mediator of E-cadherin, claudin3, and claudin4 (Senga et al., 2012; Werth et al., 2010). Additionally, GRHL2 suppresses EMT through transcriptional downregulation of the EMT transcription factor ZEB1, and ZEB1 in turn is able to directly repress GRHL2 expression, indicating the existence of a double negative feedback loop between GRHL2 and ZEB1 (Cieply et al., 2013; Werner et al., 2013). Moreover, the miR200 cluster represents another important direct target of GRHL2, providing an alternative signalling pathway to repress ZEB1-mediated EMT in cancer cells (Chung et al., 2016; Cieply et al., 2013; Gregory et al., 2011). Numerous homeodomain transcription factors are emerging as novel GRHL2 targets, including OVOL2 and the p53 family member p63 (Aue et al., 2015; Mehrazarin et al., 2015). In breast cancer, the tumour-suppressive effects of GRHL2 could be substantiated by a number of elegant functional experiments and clinical data. Ectopic expression of GRHL2 in triple-negative breast cancer cell line triggered the Mesenchymal-to-Epithelial Transition (MET), enhanced anoikis-sensitivity, and suppressed tumour initiation frequency in xenograft assays (Cieply et al., 2013; Cieply et al., 2012). Conversely, efficient knock-down of GRHL2 expression by shRNAs in human immortalized mammary epithelial cells (HMLER) resulted in enhanced tumour initiation frequency and the acquisition of a cancer stem cell phenotype (Cieply et al., 2013). Clinical data also point towards a role of GRHL2 as an EMT-suppressor. A striking loss of GRHL2 expression was detected in the leading-edge of invasive breast tumours, and a statistically significant association between loss of GRHL2 in primary breast cancers and advanced tumour stages was observed (Werner et al., 2013). Analogously, GRHL2 appeared to be specifically

downregulated in tumours representing the claudin-low breast cancer subtype (Cieply et al., 2012). This breast cancer subtype is characterized by a mesenchymal gene expression profile and a particular poor prognosis. These observations clearly justify GRHL2 to be classified as a suppressor of EMT in breast cancer and also many other epithelial-derived tumours (Frisch et al., 2017; Mlacki et al., 2015).

Interestingly, GRHL2 exhibits tumour promoting activities and acts as a proto-oncogene in non-EMT-related contexts. This could first be demonstrated by the transforming activity of GRHL2 in NIH3T3 fibroblasts where ectopic GRHL2 expression resulted in significant morphological changes, increased cell proliferation, anchorage-independent growth in semi-solid media, and tumour growth *in vivo* (Werner et al., 2013). The oncogenic effects could also be demonstrated by the ability of GRHL2 to enhance the proliferation of a wide range of cell types, including hepatocellular carcinoma cells, oral squamous cell carcinoma cells, colorectal carcinoma cells, and normal human epidermal keratinocytes (Chen et al., 2010; Kang et al., 2009; Quan et al., 2014; Quan et al., 2015; Tanaka et al., 2008). In line with this, the epidermal growth factor family member ERBB3, which plays critical role in breast cancer cell proliferation was identified as a GRHL2 target gene (Werner et al., 2013). Knocking-down GRHL2 expression in breast cancer cell lines using GRHL2-specific shRNAs or by functional repression of GRHL2 activity using a dominant-negative GRHL2 mutant led to decreased ERBB3 gene expression and a dramatic suppression of cellular proliferation (Werner et al., 2013). Data from murine models published by Xiang et al indicated that overexpression of GRHL2 may promote breast tumour growth and metastasis (Xiang et al., 2012). Likewise, analysis of large public microarrary datasets indicated that increased expression levels of GRHL2 may correlate with shorter relapse-free survival and high risk of metastasis in breast cancer patients (Xiang et al., 2012). A tumour-promoting activity of GRHL2 is further substantiated by the finding that GRHL2 could be identified as a critical mediator of telomerase reverse transcriptase (hTERT) expression in human oral squamous carcinoma cells (Kang et al., 2009). Furthermore, research carried out by Dompe et al. indicated an anti-apoptotic role of GRHL2 in breast cancer cells via regulating the expression of death receptors FAS and DR5 (Dompe et al., 2011).

In summary, published data on the putative role of GRHL2 in cancer strongly suggest that GRHL2, dependent on the cancer type and/or stage, plays dual

functions by exerting oncogenic or tumour-suppressive activities in cancer cells. For this reason, a better understanding of the mechanisms modulating GRHL2 activity in tumour cells is of uppermost interest.

2. Material

Unless otherwise specified (Table. 4), all chemicals were obtained from Sigma. Frequently used laboratory equipment (Table. 2), consumables (Table. 3), and all oligonucleotides (Table. 6-9) and antibodies (Table. 10) used in this study are listed in the appendix.

3. Methods

3.1 Cell culture

3.1.1 Cultivation and passaging of cells

Cells were propagated in culture medium at 37 °C in a 10 % CO_2, water-saturated atmosphere in a HeraCell 150 incubator (Kendro, Langenselbold). Routinely, cells were split at 1:3-1:10 ratios when reaching 80-90 % confluence by washing cells with 1 × PBS and incubation with trypsin/EDTA (0.25 % (w/v) trypsin, 1 mM EDTA in PBS) for 2-5 min at 37 °C. The activity of trypsin was quenched by addition of 3 volumes of culture medium. Following transfer to a 15-ml reaction tube and centrifugation at 1200 × g for 2 min, the supernatant was discarded and an appropriate volume of cells resuspended in culture medium was transferred to new culture flasks.

The following cell lines were used in this study:

Cell line	Description	Origin	Culture medium
COS-7	SV40 large T antigen immortalized CV-1 African green monkey kidney cells	Assmann lab	DMEM (high glucose), with 1 mM pyruvate, 2 mM glutamine, 10 % FCS
HEK-293T	Human embryonic kidney cells transfected with the SV40 large T antigen	Assmann lab	DMEM (high glucose), with 1 mM pyruvate, 2 mM glutamine, 10 % FCS

3.1.2 Freezing and thawing cells

Cells were harvested using trypsin/EDTA and centrifuged at 1200 × g for 2 min. After supernatant was removed, cells were re-suspended in 3-5 ml freshly prepared freezing medium (culture medium containing 10 % DMSO). One ml of cell suspension was transferred to a labelled cryotube (Nunc, Rockford (USA)) each. Cells were transferred to foam boxes and placed in a -80 °C freezer for 24 h. For long-term storage, frozen stocks were transferred to a liquid-nitrogen container.

Frozen cells were thawed in a 37 °C water bath for 2-3 min and were then mixed immediately with 5 ml of pre-warmed culture medium in a 15-ml reaction tube. After centrifugation at 1200 × g for 2 min, the supernatant was discarded. Cells were resuspended in culture medium and seeded into tissue culture flasks containing an appropriate volume of culture medium.

3.1.3 Transfection of cells using polyethyleneimine (PEI)

Transient transfection of cells was achieved using the polyethyleneimine (PEI; linear, ~MW 25000; Polyscienes, Warrington (UK)) transfection reagent. The day before transfection, 1.5×10^6 cells were transferred to 5 ml culture medium in a T25 flask. Immediately before transfection, complete culture medium was replaced by serum-free medium (SFM). Three µg of plasmid DNA, 30 µl of PEI reagent (1 mg/ml in H_2O) and 300 µl SFM were mixed in a 1.5-ml reaction tube and were incubated at RT for 10 min. The transfection cocktail was then added to cells. Six to eight h post transfection, transfection complexes were removed and fresh culture medium was added. Cells were analysed 24-72 h post transfection.

3.1.4 Retroviral gene transfer

For retroviral gene transfer, Phoenix-ampho packaging cells were transiently transfected using Lipofectamine 3000 (Thermo Fisher, Rockland, USA) according to the manufacturer's instructions. Forty-eight hours post transfection, the cell culture supernatant containing infectious retroviral particles was harvested, centrifuged for 5 min at 500 x g, and cleared by filtration using 0.45 µm low-protein binding filter devices (Millex Syringe Driven Filter Unit, Millipore, Schwalbach). For infection of target cells, the day before infection 2×10^5 cells were seeded into T25

cell culture flasks. Cells were infected with an infection cocktail consisting of 1 ml retroviral supernatant, 1 ml culture medium and 8 µg/ml polybrene (Sigma). Six hours post infection, the infection cocktail was replaced by culture medium. The next day, a second infection of target cells was performed after which the infection cocktail was replaced by culture medium. Infected cells were selected using 0.5 µg/ml puromycin and expression of the transgene was analysed by Western blot analysis (Section 4.7).

3.2 Molecular biology methods

3.2.1 Culturing *Escherichia coli* (*E. coli*)

For transformation and propagation of plasmid DNA, the bacterial strain DH5α was grown in sterile LB medium (1 % (w/v) Bacto-Tryptone, 0.5 % (w/v) Yeast extract, 1 % (w/v) NaCl, pH 7.5) at 37 °C and 180 rpm in a table-top shaker incubator (B. Braun Biotech, Melsungen) or on LB plates (LB medium containing 1.5 % (w/v) Bacto-Agar) at 37 °C in the Heraeus Function Line B12 incubator (Kendro). Liquid and solid medium contained ampicillin or kanamycin at a final concentration of 100 µg/ml or 25 µg/ml, respectively.

3.2.2 Transformation of *E. Coli*

For transformation of *E. coli*, chemically-competent, in-house DH5α cells were used. An aliquot of 30 µl of competent DH5α cells was thawed on ice. After adding 1-2 µl of plasmid DNA and incubation on ice for 30 min, a heat-shock at 42 °C for 30 sec was performed. Cells were then incubated on ice for 2 min and 270 µl LB medium without antibiotics was added. Following a recovery of cells at 37 °C, 180 rpm for 1 h, cells were spread on agar plates containing the appropriate antibiotic and were grown at 37 °C overnight. Starter cultures were prepared by inoculation of 5 ml LB medium with antibiotics with single colonies.

3.2.3 Small-scale DNA isolation

A single colony was inoculated into 5 ml of antibiotic LB medium and incubated at 37 °C and 180 rpm overnight. For small-scale DNA isolation, 1.5 ml starter culture was centrifuged in a 1.5 ml reaction tube at 13200 rpm (Eppendorf centrifuge 5415R) at 4°C for 5 min. The supernatant was discarded and the bacterial pellet was used for DNA isolation using the NucleoSpin Plasmid DNA Purification Kit (Macherey-Nagel, Düren). Finally, the DNA was finally dissolved in 50 µl elution buffer.

3.2.4 Large-scale DNA isolation

100 ml antibiotic LB medium containing desired antibiotic was inoculated with 10 µl starter culture and incubated at 37 °C and 180 rpm overnight. The bacterial pellet obtained by centrifugation at 5500 × g for 15 min (Multifuge Heraeus 3S-R, Kendro) was used for DNA isolation using the NucleoBond Xtra Midi Kit (Macherey-Nagel). Finally, the dried DNA was dissolved in sterile TE buffer.

3.2.5 Determination of DNA concentration and purity

The DNA concentration and purity were determined using the NanoDrop ND-1000 Spectrophotometer (Fisher Scientific, Wilmington (USA)). The OD_{260} was used to calculate the DNA concentration with an extinction of 1 corresponding to a concentration of 50 µg/ml dsDNA. The OD_{260}/OD_{280} ratio was determined to estimate DNA purity. Only DNA with an OD_{260}/OD_{280} ratio $\geq$1.8 was considered pure and was used for downstream applications.

3.2.6 Restriction endonuclease digestion of DNA

For restriction enzyme digestion of plasmid DNA or cDNA-fragments, restriction enzymes and buffers from New England Biolabs (NEB, Frankfurt) were used. DNA was digested with an excess of restriction enzyme (usually 3-5 U/µg DNA) using the appropriate restriction enzyme buffer at 37 °C for at least 3 h. Agarose gel electrophoresis was used to verify complete digestion.

3.2.7 Dephosphorylation of DNA

To avoid self-ligation of linearized plasmid DNA, 1 µg of restriction-digested plasmid DNA was dephosphorylated at the 5' ends with 1 µl Antarctic Phosphatase (5000 U/ml; NEB) in 10 x reaction buffer in a final volume of 30 µl. Reactions were carried out at 37 °C for 15 min for 5' extensions and heat-inactivated at 65 °C for 5-10 min.

3.2.8 Agarose gel electrophoresis

For agarose gel electrophoresis, a 1 % (w/v) agarose gel solution in Tris-acetate-EDTA buffer (TAE) (40 mM Tris-Acetate, 2 mM EDTA, pH 8.5) was dissolved by microwave treatment. When cooled down to about 60 °C, the agarose gel solution was poured in a horizontal electrophoresis apparatus (Bioplastics RV, Landgraaf (NL)). For visualization of DNA fragments, Ethidiumbromide (EtBr) (Sigma) was added to a final concentration of 0.1 µg/ml. Once solidified, the gel chamber was filled with 1 x TAE buffer and DNA samples containing 1/10 volume of 10 x sample buffer (50 % glycerol, 1 mM EDTA, 0.4 % (w/v) Bromophenol blue, 0.4 % (w/v) Xylencyanole, pH 8.0) and the DNA size marker HyperLadder I (BioLine, Hildesheim) were loaded. Electrophoresis was performed at 90-100 V for approximately 1 h. To visualize DNA fragments, the GeneGenius 2 (Syngene, Cambridge (UK)) documentation system equipped with the GeneSnap software were used.

3.2.9 Isolation of DNA from agarose gels

Gel slices containing desired DNA fragments were visualized by UV light, excised using a scalpel, and transferred to 1.5 ml reaction tubes. DNA was extracted using the NucleoSpin Gel and PCR clean-up Kit (Macherey-Nagel) following the manufacturer's instructions. Finally, the DNA was eluted in 15-30 µl EB buffer.

3.2.10 Ligation of DNA fragments

Ligations were performed with vector and DNA fragment at a molecular ratio of 1:3 in the presence of 1 µl T4 DNA ligase (400 U/µl) (New England Biolabs) and 1 µl 10

x T4 DNA ligase buffer in a final volume of 10 µl. The reaction was carried out at RT overnight. Amplification of recombinant plasmid DNA was achieved through transformation of chemically-competent DH5α cells (see 3.2.2).

3.2.11 Polymerase Chain Reaction (PCR) for DNA amplification

PCR reaction was performed 50-µl reaction mix containing variable amounts of template DNA (plasmid DNA or cDNA), 10 µl Q5 buffer (5 x), 10 µl Q5 Enhancer (5 x), 1 µl dNTP mix (10 mM each), 2 µl of each primer (100 ng/µl) and 0.5 µl Q5 High Fidelity DNA Polymerase (2000 U/ml) (New England Biolabs). The reaction was carried out in the Flexigene thermocycler (Techne, Staffordshire (UK)) using the following conditions:

Cycles	PCR step	Temperature	Duration
	Initial denaturation	98 °C	30 sec
12-30	Denaturation	98 °C	10 sec
	Primer annealing	58-63 °C	30 sec
	Extension	72 °C	30 sec/1 kb
	Final Extension	72 °C	2 min

3.2.12 Purification of PCR-amplification products

Following PCR amplification, DNA fragments were purified using the NucleoSpin Gel and PCR clean-up Kit (Macherey-Nagel) following the manufacturer's instructions. Finally, PCR products were eluted in 17 µl NE buffer and were then subjected to restriction endonuclease digestion (see 3.2.6).

3.2.13 Isolation of total RNA from cultured cells

Total RNA was extracted using the NucleoSpin RNA kit (Macherey-Nagel) using subconfluent cultures of breast cancer cells grown in T25 flasks according to the manufacturer’s instructions. Contaminating genomic DNA was removed by treatment of samples with rDNAse. Finally, RNA was dissolved in 40 µl of RNAse-free water and the RNA concentration and purity was determined using the NanoDrop ND-1000.

3.2.14 Reverse transcription of total RNA

The Reverse Transcription of total RNA was performed using the First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) using 1 µg of total RNA and random hexamer or oligo (dT) primers. RNA was reverse transcribed using M-MuLV Reverse Transcriptase in a 20-µl reaction according to the manufacturer's instructions.

3.2.15 Quantitative real-time PCR analysis (qRT-PCR)

For qRT-PCR analysis, 6.75 µl of cDNA (diluted 1:20) (see 3.2.14) were mixed with 0.375 µl of each primer stock solution (10 µM), 7.5 µl Maxima SYBR Green/ROX qPCR Master Mix (2x; Thermo Fisher Scientific). PCR reactions were performed using a C1000 Touch Thermal Cycler (CFX96 Real-Time System) equipped with the CFX3.1 software (BioRad, Munich) with primers listed in Table 9 (see Appendix). Real-time PCR data analysis was performed using the ΔΔ*CT* method with RPLP0 as an endogenous reference. All samples were measured in triplicates using the following PCR conditions:

Step	Temperature	Duration	Cycles
Initial denaturation	95 °C	10 min	1
Denaturation	95 °C	15 sec	
Annealing	60 °C	30 sec	40
Extension	72 °C	30 sec	

3.2.16 DNA sequence analysis

Samples containing 400-500 ng DNA were combined with 1 µl of primer (3.2 pmol/µl), 4 µl of BigDye reaction buffer (5 x) and 2 µl of BigDye mix (both from Applied Biosciences, Rockford (USA)) in a final volume of 20 µl. The reactions were cycled in the Flexigene thermal cycler (Techne) using the following parameters:

Cycles	PCR step	Temperature	Duration
	Initial denaturation	95 °C	5 min
25	Denaturation	96 °C	30 sec
	Primer annealing	50 °C	15 sec
	Extension	60 °C	1 min

The reaction products were precipitated by addition of 16 µl H_2O and 64 µl 100 % ethanol and incubation at RT for 15 min. The DNA was pelleted at 13200 rpm (Eppendorf centrifuge 5415R) for 30 min, washed in 250 µl of 70 % ethanol, centrifuged at 13200 rpm for 5 min and then dried at RT for 5-10 min. Reaction products were applied to capillary electrophoresis at the Institute of Pathology of the University Hospital Hamburg-Eppendorf. Sequence analysis was performed using FinchTV (http://www.geospiza.com/ftvdlinfo.html).

3.2.17 Site-directed mutagenesis

Site-directed mutagenesis was performed by amplifying a template plasmid with oligonucleotides containing the desired mutation using the Q5 High Fidelity DNA Polymerase (NEB). The PCR reaction was performed as described in Section 3.2.11 using 12 cycles.

PCR products were purified using the NucleoSpin Gel and PCR clean-up Kit (Macherey-Nagel) (Section 3.2.13) prior to the addition of 2 µl of CutSmart buffer (10 ×) and 1 µl of Dpn I (20 U/µl) (NEB). Following incubation at 37 °C for at least 3 h, 1 µl of Dpn I treated DNA was transformed into 30 µl competent DH5α cells (Section 3.2.2) and bacteria were plated on antibiotic-containing LB agar plates. The next day, single colonies were isolated and expanded. The introduction of a site-specific mutation was verified by DNA sequencing of plasmid DNA (Section 3.2.13) isolated from several clones.

3.3 Biochemical methods

3.3.1 Preparation of whole cell extracts (WCEs)

Cell culture medium was aspirated and the cell monolayer was washed with PBS. Cells were harvested by scraping and transferred to a reaction tube. Following centrifugation at 4 °C and 3000 rpm for 5 min, the supernatant was discarded. Immediately before use, a lysis buffer containing 1 % (w/v) Nonidet-P40, 0.5 % (w/v) sodium-deoxycholate, 0.1 % (w/v) SDS, 150 mM NaCl, 20 mM sodium phosphate, 25 mM N-ethylmaleimide (NEM), and protease and phosphatase inhibitor cocktails (both from Merck) was prepared and added to the cell pellet. Following incubation

on ice for 30 min, supernatants were collected by centrifugation at 4 °C and 13200 rpm for 10 min. The lysate was mixed with 2 x SDS sample buffer (4 % (w/v) SDS, 20 % (w/v) glycerol, 120 mM Tris-HCl (pH 6.8), 0.01 % (w/v) Bromophenol Blue, 0.2 M DTT) and denatured at 95-100 °C for 5-10 min. After incubation on ice for 2-5 min, samples were either loaded on the gel or stored at -20 °C for further use.

3.3.2 Measurement of protein concentration

Two µl of total cell extracts, 798 µl of 0.1 N NaOH and 200 µl of Bradford reagent (BioRad, Munich) were mixed in a reaction tube. The colour reaction was allowed to proceed for 1-2 min before the absorption at a wavelength of 595 nm was determined by the BioPhotometer (Eppendorf). Protein concentration of samples were calculated using the BSA-based calibration curve.

3.3.3 SDS-PAGE

SDS-PAGE was performed according to the system described by Laemmli (1970) (Laemmli, 1970). SDS-polyacrylamide gels were prepared using 1.0-mm spacers and were composed as following:

Component		6% Gel	8% Gel	15 % Gel	5% Stack gel
H_2O	[ml]	2.7	2.3	1.2	1.4
30 % Acrylamide mix	[ml]	1.0	1.3	2.5	0.33
1.5 M Tris (pH 8.8)	[ml]	1.3	1.3	1.3	-
1.0 M Tris (pH 6.8)	[ml]	-	-	-	0.25
10 % SDS	[µl]	50	50	50	20
10 % APS	[µl]	50	50	50	20
TEMED	[µl]	4	3	2	2

The gel solution was allowed to polymerize for approximately 30 min before preparing the stack gel. The SE 250 electrophoresis vertical unit (Amersham Biosciences, Buckinghamshire (UK)) was assembled and 10-50 µg of whole cell extract freshly denatured for 5-10 min at 95 °C as well as PageRuler Prestained Protein Ladder (Thermo Fisher Scientific) was loaded. Electrophoresis was

performed using 1 × Laemmli buffer (192 mM Glycin, 0.1 % (w/v) SDS, 25 mM Tris-HCl, pH 8.3) at constant current (25 mA per gel) for 1 h.

3.3.4 Semi-dry transfer

Following electrophoresis, proteins were transferred from SDS-polyacrylamide gel to membranes using the Trans-Blot SD semi-dry blotting apparatus (Bio-Rad). Membranes and Whatman 3MM filter papers (GE Healthcare, Dassel) were cut to the size of the gels. Depending on the size of the proteins, two different protocols were used:

Protocol 1:

The Nitrocellulose (NC) membrane (GE Healthcare; 0.45 µm pore size) and four filter papers were pre-equilibrated in transfer buffer containing 48 mM Tris base, 39 mM Glycine, 20 % (v/v) Methanol and 0.037 % (v/v) SDS. The transfer stack was prepared by placing two soaked filter papers, the NC membrane, the gel, and two filters onto the anode plate in the indicated order. The cathode plate was positioned on top of the stack and the transfer apparatus was assembled. The transfer was carried out at a constant current density of 0.8 mA/cm^2 for 2 h.

Protocol 2:

A discontinuous buffer system containing Anode buffer I (0.3 M Tris base, 10 % Methanol), Anode buffer II (0.025 M Tris base,10 % Methanol) and Cathode buffer (0.025 Tris base, 20 % Methanol, 0.04 M ε-Amino caproic acid) was used for transfer of small proteins (< 15 kDa) (Kyhse-Andersen, 1984). Gels were equilibrated in Cathode buffer for 15 min. The FluoroTrans W PVDF membrane (Pall Life Sciences, Dreieich; 0.2 µm pore size) was activated with methanol, distilled water, and Anode buffer II for 5 min each. The transfer stack was prepared by placing two filter papers soaked in Anode buffer I, one filter paper soaked in Anode buffer II, the activated FluoroTrans W PVDF membrane, the equilibrated gel, and three filters papers soaked in Cathode buffer II onto the anode plate in the indicated order. The cathode plate was positioned on top of the stack and the transfer apparatus was assembled. The transfer was carried out at a constant current density of 0.8 mA/cm^2 for 45 min.

3.3.5 Immunoblot analysis

Following semi-dry transfer, membranes were blocked in 5 % (w/v) milk powder in 1 × TBS-T (1 × TBS (137 mM NaCl, 20 mM Tris/HCl), 0.05 % (v/v) Tween-20) for 1 h at RT and incubated with primary antibodies diluted in 5 % (w/v) bovine serum albumin (BSA) in 1 × TBS-T overnight at 4 °C. Membranes were then washed 3 × 5 min in 1 × TBS-T and incubated with secondary antibody diluted in 5 % skim milk powder in TBS-T for 1 h at RT. Following 3 washes for 5 min in 1 × TBS-T, the membrane was incubated in a freshly prepared ECL solution consisting of a 1:1 mixture of solutions 1 (0.1 M Tris-HCl (pH 8.5), 25 mM luminol, 9 mM coumaric acid] and 2 (0.1 M Tris-HCl (pH 8.5), 0.018 % (v/v) H_2O_2). Signals were detected using an X-ray film (Fujifilm, Willich) and a Curix 60 film processor (AGFA HealthCare GmbH, Bonn).

3.3.6 Indirect immunofluorescence analysis of cells

Cells grown in 4-well chamber slides (Corning, New York) were washed twice with PBS and were then fixed for 15 min in 4 % paraformaldehyde at RT. Cells were washed three times with PBS and were permeabilised using 0.2 % Triton X-100 (Sigma) for 5 min at RT. Following three washes with PBS, non-specific binding sites were blocked by incubation for 30 min in blocking buffer (1 % BSA in PBS). Subsequently, cells were incubated with polyclonal anti-GRHL2 antibody (dilution 1:250) (Werner et al., 2013) followed by Alexa-546 goat anti-rabbit IgG (dilution 1:250) (Thermo Fisher Scientific, Rockford, USA) in blocking solution, for 90 min at RT each. After each individual antibody incubation, cells were washed four times with PBS, counterstained with DAPI (Sigma), and were finally mounted in Mowiol (Calbiochem, La Jolla, CA). Microscopic analysis was performed using an Axioplan2 microscope equipped with AxioVision SE64 (Re. 4.9.1) software (Zeiss, Jena, Germany).

3.3.7 Pull-down assay

The pull-down assay was performed essentially as described by Choi and coworkers (Choi et al., 2013). Cells co-transfected with GRHL2 and His_6-tagged SUMO expression plasmids were harvested and re-suspended in Buffer A (6 M Guanidine–HCl, 0.1 M Na_2HPO_4/NaH_2PO_4, 10 mM imidazole, pH 8.0). Following sonication at 70 % amplitude and cycle 0.5 using an ultrasonic processor (UP50H Hielscher, Teltow), the lysate was added into 50 µl of Ni–NTA agarose (QIAGEN, Hilden) pre-equilibrated in Buffer A and rotated at room temperature for 4-6 h. The beads were then washed with Buffer A, Buffer A/TI (1 vol. of Buffer A, 3 vol. of Buffer TI (25 mM Tris–HCl, 20 mM imidazole, pH 6.8)) and Buffer TI. After centrifugation for 10 sec at top seed (Eppendorf centrifuge 5415R) at RT, the supernatant was discarded. Any residual buffer trapped in agarose beads was removed using a 26 G needle. Bound proteins were released from the matrix by adding 50 µl of 2 × SDS sample buffer (see section 2.3.1) and boiling at 95-100 °C for 10 min. The supernatant was harvested after centrifugation at top speed (Eppendorf centrifuge 5415R) for 10 min at RT and was then subjected to western blot analysis.

3.3.8 Luciferase Reporter Assays

For luciferase reporter assays, 5 x 10^4 COS-7 cells were seeded into each well of a 24-well plate. The next day, cells were transiently co-transfected with 0.25 µg of expression plasmids encoding wild-type or mutant GRHL2 proteins, 0.25 µg of a reporter plasmid containing five copies of the GRHL2 consensus binding site (AACCGGTT) upstream of a minimal promoter and the *Firefly* luciferase (*luc2*) gene (unpublished), and 5 ng of the pGL4.74 normalization plasmid encoding the *Renilla* luciferase (*hRluc*) (Promega, Wisconsin, USA) using Lipofectamine 3000 according to the manufacturer's instructions. Forty-eight hours post transfection, cell extracts were prepared using the Passive Lysis Buffer (PLB) (Promega). Samples were assayed in triplicate for luciferase activities using a Dual luciferase assay kit (Promega) with the GloMax Discover Multimode plate reader (Promega). Before calculating the fold activation value, luciferase activity of each sample was normalized with respect to the activity of *Renilla* luciferase. Luciferase activity from COS-7 cells transfected with an empty vector was set arbitrarily at 1 for calculation of fold activation.

3.4 Bioinformatic analyses

3.4.1 Computer-aided prediction of SUMOylation sites by *in silico*-analysis

Several computational approaches for SUMOylation site prediction within proteins are publically available. The resources, including corresponding online webservers and references, used for prediction of possible SUMOylation sites within the human GRHL2 protein are listed below.

Publically available webservers for SUMOylation site prediction		
Resource	**Web URL**	**Reference**
ELM	http//elm.eu.org/	(Puntervoll et al., 2003)
SUMOplot	http//www.abgent.com/sumoplot/	(Abgent, I.S.D., San Diego, CA)
GPS-SUMO 2.0	http://sumosp.biocuckoo.org/online.php	(Zhao et al., 2014)
SUMOsp 2.0.4	http://bioinformatics.lcd-ustc.org/sumosp/	(Ren et al., 2009)
SUMOhydro	http://protein.cau.edu.cn/others/SUMOhydro/	(Chen et al., 2012)
JASSA	http://www.jassa.fr	(Beauclair et al., 2015)
pSUMO-CD	http://www.jci-bioinfo.cn/pSumo-CD	(Jia et al., 2016)
PCI-SUMO	http://bioinf.sce.carleton.ca/SUMO/start.php/	(Green et al., 2006)
SUMOFI	http://cbg.garvan.unsw.edu.au/sumofi/form.do	(de Castro et al., 2006)
PHOSIDA	http://www.phosida.com	(Gnad et al., 2007)

4. Results

In this study, a large number of expression constructs encoding wild-type or mutant GRHL2 as well as many other proteins (e.g., SUMO-1/2/3, PIAS-1, 2, 3, 4 etc.) were created. The expression plasmids (Table 5), primers for amplification of cDNAs (Table 6), and oligonucleotides for mutagenesis (Table 7) are listed in the Appendix. The identity of all expression plasmids was confirmed by DNA sequencing using primers listed in Table 8 (see Appendix) and functionality was verified by transient transfection into COS-7 cells, followed by Western blot analysis using appropriate antibodies. For the sake of simplicity, the generation and characterization of the expression plasmids is not detailed in the Results section.

4.1 GRHL2 is post-translationally modified by SUMOylation

Covalent conjugation of SUMO proteins to substrate can cause shifts in the molecular weight by 10-40 kDa (Wilkinson and Henley, 2010) thus enabling the detection of SUMOylated forms of a protein by a gel shift assay. To determine whether GRHL2 is modified by SUMOylation, expression plasmids encoding GRHL2 (# 1) and activated, HA-tagged SUMO-1-GG or SUMO-2-GG (# 16, 17) were transiently co-transfected into easy-to-transfect COS-7 cells and cell lysates were harvested 48 hours post-transfection. Lysates were separated by SDS-PAGE, followed by Western blot analysis using GRHL2-specific antibodies or antibodies recognizing the HA-epitope. The results presented in Fig. 4 clearly show that in addition to the non-modified GRHL2 protein with an apparent molecular weight of about 72 kDa, an additional GRHL2-specific band migrating at about 110 kDa could be detected. The detection of this slower migrating GRHL2-specific band was strictly dependent on the co-expression of activated SUMO-1-GG or SUMO-2-GG and the presence of *N*-ethylmaleimide (NEM), a potent inhibitor of cysteine (thiol) proteases required for the preservation of the SUMOylated state of substrate proteins, in the lysis buffer. These results strongly suggest that GRHL2 represents a novel substrate for SUMOylation. Moreover, the results of the gel shift assay showing the appearance of only one higher molecular weight GRHL2 species also indicate that GRHL2 most likely is modified by SUMOylation at a single lysine residue.

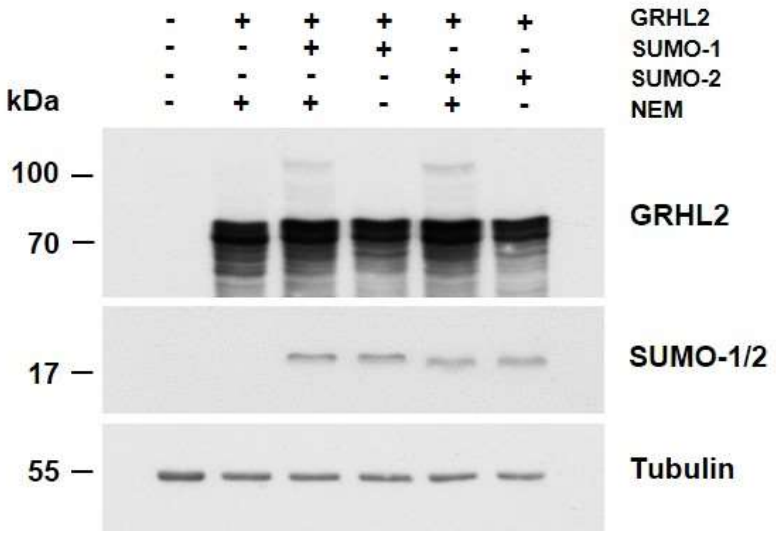

Fig. 4. GRHL2 is post-translationally modified by SUMOylation. Expression plasmids encoding GRHL2 and activated, HA-tagged SUMO-1-GG or SUMO-2-GG were transiently co-transfected into COS-7 cells and cell lysates were prepared in the presence or absence of the cysteine protease inhibitor NEM as indicated. Lysates were separated by SDS-PAGE, followed by Western blot analysis using GRHL2-specific antibodies or antibodies recognizing the HA-epitope for detection of SUMO-1/2 proteins. Equal loading was demonstrated using an antibody recognizing tubulin.

4.2 Identification of SUMOylation sites in GRHL2 using computational approaches

The human GRHL2 protein (NP_079191) contains a total of 53 lysine residues which potentially could act as acceptor sites for SUMOylation. To identify high-probability SUMOylation sites in GRHL2, we performed a computational analysis of the GRHL2 protein sequence using ten different web-based SUMOylation prediction tools, including the probably most advanced resources JASSA (Beauclair et al., 2015), GPS-SUMO 2.0 (Zhao et al., 2014), and SUMO Plot (http//www.abgent.com/sumoplot/) (Table 1). Although nineteen randomly distributed lysine residues were predicted to represent possible modification sites, only seven amino acids were predicted to be SUMOylation sites by at least two computational approaches. For this reason, only these residues (K159, K205, K366, K424, K453, K531, and K556) were considered as candidate modification sites and were selected for experimental validation by mutational analysis.

Table 1 Predicted SUMOylation sites in the human GRHL2 protein*

Resource	K 135	K 159	K 205	K 261	K 287	K 318	K 366	K 414	K 424	K 435	K 438	K 453	K 462	K 531	K 556	K 561	K 570	K 586	K 619
ELM	-	X	-	-	-	-	-	-	-	-	-	-	-	X	-	-	-	-	-
SUMO Plot	-	X 0.9	X 0.9	X 0.3	-	-	X 0.4	X 0.5	-	X 0.3	-	-	X 0.5	X 0.8	X 0.8	-	X 0.3	-	-
GPS-SUMO 2.0	-	X H	X H	-	-	-	-	-	-	-	-	-	-	X H	-	-	-	X M	X M
SUMOsp 2.0.4	X M	X H	X M	-	-	-	-	-	-	-	-	-	-	X H	-	-	-	-	-
SUMOhydro	-	X 0.7	-	-	-	-	-	-	-	-	-	-	-	X 0.7	-	-	-	-	-
JASSA	-	X H	X L	-	-	-	X H	-	X H	-	-	X L	-	X L	-	-	-	-	-
pSUMO-CD	-	X	-	-	X	X	-	-	-	-	X	-	-	X	X	X	-	-	-
PCI-SUMO	-	X 0.9	-	-	-	-	-	-	-	-	-	-	-	X 0.8	-	-	-	-	-
SUMOFI	-	X	-	-	-	-	-	-	-	-	-	-	-	X	-	-	-	-	-
PHOSIDA	-	X	-	-	-	-	X	-	X	-	-	X	-	X	-	-	-	-	-
Selected for experimental validation	-	X	X	-	-	-	X	-	X	-	-	X	-	X	X	-	-	-	-

* Putative acceptor sites for SUMOylation in the human GRHL2 protein (NP_079191) identified by web-based prediction tools listed on the left. Scores (SUMOplot, SUMOhydro) or confidence values (PCI-SUMO) are shown in red. Computational analysiswas performed using a high (H) or medium (M) threshold as indicated in blue (GPS-SUMO 2.0 and SUMOsp 2.0.4). For JASSA, best predictive scores (PS) high (H) or low (L) are shown in green. Lysine residues selected for further experimental validation are indicated below.

4.3 Identification of K159 as the major SUMOylation site in GRHL2 by site-directed mutagenesis

To identify the major SUMOylation-site within the GRHL2 protein, candidate lysine residues K159, K205, K366, K424, K453, K531, and K556 were replaced by the basic amino acid arginine (R) using site-directed mutagenesis. Expression constructs encoding wild-type or mutant GRHL2 proteins (# 1, 2, 4-9) and activated, HA-tagged SUMO-1-GG or SUMO-2-GG (# 16, 17) were transiently co-transfected into COS-7 cells and cell lysates harvested 48 hours post-transfection were then subjected to SDS-PAGE and Western blot analysis using antibodies for the detection of GRHL2 and SUMO-1/2 proteins. The results shown in Fig. 5 demonstrate that wild-type GRHL2 proteins and all but the GRHL2 K159R mutant were modified by SUMOylation, as indicated by the detection of the high-molecular weight GRHL2 proteins by Western blot analysis. This result was independent of whether SUMO-1 or SUMO-2 was co-expressed in COS-7 cells, thus further substantiating a crucial role of residue K159 in the SUMOylation of the GRHL2 protein. A Western blot analysis using antibodies recognizing SUMO-1/2 proteins and tubulin demonstrated that failure to detect SUMOylated GRHL2 (S-GRHL2) in cells transfected with the GRHL2 K159R mutant, was not the result of variable co-transfection efficiency or unequal loading of samples.

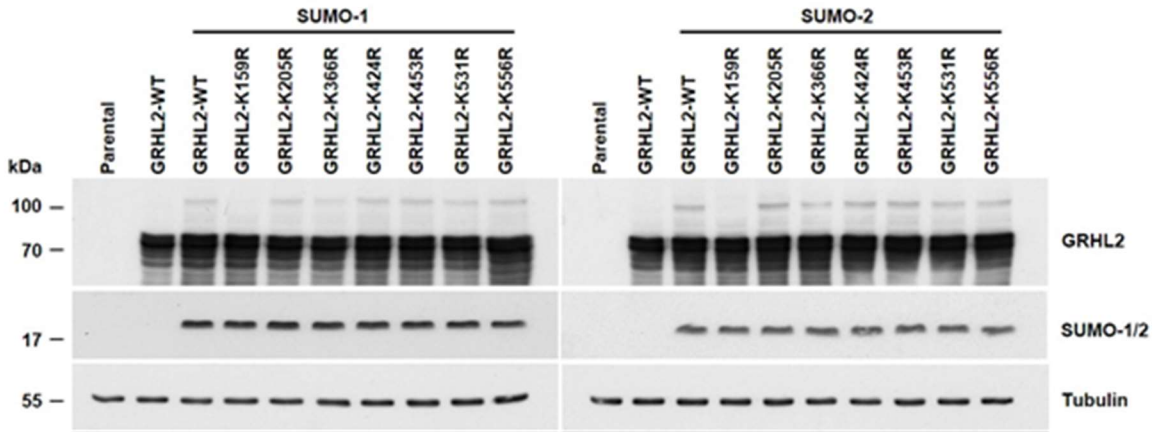

Fig. 5. Identification of K159 as the major SUMOylation site in GRHL2 by site-directed mutagenesis. Expression constructs encoding wild-type or mutant GRHL2 proteins and activated, HA-tagged SUMO-1-GG or SUMO-2-GG were transiently co-transfected into COS-7 cells and cell lysates harvested 48 hours post-transfection were then subjected to SDS-PAGE and Western blot analysis using antibodies for the detection of GRHL2 and SUMO-1/2 proteins. Equal loading was demonstrated using an antibody recognizing tubulin.

Lysine residues that are modified by SUMOylation are often embedded within a consensus motif ΨKxD/E, where Ψ is a hydrophobic amino acid, K the SUMOylated lysine residue, x any amino acid, and D/E (aspartic/glutamic acid) (Wilkinson and Henley, 2010). The major SUMOylation site identified in GRHL2 is embedded in a sequence (^{158}VKAE^{161}) which conforms to this consensus motif. Still, to obtain additional evidence for an involvement of K159 in SUMOylation of GRHL2, a GRHL2 mutant (GRHL2 E161A) (# 3) was generated in which the indispensable residue E161 of the consensus motif ΨKxD/E was replaced by alanine (A). Using an identical experimental approach as described above, a critical role of E161 in SUMOylation of GRHL2 could be demonstrated (Fig. 6). Disruption of the target sequence for SUMOylation in GRHL2 (^{158}V<u>K</u>A<u>E</u>^{161}) at two distinct positions uniformly abolished SUMOylation of GRHL2, as demonstrated by gel shift analysis. These results further demonstrate that the acceptor lysine for SUMOylation of GRHL2 is the K159 residue. Also, these findings strongly suggest that high molecular weight GRHL2 proteins migrating at about 110 kDa represent SUMOylated GRHL2 (S-GRHL2) and are not the result of a modification of K159 by another post-translational modification requiring a lysine as an acceptor residue (e.g., acetylation, ubiquitination etc.).

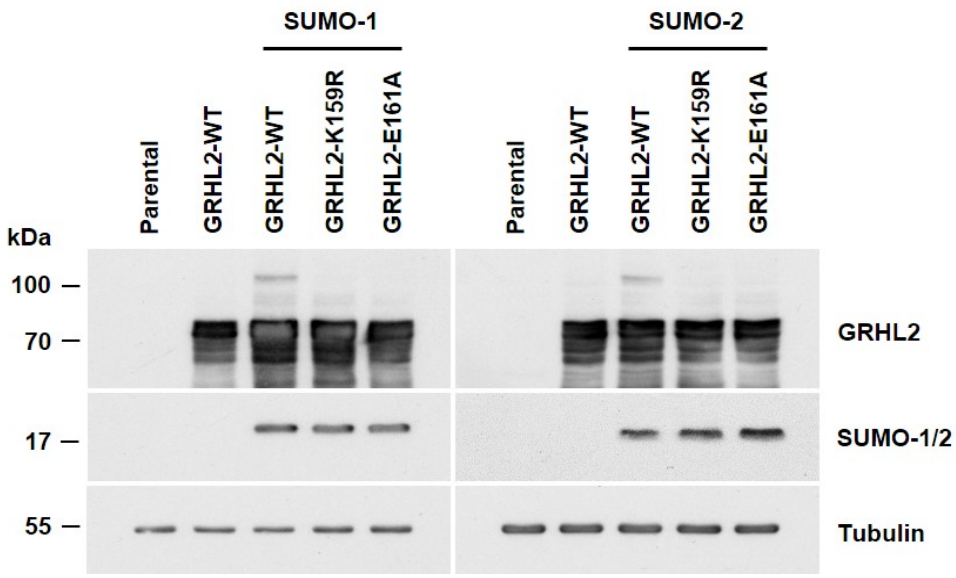

Fig. 6. SUMOylation of GRHL2 at K159 occurs at a ΨKxD/E consensus motif. Expression constructs encoding wild-type or mutant GRHL2 proteins and activated, HA-tagged SUMO-1-GG or SUMO-2-GG were transiently co-transfected into COS-7 cells and cell lysates harvested 48 hours post-transfection were then subjected to SDS-PAGE and Western blot analysis using antibodies for the detection of GRHL2 and SUMO-1/2 proteins. An antibody detecting tubulin served as a loading control in these experiments.

To obtain further evidence for K159 as the site of modification by the SUMOylation pathway, pull-down experiments for the specific enrichment of SUMOylated proteins was performed. To this end, expression plasmids encoding wild-type GRHL2 or mutant GRHL2 (K159R) proteins (# 1, 2) and were co-transfected with expression plasmids coding for His_6-tagged, activated SUMO-1/2 proteins (# 18, 19) into COS-7 cells. Forty-eight hours post-transfection, cells were harvested using denaturing conditions, total SUMOylated proteins were recovered by nickel-affinity chromatography, and specific proteins probed for by Western blot analysis. The results shown in Fig. 7 demonstrate that the high molecular weight GRHL2 protein representing S-GRHL2 specifically could be enriched using this pull-down experiment. The specificity of this approach is demonstrated by the absence of the 110 kDa band representing S-GRHL2 in lysates from cells expressing mutant GRHL2 (K159R) proteins and also by the absence of non-modified GRHL2 proteins migrating at about 72 kDa in all cell lysates.

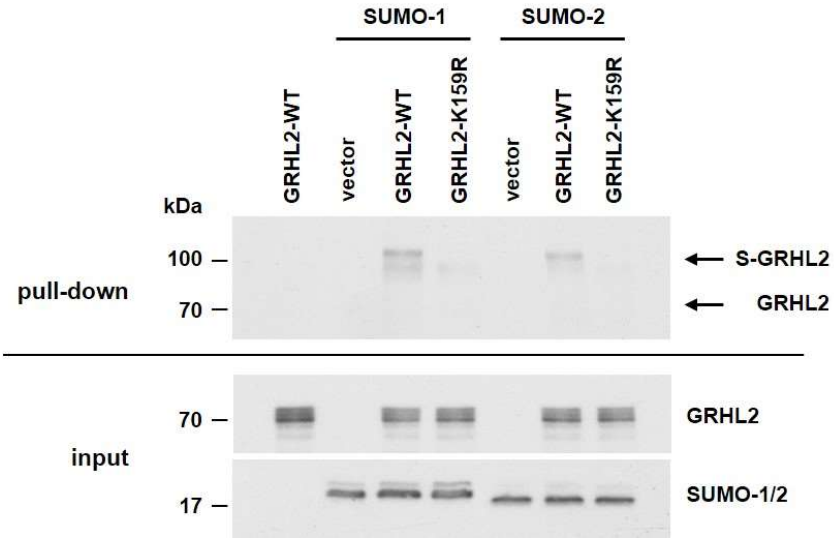

Fig. 7. Enrichment of SUMOylated GRHL2 proteins using a pull-down approach. Expression constructs encoding wild-type or mutant GRHL2 proteins and activated, His_6-tagged SUMO-1-GG or SUMO-2-GG were transiently co-transfected into COS-7 cells. Forty-eight hours post-transfection, cell lysates were prepared using denaturing conditions. Total SUMOylated proteins were recovered by Nickel-affinity chromatography, and specific proteins were probed for by Western blot analysis using GRHL2-specific antibodies (*top*). Bands representing non-SUMOylated and SUMOylated GRHL2 (S-GRHL2) are marked with arrows. To demonstrate equal transfection efficiency, lysates used for pull-down assay (*input*) were also subjected to Western blot analysis using antibodies specific for GRHL2 or SUMO-1/SUMO-2 (*bottom*).

In summary, these results strongly suggest that the GRHL2 protein is modified by SUMOylation. GRHL2 appears to be SUMOylated at residue K159 which is located in the hinge region linking the N-terminal transactivation domain (TAD; residues 1-135) with the DNA-binding region (DBD; residues 245-494) (Fig. 8).

Fig. 8. Localisation of the major SUMOylation-site of GRHL2. The major SUMOylation site in GRHL2 (K159) is located in the hinge region linking the N-terminal transactivation domain (TAD; residues 1-135) with the DNA-binding region (DBD; residues 245-494). For a detailed description of the GRHL2 domain structure see Section 1.3.2.

4.4 SUMOylation of GRHL2 is enhanced by PIAS proteins

SUMOylation of substrate proteins can be enhanced in the presence of SUMO E3 ligases. This relatively small family of structurally unrelated proteins also includes the PIAS proteins encoded by four genes, namely PIAS1, PIAS2, PIAS3, and PIAS4 (Rytinki et al., 2009). Interestingly, PIAS3 was identified as candidate interaction partner of GRHL2 in a Yeast-Two-Hybrid based screen using human full-length GRHL2 as a bait (unpublished observation). Furthermore, an entry in the INTact database containing experimentally validated protein-protein interactions suggested a possible interaction between GRHL2 and PIAS2 (https://www.ebi.ac.uk/intact/), implying that GRHL2 might be able to interact with several members of the PIAS family of proteins.

To investigate whether PIAS proteins are able to promote SUMO conjugation to GRHL2, expression constructs encoding wild-type GRHL2 (# 10) or mutant GRHL2 (K159R) (#11) proteins, PIAS1-4 proteins (# 20-23), and SUMO1 or SUMO2 proteins (#16, 17) were co-transfected into COS-7 cells. Ectopic expression of all four PIAS protein markedly enhanced SUMO modification of GRHL2 by both SUMO1 and SUMO2, suggesting that all four PIASs exhibit SUMO E3 activity in the GRHL2 SUMOylation pathway (Fig. 9). Gradual differences in SUMOylation induced by individual PIAS proteins fail to reach significance and possibly just reflect

variability in co-transfection experiments as well as differences in PIAS 1-4 expression levels (Fig. 9).

Taken together, these results suggest that SUMOylation of GRHL2 is modulated by all four members of the PIAS family of proteins. At least under experimental conditions in which both GRHL2 and PIAS proteins are transiently overexpressed in COS-7 cells, there seems to be no preference in the interaction between GRHL2 and the SUMO E3 ligases PIAS1-4.

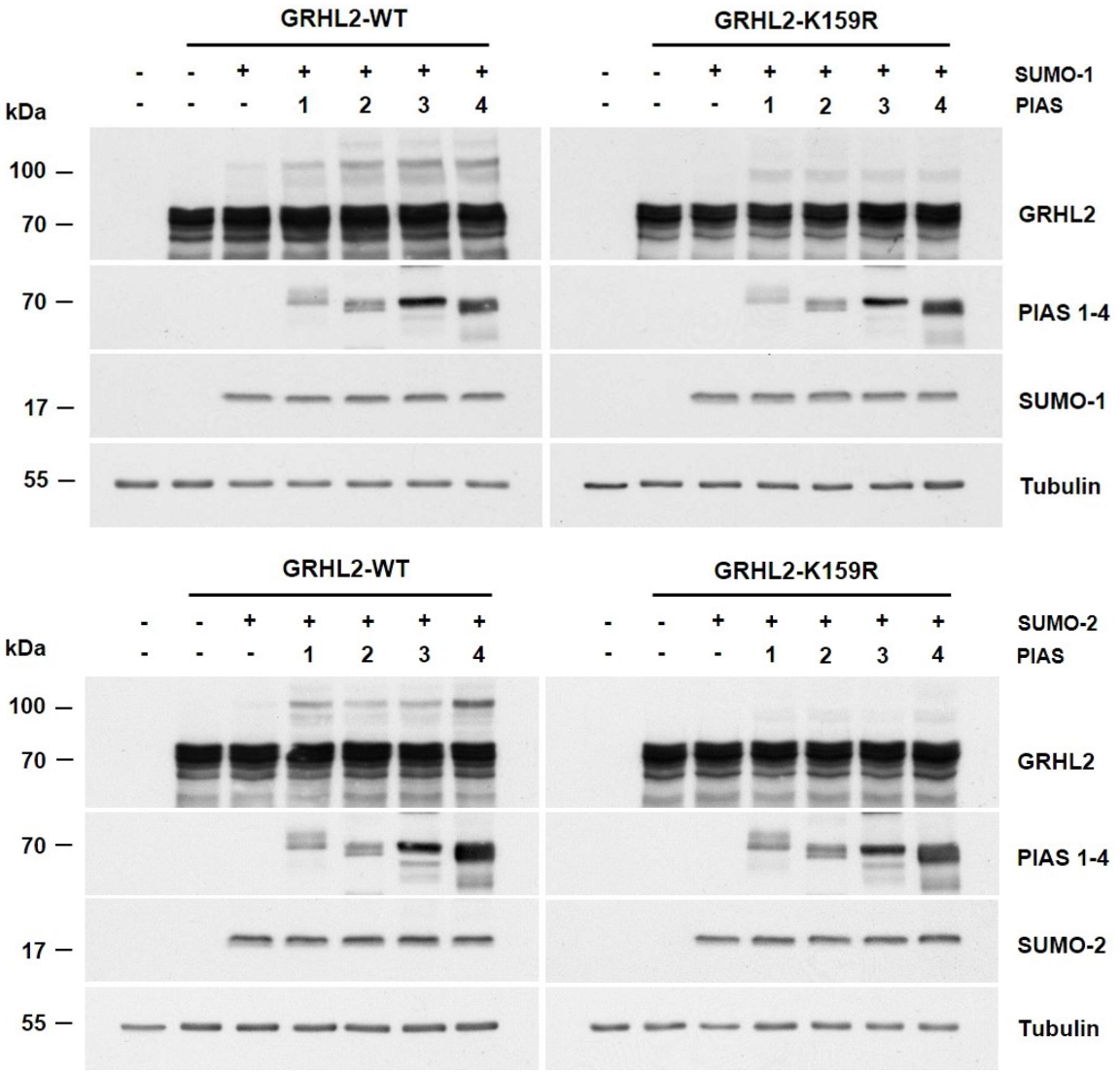

Fig. 9. SUMOylation of GRHL2 is enhanced by PIAS proteins. Expression constructs encoding wild-type GRHL2 or mutant GRHL2 (K159R) proteins, HA-tagged PIAS1-4 proteins, and SUMO-1 or SUMO-2 proteins were transiently co-transfected into COS-7 cells as indicated. Forty-eight hours post-transfection, cell lysates were prepared and were then subjected to SDS-PAGE and Western blot analysis using antibodies for the detection of GRHL2, HA-tagged PIAS 1-4 proteins, and HA-tagged SUMO-1 (*top*) and SUMO-2 proteins (*bottom*). An antibody detecting tubulin served as a loading control in these experiments.

4.5 The subnuclear localisation of GRHL2 is not influenced by SUMOylation

SUMO proteins covalently attached to substrate proteins can act a molecular glue enabling protein-protein interactions and the formation of higher-order structures. Not surprisingly, SUMOylation of target proteins therefore often influences the distribution of proteins and their association with subcellular structures. To test whether SUMOylation affects the subnuclear localisation of GRHL2, expression plasmids encoding wild-type GRHL2 (# 10) as well as mutant GRHL2 (K159R) and GRHL2 (E161A) proteins (# 11, 12) were transiently transfected into COS-7 cells and the distribution of GRHL2 proteins was determined by indirect immunofluorescence analysis using GRHL2-specific antibodies.

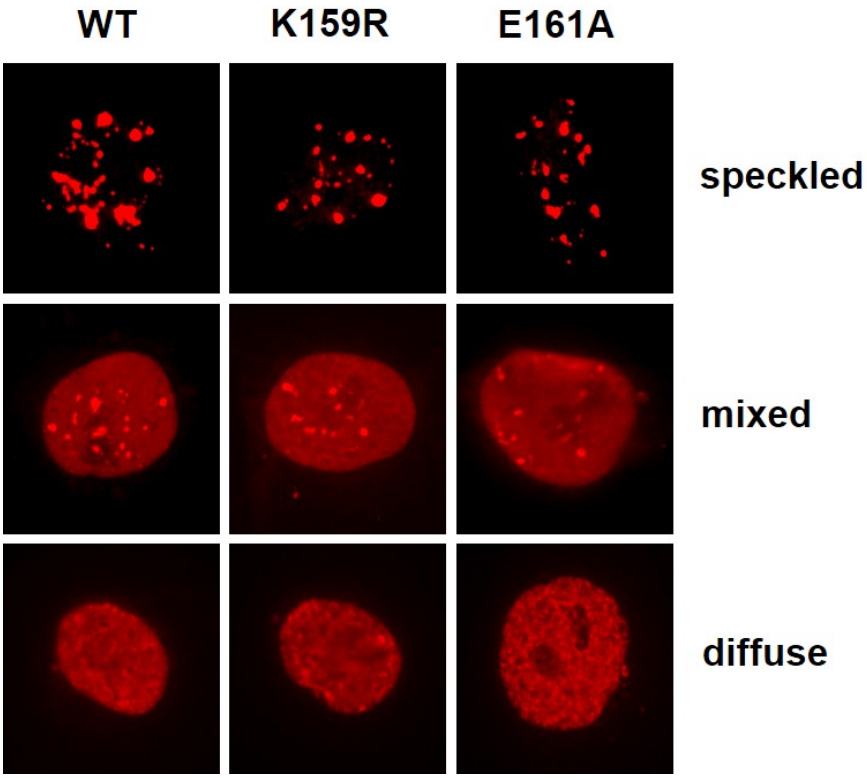

Fig. 10. The subnuclear localisation of GRHL2 is not influenced by SUMOylation. Expression plasmids encoding wild-type GRHL2 as well as mutant GRHL2 (K159R) and GRHL2 (E161A) proteins were transiently transfected into COS-7 cells and the distribution of GRHL2 proteins was determined by indirect immunofluorescence analysis using GRHL2-specific antibodies. Depending on the subnuclear distribution of GRHL2 proteins, distinct patterns classified as speckled, mixed, or diffuse could be observed. Microscopic analysis was performed using an Axioplan2 microscope equipped with AxioVision SE64 software. Original magnification 400x.

Interestingly, distinct patterns of staining could be detected in COS-7 cells transfected with wild-type or SUMOylation-deficient GRHL2 proteins. In general, transfected cells showed either a diffuse-type, a mixed-type, or a granular-type of distribution of GRHL2 proteins (Fig. 10). Irrespective of the expressed GRHL2 protein (wild-type vs. mutant) about 80-90 % of cells showed a mixed/diffuse staining pattern and only relatively rarely a speckled staining. A high variability in size and number of granules ranging from about 5-30 could be observed.

In conclusion, these results did not reveal striking differences between wild-type and mutant GRHL2 proteins, implying that SUMOylation does not influence the subnuclear distribution of GRHL2 proteins in COS-7 cells. As co-expression of EGFP-tagged, activated SUMO-1-GG also did not influence the localisation of GRHL2 in the nucleus (data not shown), it can be concluded that distinct GRHL2 staining patterns are not due to a differential modification of GRHL2 by SUMOylation.

4.6 Regulation of the GRHL2 subnuclear localisation through association with PIAS proteins

To investigate whether the interaction of GRHL2 with members of the PIAS family of proteins could influence the subnuclear localisation of GRHL2 in COS-7 cells another set of immunofluorescence analyses were performed. Wild-type GRHL2 proteins (# 10) together with EGFP-tagged PIAS1-4 proteins (# 24-27) were introduced into COS-7 cells by transient transfection of plasmid DNA and the distribution of GRHL2 and PIAS1-4 proteins was determined by indirect immunofluorescence analysis using a GRHL2-specific antibody. Remarkably, all four PIAS proteins drastically changed the subnuclear distribution of GRHL2 in that about 70-80 % of co-transfected cells showed a striking speckled staining pattern (Fig. 11). The percentage of cells showing a speckled GRHL2 staining thus is significantly higher than in COS-7 cells expressing only GRHL2 (see 4.5).

Significant differences between distinct PIAS proteins in dragging GRHL2 into granular structures could not be observed. Most importantly, all PIAS proteins clearly co-localised with GRHL2 in these granules, thus further confirming a physical and functional association between GRHL2 and this class of E3 SUMO ligases.

These results clearly establish PIAS1-4 proteins as novel modulators of GRHL2 localisation and function in cells.

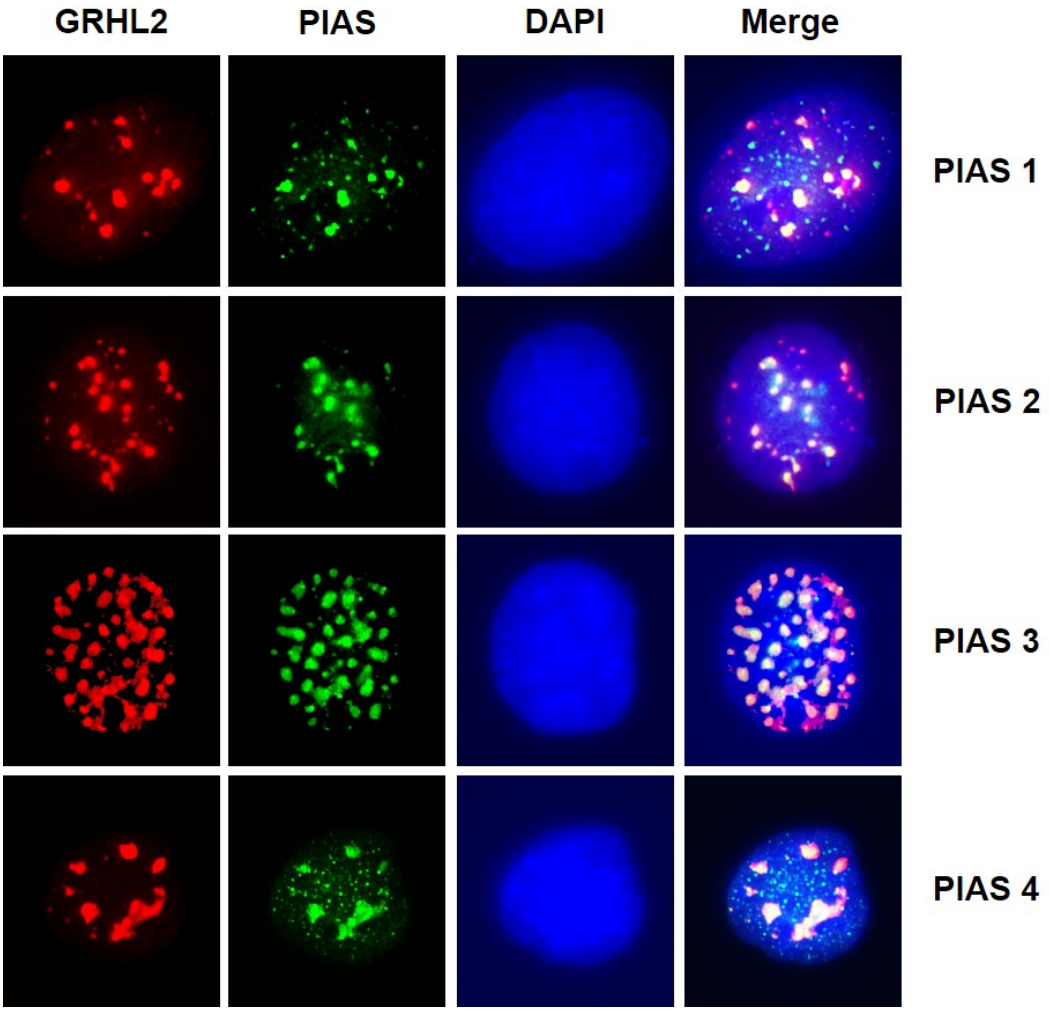

Fig. 11. Co-localisation of GRHL2 with PIAS proteins in COS-7 cells. Wild-type GRHL2 proteins and EGFP-tagged PIAS1-4 proteins were introduced into COS-7 cells by transient transfection of plasmid DNA. The distribution of GRHL2 and PIAS1-4 proteins was determined by indirect immunofluorescence analysis using a GRHL2-specific antibody. Nuclei were visualized using DAPI. Microscopic analysis was performed using an Axioplan2 microscope equipped with AxioVision SE64 software. Original magnification 400x.

4.7 SUMOylation positively regulates GRHL2 transcriptional activity

SUMOylation can lead to activation or suppression of the activity of transcriptional regulators. To analyse how SUMOylation affects GRHL2 activity, two distinct but complementary approaches were employed.

For luciferase reporter assays, COS-7 cells were transiently co-transfected with expression plasmids encoding wild-type GRHL2 (# 10) or mutant GRHL2 (K159R) and GRHL2 (E161A) proteins (# 11, 12), a reporter plasmid containing five copies of the GRHL2 consensus binding site (AACCGGTT) upstream of a minimal promoter and the *Firefly* luciferase (luc2) gene, and a normalisation plasmid encoding the *Renilla* luciferase (hRluc) using Lipofectamine 3000. Forty-eight hours post transfection, cell extracts were assayed in triplicate for luciferase activities using a Dual luciferase assay kit. Before calculating the fold activation value, luciferase activity of each sample was normalised with respect to the activity of Renilla luciferase. Luciferase activity from COS-7 cells transfected with an empty vector was set arbitrarily at 1 for calculation of fold activation.

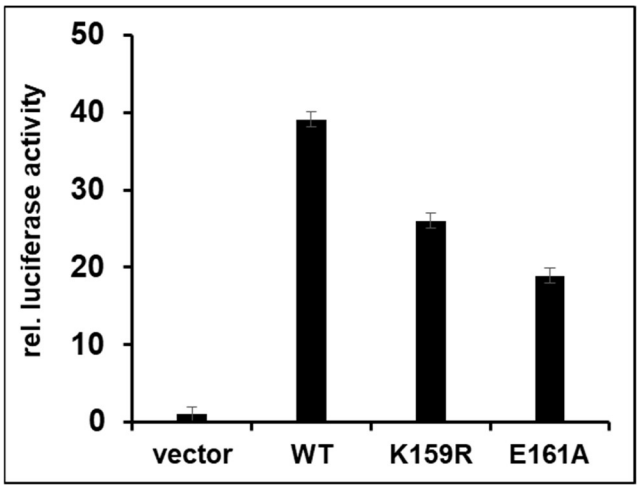

Fig. 12. SUMOylation positively regulates GRHL2 transcriptional activity. Cells were transiently co-transfected with expression plasmids encoding wild-type or mutant GRHL2 proteins, a reporter plasmid containing five copies of the GRHL2 consensus binding site (AACCGGTT) upstream of a minimal promoter and the Firefly luciferase (luc2) gene, and the pGL4.74 normalization plasmid encoding the Renilla luciferase (hRluc). Forty-eight hours post transfection, cell extracts were prepared and were assayed in triplicate for luciferase activities using a Dual luciferase assay kit with the GloMax Discover Multimode plate reader. Before calculating the fold activation value, luciferase activity of each sample was normalised with respect to the activity of Renilla luciferase. Luciferase activity from COS-7 cells transfected with an empty vector was set arbitrarily at 1 for calculation of fold activation. *Error bars*, S.D.

The results included in Fig. 12 show that wild-type GRHL2 induced an about 38-fold induction of the *Firefly* luciferase reporter gene as compared to empty-vector control. Interestingly, GRHL2 mutant proteins with disrupted SUMOylation sites (GRHL2 K159R and GRHL2 E161A) showed a significantly reduced ability to transactivate reporter gene expression. Differences in the ability to induce reporter gene expression between the two SUMOylation-deficient GRHL2 mutants most likely are attributable to structural changes introduced by mutational analysis. Still, these results indicate that SUMOylation enhances GRHL2 transcriptional activity in luciferase reporter assays.

To further substantiate this finding, retroviral expression constructs encoding wild-type GRHL2 (# 13) or mutant GRHL2 (K159R) and GRHL2 (E161A) proteins (# 14, 15) were generated and introduced into GRHL2-deficient MDA-MB-231 breast cancer cells by means of retroviral gene transfer. Infected cells were selected with puromycin and pooled clones were analysed for GRHL2 expression by Western blot analysis (Fig. 13). The results presented in Fig. 13 demonstrate GRHL2 protein expression in MDA-MB-231 cells infected with GRHL2-expressing retroviruses, but not in parental cells or cells infected with an empty vector. Thus, the expression analysis demonstrates the successful establishment of a model system for studying SUMOylation-dependent GRHL2 target gene expression.

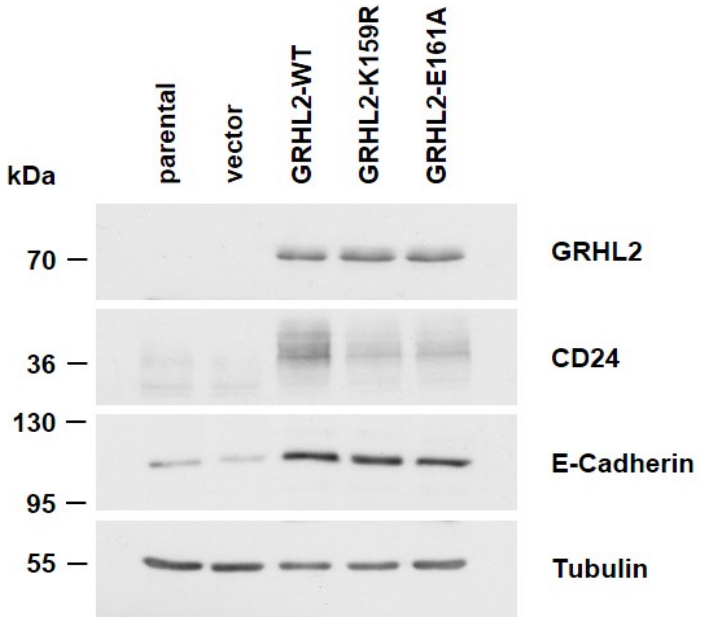

Fig. 13. Establishment of a model system for the analysis of the SUMOylation-dependent regulation of GRHL2 target gene expression. Retroviral expression constructs encoding wild-type GRHL2 or mutant GRHL2 (K159R) and GRHL2 (E161A) proteins were generated and introduced into GRHL2-deficient MDA-MB-231 breast cancer cells by means of retroviral gene transfer. Infected cells were selected with puromycin and pooled clones were analysed for GRHL2 expression by Western blot analysis using GRHL2-specific antibodies. Expression of selected GRHL2 target genes was analysed using antibodies for CD24 and E-Cadherin, respectively. An antibody detecting tubulin served as a loading control in these experiments.

Next, expression of selected GRHL2 target genes CD24 and E-Cadherin (CDH1) was analysed by Western blot analysis. Consistent with results published by Werner (2013) (Werner et al., 2013), GRHL2 induced expression of CD24 and E-Cadherin in MDA-MB-231 cells (Fig. 13). Interestingly, expression levels in cells expressing GRHL2 mutants GRHL2 K159R and GRHL2 E161A was significantly lower as in cells expressing wild-type GRHL2 proteins, suggesting that SUMOylation enhances transcriptional activity of GRHL2.

A qRT-PCR analysis of GRHL2, CD24, and E-Cadherin gene expression in the different MDA-MB-231 cell clones using primers listed in Table 9 (see Appendix) was also performed. The results shown in Fig. 14 confirm the data obtained by Western blot analysis. Although GRHL2 K159R and GRHL2 E161A mutant proteins are expressed at a slightly higher level (about 1.1- or 1.2-fold, respectively) compared to wild-type GRHL2, CD24 and E-Cadherin/CDH1 target gene expression still was reduced as compared to wild-type GRHL2 expressing cells.

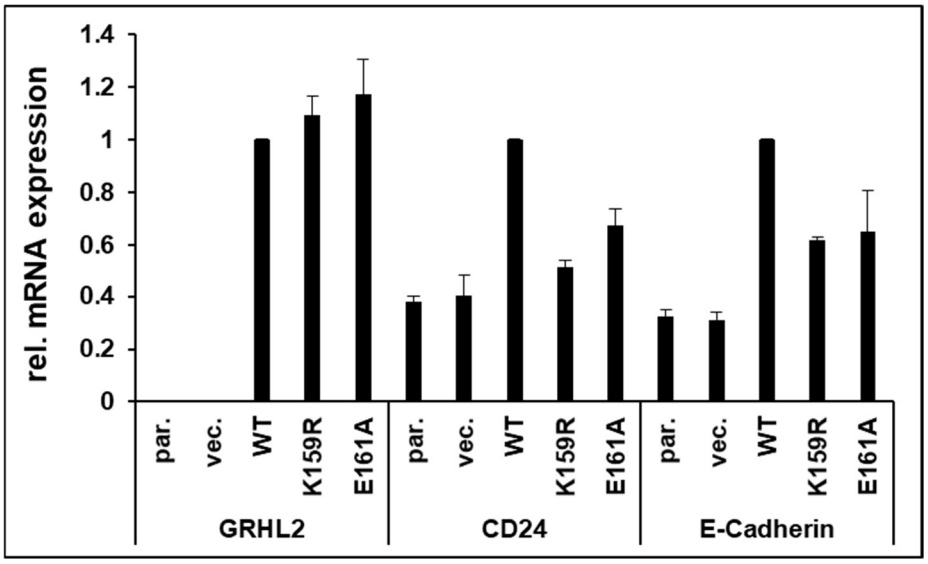

Fig. 14. qRT-PCR analysis of GRHL2 target genes in MDA-MB-231 breast cancer cells. Retroviral expression constructs encoding wild-type GRHL2 or mutant GRHL2 (K159R) and GRHL2 (E161A) proteins were generated and introduced into GRHL2-deficient MDA-MB-231 breast cancer cells by means of retroviral gene transfer. Infected cells were analysed for GRHL2, CD24, and E-Cadherin mRNA expression by qRT-PCR analysis. Expression of GRHL2, CD24, and E-Cadherin mRNAs in cells expressing wild-type GRHL2 was set arbitrarily at 1, respectively.

These results together with those obtained by luciferase reporter assays therefore strongly suggest that SUMOylation enhances transcriptional activity of the GRHL2 transcription factor.

5. Discussion

Breast cancer metastasis is an extremely complex process and remains to be a major medical challenge until now. Limited clinical strategy raises urgent demands for a better understanding of the molecular mechanism driving metastatic process. The epithelial-to-mesenchymal transition (EMT) is a morphogenetic program that has been shown to play a crucial role in cancer progression and metastasis (Kalluri and Weinberg, 2009). Several independent groups identified the largely uncharacterized developmental transcription factor Grainyhead-like 2 (GRHL2) as a potent suppressor of EMT and breast cancer metastasis (Cieply et al., 2013; Cieply et al., 2012; Werner et al., 2013). Although transcription factors in general can be regulated by various posttranslational modifications (PTMs), nothing is known about possible PTMs regulating GRHL2 activity in breast cancer cells. Preliminary results obtained by Yeast-Two-Hybrid (Y2H) protein-protein interaction screen suggest an interaction of GRHL2 with various components of SUMOylation machinery, implying that SUMOylation potentially could represent a novel regulatory mechanism of GRHL2 in breast cancer cells. Major aim of this project therefore is to investigate whether GRHL2 activity can be modulated by SUMOylation in breast cancer cells.

GRHL2 is post-translationally modified by SUMOylation

The post-translational modification SUMOylation is defined as the covalent and reversible conjugation of a member of the SUMO (Small Ubiquitin-like Modifier) family of proteins to a lysine residue in the substrate proteins (Wilkinson and Henley, 2010). SUMOylation is indispensable for the normal function of all eukaryotic cells and contributes to a wide variety of cellular processes, such as maintenance of protein stability, regulation of protein activity, and protein subcellular localisation (Hay, 2005; Heun, 2007; Johnson, 2004). Four SUMO paralogues have been identified in mammals, designated SUMO1 to SUMO4, whereas SUMO2 and SUMO3, due to a 95 % sequence similarity, are often collectively referred to as SUMO2/3 (Wilkinson and Henley, 2010). Immature SUMO proteins are processed by SUMO-specific proteases (SENP) and undergo C-terminal cleavage. The mature form of SUMOs are then engaged in the SUMOylation cycle and covalently bind to

target proteins via an enzymatic cascade, including activation by E1 enzyme, conjugation by E2 enzyme, and ligation by E3 enzyme. Subsequently, de-SUMOylation process is again mediated by SENP enzymes to release unmodified substrate protein and mature SUMO moiety (Fig.15) (Wilkinson and Henley, 2010).

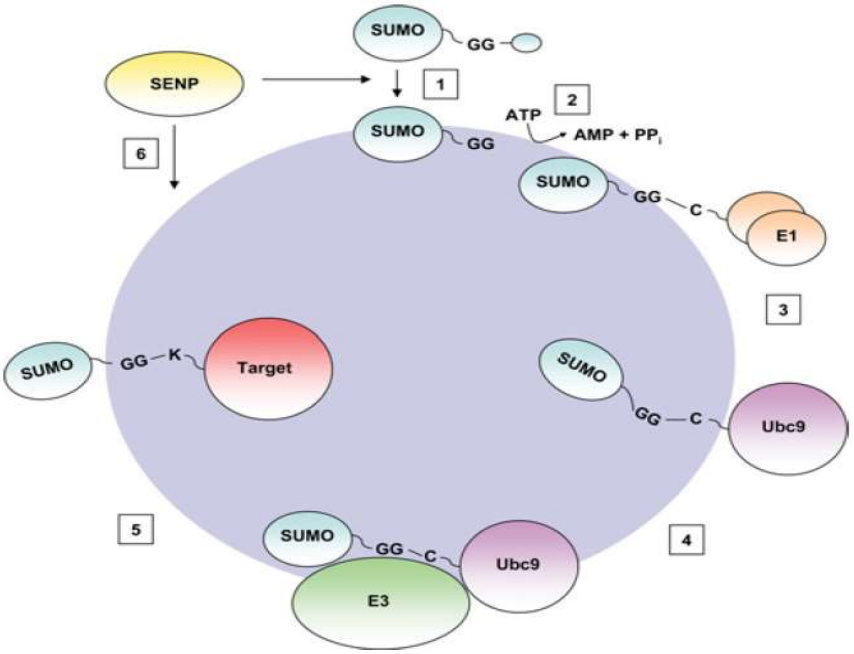

Fig 15. The SUMOylation cycle. (1) The C-terminal tail of SUMO precursor is cleaved off by members of SENP family to expose a di-glycine motif. (2) Processed SUMO is then activated by activating E1 enzyme, a heterodimer of SAE1 and SAE2, in an ATP-dependent manner resulting in the formation of a thioester bond with the active site of E1. (3) Activated SUMO is then shifted to the catalytic cysteine residue of conjugating E2 enzyme, UBC9. (4, 5) E2 enzyme catalyses the transfer of SUMO to substrate protein, often in connection with an E3 enzyme. During this step SUMO covalently binds to the substrate protein via an isopeptide bond between C-terminal glycine residue of SUMO and a lysine residue in the substrate protein. (6) The SENPs subsequently de-conjugate the SUMO modified substrate and release mature SUMO moiety for further rounds of SUMOylation (Wilkinson and Henley, 2010).

In this study, it is demonstrated for the first time that GRHL2 represents a novel substrate protein for SUMOylation. GRHL2 therefore represents the first known SUMOylated protein of the GRHL2 family of transcription factors. Using COS-7 cells as a model system, it could be shown that in addition to the 70 kDa protein representing the GRHL2 protein, a GRHL2 protein species migrating at about 110 kDa could be detected by Western blot analysis. A detailed molecular analysis demonstrated that this high molecular weight GRHL2 protein represents

SUMOylated GRHL2. Although the molecular weight of a SUMO moiety is approximately 11 kDa, the conjugation of one SUMO moiety to substrate proteins often causes band shifts of about 10-40 kDa (Wilkinson and Henley, 2010). This is thought to be due to the branched nature of modification of proteins by SUMOylation (Girach et al., 2013; Martin et al., 2007).

The detection of SUMOylated GRHL2 proteins by Western blot analysis was experimentally challenging in that SENP family proteases mediating de-SUMOylation and therefore rapid turnover of this modification within cells, but also during cell lysis need to be inhibited. This is a well-known problem in any SUMOylation study and therefore many strategies have been developed to inhibit the activities of SENP proteases to maintain SUMO conjugation (Drag and Salvesen, 2008). NEM is able to block the catalytic activity of all cysteine proteases and thus is commonly used as SENP protease inhibitor (Albrow et al., 2011). In this study, 25 mM N-ethylmaleimide (NEM) was used during cell lysis to preserve the SUMOylation state of GRHL2 (Meulmeester et al., 2008). However, to exert full inhibitory capacity, lysis buffers containing NEM should have a pH 6.5-7.5 and also should be free of reducing reagents, such as DTT, DTE, and β-mercaptoethanol. Besides, an amine-free buffer system also is highly recommended. Apart from requirements raised by NEM, the addition of ionic and non-ionic detergents to efficiently extract SUMOylated proteins from nuclear substructures is equally important (Heun, 2007; Johnson, 2004). Thus, many different cell lysis buffers were tested for their ability to preserve the SUMOylated state of GRHL2. The lysis buffer described in this study gave best results and allowed the robust detection of SUMOylated GRHL2 proteins by Western blot analysis. The specificity of the results could be demonstrated by experiments in which NEM was omitted from the lysis buffer. Under these experimental conditions SUMOylated GRHL2 proteins could not be detected by Western blot analysis anymore. The findings described in this study showed that GRHL2 contains one major SUMOylation-site, resulting the detection of an additional GRHL2 proteins with an apparent molecular weight of approximately 110 kDa.

GRHL2 is SUMOylated at position K159

To identify the acceptor lysine residue for SUMOylation within the GRHL2 protein, a computational analysis of the GRHL2 protein sequence was performed. This analysis yielded a large number of predicted acceptor sites within the GRHL2 protein. Residues predicted to be modified by SUMOylation by at least three distinct algorithms (n=7) were selected for further analysis. Lysine residues (K159, K205, K366, K424, K453, K531, and K556) were replaced by arginine and expression constructs encoding wild-type and mutant GRHL2 proteins were co-transfected with SUMO-1 or SUMO-2 into COS-7 cells. Site-directed mutagenesis of candidate lysine residues followed by gel shift assay represents a common strategy to identify SUMO conjugation sites within proteins (Tatham et al., 2009). Since mutation that changes a target lysine to arginine abolishes SUMO conjugation at that site, the absence of a SUMO-conjugation band for a certain lysine-arginine mutant is a powerful indication as to the site of SUMO modification. This experiments clearly revealed that only K159 served as an acceptor site for SUMOylation of the GRHL2 protein in that mutation of this residue completely abolished SUMOylation, as revealed by Western blot analysis of COS-7 cell lysates.

Lysine residues that are modified by SUMOylation are often embedded within a consensus motif comprises ΨKxD/E, where Ψ is a hydrophobic amino acid, K SUMOylated lysine residue, x any amino acid, and D/E (aspartic/glutamic acid). It is estimated that 75 % of known SUMO targets are modified within a consensus motif (Wilkinson and Henley, 2010). In addition, extended variants of consensus motif are also reported, including inverted consensus motif (ICM), phosphorylation-dependent SUMO motif (PDSM), negatively-charged amino acid-depend SUMO motif (NDSM), and hydrophobic cluster SUMO motif (HCSM) (Beauclair et al., 2015). SUMOylation can also occur at non-consensus sites which do not conform to any of these SUMOylation motifs (Beauclair et al., 2015). It is important to note that the K159 residue embedded in the VKAE sequence of GRHL2 (residues 158-161) perfectly matches the ΨKxD/E consensus motif for SUMOylation of proteins (Beauclair et al., 2015).

Although mutation at K159 caused disappearance of SUMO conjugation to GRHL2, this is not sufficient to actually prove an involvement of K159 in SUMOylation of the GRHL2 protein. It is now well accepted that mutation of both the hydrophobic or aspartic/glutamic amino acid, but not the ‘x’ residue, within a SUMO consensus motif

also affects SUMO conjugation of the target lysine (Rodriguez et al., 2001; Sampson et al., 2001). Mutation to these residues served as an additional strategy for SUMO conjugation site analysis. Even though to a lesser degree, it provides more rigorous and solid proof that a suspected lysine is subjected to SUMO modification. To provide further evidence for an involvement of K159 in SUMOylation of GRHL2, residue E161 was replaced by A (alanine) to disrupt the putative VKAE consensus motif in GRHL2. Replacing E161 by A completely abolished conjugation of SUMO-1 or -2 to GRHL2, thus confirming that K159 indeed represents the residue used for covalent attachment of SUMO-1/2 proteins to GRHL2.

Target residue K159 is located between the transactivation domain and DNA binding domain of GRHL2 in a structurally disordered region of the protein, consistent with the typical localisation of SUMO conjugation sites within poorly-structured stretches of polypeptides (Diella et al., 2008). Of interest, K159 may be part of a phosphorylation-dependent SUMO motif (PDSM), a variation of the SUMOylation consensus motif which comprises a SUMOylation consensus site and a proline-directed phosphorylation site, separated by two any amino acids (ΨKxD/EXXS/TP) (Hietakangas et al., 2006). PDSMs were identified in a large number of target proteins and induce phosphorylation-mediated enhancement of SUMOylation of target proteins (Hietakangas et al., 2006). The sequence VKAEDFTP (residues 158-165) in GRHL2 perfectly matches a PSDM, raising the possibility that phosphorylation T164 may influence SUMOylation of GRHL2 at K159. However, this possibility needs to be investigated in future experiments.

To provide further evidence for an involvement of K159 in SUMOylation of GRHL2, pull-down experiments using His_6-tagged SUMO-1/2 proteins and nickel affinity chromatography was employed. The protocol uses 6 M guanidine-HCl to completely denature proteins during cell lysis (Choi et al., 2013). Under these experimental conditions, non-covalent SUMO-binding via SIMs is effectively abolished, thus removing most unspecific contaminants. Denaturing conditions also result in the rapid inactivation of SENP family of proteases preventing de-conjugation of SUMO from substrate proteins (Da Silva-Ferrada et al., 2012). In this assay, GRHL2 proteins modified by covalent attachment of His_6-tagged SUMO-1/2, but not the non-modified GRHL2 proteins could be specifically enriched and detected by Western blot analysis. Again, the detection of SUMOylated GRHL2 was strictly dependent on the presence of K159 or E161 residues, as mutation of any of these

two residues completely abolished enrichment of SUMOylated GRHL2 proteins from cell lysates.

Taken together, these results demonstrate that GRHL2 represents a novel SUMOylation substrate. Evidence is provided that SUMO-1/2 proteins covalently become attached to K159 residue within a SUMO consensus motif. Whether this sequence also represents a phosphorylation-dependent SUMO consensus motif (PSDM) remains to be experimentally confirmed.

PIAS proteins modulate SUMOylation and localisation of GRHL2 proteins

SUMOylation of substrate proteins can be enhanced in the presence of SUMO E3 ligases. In contrast to ubiquitin system, where hundreds of proteins were reported to possess ubiquitin E3 activity, only a few SUMO E3 ligases were identified in the SUMOylation pathway. The SUMO E3 ligases identified thus far include, for example, PIASs, RanBP2, Pc2, Mms21, HDAC4, HDAC7, MUL1, Rhes, TLS, TRAF7, and TOPORS (Wilkinson and Henley, 2010). Among them, the PIAS (protein inhibitor of activated STAT) family of proteins are capable of stimulating SUMO conjugation of a wide variety of mammalian SUMOylation targets and therefore represent strong candidates for E3 ligases in many SUMOylation studies (Rytinki et al., 2009). The PIAS proteins are encoded by four genes, namely PIAS1, PIAS2, PIAS3, and PIAS4. Splice variants have been functionally described for PIAS2, PIAS3, and PIAS4, but not yet for PIAS1 (Rytinki et al., 2009). The four members of PIAS family share high sequence similarity and conserved domain structure. In the central area of PIASs, there is a cysteine-rich region forming a RING finger type of motif termed SP-RING, which shares structural similarities with the RING domain found in ubiquitin E3 ligases (Rytinki et al., 2009). Analogous to the function of RING-type ubiquitin E3s, PIASs do not directly bind to SUMO through a thioester linkage. Instead, they act as adaptors bringing SUMO-loaded UBC9 into attachment with the substrate targets or maintaining the SUMO-UBC9 thioester in a conformation conducive to SUMO conjugation (Wilkinson and Henley, 2010).

Interestingly, PIAS3 was identified as candidate interaction partner of GRHL2 in a Yeast-Two-Hybrid based screen using human full-length GRHL2 as a bait (unpublished observation). Furthermore, an entry in the INTact database containing experimentally validated protein-protein interactions suggested a possible

interaction between GRHL2 and PIAS2 (https://www.ebi.ac.uk/intact/), implying that GRHL2 might be able to interact with several members of the PIAS family of proteins. To investigate whether PIAS proteins are able to promote SUMO conjugation to GRHL2, expression constructs encoding individual PIAS proteins, wild-type GRHL2, and SUMO1 or SUMO2 proteins were co-transfected into COS-7 cells. Ectopic expression of all four PIAS protein markedly enhanced SUMO modification of GRHL2 by both SUMO1 and SUMO2, suggesting that all four PIASs exhibit SUMO E3 activity in the GRHL2 SUMOylation pathway.

It is well known that many substrates are able to interact with more than one member of the PIAS family of protein which, in turn, enhance their SUMOylation. For example, Estruch et al. also reported an interaction of the FOXP2 transcription factor with all four members of the PIAS family of proteins (Estruch et al., 2016). In contrast, Rytinki et al. rather describe a differential interaction of the RIP140 transcriptional regulator with PIAS proteins in that an interaction of RIP140 with PIAS2 and 4 but not PIAS 1 and 3 was observed (Rytinki and Palvimo, 2008). It seems that PIAS proteins exhibit slightly different substrate preferences which, dependent on the experimental conditions, may become obvious or remain obscured, especially if substrate proteins and PIASs were co-expressed in cells at very high levels. Further studies will be necessary to identify PIAS proteins present in breast cancer cells and to illuminate those which interact with GRHL2 under physiological conditions.

SUMOylation of substrate proteins is a molecular glue which enables modified proteins to interact with other SIM-containing proteins which in turn may result in the formation of higher-order structures (Rytinki et al., 2009). To investigate whether SUMOylation of GRHL2 influences the subnuclear distribution of GRHL2 in COS-7 cells, extensive immunofluorescence studies were performed. GRHL2 wild-type as well SUMOylation-deficient proteins (GRHL2 K159R and E161A) predominantly showed a diffuse staining pattern with little granular staining when transiently expressed in COS-7 cells, suggesting that SUMOylation does not influence the localisation of the GRHL2 transcription factor. Notably, co-expression of GRHL2 with individual members of the PIAS family of proteins uniformly caused a striking change in the distribution of GRHL2. Under these experimental conditions, GRHL2 was mostly localised in a highly variable number of nuclear speckles. Moreover, a co-localisation of GRHL2 with all PIAS proteins in nuclear foci clearly was observed,

thus confirming a physical and functional interplay between GRHL2 and PIAS proteins in cells. It is unclear, however, whether re-localisation of GRHL2 to nuclear speckles is dependent on SUMOylation or on the interaction with PIAS proteins or even both. In fact, SP-RING mediated interaction between PIAS4 and FIP200 caused translocation of FIP200 from cytoplasm to nucleus, even though FIP200 is not a *bona fide* SUMO substrate (Martin et al., 2008). Also, the proper localisation of Msx1 requires interaction between PIAS1 and Msx1, in this case, neither the SUMO ligase activity of PIAS1 nor SUMOylation of Msx1 is needed (Lee et al., 2006a). Due to rapid turnover of SUMOylation of substrate proteins and therefore the extremely low percentage of SUMOylated GRHL2 in cells, it is very difficult to establish whether SUMOylation of GRHL2 is necessary for re-localisation of the transcription factor. Using PIAS mutants may help to dissect the molecular mechanisms determining the subnuclear localisation and possibly also activity of the GRHL2 transcription factor in cancer cells.

Taken together, in this study clearly PIAS proteins were identified as novel modifiers of GRHL2 activity and localisation in cancer cells.

SUMOylation increases transcriptional activity of GRHL2

Evidence is accumulating that SUMOylation plays an important role in transcriptional regulation of gene expression by influencing protein stability, protein-protein interaction, and sub-nuclear localisation of transcription factors and/or transcriptional co-regulators (Hay, 2005; Heun, 2007; Johnson, 2004) (Lyst and Stancheva, 2007; Verger et al., 2003). In fact, a growing number of transcription factors and nuclear receptors have been shown to be modified by SUMOylation, resulting in either enhanced or suppressed gene transcription activities (Lyst and Stancheva, 2007; Verger et al., 2003).

In this study, results are presented showing that SUMOylation rather increases the activity of GRHL2 to initiate transcription of target genes. This finding is mainly based on luciferase-reporter assays in conjunction with results obtained using MDA-MB-231 breast cancer cells as a model system. The activation of GRHL2 activity through SUMOylation is rather unusual in that in most cases SUMO-modified transcription factors are inactivated (Gill, 2005; Verger et al., 2003). For example, SUMO conjugation to transcription factors ZEB2, C-JUN, C-MYB, and ELK1 were

reported to negatively influence their transcriptional activity (Bies et al., 2002; Long et al., 2005; Muller et al., 2000; Yang et al., 2003). However, consistent with our observation, an increasing body of evidence indicates that SUMOylation of transcription factors can also lead to an upregulation transcriptional activity. For example, SUMO attachment to heat shock transcription factor HSF1 promotes its DNA-binding ability, thereby presumably promoting its activity to regulate target gene expression in response to stress (Hong et al., 2001).

Given the molecular and biological consequences of SUMO modification, different mechanisms by which SUMOylation promotes GRHL2 transcriptional activity can be considered. Firstly, in this study unpublished data indicated that SUMOylation increases GRHL2 protein stability and expression. This raises the possibility that SUMOylation of K159 might block ubiquitination of the same residue and subsequent proteasomal degradation of GRHL2 proteins. A similar mechanism was found in transcriptional regulation by p53, where SUMOylation of p53 promotes its capacity to activate reporter genes by antagonizing Mdm2-mediated ubiquitination required for p53 for degradation (Gostissa et al., 1999; Rodriguez et al., 1999). Alternatively, SUMOylation may target GRHL2 proteins to specific subnuclear compartments, such as the nuclear speckles induced by PIAS proteins, and thus prevents degradation of GRHL2 by the proteasome. Both mechanisms potentially could lead to increased levels GRHL2 expression and activity.

Another possibility is that the covalent attachment of the SUMO moiety leads to general conformational changes or specific changes at crucial interfaces, thereby influencing the interaction of GRHL2 with other proteins participating in GRHL2-driven gene expression. In line with this, SUMOylation of the transcription factor Ikaros restrained its interaction with the Sin3 and NuRD co-repressor complex, thereby inhibiting Ikaros-mediated repression of transcription (Gomez-del Arco et al., 2005).

Additionally, it could also be possible that SUMO conjugation activates GRHL2 transcriptional activity by altering GRHL2 sub-nuclear localisation and creation of proximity to target genes. Consistent with this hypothesis, SUMO modification of promyelocytic leukemia protein (PML) plays a critical role in the formation of PML oncogenic nuclear bodies and in the recruitment of additional nuclear body-associated factors, such as Sp100 and Daxx. The PML body is associated

with the nuclear matrix and governs transcriptional activity, cellular cycle, and viral infection (Seeler and Dejean, 2001).

Although SUMOylation clearly enhances the activity of the GRHL2 transcription factor, the underlying molecular mechanisms still remain to be elucidated.

Regulation of Epithelial-Mesenchymal transition through SUMOylation

The epithelial-to-mesenchymal transition (EMT) and its reverse program mesenchymal-to-epithelial transition (MET) are orchestrated by a large variety of transcription factors. Interestingly, SUMO modification can have broad effects on the functional activities of many of these transcriptional regulators, mediating transition between epithelial and mesenchymal states (Fig. 16) (Bogachek et al., 2015).

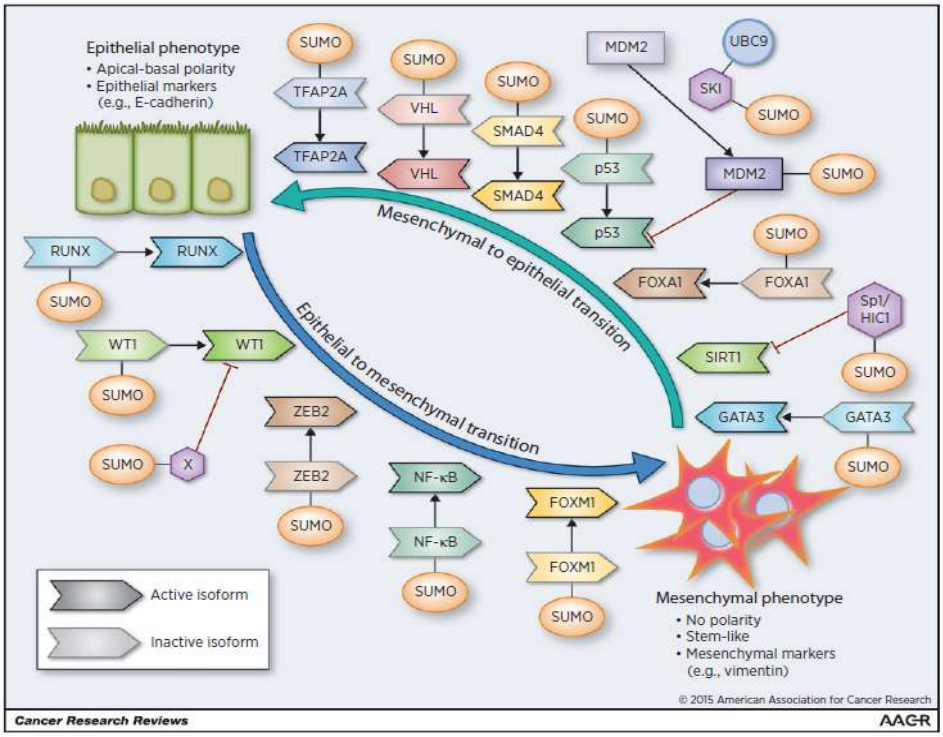

Fig 16. Schematic representation EMT/MET transcription factors regulated by SUMO modification (Bogachek et al., 2015).

For example, the zinc finger E-box transcription factor ZEB2 is one of the pioneer EMT inducers. ZEB2 was identified as direct suppressor of E-cadherin (CDH1) expression as it binds to the CDH1 promoter in the presence of the corepressor CtBP (Park et al., 2008). Notably, SUMO modification of ZEB2 decreases its transcriptional repression of E-cadherin expression by disrupting recruitment of CtBP proteins (Long et al., 2005). Another EMT-inducing transcription factor represents the forkhead box transcription factor FOXM1 which has been shown to promote the mesenchymal phenotype through its ability to suppress miR200b and to promote Slug expression (Bao et al., 2011; Yang et al., 2013). Results published by Myatt indicate that SUMOylation of FOXM1 suppressed its transcriptional activity via enhanced translocation to cytoplasm and subsequent ubiquitination-dependent proteasomal degradation (Myatt et al., 2014). In contrast, Wang demonstrated that SUMOylation of FOXM1b, the transcriptionally active isoform of FOXM1, was essential for transcriptional activation of target genes, such as JNK1, and transcriptional repression of miR200 (Wang et al., 2014). Hence, SUMO modification exhibits complicated or even contradictory effects on transcription factors promoting EMT program.

However, SUMOylation of transcription factors not only influences EMT but also the reverse process, the mesenchymal-to-epithelial transition (MET). A good example represents the GATA3 transcription factor, the expression of which is related with increased E-cadherin levels, ER expression and induction of epithelial phenotype in breast cancer cells (Yan et al., 2010). Here, SUMO modification leads to suppression of GATA3 transcriptional activity and GATA3-induced epithelial phenotype (Chun et al., 2003).

In this study, the MET-inducing transcription factor GRHL2 was identified as a novel target for SUMOylation. GRHL2, however, stands out among the EMT/MET transcription factors in that SUMOylation of GRHL2 rather seems to promote its transcriptional activity. Although the precise molecular mechanisms of the SUMOylation-dependent regulation of GRHL2 transcriptional activity still are largely unclear, a critical role of this novel regulatory mechanism in EMT/MET induction is obvious. A detailed functional analysis of SUMOylation-dependent regulation of GRHL2 transcriptional activity is of uppermost interest for a better understanding of the GRHL2-mediated stabilization of the hybrid E/M phenotype and the metastatic spread of breast cancer cells.

Conclusion

In this study, it was demonstrated for the first time that the GRHL2 transcription factor is posttranslationally modified and positively regulated by SUMOylation. Also, many details of this novel regulatory mechanism could already be discovered. For example, it could be shown that there is a functional interaction between GRHL2 and PIAS proteins leading to increased SUMOylation at lysine 159 which in turn enhances GRHL2 transcriptional activity in cancer cells. Still, several aspects of this novel regulatory pathway still deserve an in-depth molecular analysis. GRHL2 has been shown to interact with SUMO-1 in protein-protein interaction studies, suggesting that GRHL2 in addition to the SUMOylation site at K159 also contains SUMO-interaction motifs (SIMs) known to mediate a non-covalent interactions of a protein with SUMOs (Wilkinson and Henley, 2010). The identification of SIM-sites in the GRHL2 protein, albeit experimentally challenging, would be very interesting as this would add another layer of regulation for GRHL2 activity. Furthermore, the role of PIAS proteins in regulating GRHL2 activity and subnuclear localisation needs to be analysed in more detail. Although this study has shown that PIAS proteins drag GRHL2 into nuclear speckles, it is unknown, however, whether this process is mediated by PIAS-induced SUMOylation of GRHL2 or by the direct physical interaction of GRHL2 with PIAS proteins serving as scaffold proteins for the formation of nuclear speckles. To demonstrate a pathophysiological relevance of the findings described in this study, it certainly would be of great interest to determine the distribution of GRHL2 proteins in cultured breast cancer cells or in primary tumours and to investigate whether differences could be related to clinical outcome. Still, the results included in this study already paved the way for a better understanding of the molecular mechanisms regulating GRHL2 activity and the hybrid E/M phenotype which currently is regarded as the driving force in breast cancer metastasis.

6. Summary/Zusammenfassung

In recent years, SUMOylation of proteins has emerged as an important regulatory mechanism of transcription factors in cancer. Aim of this project therefore was to investigate a possible SUMOylation-dependent regulation of the EMT suppressor GRHL2 in breast cancer.

Major results of this study can be summarized as follows:

1. GRHL2 for the first time could be identified as a novel substrate for SUMO modification.
2. By combining mutational analyses, biochemical assays, and computational approaches lysine K159 could be identified as the major acceptor residue for SUMOylation in the GRHL2 protein.
3. It could be demonstrated that members of PIAS family of proteins (PIAS1-4) enhance SUMOylation of the GRHL2 transcription factor.
4. Indirect immunofluorescence studies revealed that GRHL2 protein exhibits distinct staining patterns (granular, mixed, diffuse) when transiently overexpressed in COS-7 cells.
5. The interaction with PIAS proteins but not SUMOylation alone causes GRHL2 to accumulate in granular structures within the nucleus.
6. Using luciferase-reporter assays it could be demonstrated that SUMOylation enhance the transcriptional activity of GRHL2 in COS-7 cells.
7. A model system for the analysis of SUMOylation-dependent regulation of GRHL2 target genes was generated using MDA-MB-231 breast cancer cells. Analysis of the expression of selected GRHL2 target genes (CD24, E-Cadherin) showed that SUMOylation stimulates GRHL2-driven gene expression also in breast cancer cells.

The results presented in this study for the first time demonstrate that the GRHL2 transcription factor is posttranslationally regulated by SUMOylation. Hopefully, the discovery of this novel regulatory mechanism contributes to a better understanding of the role of the EMT suppressor GRHL2 in breast cancer invasion and metastasis.

6. Zusammenfassung

In den letzten Jahren hat sich die SUMOylierung von Proteinen als wichtiger Regulationsmechanismus für Transkriptionsfaktoren in Tumorzellen etabliert. Ziel dieses Projektes war es daher, eine mögliche SUMOylierungs-abhängige Regulation des EMT-Suppressors GRHL2 in Mammakarzinomzellen zu untersuchen.

Die wichtigsten Ergebnisse dieser Studie lassen sich wie folgt zusammenfassen:

1. GRHL2 konnte erstmals als neuartiges Substrat für die SUMO-Modifikation identifiziert werden.
2. Durch Kombination von Mutationsanalysen, biochemischen Assays und bioinformatischen Ansätzen konnte Lysin K159 als die hauptsächlich durch SUMOylierung modifizierte Aminosäure im GRHL2-Protein identifiziert werden.
3. Es konnte gezeigt werden, dass Mitglieder der PIAS-Proteinfamilie (PIAS1-4) die SUMOylierung des GRHL2-Transkriptionsfaktors verstärken.
4. Indirekte Immunfluoreszenzstudien zeigten, dass das GRHL2-Protein bei transienter Überexpression in COS-7-Zellen unterschiedliche Färbungsmuster (granulär, gemischt, diffus) ergibt.
5. Die Wechselwirkung mit PIAS-Proteinen, nicht jedoch SUMOylation alleine, bewirkt, dass sich GRHL2 deutlich vermehrt in granulären Strukturen im Kern ansammelt.
6. Mit Hilfe von Luciferase-Reporter-Assays konnte gezeigt werden, dass die SUMOylierung die transkriptionelle Aktivität von GRHL2 in COS-7-Zellen erhöht.
7. Ein Modellsystem zur Analyse der SUMOylierungs-abhängigen Regulation von GRHL2-Zielgenen wurde unter Verwendung von MDA-MB-231 Mammakarzinomzellen erzeugt. Die Analyse der Expression ausgewählter GRHL2-Zielgene (CD24, E-Cadherin) ergab, dass eine GRHL2-vermittelte Genexpression auch in Mammakarzinomzellen durch SUMOylierung stimuliert wird.

   Die in dieser Studie vorgestellten Ergebnisse zeigen erstmals, dass der GRHL2-Transkriptionsfaktor durch SUMOylierung posttranslational reguliert wird. Es ist

zu hoffen, dass die Entschlüsselung dieses neuartigen Regulations-mechanismus‘ zu einem besseren Verständnis der Rolle des EMT-Suppressors GRHL2 bei der Invasion und Metastasierung des Mammakarzinoms beiträgt.

7. References

Aceto, N., Bardia, A., Miyamoto, D.T., Donaldson, M.C., Wittner, B.S., Spencer, J.A., Yu, M., Pely, A., Engstrom, A., Zhu, H., Brannigan, B.W., Kapur, R., Stott, S.L., Shioda, T., Ramaswamy, S., Ting, D.T., Lin, C.P., Toner, M., Haber, D.A., Maheswaran, S. (2014). Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. Cell 158, 1110-1122.

Al Habyan, S., Kalos, C., Szymborski, J., McCaffrey, L. (2018). Multicellular detachment generates metastatic spheroids during intra-abdominal dissemination in epithelial ovarian cancer. Oncogene 37, 5127-5135.

Albrow, V.E., Ponder, E.L., Fasci, D., Bekes, M., Deu, E., Salvesen, G.S., Bogyo, M. (2011). Development of small molecule inhibitors and probes of human SUMO deconjugating proteases. Chem Biol 18, 722-732.

Alix-Panabieres, C., Pantel, K. (2016). Clinical Applications of Circulating Tumor Cells and Circulating Tumor DNA as Liquid Biopsy. Cancer Discov 6, 479-491.

Auden, A., Caddy, J., Wilanowski, T., Ting, S.B., Cunningham, J.M., Jane, S.M. (2006). Spatial and temporal expression of the Grainyhead-like transcription factor family during murine development. Gene Expr Patterns 6, 964-970.

Aue, A., Hinze, C., Walentin, K., Ruffert, J., Yurtdas, Y., Werth, M., Chen, W., Rabien, A., Kilic, E., Schulzke, J.D., Schumann, M., Schmidt-Ott, K.M. (2015). A Grainyhead-Like 2/Ovo-Like 2 Pathway Regulates Renal Epithelial Barrier Function and Lumen Expansion. J Am Soc Nephrol 26, 2704-2715.

Bao, B., Wang, Z., Ali, S., Kong, D., Banerjee, S., Ahmad, A., Li, Y., Azmi, A.S., Miele, L., Sarkar, F.H. (2011). Over-expression of FoxM1 leads to epithelial-mesenchymal transition and cancer stem cell phenotype in pancreatic cancer cells. J Cell Biochem 112, 2296-2306.

Bardelli, A., Pantel, K. (2017). Liquid Biopsies, What We Do Not Know (Yet). Cancer Cell 31, 172-179.

Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J., Garcia De Herreros, A. (2000). The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat Cell Biol 2, 84-89.

Bax, N.A., Pijnappels, D.A., van Oorschot, A.A., Winter, E.M., de Vries, A.A., van Tuyn, J., Braun, J., Maas, S., Schalij, M.J., Atsma, D.E., Goumans, M.J., Gittenberger-de Groot, A.C. (2011). Epithelial-to-mesenchymal transformation

alters electrical conductivity of human epicardial cells. J Cell Mol Med 15, 2675-2683.

Beauclair, G., Bridier-Nahmias, A., Zagury, J.F., Saib, A., Zamborlini, A. (2015). JASSA: a comprehensive tool for prediction of SUMOylation sites and SIMs. Bioinformatics 31, 3483-3491.

Bielenberg, D.R., Zetter, B.R. (2015). The Contribution of Angiogenesis to the Process of Metastasis. Cancer J 21, 267-273.

Bies, J., Markus, J., Wolff, L. (2002). Covalent attachment of the SUMO-1 protein to the negative regulatory domain of the c-Myb transcription factor modifies its stability and transactivation capacity. J Biol Chem 277, 8999-9009.

Bischof, P., Aplin, J.D., Bentin-Ley, U., Brannstrom, M., Casslen, B., Castrillo, J.L., Classen-Linke, I., Critchley, H.O., Devoto, L., D'Hooghe, T., Horcajadas, J.A., Groothuis, P., Ivell, R., Pongrantz, I., Macklon, N.S., Sharkey, A., Vicovac, L., White, J.O., Winterhager, E., von Wolff, M., Simon, C., Stavreus-Evers, A. (2006). Implantation of the human embryo: research lines and models. From the implantation research network 'Fruitful'. Gynecol Obstet Invest 62, 206-216.

Bogachek, M.V., De Andrade, J.P., Weigel, R.J. (2015). Regulation of epithelial-mesenchymal transition through SUMOylation of transcription factors. Cancer Res 75, 11-15.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin.

Bray, S.J., Burke, B., Brown, N.H., Hirsh, J. (1989). Embryonic expression pattern of a family of Drosophila proteins that interact with a central nervous system regulatory element. Genes Dev 3, 1130-1145.

Bray, S.J., Kafatos, F.C. (1991). Developmental function of Elf-1: an essential transcription factor during embryogenesis in Drosophila. Genes Dev 5, 1672-1683.

Butz, H., Szabo, P.M., Nofech-Mozes, R., Rotondo, F., Kovacs, K., Mirham, L., Girgis, H., Boles, D., Patocs, A., Yousef, G.M. (2014). Integrative bioinformatics analysis reveals new prognostic biomarkers of clear cell renal cell carcinoma. Clin Chem 60, 1314-1326.

Chambers, A.F., Groom, A.C., MacDonald, I.C. (2002). Dissemination and growth of cancer cells in metastatic sites. Nat Rev Cancer 2, 563-572.

Chen, W., Dong, Q., Shin, K.H., Kim, R.H., Oh, J.E., Park, N.H., Kang, M.K. (2010). Grainyhead-like 2 enhances the human telomerase reverse transcriptase gene expression by inhibiting DNA methylation at the 5'-CpG island in normal human keratinocytes. J Biol Chem 285, 40852-40863.

Chen, W., Yi, J.K., Shimane, T., Mehrazarin, S., Lin, Y.L., Shin, K.H., Kim, R.H., Park, N.H., Kang, M.K. (2016). Grainyhead-like 2 regulates epithelial plasticity and stemness in oral cancer cells. Carcinogenesis 37, 500-510.

Chen, Y.Z., Chen, Z., Gong, Y.A., Ying, G. (2012). SUMOhydro: a novel method for the prediction of sumoylation sites based on hydrophobic properties. PLoS One 7, e39195.

Cheung, K.J., Padmanaban, V., Silvestri, V., Schipper, K., Cohen, J.D., Fairchild, A.N., Gorin, M.A., Verdone, J.E., Pienta, K.J., Bader, J.S., Ewald, A.J. (2016). Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. Proc Natl Acad Sci U S A 113, E854-863.

Choi, S.Y., Jang, H., Roe, J.S., Kim, S.T., Cho, E.J., Youn, H.D. (2013). Phosphorylation and ubiquitination-dependent degradation of CABIN1 releases p53 for transactivation upon genotoxic stress. Nucleic Acids Res 41, 2180-2190.

Chun, T.H., Itoh, H., Subramanian, L., Iniguez-Lluhi, J.A., Nakao, K. (2003). Modification of GATA-2 transcriptional activity in endothelial cells by the SUMO E3 ligase PIASy. Circ Res 92, 1201-1208.

Chung, V.Y., Tan, T.Z., Tan, M., Wong, M.K., Kuay, K.T., Yang, Z., Ye, J., Muller, J., Koh, C.M., Guccione, E., Thiery, J.P., Huang, R.Y. (2016). GRHL2-miR-200-ZEB1 maintains the epithelial status of ovarian cancer through transcriptional regulation and histone modification. Sci Rep 6, 19943.

Cieply, B., Farris, J., Denvir, J., Ford, H.L., Frisch, S.M. (2013). Epithelial-mesenchymal transition and tumor suppression are controlled by a reciprocal feedback loop between ZEB1 and Grainyhead-like-2. Cancer Res 73, 6299-6309.

Cieply, B., Riley, P.t., Pifer, P.M., Widmeyer, J., Addison, J.B., Ivanov, A.V., Denvir, J., Frisch, S.M. (2012). Suppression of the epithelial-mesenchymal transition by Grainyhead-like-2. Cancer Res 72, 2440-2453.

Da Silva-Ferrada, E., Lopitz-Otsoa, F., Lang, V., Rodriguez, M.S., Matthiesen, R. (2012). Strategies to Identify Recognition Signals and Targets of SUMOylation. Biochem Res Int 2012, 875148.

de Castro, E., Sigrist, C.J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch, A., Hulo, N. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34, W362-365.

De Craene, B., Berx, G. (2013). Regulatory networks defining EMT during cancer initiation and progression. Nat Rev Cancer 13, 97-110.

Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N.P., Trave, G., Gibson, T.J. (2008). Understanding eukaryotic linear motifs and their role in cell signaling and regulation. Front Biosci 13, 6580-6603.

Diepenbruck, M., Christofori, G. (2016). Epithelial-mesenchymal transition (EMT) and metastasis: yes, no, maybe? Curr Opin Cell Biol 43, 7-13.

Dompe, N., Rivers, C.S., Li, L., Cordes, S., Schwickart, M., Punnoose, E.A., Amler, L., Seshagiri, S., Tang, J., Modrusan, Z., Davis, D.P. (2011). A whole-genome RNAi screen identifies an 8q22 gene cluster that inhibits death receptor-mediated apoptosis. Proc Natl Acad Sci U S A 108, E943-951.

Drag, M., Salvesen, G.S. (2008). DeSUMOylating enzymes--SENPs. IUBMB Life 60, 734-742.

Dynlacht, B.D., Attardi, L.D., Admon, A., Freeman, M., Tjian, R. (1989). Functional analysis of NTF-1, a developmentally regulated Drosophila transcription factor that binds neuronal cis elements. Genes Dev 3, 1677-1688.

Estruch, S.B., Graham, S.A., Deriziotis, P., Fisher, S.E. (2016). The language-related transcription factor FOXP2 is post-translationally modified with small ubiquitin-like modifiers. Sci Rep 6, 20911.

Faltas, B. (2012). Cornering metastases: therapeutic targeting of circulating tumor cells and stem cells. Front Oncol 2, 68.

Felipe Lima, J., Nofech-Mozes, S., Bayani, J., Bartlett, J.M. (2016). EMT in Breast Carcinoma-A Review. J Clin Med 5.

Frisch, S.M., Farris, J.C., Pifer, P.M. (2017). Roles of Grainyhead-like transcription factors in cancer. Oncogene 36, 6067-6073.

Fustaino, V., Presutti, D., Colombo, T., Cardinali, B., Papoff, G., Brandi, R., Bertolazzi, P., Felici, G., Ruberti, G. (2017). Characterization of epithelial-mesenchymal transition intermediate/hybrid phenotypes associated to resistance to

EGFR inhibitors in non-small cell lung cancer cell lines. Oncotarget 8, 103340-103363.

Gao, X., Vockley, C.M., Pauli, F., Newberry, K.M., Xue, Y., Randell, S.H., Reddy, T.E., Hogan, B.L. (2013). Evidence for multiple roles for grainyhead-like 2 in the establishment and maintenance of human mucociliary airway epithelium.[corrected]. Proc Natl Acad Sci U S A 110, 9356-9361.

Gill, G. (2005). Something about SUMO inhibits transcription. Curr Opin Genet Dev 15, 536-541.

Girach, F., Craig, T.J., Rocca, D.L., Henley, J.M. (2013). RIM1alpha SUMOylation is required for fast synaptic vesicle exocytosis. Cell Rep 5, 1294-1301.

Giuliano, M., Shaikh, A., Lo, H.C., Arpino, G., De Placido, S., Zhang, X.H., Cristofanilli, M., Schiff, R., Trivedi, M.V. (2018). Perspective on Circulating Tumor Cell Clusters: Why It Takes a Village to Metastasize. Cancer Res 78, 845-852.

Gnad, F., Ren, S., Cox, J., Olsen, J.V., Macek, B., Oroshi, M., Mann, M. (2007). PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. Genome Biol 8, R250.

Goldman, A., Majumder, B., Dhawan, A., Ravi, S., Goldman, D., Kohandel, M., Majumder, P.K., Sengupta, S. (2015). Temporally sequenced anticancer drugs overcome adaptive resistance by targeting a vulnerable chemotherapy-induced phenotypic transition. Nat Commun 6, 6139.

Gomez-del Arco, P., Koipally, J., Georgopoulos, K. (2005). Ikaros SUMOylation: switching out of repression. Mol Cell Biol 25, 2688-2697.

Gostissa, M., Hengstermann, A., Fogal, V., Sandy, P., Schwarz, S.E., Scheffner, M., Del Sal, G. (1999). Activation of p53 by conjugation to the ubiquitin-like protein SUMO-1. EMBO J 18, 6462-6471.

Green, J., Dmochowski, G., Golshani, A., 2006. Prediction Of Protein Sumoylation Sites Via Parallel Cascade Identification.

Gregory, P.A., Bracken, C.P., Smith, E., Bert, A.G., Wright, J.A., Roslan, S., Morris, M., Wyatt, L., Farshid, G., Lim, Y.Y., Lindeman, G.J., Shannon, M.F., Drew, P.A., Khew-Goodall, Y., Goodall, G.J. (2011). An autocrine TGF-beta/ZEB/miR-200 signaling network regulates establishment and maintenance of epithelial-mesenchymal transition. Mol Biol Cell 22, 1686-1698.

Hanahan, D., Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell 144, 646-674.

Hay, E.D. (1995). An overview of epithelio-mesenchymal transformation. Acta Anat (Basel) 154, 8-20.

Hay, R.T. (2005). SUMO: a history of modification. Mol Cell 18, 1-12.

Heun, P. (2007). SUMOrganization of the nucleus. Curr Opin Cell Biol 19, 350-355.

Hietakangas, V., Anckar, J., Blomster, H.A., Fujimoto, M., Palvimo, J.J., Nakai, A., Sistonen, L. (2006). PDSM, a motif for phosphorylation-dependent SUMO modification. Proc Natl Acad Sci U S A 103, 45-50.

Hong, Y., Rogers, R., Matunis, M.J., Mayhew, C.N., Goodson, M.L., Park-Sarge, O.K., Sarge, K.D. (2001). Regulation of heat shock transcription factor 1 by stress-induced SUMO-1 modification. J Biol Chem 276, 40263-40267.

Huang, R.Y., Guilford, P., Thiery, J.P. (2012). Early events in cell adhesion and polarity during epithelial-mesenchymal transition. J Cell Sci 125, 4417-4422.

Jia, J., Zhang, L., Liu, Z., Xiao, X., Chou, K.C. (2016). pSumo-CD: predicting sumoylation sites in proteins with covariance discriminant algorithm by incorporating sequence-coupled effects into general PseAAC. Bioinformatics 32, 3133-3141.

Johnson, E.S. (2004). Protein modification by SUMO. Annu Rev Biochem 73, 355-382.

Jolly, M.K., Boareto, M., Debeb, B.G., Aceto, N., Farach-Carson, M.C., Woodward, W.A., Levine, H. (2017). Inflammatory breast cancer: a model for investigating cluster-based dissemination. NPJ Breast Cancer 3, 21.

Jolly, M.K., Boareto, M., Huang, B., Jia, D., Lu, M., Ben-Jacob, E., Onuchic, J.N., Levine, H. (2015). Implications of the Hybrid Epithelial/Mesenchymal Phenotype in Metastasis. Front Oncol 5, 155.

Jolly, M.K., Mani, S.A., Levine, H. (2018). Hybrid epithelial/mesenchymal phenotype(s): The 'fittest' for metastasis? Biochim Biophys Acta Rev Cancer 1870, 151-157.

Kalluri, R., Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. J Clin Invest 119, 1420-1428.

Kang, X., Chen, W., Kim, R.H., Kang, M.K., Park, N.H. (2009). Regulation of the hTERT promoter activity by MSH2, the hnRNPs K and D, and GRHL2 in human oral squamous cell carcinoma cells. Oncogene 28, 565-574.

Kokoszynska, K., Ostrowski, J., Rychlewski, L., Wyrwicz, L.S. (2008). The fold recognition of CP2 transcription factors gives new insights into the function and evolution of tumor suppressor protein p53. Cell Cycle 7, 2907-2915.

Kyhse-Andersen, J. (1984). Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. J Biochem Biophys Methods 10, 203-209.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Lamouille, S., Xu, J., Derynck, R. (2014). Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol 15, 178-196.

Lee, H., Quinn, J.C., Prasanth, K.V., Swiss, V.A., Economides, K.D., Camacho, M.M., Spector, D.L., Abate-Shen, C. (2006a). PIAS1 confers DNA-binding specificity on the Msx1 homeoprotein. Genes Dev 20, 784-794.

Lee, J.M., Dedhar, S., Kalluri, R., Thompson, E.W. (2006b). The epithelial-mesenchymal transition: new insights in signaling, development, and disease. J Cell Biol 172, 973-981.

Lee, J.Y., Kong, G. (2016). Roles and epigenetic regulation of epithelial-mesenchymal transition and its transcription factors in cancer initiation and progression. Cell Mol Life Sci 73, 4643-4660.

Lim, J., Thiery, J.P. (2012). Epithelial-mesenchymal transitions: insights from development. Development 139, 3471-3486.

Long, J., Zuo, D., Park, M. (2005). Pc2-mediated sumoylation of Smad-interacting protein 1 attenuates transcriptional repression of E-cadherin. J Biol Chem 280, 35477-35489.

Lyst, M.J., Stancheva, I. (2007). A role for SUMO modification in transcriptional repression and activation. Biochem Soc Trans 35, 1389-1392.

Martin, N., Schwamborn, K., Urlaub, H., Gan, B., Guan, J.L., Dejean, A. (2008). Spatial interplay between PIASy and FIP200 in the regulation of signal transduction and transcriptional activity. Mol Cell Biol 28, 2771-2781.

Martin, S., Nishimune, A., Mellor, J.R., Henley, J.M. (2007). SUMOylation regulates kainate-receptor-mediated synaptic transmission. Nature 447, 321-325.

McNiven, M.A. (2013). Breaking away: matrix remodeling from the leading edge. Trends Cell Biol 23, 16-21.

Mehrazarin, S., Chen, W., Oh, J.E., Liu, Z.X., Kang, K.L., Yi, J.K., Kim, R.H., Shin, K.H., Park, N.H., Kang, M.K. (2015). The p63 Gene Is Regulated by Grainyhead-like 2 (GRHL2) through Reciprocal Feedback and Determines the Epithelial Phenotype in Human Keratinocytes. J Biol Chem 290, 19999-20008.

Meulmeester, E., Kunze, M., Hsiao, H.H., Urlaub, H., Melchior, F. (2008). Mechanism and consequences for paralog-specific sumoylation of ubiquitin-specific protease 25. Mol Cell 30, 610-619.

Miles, F.L., Pruitt, F.L., van Golen, K.L., Cooper, C.R. (2008). Stepping out of the flow: capillary extravasation in cancer metastasis. Clin Exp Metastasis 25, 305-324.

Mlacki, M., Kikulska, A., Krzywinska, E., Pawlak, M., Wilanowski, T. (2015). Recent discoveries concerning the involvement of transcription factors from the Grainyhead-like family in cancer. Exp Biol Med (Maywood) 240, 1396-1401.

Mohme, M., Riethdorf, S., Pantel, K. (2017). Circulating and disseminated tumour cells - mechanisms of immune surveillance and escape. Nat Rev Clin Oncol 14, 155-167.

Mooney, S.M., Talebian, V., Jolly, M.K., Jia, D., Gromala, M., Levine, H., McConkey, B.J. (2017). The GRHL2/ZEB Feedback Loop-A Key Axis in the Regulation of EMT in Breast Cancer. J Cell Biochem 118, 2559-2570.

Moreno-Bueno, G., Portillo, F., Cano, A. (2008). Transcriptional regulation of cell polarity in EMT and cancer. Oncogene 27, 6958-6969.

Muller, S., Berger, M., Lehembre, F., Seeler, J.S., Haupt, Y., Dejean, A. (2000). c-Jun and p53 activity is modulated by SUMO-1 modification. J Biol Chem 275, 13321-13329.

Myatt, S.S., Kongsema, M., Man, C.W., Kelly, D.J., Gomes, A.R., Khongkow, P., Karunarathna, U., Zona, S., Langer, J.K., Dunsby, C.W., Coombes, R.C., French, P.M., Brosens, J.J., Lam, E.W. (2014). SUMOylation inhibits FOXM1 activity and delays mitotic transition. Oncogene 33, 4316-4329.

Navarro, C., Nola, S., Audebert, S., Santoni, M.J., Arsanto, J.P., Ginestier, C., Marchetto, S., Jacquemier, J., Isnardon, D., Le Bivic, A., Birnbaum, D., Borg, J.P. (2005). Junctional recruitment of mammalian Scribble relies on E-cadherin engagement. Oncogene 24, 4330-4339.

Nieto, M.A., Huang, R.Y., Jackson, R.A., Thiery, J.P. (2016). EMT: 2016. Cell 166, 21-45.

Nusslein-Volhard, C., Wieschaus, E., Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle inDrosophila melanogaster : I. Zygotic loci on the second chromosome. Wilehm Roux Arch Dev Biol 193, 267-282.

O'Shaughnessy, J. (2005). Extending survival with chemotherapy in metastatic breast cancer. Oncologist 10 Suppl 3, 20-29.

Pantel, K., Hayes, D.F. (2018). Disseminated breast tumour cells: biological and clinical meaning. Nat Rev Clin Oncol 15, 129-131.

Park, S.M., Gaur, A.B., Lengyel, E., Peter, M.E. (2008). The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. Genes Dev 22, 894-907.

Pastushenko, I., Brisebarre, A., Sifrim, A., Fioramonti, M., Revenco, T., Boumahdi, S., Van Keymeulen, A., Brown, D., Moers, V., Lemaire, S., De Clercq, S., Minguijon, E., Balsat, C., Sokolow, Y., Dubois, C., De Cock, F., Scozzaro, S., Sopena, F., Lanas, A., D'Haene, N., Salmon, I., Marine, J.C., Voet, T., Sotiropoulou, P.A., Blanpain, C. (2018). Identification of the tumour transition states occurring during EMT. Nature 556, 463-468.

Pecina-Slaus, N. (2003). Tumor suppressor gene E-cadherin and its role in normal and malignant cells. Cancer Cell Int 3, 17.

Peinado, H., Olmeda, D., Cano, A. (2007). Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat Rev Cancer 7, 415-428.

Peters, L.M., Anderson, D.W., Griffith, A.J., Grundfast, K.M., San Agustin, T.B., Madeo, A.C., Friedman, T.B., Morell, R.J. (2002). Mutation of a transcription factor, TFCP2L3, causes progressive autosomal dominant hearing loss, DFNA28. Hum Mol Genet 11, 2877-2885.

Puntervoll, P., Linding, R., Gemund, C., Chabanis-Davidson, S., Mattingsdal, M., Cameron, S., Martin, D.M., Ausiello, G., Brannetti, B., Costantini, A., Ferre, F., Maselli, V., Via, A., Cesareni, G., Diella, F., Superti-Furga, G., Wyrwicz, L., Ramu,

C., McGuigan, C., Gudavalli, R., Letunic, I., Bork, P., Rychlewski, L., Kuster, B., Helmer-Citterich, M., Hunter, W.N., Aasland, R., Gibson, T.J. (2003). ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. Nucleic Acids Res 31, 3625-3630.

Pyrgaki, C., Liu, A., Niswander, L. (2011). Grainyhead-like 2 regulates neural tube closure and adhesion molecule expression during neural fold fusion. Dev Biol 353, 38-49.

Qin, Y., Capaldo, C., Gumbiner, B.M., Macara, I.G. (2005). The mammalian Scribble polarity protein regulates epithelial cell adhesion and migration through E-cadherin. J Cell Biol 171, 1061-1071.

Quan, Y., Jin, R., Huang, A., Zhao, H., Feng, B., Zang, L., Zheng, M. (2014). Downregulation of GRHL2 inhibits the proliferation of colorectal cancer cells by targeting ZEB1. Cancer Biol Ther 15, 878-887.

Quan, Y., Xu, M., Cui, P., Ye, M., Zhuang, B., Min, Z. (2015). Grainyhead-like 2 Promotes Tumor Growth and is Associated with Poor Prognosis in Colorectal Cancer. J Cancer 6, 342-350.

Ren, J., Gao, X., Jin, C., Zhu, M., Wang, X., Shaw, A., Wen, L., Yao, X., Xue, Y. (2009). Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0. Proteomics 9, 3409-3412.

Riethdorf, S., Frey, S., Santjer, S., Stoupiec, M., Otto, B., Riethdorf, L., Koop, C., Wilczak, W., Simon, R., Sauter, G., Pantel, K., Assmann, V. (2016). Diverse expression patterns of the EMT suppressor grainyhead-like 2 (GRHL2) in normal and tumour tissues. Int J Cancer 138, 949-963.

Rodriguez, M.S., Dargemont, C., Hay, R.T. (2001). SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. J Biol Chem 276, 12654-12659.

Rodriguez, M.S., Desterro, J.M., Lain, S., Midgley, C.A., Lane, D.P., Hay, R.T. (1999). SUMO-1 modification activates the transcriptional response of p53. Embo j 18, 6455-6461.

Rytinki, M.M., Kaikkonen, S., Pehkonen, P., Jaaskelainen, T., Palvimo, J.J. (2009). PIAS proteins: pleiotropic interactors associated with SUMO. Cell Mol Life Sci 66, 3029-3041.

Rytinki, M.M., Palvimo, J.J. (2008). SUMOylation modulates the transcription repressor function of RIP140. J Biol Chem 283, 11586-11595.

Sampson, D.A., Wang, M., Matunis, M.J. (2001). The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. J Biol Chem 276, 21664-21669.

Santa-Maria, C.A., Gradishar, W.J. (2015). Changing Treatment Paradigms in Metastatic Breast Cancer: Lessons Learned. JAMA Oncol 1, 528-534; quiz 549.

Sato, R., Semba, T., Saya, H., Arima, Y. (2016). Concise Review: Stem Cells and Epithelial-Mesenchymal Transition in Cancer: Biological Implications and Therapeutic Targets. Stem Cells 34, 1997-2007.

Seeler, J.S., Dejean, A. (2001). SUMO: of branched proteins and nuclear bodies. Oncogene 20, 7243-7249.

Senga, K., Mostov, K.E., Mitaka, T., Miyajima, A., Tanimizu, N. (2012). Grainyhead-like 2 regulates epithelial morphogenesis by establishing functional tight junctions through the organization of a molecular network among claudin3, claudin4, and Rab25. Mol Biol Cell 23, 2845-2855.

Senkus, E., Lacko, A. (2017). Over-treatment in metastatic breast cancer. Breast 31, 309-317.

Serrano-Gomez, S.J., Maziveyi, M., Alahari, S.K. (2016). Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications. Mol Cancer 15, 18.

Smid, M., Wang, Y., Zhang, Y., Sieuwerts, A.M., Yu, J., Klijn, J.G., Foekens, J.A., Martens, J.W. (2008). Subtypes of breast cancer show preferential site of relapse. Cancer Res 68, 3108-3114.

Somarelli, J.A., Shetler, S., Jolly, M.K., Wang, X., Bartholf Dewitt, S., Hish, A.J., Gilja, S., Eward, W.C., Ware, K.E., Levine, H., Armstrong, A.J., Garcia-Blanco, M.A. (2016). Mesenchymal-Epithelial Transition in Sarcomas Is Controlled by the Combinatorial Expression of MicroRNA 200s and GRHL2. Mol Cell Biol 36, 2503-2513.

St Johnston, D., Ahringer, J. (2010). Cell polarity in eggs and epithelia: parallels and diversity. Cell 141, 757-774.

Stone, R.C., Pastar, I., Ojeh, N., Chen, V., Liu, S., Garzon, K.I., Tomic-Canic, M. (2016). Epithelial-mesenchymal transition in tissue repair and fibrosis. Cell Tissue Res 365, 495-506.

Tan, T.Z., Miow, Q.H., Miki, Y., Noda, T., Mori, S., Huang, R.Y., Thiery, J.P. (2014). Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. EMBO Mol Med 6, 1279-1293.

Tanaka, Y., Kanai, F., Tada, M., Tateishi, R., Sanada, M., Nannya, Y., Ohta, M., Asaoka, Y., Seto, M., Shiina, S., Yoshida, H., Kawabe, T., Yokosuka, O., Ogawa, S., Omata, M. (2008). Gain of GRHL2 is associated with early recurrence of hepatocellular carcinoma. J Hepatol 49, 746-757.

Tatham, M.H., Rodriguez, M.S., Xirodimas, D.P., Hay, R.T. (2009). Detection of protein SUMOylation in vivo. Nat Protoc 4, 1363-1371.

Thiery, J.P., Acloque, H., Huang, R.Y., Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell 139, 871-890.

Thiery, J.P., Sleeman, J.P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. Nat Rev Mol Cell Biol 7, 131-142.

Ting, S.B., Wilanowski, T., Cerruti, L., Zhao, L.L., Cunningham, J.M., Jane, S.M. (2003). The identification and characterization of human Sister-of-Mammalian Grainyhead (SOM) expands the grainyhead-like family of developmental transcription factors. Biochem J 370, 953-962.

Torres-Reyes, L.A., Alvarado-Ruiz, L., Pina-Sanchez, P., Martinez-Silva, M.G., Ramos-Solano, M., Olimon-Andalon, V., Ortiz-Lazareno, P.C., Hernandez-Flores, G., Bravo-Cuellar, A., Aguilar-Lemarroy, A., Jave-Suarez, L.F. (2014). Expression of transcription factor grainyhead-like 2 is diminished in cervical cancer. Int J Clin Exp Pathol 7, 7409-7418.

Vanharanta, S., Massague, J. (2013). Origins of metastatic traits. Cancer Cell 24, 410-421.

Varma, S., Cao, Y., Tagne, J.B., Lakshminarayanan, M., Li, J., Friedman, T.B., Morell, R.J., Warburton, D., Kotton, D.N., Ramirez, M.I. (2012). The transcription factors Grainyhead-like 2 and NK2-homeobox 1 form a regulatory loop that coordinates lung epithelial cell morphogenesis and differentiation. J Biol Chem 287, 37282-37295.

Verger, A., Perdomo, J., Crossley, M. (2003). Modification with SUMO. A role in transcriptional regulation. EMBO Rep 4, 137-142.

Vona, B., Nanda, I., Neuner, C., Muller, T., Haaf, T. (2013). Confirmation of GRHL2 as the gene for the DFNA28 locus. Am J Med Genet A 161a, 2060-2065.

Wang, C.M., Liu, R., Wang, L., Nascimento, L., Brennan, V.C., Yang, W.H. (2014). SUMOylation of FOXM1B alters its transcriptional activity on regulation of MiR-200 family and JNK1 in MCF7 human breast cancer cells. Int J Mol Sci 15, 10233-10251.

Wang, S., Samakovlis, C. (2012). Grainy head and its target genes in epithelial morphogenesis and wound healing. Curr Top Dev Biol 98, 35-63.

Werner, S., Frey, S., Riethdorf, S., Schulze, C., Alawi, M., Kling, L., Vafaizadeh, V., Sauter, G., Terracciano, L., Schumacher, U., Pantel, K., Assmann, V. (2013). Dual roles of the transcription factor grainyhead-like 2 (GRHL2) in breast cancer. J Biol Chem 288, 22993-23008.

Werner, S., Pantel, K. (2017). Tracing the Seeds in the Soil. Clin Chem 63, 1764-1765.

Werth, M., Walentin, K., Aue, A., Schonheit, J., Wuebken, A., Pode-Shakked, N., Vilianovitch, L., Erdmann, B., Dekel, B., Bader, M., Barasch, J., Rosenbauer, F., Luft, F.C., Schmidt-Ott, K.M. (2010). The transcription factor grainyhead-like 2 regulates the molecular composition of the epithelial apical junctional complex. Development 137, 3835-3845.

Wilanowski, T., Tuckfield, A., Cerruti, L., O'Connell, S., Saint, R., Parekh, V., Tao, J., Cunningham, J.M., Jane, S.M. (2002). A highly conserved novel family of mammalian developmental transcription factors related to Drosophila grainyhead. Mech Dev 114, 37-50.

Wilkinson, K.A., Henley, J.M. (2010). Mechanisms, regulation and consequences of protein SUMOylation. Biochem J 428, 133-145.

Xiang, J., Fu, X., Ran, W., Chen, X., Hang, Z., Mao, H., Wang, Z. (2013). Expression and role of grainyhead-like 2 in gastric cancer. Med Oncol 30, 714.

Xiang, J., Fu, X., Ran, W., Wang, Z. (2017). Grhl2 reduces invasion and migration through inhibition of TGFbeta-induced EMT in gastric cancer. Oncogenesis 6, e284.

Xiang, X., Deng, Z., Zhuang, X., Ju, S., Mu, J., Jiang, H., Zhang, L., Yan, J., Miller, D., Zhang, H.G. (2012). Grhl2 determines the epithelial phenotype of breast cancers and promotes tumor progression. PLoS One 7, e50781.

Yan, W., Cao, Q.J., Arenas, R.B., Bentley, B., Shao, R. (2010). GATA3 inhibits breast cancer metastasis through the reversal of epithelial-mesenchymal transition. J Biol Chem 285, 14042-14051.

Yang, C., Chen, H., Tan, G., Gao, W., Cheng, L., Jiang, X., Yu, L., Tan, Y. (2013). FOXM1 promotes the epithelial to mesenchymal transition by stimulating the transcription of Slug in human breast cancer. Cancer Lett 340, 104-112.

Yang, S.H., Jaffray, E., Hay, R.T., Sharrocks, A.D. (2003). Dynamic interplay of the SUMO and ERK pathways in regulating Elk-1 transcriptional activity. Mol Cell 12, 63-74.

Yeung, K.T., Yang, J. (2017). Epithelial-mesenchymal transition in tumor metastasis. Mol Oncol 11, 28-39.

Yilmaz, M., Christofori, G. (2009). EMT, the cytoskeleton, and cancer cell invasion. Cancer Metastasis Rev 28, 15-33.

Yilmaz, M., Christofori, G. (2010). Mechanisms of motility in metastasizing cells. Mol Cancer Res 8, 629-642.

Zhao, Q., Xie, Y., Zheng, Y., Jiang, S., Liu, W., Mu, W., Liu, Z., Zhao, Y., Xue, Y., Ren, J. (2014). GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Res 42, W325-330.

8. Appendix

Table 2 Frequently used equipment

Equipment	Manufacturer
Analytical Balance Sartorius BP610	MS Laborgeräte, Heidelberg
Analytical Balance Sartorius CP2245	MS Laborgeräte, Heidelberg
Biofuge pico Heraeus	Kendro, Langenselbold
BioPhotometer	Eppendorf AG, Hamburg
Biosafety Cabinet HeraSafe 150	Kendro, Langenselbold
Cast Stand Hoefer	Amersham Biosciences, Buckinghamshire (UK)
Centrifuge Rotofix 32	Hettich, Villingen-Schwenningen
CFX96 Touch Real-Time System	Bio-Rad, Munich
CO_2-Incubator Heracell 150	Kendro, Langenselbold
Curix 60 Film Processor	AGFA HealthCare GmbH, Bonn
Dri-Block Heater DB-2A	Techne, Staffordshire (UK)
Eppendorf Centrifuge 5415R	Eppendorf AG, Hamburg
Flexigene Thermocycler	Techne, Staffordshire (UK)
Gel Apparatus ComPhor L Mini/Midi	Biolastics RV, Landgraaf (NL)
GeneGenius 2 Documentation System	Syngene, Cambridge (UK)
Heraeus Function Line B12 Incubator	Kendro, Langenselbold
Horizontal Electrophoresis Apparatus	Bioplastics RV, Landgraaf (NL)
Inverted Microscope DM IL LED	Leica Mikrosysteme Vertrieb GmbH, Wetzlar
Multifuge Heraeus 3S-R	Kendro, Langenselbold
NanoDrop ND-1000	Fisher Scientific, Wilmington (USA)
Pipetus	Hirschmann Laborgeräte, Eberstadt
Powersupply E143	Consort, Turnhout (BE)
Powersupply E835	Consort, Turnhout (BE)
Refrigerated Centrifuge 5415R	Eppendorf AG, Hamburg
Roller Mixer Stuart bSRT1	Bibby Scientific, Staffordshire (UK)
SE250 Electrophoresis Vertical Unit	Amersham Biosciences, Buckinghamshire (UK)
Semi-dry Blotting Apparatus	Bio-Rad, Müchen
Shaker	GmbH für Labortechnik, Burgwedel
Single Channel Pipette 2,5 µl, 10 µl, 100 µl, 200 µl, 1000 µl	Eppendorf AG, Hamburg
Table-top Shaker Incubator Certomat R	B. Braun Biotech, Melsungen

Ultrasonic Processor	UP50H Hielscher, Teltow
Vortex-Genie 2	Scientific Industries, New York (USA)
Water Bath GFL-1003	GmbH für Labortechnik, Burgwedel

Table 3 Selected consumables

Name	Manufacturer
Centrifuge Tube 15 ml, 50 ml	Greiner Bio-One, Frickenhausen
Cryotube	Nunc, Rockford (USA)
FluoroTrans W PVDF membrane (0.2 µm pore size)	Pall Life Sciences, Dreieich
Hard-Shell PCR Plate (96-well, thin-wall)	BioRad, Munich
Nitrocellulose membrane (0.45 µm pore size)	GE Healthcare, Dassel
PCR Plate Sealing Film Microseal	Bio-Rad, Munich
Pipette Tips (10 µl, 200 µl, 1000 µl)	Sarstedt, Nümbrecht
Serological Pipette (2ml, 5ml, 10ml)	Sarstedt, Nümbrecht
Syringe 5 ml	Becton Dickinson, Heidelberg
Syringe Needle 26G	Becton Dickinson, Heidelberg
TC-Flask T25, T7	Sarstedt, Nümbrecht
UV Transparent Disposable Cuvette	Sarstedt, Nümbrecht
Whatman 3MM filter papers	GE Healthcare, Dassel
X-ray Film Super RX	Fujifilm, Willich

Table 4 Chemicals and Reagents

Name	Manufacturer
Ampicillin	Roth, Karlsruhe
Bacto-Agar	Becton Dickinson, Heidelberg
Bacto-Trypton	Becton Dickinson, Heidelberg
Big Dye (Applied Biosystems)	Thermo Fisher Scientific, Rockford (USA)
Bovine Serum Albumin (BSA)	Biomol, Hamburg
Bradford Reagent	BioRad, Munich
Bromophenol Blue	Sigma-Aldrich, Deisenhof
Dimethylsulfoxide (DMSO)	SERVA, Heidelberg
Ethanol	Merck, Darmstadt
Ethidiumbromide (EtBr)	Sigma-Aldrich, Deisenhof
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich, Deisenhof
Fetal Calf Serum (FCS)	Gibco, Eggenstein
Hydrogen Peroxide 30%	Merck, Darmstadt
Isopropanol	Merck, Darmstadt
LE Agarose	Genaxxon Bioscience, Ulm
Luminol	Sigma-Aldrich, Deisenhof
Methanol	J. T. Baker, Deventer (NE)
Ni–NTA agarose	QIAGEN, Hilden
Nonident-P-40	Roche Diagnostics, Basel
Phosphate Buffered Saline (PBS)	Gibco, Eggenstein
Polyethyleneimine (PEI, linear, ~MW 25000)	Polyscienes, Warrington (UK)
Powdered Milk	Roth, Karlsruhe
ProtoGel	National Diagnostics, Hessisch Oldendorf
P-Coumaric acid	Sigma-Aldrich, Deisenhof
Sodiumdodecylsulfate (SDS)	Roth, Karlsruhe
SYBRGreen	Thermo Fisher Scientific, Rockford (USA)
Tetramethylendiamin (TEMED)	Fluka, Buchs
Tween-20	Merck, Darmstadt
Yeast Extract	Becton Dickinson, Heidelberg

Table 5 Expression plasmids

Nr.	Code	Name	Primer up	Primer do	Section
1	VA-129	pCMX3b-FLAG-GRHL2-WT	1	2	4.1, 4.3
2	VA-130	pCMX3b-FLAG-GRHL2-K159R	1	2	4.3
3	VA-111	pCMX3b-FLAG-GRHL2-E161A	1	2	4.3
4	VA-110	pCMX3b-FLAG-GRHL2-K205R	1	2	4.3
5	VA-147	pCMX3b-FLAG-GRHL2-K366R	1	2	4.3
6	VA-148	pCMX3b-FLAG-GRHL2-K424R	1	2	4.3
7	VA-157	pCMX3b-FLAG-GRHL2-K453R	1	2	4.3
8	VA-131	pCMX3b-FLAG-GRHL2-K531R	1	2	4.3
9	VA-149	pCMX3b-FLAG-GRHL2-K556R	1	2	4.3
10	32	pSF-1-GRHL2-WT/SG-HAtag	3	4	4.4, 4.5, 4.6, 4.7
11	VA-210	pSF-1-GRHL2-K159R/SG-HAtag	3	4	4.4, 4.5, 4.7
12	VA-211	pSF-1-GRHL2-E161A/SG-HAtag	3	4	4.5, 4.7
13	36	pMXs-IP-GRHL2-WT/SG-HAtag	5	4	4.7
14	VA-190	pMXs-IP-GRHL2-K159R/SG-HAtag	5	4	4.7
15	VA-205	pMXs-IP-GRHL2-E161A/SG-HAtag	5	4	4.7
16	45	phCMV2-SUMO-1	6	7	4.1, 4.3, 4.4
17	46	phCMV2-SUMO-2	8	9	4.1, 4.3, 4.4
18	48	phCMV4-SUMO-1	6	7	4.3
19	49	phCMV4-SUMO-2	8	9	4.3
20	73	phCMV3-PIAS-1	10	11	4.4
21	70	phCMV3-PIAS-2	12	13	4.4
22	84	phCMV3-PIAS-3	14	15	4.4
23	72	phCMV3-PIAS-4	16	17	4.4
24	85	phCMV2-EGFP-PIAS-1	18	19	4.6
25	86	phCMV2-EGFP-PIAS-2	20	21	4.6
26	87	phCMV2-EGFP-PIAS-3	22	23	4.6
27	88	phCMV2-EGFP-PIAS-4	24	25	4.6

Table 6 Primers used for cDNA amplification

Nr.	Code	up/ do	Sequence	Enzyme	Plasmid
1	VA-GRHL2-77	up	GCGCATCGGATCCATTGGATCAAACAT GTCACAAGAGTCGG	BamH I	1-9
2	VA-GRHL2-78	do	GCGCATCAGATCTCTAGATTTCCATGAG CGTGACCTTGAAGC	Bgl II	1-9
3	GRHL2-11	up	GCGCATCGTAACGCGTGGATCAAACAT GTCACAAGAGTCG	Mlu I	10-12
4	GRHL2-12	do	GCGCTACAGCGGCCGCTTAAGCGTAAT CCGGAACATCGTATGGGTAGGATCCGA TTTCCATGAGCGTGACCTTGAAGC	Not I	10-12, 13-15
5	GRHL2-13	up	GCGCATCGTAAGATCTGGATCAAACATG TCACAAGAGTCG	Bgl II	13-15
6	SUMO1 up	up	GCGCATCAGATCTGTCTGACCAGGAGG CAAAACCTTC	Bgl II	16,18
7	SUMO1 do	do	GCGCATCGCGGCCGCTAACCCCCCGTT TGTTCCTGATAAAC	Not I	16,18
8	SUMO2 up	up	GCGCATCAGATCTGGCCGACGAAAAGC CCAAGGAAGG	Bgl II	17,19
9	SUMO2 do	do	GCGCATCGCGGCCGCTAACCTCCCGTC TGCTGTTGGAACAC	Not I	17,19
10	PIAS1 up3	up	GCGCATCAGATCTCGAAGTTCACTGCG CTTGCGCTG	Bgl II	20
11	PIAS1 do2	do	GCGCATCGAATTCGATCCGTCCAATGAA ATAATGTCTGG	EcoR I	20
12	PIAS2 up2	up	GCGCATCAGATCTGGTGGTATAAAATG GCGGATTTCGAAGAGTTGAGG	Bgl II	21
13	PIAS2 do2	do	GCGCATCGAATTCGATCCCTGTTGCACA GTATCAGAAGATGTTCC	EcoR I	21
14	PIAS3 up4	up	GCGCATCAGATCTTCGAGGCCACCATG GCGGAGCTGGGCGAATTAAAGC	Bgl II	22
15	PIAS3 do2	do	GCGCATCGAATTCGATCCGTCCAGGGA AATGATGTCTGACC	EcoR I	22
16	PIAS4 up2	up	GCGCATCAGATCTGCTGGTGACCAAGA TGGCGGCGGAGCTGGTGGAGG	Bgl II	23
17	PIAS4 do2	do	GCGCATCGAATTCGATCCGCAGGCCGG CACCAGGCCCTTCTGG	EcoR I	23
18	PIAS1 up4	up	GCGCATCGAATTCGCGGACAGTGCGGA ACTAAAGC	EcoR I	24

19	PIAS1 do	do	GCGCATCGCGGCCGCTCAGTCCAATGAAATAATGTCTGG	Not I	24
20	PIAS2 up3	up	GCGCATCGAATTCGCGGATTTCGAAGAGTTGAGG	EcoR I	25
21	PIAS2 do	do	GCGCATCGCGGCCGCTCACTGTTGCACAGTATCAGAAGATG	Not I	25
22	PIAS3 up3	up	GCGCATCGAATTCGCGGAGCTGGGCGAATTAAAGC	EcoR I	26
23	PIAS3 do	do	GCGCATCGCGGCCGCTCAGTCCAGGGAAATGATGTCTGACC	Not I	26
24	PIAS4 up3	up	GCGCATCGAATTCGCGGCGGAGCTGGTGGAGGCC	EcoR I	27
25	PIAS4 do	do	GCGCATCGCGGCCGCGCACCAGGAAAGTCGAGTGTGC	Not I	27

Table 7 Primers used for site-directed mutagenesis

Mutation	Code	up/do	Sequence
GRHL2 K159R	95 B (Lea)	up	GGAATCACGGTGGTGAGAGCTGAAGATTTCACACC
	96 (Lea)	do	GGTGTGAAATCTTCAGCTCTCACCACCGTGATTCC
GRHL2 E161A	VA-GRHL2-81	up	GAAAGCTGCAGATTTCACACCAGTTTTCATGGCCCCACCTG
	VA-GRHL2-82	do	TGAAATCTGCAGCTTTCACCACCGTGATTCCCGACACCGG
GRHL2 K205R	VA-GRHL2-79	up	CTATCTCAGAGACGACCAGCGCAGCACTCCGGACAGCAC
	VA-GRHL2-80	do	GGTCGTCTCTGAGATAGGCGCTGTGGGTGGCCAGCGAGG
GRHL2 K366R	VA-GRHL2-95	up	AAGAGGCGAGGATTTTCATCACCGTGAATTGCTTGAGCACAG
	VA-GRHL2-96	do	TGAAAATCCTCGCCTCTTCATTCACGTCCCAGGTAAAGGAAACAGC
GRHL2 K424R	VA-GRHL2-97	up	CAGAAAGAAGAATCCGAGATGAAGAGCGGAAGCAGAACAGG
	VA-GRHL2-98	do	TCGGATTCTTCTTTCTGCTCCTTTGTCACAGAAGACCTTG
GRHL2 K453R	VA-GRHL2-111	up	TGATGGGAGATTGGCTGCCATACCTTTACAGAAGAAGAGTGAC
	VA-GRHL2-112	do	CAGCCAATCTCCCATCAGAGGAGCTGTTGCATTGAGTTTGGGAG
GRHL2 K531R	97 (Lea)	up	CCTTCAAAGCAGATGAGAGAAGAAGGGACAAAGC
	98 (Lea)	do	GCTTTGTCCCTTCTTCTCTCATCTGCTTTGAAGG
GRHL2 K556R	VA-GRHL2-99	up	GATGTTGAGGTCTCCCACAGTGAAGGGCCTGATGGAAGCG
	VA-GRHL2-100	do	TGGGAGACCTCAACATCAATGCATCGAACACATCGTCAGTCTCC

Table 8 Primers used for DNA sequencing

Nr.	Code	up/do	Position	Sequence
1	GRHL2-Seq0	up	345-367	GGACAATAATAAAAGACTAGTGG
2	GRHL2-Seq1	up	837-855	CCCACCTGTGCACTATCCC
3	GRHL2-Seq2	do	1685-1702	GGTATGGCAGCCAACTTCC
4	GRHL2-Seq3	up	1685-1702	GGAAGTTGGCTGCCATACC

Table 9 Primers used for quantitative real-time PCR

Nr.	Code	up/do	Sequence
1	RPLP0-f	up	ACCCAGCTCTGGAGAAACTGC
2	RPLP0-r	do	TGAGGTCCTCCTTGGTGAACA
3	GRHL2-f-exo	up	CATGCCTGATCTCCACTCACAG
4	GRHL2-r-exo	do	CTGCCACCTTCTCGTTCATCA
5	CD24-f	up	CACTGCTCCTACCCACGCAGAT
6	CD24-r	do	CTTGGTGGTGGCATTAGTTGG
7	E-CADHERIN-f	up	CAGGAACCTCTGTGATGGAG
8	E-CADHERIN-r	do	CACTGATGACTCCTGTGTTCCTG

Table 10 Primary antibodies

Antibody	M* / P*	Species	Dilution	Manufacturer
GRHL2	P	Rabbit	1:1000	Werner et.al. (Werner et al., 2013)
Hemagglutinin (HA)	P	Rabbit	1:2000	Sigma-Aldrich, Munich
Tublin	P	Rabbit	1:4000	Cell Signaling, Frankfort
SUMO1/2	P	Rabbit	1:1000	Cell Signaling, Frankfort

M* Monoclonal, P* Polyclonal

Table 11 Abbreviations

A	Alanine
Abb.	Abbreviation
APS	Ammonium persulfate
BSA	Bovine serum albumin
cDNA	Complementary DNA
D	Aspartic acid
DBD	DNA binding domain
DD	Dimerization domain
DMEM	Dulbecco's modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
E	Glutamic acid
E-cadherin	Epithelial cadherin
E.coli	Escherichia coli
ECL	Enhanced chemiluminescence
ECM	Extracellular matrix
EGF	Epidermal growth factor
EMT	Epithelial-Mesenchymal transition
EMT-TFs	EMT-inducing transcriptional factors
F	Phenylalanine
FCS	Fetal Calf Serum
GRHL	Grainyhead-like
h	hour
HMLER	Human immortalized mammary epithelial cells
HRP	Horseradish peroxidase
hTERT	Telomerase reverse transcriptase
K	Lysine
kDa	Kilo-Dalton
LB-Medium	Lysogeny Broth Medium
MBC	Metastatic breast cancer
MET	Mesenchymal-Epithelial transition
min	Minute

MRD	Minimal residual disease
mRNA	Messenger ribonucleic acid
NEM	N-Ethylmaleimide
OD	Optical density
P	Proline
PAGE	Polyacrylamide gel electrophoresis
PAR	Partitioning-defective
PBS	Phosphate Buffered Saline
PCR	Polymerase chain reaction
PIAS	Protein inhibitor of STAT
PTMs	Post-translational modifications
PVDF	Polyvinylidene fluoride
R	Arginine
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Rounds per Minute
RT	Room temperature
RT-PCR	Real-time polymerase chain reaction
S	Serine
SDS	Sodiumdodecylsulfate
sec	Second
SENP	SUMO-specific proteases
SIMs	SUMO interacting motif
SUMO	Small Ubiquitin-like Modifier
T	Threonine
TAD	Transactivation domain
TBS	Tris-buffered saline
TEMED	Tetramethylethylenediamine
TGF-β	Transforming Growth Factor-β
V	Volt
V	Valine
v/v	Volume per Volume
w/v	Mass per Volume

9. Acknowledgement

First of all, I would like to express my deepest appreciation to Prof. Dr. Klaus Pantel for inviting me to join his excellent research team and for providing me with the opportunity to do my first steps in experimental metastasis research. During the time at the Institute of Tumor Biology, I obtained completely new insights into the biology of cancer metastasis which, I am sure, will tremendously influence my future research activities.

I also would like to thank my enthusiastic supervisor, Dr. Volker Assmann, who has walked me through all stages of my medical doctoral career. He provided me with unlimited academic support and helped me a great deal to grow and mature into a better and more independent medical scientist. Special thanks to his scientific guidance, personal help, countless encouragement, and great efforts in correcting this dissertation.

Furthermore, I am most appreciatively to my lab mates and friends Johanna Keller and Sonja Santjer and also to Dr. Sabine Riethdorf not only for their continuous support in the lab, but also for their valuable help in my personal life.

I'm also very grateful to all members of the Institute of Tumor Biology for their support, friendship and a fabulous working environment throughout the past two years. You have all made my time in the lab a truly wonderful and memorable experience.

Thanks should also go to my family for their support, encouragement and understanding throughout the years. Undertaking a medical doctoral career has been a challenging experience in my life. You have always been there for me and I certainly wouldn't have made it without all your help.

Last but not least, a great amount of thanks to the China Scholarship Council. With the scholarship, I was able to conduct my medical doctoral research project in Germany without any financial worry or anxiety. This thesis is dedicated to you all.

10. Curriculum Vitae

Personal Information

Name: Yang, Xu

Date of Birth: 19/07/1989

Gender: Female

Nationality: Chinese

Language: Chinese, English

E-mail: xxxxuyang@126.com

Phone: 0176 4327 3969

Educational Background

2017- Present	Doktor der Medizin (Dr. Med.)
	Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany
2013 – 2016	Master of Clinical Medicine
	Harbin Medical University, Heilongjiang, China
2008 – 2013	Bachelor of Clinical Medicine
	Harbin Medical University, Heilongjiang, China

Awards and Honors

09/2015	Third Prize in the innovation competition in Heilongjiang, China
02/2015	Second Prize in Medical Health and Science & Technology in Heilongjiang, China
12/2014	Occupational physician credential of People's Republic of China
10/2010	Medical Scholarship of Daiichi-Sankyo Corporation, Japan

11. Eidesstattliche Erklärung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.
Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.
Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift: .....